Facultad de Ciencias Exactas, Ingeniería y Agrimensura

Universidad Nacional de Rosario

ÁLGEBRA LINEAL

Adriana Mateu - José Semitiel

Departamento de Matemática - Escuela de Formación Básica

MATERIAL DIDÁCTICO DE USO INTERNO - 2024

Presentación

El Álgebra Lineal es una rama de la Matemática requerida por diferentes áreas de la Ingeniería, que facilita la modelización y simplifica los cálculos, en el estudio del comportamiento de algunos sistemas complejos. El principal objetivo de un curso de esta disciplina, en carreras de ingeniería, es brindar a los estudiantes los conocimientos básicos que aplicarán posteriormente en las diferentes especialidades.

En el Álgebra Lineal se conjugan dos facetas fundamentales de la Matemática: la abstracción y la aplicación. En nuestro curso el énfasis está puesto en la comprensión de los conceptos, mientras que las aplicaciones se dejan prioritariamente para el ciclo superior, donde los conceptos son utilizados desde diferentes perspectivas, atendiendo a las necesidades de cada especialidad.

Este texto contiene el material mínimo correspondiente a la asignatura *Álgebra Lineal* de las carreras de Ingeniería (Plan 2014) de la Facultad de Ciencias Exactas, Ingeniería y Agrimensura de la Universidad Nacional de Rosario. Fue escrito principalmente con el objetivo de ordenar, focalizar y hacer más accesible, en un mismo texto, los temas principales de esta disciplina correspondientes al programa vigente de la asignatura. Cabe aclarar que éste no es un trabajo finalizado y que se encuentra en proceso revisión. Debe considerarse como borrador y material de uso interno para la cátedra.

Este material de estudio está organizado en siete capítulos. Los capítulos 1, 2 y 3 corresponden a sistemas de ecuaciones lineales y matrices que son temas fundamentales para el Álgebra Lineal a partir de los cuales se desarrolla este curso. Los capítulos 4, 5 , 6 y 7 constituyen el corazón de la asignatura: espacios vectoriales, espacios con producto interno transformaciones lineales, autovalores y autovectores y diagonalización.

Cada capítulo de este texto se divide en diferentes secciones. Al final de cada una de éstas se presenta la ejercitación correspondiente y se indican recuadrados los ejercicios y problemas que consideramos una práctica mínima para la comprensión de los diferentes temas.

Adriana Mateu y *José Semitiel*

Índice general

Capítulo 1

SISTEMAS DE ECUACIONES LINEALES

1.1. Ecuaciones lineales con n incógnitas

Definición (Ecuación lineal). Una *ecuación lineal* en las n variables o incógnitas x_1, x_2,..., x_n es una ecuación que puede escribirse en la forma

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b, \tag{1.1}$$

llamada forma canónica o estándar de la ecuación lineal, donde a_1,a_2,...,a_n y b son constantes reales. Los a_i con $i = 1, 2, ..., n$ son los *coeficientes* de la ecuación y b es el *término constante.*

Si $b = 0$ la ecuación se denomina *homogénea.*

Considerando las variables ordenadas de izquierda a derecha, la primer variable cuyo coeficiente es distinto de cero, recibe el nombre de *variable delantera o principal.* Las demás variables se llaman *variables libres.*

Ejemplo 1.1. La ecuación

$$1 + 2x_1 + x_2 - x_3 + 3x_4 = 2x_1 + 2 + x_3,$$

es lineal no homogonéa en cuatro variables x_1, x_2, x_3 y x_4. Su forma canónica o estándar es:

$$0x_1 + x_2 - 2x_3 + 3x_4 = 1.$$

Los coeficientes (en orden) son 0, 1, -2 y 3 y el término constante es 1. La variable delantera es x_2 y x_1, x_3 y x_4 son las variables libres.◀

Ejemplo 1.2. Las ecuaciones

$$2xy + 2x = 3 \quad , \quad x^2 - y = 1 \quad , \quad \sqrt{x} + y - 3 = 0,$$

no son lineales.◀

Ejemplo 1.3. La ecuación

$$2x + 3y - z = 0,$$

es lineal homogénea en las variables x, y, z. Los coeficientes (en orden) son 2, 3 y -1 y el término constante es 0. La variable delantera es x y las variables libres son y, z. Notemos que esta es la ecuación de un plano que contiene al origen de coordenadas.

En general, toda ecuación lineal en tres variables

$$ax + by + cz = d,$$

con al menos un coeficiente no nulo, es la ecuación de un **plano en el espacio**.◀

Ejemplo 1.4. Toda ecuación lineal en dos variables

$$ax + by = c,$$

con al menos uno de los coeficientes no nulos, es la ecuación de una **recta en el plano**.◀

Definición (Solución de una ecuación lineal). Una *solución (particular)* de la ecuación lineal

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b,$$

es una $n-$ upla odenada de números reales $(c_1, c_2, ..., c_n)$ tales que cuando se sustituye x_i por c_i para todo $i = 1, 2, ..., n$ la ecuación se convierte en una identidad.

El conjunto de todas las soluciones de la ecuación lineal se llama *conjunto solución.*

Resolver una ecuación lineal es hallar su conjunto solución. Para obtenerlo se despeja la variable delantera en función de las variables libres, haciendo que cada variable libre asuma cualquier valor real. Con esto se llega a un elemento genérico del conjunto solución que es la *solución general* de la ecuación.

Ejemplo 1.5. Una solución de la ecuación lineal del Ejemplo 1.1

$$0x_1 + x_2 - 2x_3 + 3x_4 = 1,$$

es $(1, 5, 2, 0)$. Es más, cualquier $4-$upla ordenada de la forma $(a, 5, 2, 0)$ es solución.

Para obtener todas las soluciones, despejando x_2 de la ecuación se obtiene

$$x_2 = 1 + 0x_1 + 2x_3 - 3x_4.$$

Las variables libres pueden tomar cualquier valor real, así podemos considerar $x_1 = t$, $x_3 = s$ y $x_4 = r$. Por consiguiente la solución general se expresa como:

$$x_1 = t, \quad x_2 = 1 + 0t + 2s - 3r = 1 + 2s - 3r, \quad x_3 = s \quad \text{y} \quad x_4 = r, \quad \text{para todo} \quad t, s, r \in \mathbb{R},$$

o en forma de 4-upla

$$(t, 1 + 2s - 3r, s, r)$$

para todo $t, s, r \in \mathbb{R}$. Las letras t, s y r reciben el nombre de **parámetros**.

El conjunto solución es

$$\{(t, 1 + 2s - 3r, s, r) : t, s, r \in \mathbb{R}\}. \blacktriangleleft$$

Actividad 1.1. Obtenga el conjunto solución de la ecuación lineal

$$-3x_1 + 2x_2 - x_3 = -x_1 + x_4 - 1.$$

Definición (Ecuaciones consistentes). Una ecuación lineal se dice *consistente* si tiene al menos una solución y es *inconsistente* si no tiene solución.

La ecuación del ejemplo anterior es consistente.

Ejemplo 1.6. Ecuación lineal homogénea. Una ecuación lineal homogénea

$$a_1x_1 + a_2x_2 + ... + a_nx_n = 0,$$

siempre es consistente, tiene al menos la llamada **solución trivial** $x_1 = x_2 = ... = x_n = 0.\blacktriangleleft$

Ejemplo 1.7. Sea la ecuación lineal homogénea $2x + 3y = 0$. La solución general es $x = -\frac{3}{2}t$, $y = t$ con $t \in \mathbb{R}$. Su conjunto solución es

$$\left\{\left(-\frac{3}{2}t, t\right) : t \in \mathbb{R}\right\}. \blacktriangleleft$$

Ejemplo 1.8. La ecuación lineal $\frac{7}{3}x_1+2x_2-5 = \frac{7}{3}x_1+2x_2-5$ se satisface para cualquier par ordenado de números reales. Su forma canónica es

$$0x_1 + 0x_2 = 0,$$

y su conjunto solución es $\{(x_1, x_2) : x_1, x_2 \in \mathbb{R}\}.\blacktriangleleft$

Ejemplo 1.9. Claramente, la ecuación $x_1 + 2x_2 - 5x_3 = 8 + x_1 + 2x_2 - 5x_3$ es inconsistente. Su forma canónica es

$$0x_1 + 0x_2 + 0x_3 = 8.$$

Ninguna terna de números reales la satisface y por lo tanto su conjunto solución es vacío.$\blacktriangleleft$

En la ecuación del ejemplo anterior, se observa que todos los coeficientes son iguales a cero mientras que el término constante es distinto de cero. Se verá que ésta es la forma que tienen las ecuaciones lineales inconsistentes.

Teorema 1.1 (Inconsistencia de una ecuación lineal)**.** Una ecuación lineal en n variables

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b,$$

es inconsistente si y sólo si $b \neq 0$ y $a_1 = a_2 = ... = a_n = 0$.

Demostración. En efecto: $0x_1 + 0x_2 + ... + 0x_n = b$ con $b \neq 0$ es inconsistente.

Recíprocamente:

- Si $b = 0$ la ecuación es homogénea y por lo tanto es consistente cualesquiera sean los coeficientes.
- Si algún coeficiente es distinto de cero, la ecuación es consistente cualquiera sea b puesto que habrá una variable delantera que podrá "despejarse".

□

Nota: Una ecuación lineal en una variable $ax = b$ tiene una única solución, infinitas soluciones o es inconsistente. Si la ecuación lineal tiene dos o más variables, es inconsistente o bien tiene un número infinito de soluciones.

Actividad 1.2. Encuentre para qué valores de a y b, la ecuación lineal en una variable $ax = b$, tiene una única solución, infinitas soluciones o ninguna solución.

Ejercicios 1.1

1. Para cada una de las siguientes ecuaciones:

   a) $2x - 5 = x - 3y + z$

   b) $xy - 4x + y = 0$

   c) $x_1 - 2x_2 + 3 = x_2 + 5x_3 - x_4 + 3$

   d) $x_1 + x_2 - x_3 = x_1 + x_2 - x_3 + 9$

   determine si es o no lineal. En caso de ser lineal, indique: si es homogénea, sus coeficientes, su término constante, la variable delantera, las variables libres y obtenga su conjunto solución.

2. Determine todos los valores de a de modo que cada una de las ecuaciones: (i) tenga única solución, (ii) tenga infinitas soluciones, (iii) no tenga solución.

   a) $a^2x - 4 = 16x + a$

   b) $ax_1 - a^2x_2 = 2a$

   c) $a^2(x + y) - x - y = a - 1$

   d) $a(x + y) - x - y - a + 1 = 0$

1.2. Sistemas de ecuaciones lineales

Muchos problemas en Matemática, Ingeniería, Economía, Física y otras ciencias se reducen a resolver sistemas de ecuaciones lineales. Un ejemplo sencillo de un sistema lineal se presenta en el siguiente problema:

Un día una tienda vendió 45 *camisetas. Las lisas tenían un precio de venta de* \$110 *mientras que las estampadas* \$135. *Ese día se recaudó un total de* \$5575 *en ventas de camisetas. ¿Cuántas camisetas estampadas y cuántas lisas se vendieron?*

Llamando x al número de camisetas lisas y y al número de camisetas estampadas, el problema consiste en hallar los valores de x y y que satisfacen simultáneamente las ecuaciones lineales:

$$\begin{aligned} x+y &= 45 \\ 110x+135y &= 5575 \end{aligned}$$

El conjunto formado por estas dos ecuaciones lineales recibe el nombre de s*sistema lineal.* Más precisamente:

Definición (Sistema lineal). Un *sistema lineal* de m ecuaciones con n incógnitas o variables x_1, x_2,..., x_n es un conjunto de m ecuaciones lineales en dichas incógnitas. Su forma canónica o estándar es:

$$\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ... + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + ...a_{mn}x_n = b_m \end{array}\right. \tag{1.2}$$

Los números a_{ij} con $i = 1, 2, .., m$ y $j = 1, 2, ..., n$ son los *coeficientes* del sistema y b_i con $i = 1, 2, ..., m$ son los *términos constantes.*

Si $b_i = 0$ para todo $i = 1, 2, ..., m$ el sistema lineal se denomina *homogéneo.*

Una única ecuación lineal es un sistema lineal ($m = 1$).

Ejemplo 1.10. El sistema

$$\left\{\begin{array}{l} x_1 - 2x_2 + x_3 = x_4 + 2 \\ 2x_1 = x_3 - 2x_4 \\ -x_2 + x_3 - x_4 = 1 \end{array}\right.$$

es lineal no homogéno de 3 ecuaciones con 4 incógnitas x_1, x_2, x_3 y x_4. Su forma canónica o estándar es:

$$\left\{\begin{array}{ccccccc} x_1 & -2x_2 & +x_3 & -x_4 & = & 2 \\ 2x_1 & +0x_2 & -x_3 & +2x_4 & = & 0 \\ 0x_1 & -x_2 & +x_3 & -x_4 & = & 1 \end{array}\right.$$

aunque suelen omitirse los términos con coeficientes cero escribiéndose:

$$\left\{\begin{array}{ccccccc} x_1 & -2x_2 & +x_3 & -x_4 & = & 2 \\ 2x_1 & & -x_3 & +2x_4 & = & 0 \\ & -x_2 & +x_3 & -x_4 & = & 1 \end{array}\right.$$

Los coeficientes son: $a_{11} = 1$, $a_{12} = -2$, $a_{13} = 1$, $a_{14} = -1$, $a_{21} = 2$, $a_{22} = 0$, $a_{23} = -1$, $a_{24} = 2$, $a_{31} = 0$, $a_{32} = -1$, $a_{33} = 1$ y $a_{34} = -1$. Los términos constantes son $b_1 = 2$, $b_2 = 0$ y $b_3 = 1$.◀

Definición (Solución de un sistema lineal). Una *solución (particular)* del sistema lineal (1.2) es una $n-$upla ordenada de números reales $(c_1, c_2, ..., c_n)$ que es solución de todas las ecuaciones del sistema. El conjunto de todas las soluciones se llama *conjunto solución.*

Resolver un sistema lineal es hallar su conjunto solución. Un elemento genérico del conjunto solución es una *solución general del sistema.*

Ejemplo 1.11. La 4-upla $(-3,-4,0,3)$ es una solución del sistema del Ejemplo 1.10. En efecto,

$$\begin{cases} -3-2(-4)+0-3 &= 2 \\ 2(-3)+0+0+2(3) &= 0 \\ 0-(-4)+0-3 &= 1 \end{cases}$$

Entonces dicho sistema tiene al menos una solución, ¿será la única solución o admitirá más soluciones?. Para cualquier $t \in \mathbb{R}$, la 4-upla $(-t,-1-t,0,t)$ es una solución del sistema (verifíquelo) y por lo tanto este tiene infinitas soluciones. Puede observarse que si $t=3$ se obtiene la solución particular previamente considerada.◀

Definición (Sistemas consistentes). Un sistema lineal se dice *consistente* si tiene al menos una solución e *inconsistente* si no tiene solución.

Nota: Se puede probar que un sistema lineal puede tener una, infinitas o ninguna solución. En la sección 1.4 se profundizarán estas ideas.

Ejemplo 1.12. Sistemas lineales de dos ecuaciones con dos incógnitas.

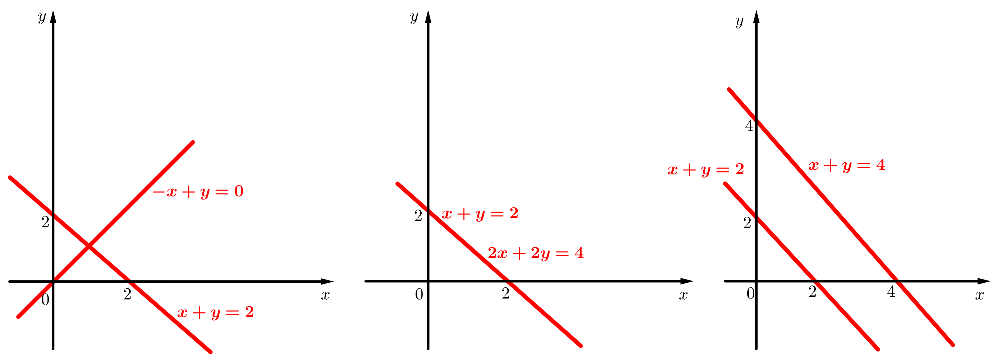


Figura 1.1: Resolución gráfica

Los sistemas lineales:

$$\begin{cases} x+y=2 \\ -x+y=0 \end{cases} \quad \begin{cases} x+y=2 \\ 2x+2y=4 \end{cases} \quad \text{y} \quad \begin{cases} x+y=2 \\ x+y=4 \end{cases}$$

son: consistente con única solución, consistente con infinitas soluciones e inconsistente, respectivamente. En efecto, si se recurre a la Geometría Analítica, gráficamente el conjunto solución de cada uno de estos sistemas es el conjunto de puntos del plano que pertenecen a ambas rectas como se observa en la Figura 1.1◀

Ejemplo 1.13. Sistema lineal homogéneo. Un sistema lineal homogéneo

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ... + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + ...a_{mn}x_n = 0 \end{cases}$$

siempre es consistente. En efecto: $x_1 = x_2 = ... = x_n = 0$ siempre es una solución, es la llamada **solución trivial** o **solución cero**. Cualquier otra solución es una **solución no trivial**.

Por ejemplo, $x = 1$, $y = 1$ es una solución no trivial del sistema homogéneo

$$\begin{cases} x - y = 0 \\ -2x + 2y = 0 \end{cases} .$$

y $x = 0$, $y = 0$ es la solución trivial del mismo.◀

Resolviendo un sistema de ecuaciones lineales

Claramente un sistema lineal que tiene una ecuación sin solución $(0x_1 + 0x_2 + ... + 0x_n = b,\ b \neq 0)$ es un sistema inconsistente. En estos casos se detecta a simple vista la inconsistencia del sistema, pero ésta no es la forma usual en la que se presenta un sistema inconsistente. Por lo general todas las ecuaciones que componen un sistema lineal tienen solución y por lo tanto no se detecta rápidamente si el sistema es o no consistente.

Hay sistemas de ecuaciones lineales en los que, por su forma, se hace evidente la consistencia o inconsistencia de los mismos y que, si son consistentes, son simples de resolver. Éstos son los llamados **sistemas escalonados** o **triangulares**. En ellos la variable delantera de cada ecuación se presenta más a la derecha de la variable delantera de la ecuación anterior (inmediata superior).

En un sistema *escalonado* se llama **variable delantera del sistema** a una variable que es delantera en alguna ecuación. Es decir, son variables delanteras aquellas que son las primeras cuyos coeficientes son distintos de cero. Las restantes variables se llaman **variables libres**.

Por ejemplo, el sistema lineal

$$\begin{cases} x_1 & -2x_2 & +x_3 & -x_4 & = & 2 \\ & 4x_2 & -3x_3 & +4x_4 & = & -4 \\ & & \frac{1}{4}x_3 & & = & 0 \end{cases} \tag{1.3}$$

es escalonado, sus variables delanteras son x_1, x_2 y x_3, la variable x_4 es libre.

El sistema

$$\begin{cases} 2x_1 & +x_2 & -5x_3 & +x_4 & -2x_5 & = & 2 \\ & & 3x_3 & & +6x_5 & = & 12 \end{cases}$$

es triangular. Sus variables delanteras son x_1 y x_3, mientras que x_2, x_4 y x_5 son las variables libres.

Un sistema escalonado que no tiene ecuaciones inconsistentes es siempre un sistema consistente como lo muestra el método llamado **método de sustitución hacia atrás** que conduce a la solución general de un sistema escalonado.

Este método de resolución comienza en la última ecuación y avanza hacia arriba. Se comienza despejando la variable delantera de la última ecuación en término de las libres (si las hay) y las constantes. Luego se sustituye la expresión encontrada para dicha variable en la ecuación anterior y se despeja la variable delantera de dicha ecuación. Así, continuando de este modo, hasta despejar la variable delantera de la primera ecuación.

Ejemplo 1.14. En el sistema escalonado (1.3), de la última ecuación se observa que necesariamente

$$x_3 = 0.$$

Sustituyendo este valor en la ecuación anterior se obtiene $4x_2+4x_4 = -4$ por lo cual necesariamente es $x_2 = -1 - x_4$ pudiendo tomar x_4 cualquier valor real. Así

$$x_4 = t,\ t \in \mathbb{R},$$

y

$$x_2 = -1 - t.$$

Reemplazando en la primera ecuación $x_2 = -1-t$, $x_3 = 0$ y $x_4 = t$, se tiene que $x_1-2(-1-t)-t = 2$ es decir

$$x_1 = -t.$$

Por consiguiente la solución general del sistema lineal (1.3) se expresa como:

$$x_1 = -t, \quad x_2 = -1 - t, \quad x_3 = 0 \quad \text{y} \quad x_4 = t \quad \text{para todo } t \in \mathbb{R},$$

o en forma de 4-upla

$$(-t, -1 - t, 0, t) \quad \text{para todo } t \in \mathbb{R},$$

por lo cual el sistema es consistente con infinitas soluciones.

El conjunto solución

$$\{(-t, -1 - t, 0, t) : t \in \mathbb{R}\},$$

es uniparamétrico infinito. ◄

Actividad 1.3. Resuelva el sistema lineal escalonado: $\left\{\begin{aligned} 2x_1 - x_2 + x_3 &= 0 \\ 3x_2 - x_3 &= 1 \\ 2x_3 &= 4 \end{aligned}\right.$.

El sistema lineal del Ejemplo 1.10 no es escalonado. Si bien es posible resolverlo haciendo apropiadas sustituciones, es más sencilla su resolución hallando un sistema escalonado que sea *equivalente* a él.

Definición (Sistemas equivalentes). Dos sistemas lineales son *equivalentes* si tienen el mismo conjunto solución.

Un sistema lineal puede transformarse en un sistema equivalente mediante ciertas *operaciones entre ecuaciones*:

Operaciones elementales en ecuaciones

Las operaciones elementales en ecuaciones de un sistema lineal son:

- Intercambio. Intercambia ecuaciones: $E_i \leftrightarrow E_j$.
- Escalamiento. Multiplica una ecuación por una constante $c \neq 0$: $cE_i \to E_i$.
- Eliminación. Suma a una ecuación un múltiplo escalar de otra ecuación: $E_i + cE_j \to E_i$.

Teorema 1.2 (Operaciones que conducen a sistemas equivalentes)**.** La aplicación de cualquier operación elemental en ecuaciones de un sistema lineal da como resultado un sistema equivalente.

Demostración. Claramente si se intercambian dos ecuaciones de un sistema no se alteran las soluciones del mismo.

También, si se multiplica una ecuación por una constante distinta de cero, se obtiene una ecuación equivalente y por lo tanto no se alteran las soluciones del sistema.

Se deja como ejercicio probar que una operación de eliminación no cambia las soluciones del sistema (Ejercicio 4).

□

Una apropiada aplicación de estas operaciones permite encontrar un sistema escalonado equivalente a uno dado. En la sección siguiente se justificará la existencia de un sistema lineal escalonado equivalente a uno dado.

Ejemplo 1.15. Para resolver el sistema lineal (no escalonado) del Ejemplo 1.10 se usará la técnica de eliminación para transformarlo en uno escalonado equivalente. Para ello, se eliminan algunas de las incógnitas sumando un múltiplo de una ecuación a otra ecuación.

$$\left\{\begin{array}{ccccccl} x_1 & -2x_2 & +x_3 & -x_4 & = & 2 & (E_1) \\ 2x_1 & & -x_3 & +2x_4 & = & 0 & (E_2) \\ & -x_2 & +x_3 & -x_4 & = & 1 & (E_3) \end{array}\right.$$

Si se sustituye la ecuación E_2 por E_2 menos dos veces la ecuación E_1 $(E_2 + (-2)E_1 \to E_2)$ se elimina x_1 de la segunda ecuación, quedando el sistema equivalente:

$$\left\{\begin{array}{ccccccl} x_1 & -2x_2 & +x_3 & -x_4 & = & 2 & (E_1) \\ & 4x_2 & -3x_3 & +4x_4 & = & -4 & (E_2) \\ & -x_2 & +x_3 & -x_4 & = & 1 & (E_3) \end{array}\right.$$

Para eliminar la variable x_2 de la tercera ecuación, se sustituye la ecuación E_3 por E_3 más $\frac{1}{4}$ de la ecuación E_2 $\left(E_3 + \frac{1}{4}E_2 \to E_3\right)$. Se obtiene así el sistema escalonado equivalente:

$$\left\{\begin{array}{cccccc} x_1 & -2x_2 & +x_3 & -x_4 & = & 2 \\ & 4x_2 & -3x_3 & +4x_4 & = & -4 \\ & & \frac{1}{4}x_3 & & = & 0 \end{array}\right.$$

que fue resuelto en el Ejemplo 1.14 y su conjunto solución es:

$$\{(-t, -1-t, 0, t) : t \in \mathbb{R}\} . \blacktriangleleft$$

Ejemplo 1.16. Un sistema inconsistente. El sistema lineal

$$\left\{\begin{array}{ccccc} 2x & +3y & +z & = & 2 \\ 8x & +4y & +3z & = & 4 \\ 12x & +10y & +5z & = & 12 \end{array}\right.$$

es inconsistente. En efecto, llevándolo a uno escalonado equivalente:

$$\left\{\begin{array}{rrrcrl} 2x & +3y & +z & = & 2 & (E_1)\\ 8x & +4y & +3z & = & 4 & (E_2)\\ 12x & +10y & +5z & = & 12 & (E_3)\end{array}\right. \xrightarrow[\substack{E_2-4E_1\to E_2\\ E_3-6E_1\to E_3}]{} \left\{\begin{array}{rrrcrl} 2x & +\ 3y & +\ z & = & 2 & (E_1)\\ & -\ 8y & -\ z & = & -4 & (E_2)\\ & -\ 8y & -\ z & = & 0 & (E_3)\end{array}\right. \to$$

$$\xrightarrow[E_3-E_2\to E_3]{} \left\{\begin{array}{rrcrl} 2x\ +3y & +z & = & 2 & (E_1)\\ -8y & -z & = & -8 & (E_2)\\ & 0z & = & 4 & (E_3)\end{array}\right.$$

Observemos que todas las ecuaciones, tanto las del sistema original como las del que le sigue, son consistentes pero el sistema no lo es; con algo de atención, en el segundo sistema se ve la inconsistencia ya que las ecuaciones E_2 y E_3 no tienen una solución común. La inconsistencia es evidente en el último sistema equivalente (que es escalonado) pues su última ecuación no tiene solución.◀

Como ya se ha visto, un sistema lineal escalonado es consistente si y solo si todas sus ecuaciones son consistentes, es decir, si y solo si no tiene una ecuación de la forma

$$0x_1 + 0x_2 + ... + 0x_n = b, \quad b \neq 0.$$

Con respecto a la consistencia de los sistemas lineales en general, se tiene el siguiente resultado:

Teorema 1.3. Un sistema lineal es consistente si y solo cualquier sistema equivalente en forma escalonada no tiene una ecuación de la forma

$$0x_1 + 0x_2 + ... + 0x_n = b, \quad b \neq 0.$$

Ejercicios 1.2

1. Escriba el sistema lineal

$$\left\{\begin{array}{c} -3x_1 + x_2 = -x_3 + x_4 + 3\\ 2x_1 - x_4 + 9 = 0\\ x_2 + x_3 = x_4 - 3 - x_2\\ 2x_1 - x_2 + 5 = x_3 + x_4 - 1\end{array}\right.$$

en forma canónica (estándar) y verifique que la 4−upla $(-3, 3, -6, 3)$ es una solución.

2. Obtenga los valores de a y b de manera que la terna $(1, 1, 1)$ sea una solución del sistema lineal

$$\left\{\begin{array}{rcl} ax - y + 2z & = & b\\ x + by + z & = & 3a\end{array}\right. .$$

3. Considere el sistema homogéneo $\left\{\begin{array}{rcl} a_{11}x + a_{12}y & = & 0\\ a_{21}x + a_{22}y & = & 0\end{array}\right.$

   a) Demuestre que si (x_0, y_0) es una solución para tal sistema entonces también lo es (kx_0, ky_0) para cualquier k constante.

b) Demuestre que si (x_0, y_0) y (x_1, y_1) son soluciones del sistema lineal homogéneo entonces $(x_0 + x_1, y_0 + y_1)$ también es solución.

4. Demuestre que el sistema lineal que se obtiene al reemplazar una ecuación por ella misma más un múltiplo escalar de otra tiene las mismas soluciones que el sistema original (Teorema 1.2).

5. Resuelva el problema planteado al inicio de esta sección.

$\boxed{6.}$ Resuelva los siguientes sistemas:

a) $\begin{cases} -x + 2y + z = 1 \\ y - 3z = 0 \\ -2z = 16 \end{cases}$

b) $\begin{cases} x_1 - 3x_2 + 5x_3 + x_4 - 1 = 0 \\ 2x_2 = x_3 + 9 \\ -x_4 = 0 \end{cases}$

c) $\begin{cases} x + 3y - z = 0 \\ 2y + 9z = 0 \\ 3z = 0 \end{cases}$

d) $\begin{cases} x_1 - x_2 + x_3 - x_4 - 5 = 0 \\ x_3 = 8 - x_4 \end{cases}$

$\boxed{7.}$ Demuestre que los sistemas lineales

$$\left\{\begin{array}{ccccccc} x_1 & +2x_2 & -x_3 & +x_4 & = & 0 \\ 2x_1 & -x_2 & +x_3 & -x_4 & = & 1 \\ & 2x_2 & +x_3 & +x_4 & = & 2 \end{array}\right. \quad \text{y} \quad \left\{\begin{array}{ccccccc} -x_1 & -2x_2 & +x_3 & -x_4 & = & 0 \\ & x_2 & +\frac{1}{2}x_3 & +\frac{1}{2}x_4 & = & 1 \\ & -5x_2 & +3x_3 & -3x_4 & = & 1 \end{array}\right.$$

son equivalentes.

8. Indique cuáles son las operaciones elementales realizadas que permiten obtener los sucesivos sistemas equivalentes. Resuelva el sistema original.

$$\left\{\begin{array}{cccccc} 2x & -y & +z & = & 4 \\ x & +y & -2z & = & 0 \\ -x & +y & +z & = & 2 \end{array}\right. \rightarrow \left\{\begin{array}{cccccc} x & +y & -2z & = & 0 \\ 2x & -y & +z & = & 4 \\ -x & +y & +z & = & 2 \end{array}\right. \rightarrow \left\{\begin{array}{cccccc} x & +y & -2z & = & 0 \\ & -3y & +5z & = & 4 \\ & 2y & -z & = & 2 \end{array}\right. \rightarrow$$

$$\rightarrow \left\{\begin{array}{cccccc} x & +y & -2z & = & 0 \\ & -3y & +5z & = & 4 \\ & & \frac{7}{3}z & = & \frac{14}{3} \end{array}\right. \rightarrow \left\{\begin{array}{cccccc} x & +y & -2z & = & 0 \\ & -3y & +5z & = & 4 \\ & & 7z & = & 14 \end{array}\right. .$$

1.3. Eliminación gaussiana y de Gauss-Jordan

En la sección anterior introducimos un método para la solución de sistemas lineales. Se estudiará este procedimiento con mayor profundidad comenzando con la noción de *matriz*.

Definición (Matriz). Si m y n son enteros positivos, una *matriz* A de $m \times n$ es un arreglo rectangular de mn números dispuestos en m renglones (o filas) y n columnas.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$$

El número a_{ij} es el *elemento* o *componente* de A que está en el renglón i, columna j. Una matriz de m renglones y n columnas es de *tamaño* u *orden* $m \times n$.

Si $m = n$ la matriz es **cuadrada** de orden n, los elementos a_{11}, a_{22},...,a_{nn} son los elementos de la **diagonal principal**.

En este texto los elementos a_{ij} serán números reales salvo otra indicación.

Se usan letras mayúsculas imprentas como A, B, C, etc. para denotar una matriz. También pueden denotarse mediante un elemento genérico escrito entre corchetes, por ejemplo: $[a_{ij}]$, $[b_{ij}]$, etc. Así, para abreviar la escritura, la matriz A puede denotarse simplemente por

$$A = [a_{ij}]$$

indicando además que es de tamaño $m \times n$.

Por ejemplo:

$$A = \begin{bmatrix} 1 & -2 & 5 \\ 3 & 1 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad C = \begin{bmatrix} -1 & 2 & 4 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 2 & \frac{3}{7} \\ 3 & 8 \end{bmatrix}.$$

A es una matriz 2×3, sus elementos son $a_{11} = 1$, $a_{12} = -2$, $a_{13} = 5$, $a_{21} = 3$, $a_{22} = 1$ y $a_{23} = 6$. B es una matriz 2×1 y se dice que es una **matriz columna** o un **vector columna**. C es una matriz de 1×3, es una **matriz renglón**. Se suelen utilizar letras minúsculas negritas para denotar matrices columna o renglón. Por ejemplo:

$$\mathbf{b} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Un **renglón cero** de una matriz solo incluye ceros. Un **renglón no cero** tiene al menos un elemento distinto de cero. El primer elemento no cero de un renglón no cero se denomina **elemento delantero** o **principal**. Si el elemento delantero es un 1 se lo llama **1-delantero**. Por ejemplo, la matriz de 4×5:

$$\begin{bmatrix} 1 & 2 & 4 & 0 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ -8 & 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

tiene el último renglón cero. Sus elementos delanteros son 1, 3 y -8.

Mediante el uso de matrices, se puede abreviar la escritura de un sistema lineal anotando solo sus coeficientes y/o términos constantes, siempre que están especificados los nombres y el orden de las variables. Consideremos el sistema lineal

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ... + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + ...a_{mn}x_n = b_m \end{cases}$$

El arreglo rectangular de los coeficientes del sistema es la **matriz de coeficientes**:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}.$$

Si a esta matriz se le agrega como última columna el vector de las constantes

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

se obtiene la llamada **matriz aumentada** o **ampliada** del sistema:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} & b_m \end{bmatrix}.$$

Para hacer énfasis de que se trata de una matriz aumentada, se la suele representar de la siguiente manera:

$$[A : \mathbf{b}] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & : & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & : & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} & : & b_m \end{bmatrix}.$$

Ejemplo 1.17. Para el sistema lineal del Ejemplo 1.10

$$\begin{cases} x_1 & -2x_2 & +x_3 & -x_4 & = & 2 \\ 2x_1 & & -x_3 & +2x_4 & = & 0 \\ & -x_2 & +x_3 & -x_4 & = & 1 \end{cases}$$

la matriz de coeficientes es

$$A = \begin{bmatrix} 1 & -2 & 1 & -1 \\ 2 & 0 & -1 & 2 \\ 0 & -1 & 1 & -1 \end{bmatrix}$$

y su matriz aumentada es

$$[A : \mathbf{b}] = \begin{bmatrix} 1 & -2 & 1 & -1 & : & 2 \\ 2 & 0 & -1 & 2 & : & 0 \\ 0 & -1 & 1 & -1 & : & 1 \end{bmatrix}. \blacktriangleleft$$

Actividad 1.4. Escriba el sistema lineal cuya matriz aumentada es $\begin{bmatrix} 1 & -2 & : & 1 \\ 2 & 0 & : & -1 \\ 0 & -1 & : & 1 \\ 1 & 1 & : & 0 \end{bmatrix}$.

Como la matriz aumentada de un sistema lineal es una representación abreviada del mismo, se puede ahorrar tiempo y trabajar en forma mas eficiente si se utiliza esta notación en el proceso de resolución del mismo.

Las tres operaciones de ecuaciones que se pueden utilizar en un sistema lineal para generar un sistema equivalente, se corresponden con ciertas *operaciones entre renglones* de la matriz aumentada del sistema.

Definición (Operaciones elementales de renglón). Las operaciones elementales de renglón de una matriz son las siguientes:

- Intercambio. Intercambia renglones: $R_i \leftrightarrow R_j$.
- Escalamiento. Multiplica un renglón por una constante $c \neq 0$: $cR_i \to R_i$.
- Eliminación. Suma a un renglón un múltiplo escalar de otro renglón: $R_i + cR_j \to R_i$.

Ejemplo 1.18. Retomando el sistema lineal del Ejemplo 1.15 se muestra a la derecha su resolución trabajando en forma abreviada con la matriz aumentada del sistema, haciendo uso de las operaciones elementales de renglón correspondientes a las operaciones en ecuaciones realizadas:

$$\left\{\begin{array}{rrrrcr} x_1 & -2x_2 & +x_3 & -x_4 & = & 2 \\ 2x_1 & & -x_3 & +2x_4 & = & 0 \\ & -x_2 & +x_3 & -x_4 & = & 1 \end{array}\right. \qquad \begin{bmatrix} 1 & -2 & 1 & -1 & : & 2 \\ 2 & 0 & -1 & 2 & : & 0 \\ 0 & -1 & 1 & -1 & : & 1 \end{bmatrix}$$

$$E_2 - 2E_1 \to E_2 \qquad R_2 - 2R_1 \to R_2$$

$$\left\{\begin{array}{rrrrcr} x_1 & -2x_2 & +x_3 & -x_4 & = & 2 \\ & 4x_2 & -3x_3 & +4x_4 & = & -4 \\ & -x_2 & +x_3 & -x_4 & = & 1 \end{array}\right. \qquad \begin{bmatrix} 1 & -2 & 1 & -1 & : & 2 \\ 0 & 4 & -3 & 4 & : & -4 \\ 0 & -1 & 1 & -1 & : & 1 \end{bmatrix}$$

$$E_3 + \tfrac{1}{4}E_2 \to E_3 \qquad R_3 + \tfrac{1}{4}R_2 \to R_3$$

$$\left\{\begin{array}{rrrrcr} x_1 & -2x_2 & +x_3 & -x_4 & = & 2 \\ & 4x_2 & -3x_3 & +4x_4 & = & -4 \\ & & \frac{1}{4}x_3 & & = & 0 \end{array}\right. \qquad \begin{bmatrix} 1 & -2 & 1 & -1 & : & 2 \\ 0 & 4 & -3 & 4 & : & -4 \\ 0 & 0 & \frac{1}{4} & 0 & : & 0 \end{bmatrix}$$

y su conjunto solución (véase Ejemplo 1.15) es

$$\{(-t, -1-t, 0, t) : t \in \mathbb{R}\}. \blacktriangleleft$$

Una operación elemental de renglón aplicada a la matriz aumentada de un sistema lineal produce la matriz aumentada de un sistema lineal equivalente. Esto, y dado que toda matriz es la matriz aumentada de un sistema lineal, da origen a la siguiente definición:

Definición (Matrices equivalentes por renglones). Dos matrices A y B son *equivalentes por renglones* y se denota

$$A \sim B$$

si una de ellas puede obtenerse a partir de la otra mediante una sucesión finita de operaciones elementales entre renglones.

Por simplicidad, a lo largo de este texto llamaremos simplemente **matrices equivalentes** a las matrices equivalentes por renglones.

Ejemplo 1.19. Las matrices

$$A = \begin{bmatrix} 1 & -2 & 1 & 1 \\ -1 & 3 & 1 & -1 \\ 4 & 2 & -6 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & -2 & 1 & 1 \\ 0 & 1 & 2 & 0 \\ 2 & 1 & -3 & 0 \end{bmatrix}$$

son equivalentes. En efecto: B se obtiene de A mediante una operación de eliminación seguida por un escalamiento.

$$A = \begin{bmatrix} 1 & -2 & 1 & 1 \\ -1 & 3 & 1 & -1 \\ 4 & 2 & -6 & 0 \end{bmatrix} \xrightarrow[R_2 + R_1 \to R_2]{} \begin{bmatrix} 1 & -2 & 1 & 1 \\ 0 & 1 & 2 & 0 \\ 4 & 2 & -6 & 0 \end{bmatrix} \xrightarrow[\frac{1}{2}R_3 \to R_3]{} \begin{bmatrix} 1 & -2 & 1 & 1 \\ 0 & 1 & 2 & 0 \\ 2 & 1 & -3 & 0 \end{bmatrix} = B. \blacktriangleleft$$

Las operaciones elementales de renglón son revertibles, o inversibles, en el sentido de que cada operación elemental tiene una operación elemental inversa que revierte el resultado.

Actividad 1.5. Verifique que la matriz A del ejemplo anterior puede obtenerse a partir de la matriz B revirtiendo la sucesión de operaciones elementales.

Actividad 1.6. ¿Cuál es la operación elemental inversa de $cR_i \to R_i$ $(c \neq 0)$?. ¿Y cuál es la operación elemental inversa de $R_i + cR_j \to R_i$?

Se han resuelto sistemas lineales llevándolos a uno escalonado equivalente. Los sistemas escalonados son aquellos cuya matriz aumentada es de *forma escalonada*. Precisamente:

Definición (Matrices en forma escalonada por renglones y en forma escalonada por renglones reducida). Una matriz está en *forma escalonada por renglones* si verifica las siguientes condiciones:

- Todos los renglones cero (si los tiene) están en la parte inferior de la matriz.
- El elemento delantero de cada renglón no cero después del primero está a la derecha del elemento delantero del renglón anterior.

La matriz está en *forma escalonada por renglones reducida* si además verifica las dos siguientes condiciones:

- El elemento delantero de cada renglón no cero es 1 (1-delantero).
- Toda columna que contiene un 1-delantero, tiene ceros en sus restantes posiciones.

Por simplicidad diremos que una matriz está en "forma escalonada" o en "forma escalonada reducida" para indicar que está en forma escalonada por renglones o en forma escalonada por renglones reducida, respectivamente.

Por ejemplo, sean las matrices:

$$A = \begin{bmatrix} \boxed{1} & 2 & 5 \\ 0 & \boxed{1} & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 0 & 9 \\ 0 & 2 & 1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} \boxed{2} & 2 & 4 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{-1} \end{bmatrix}, \quad D = \begin{bmatrix} \boxed{1} & 3 \\ 0 & 0 \end{bmatrix},$$

$$E = \begin{bmatrix} \boxed{1} & 0 & 3 & 0 \\ 0 & \boxed{1} & 2 & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix}, \quad G = \begin{bmatrix} \boxed{1} & 2 & 5 \end{bmatrix} \quad \text{y} \quad H = \begin{bmatrix} 0 & \boxed{2} & 5 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \end{bmatrix}.$$

Las matrices A, C, D E, F G y H están en forma escalonada y se han recuadrado sus elementos delanteros. La matriz B no está en forma escalonada. Las matrices D, E, F y G están en forma escalonada reducida.

Teorema 1.4. Toda matriz es equivalente a una matriz en forma escalonada.

Demostración. La prueba de este teorema es el *algoritmo de Gauss o de eliminación gaussiana.* □

El algoritmo de Gauss o método de eliminación (o reducción) gaussiana es un procedimiento que permite convertir cualquier matriz en una equivalente en forma escalonada. El método indica las operaciones elementales de renglón que deben ir aplicándose sucesivamente. Los pasos son los siguientes:

Algoritmo de eliminación gaussiana

1. Vaya a la columna no cero extremo izquierda.
2. Si el primer renglón tiene un cero en la columna del paso 1, intercámbielo con uno que tenga un elemento no cero en la misma columna. Se obtiene así una matriz equivalente con el primer renglón no cero cuyo elemento delantero está en la primera columna no cero.
3. Obtenga ceros abajo de dicho elemento delantero sumando múltiplos adecuados del primer renglón a los renglones debajo de él.
4. Cubra (o ignore) el renglón superior y repita el mismo procedimiento comenzando con el paso 1 aplicado a la matriz resultante.

Ejemplo 1.20. Para obtener una matriz en forma escalonada equivalente a la matriz

$$A = \begin{bmatrix} 0 & 2 & 3 & 3 & 0 \\ 1 & 1 & 2 & -1 & 1 \\ 2 & 0 & 1 & 0 & 12 \\ 3 & 1 & 3 & -1 & 13 \end{bmatrix}$$

se procede, siguiendo el algoritmo de eliminación gaussiana, de la siguiente manera:

$$\begin{bmatrix} 0 & 2 & 3 & 3 & 0 \\ 1 & 1 & 2 & -1 & 1 \\ 2 & 0 & 1 & 0 & 12 \\ 3 & 1 & 3 & -1 & 13 \end{bmatrix} \xrightarrow[R_1 \leftrightarrow R_2]{} \begin{bmatrix} \boxed{1} & 1 & 2 & -1 & 1 \\ 0 & 2 & 3 & 3 & 0 \\ 2 & 0 & 1 & 0 & 12 \\ 3 & 1 & 3 & -1 & 13 \end{bmatrix} \xrightarrow[\substack{R_4 - 3R_1 \to R_4 \\ R_3 - 2R_1 \to R_3}]{} \begin{bmatrix} \boxed{1} & 1 & 2 & -1 & 1 \\ 0 & \boxed{2} & 3 & 3 & 0 \\ 0 & -2 & -3 & 2 & 10 \\ 0 & -2 & -3 & 2 & 10 \end{bmatrix} \to$$

$$\xrightarrow[\substack{R_4 + R_2 \to R_4 \\ R_3 + R_2 \to R_3}]{} \begin{bmatrix} \boxed{1} & 1 & 2 & -1 & 1 \\ 0 & \boxed{2} & 3 & 3 & 0 \\ 0 & 0 & 0 & \boxed{5} & 10 \\ 0 & 0 & 0 & 5 & 10 \end{bmatrix} \xrightarrow[R_4 - R_3 \to R_4]{} \begin{bmatrix} \boxed{1} & 1 & 2 & -1 & 1 \\ 0 & \boxed{2} & 3 & 3 & 0 \\ 0 & 0 & 0 & \boxed{5} & 10 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

La matriz obtenida

$$E = \begin{bmatrix} 1 & 1 & 2 & -1 & 1 \\ 0 & 2 & 3 & 3 & 0 \\ 0 & 0 & 0 & 5 & 10 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

está en forma escalonada.◀

Teorema 1.5. Toda matriz es equivalente a una matriz en forma escalonada reducida.

Demostración. Su prueba es el *algoritmo de Gauss-Jordan.* □

El algoritmo de Gauss-Jordan continúa el proceso de eliminación gaussiana hasta obtener una matriz equivalente en forma escalonada reducida. Se procede de la siguiente manera:

Algoritmo de Gauss-Jordan

I. Si la matriz está en forma escalonada pase al paso II. Si la matriz no está en forma escalonada aplique los pasos 1, 2, 3 y 4 del algoritmo de Gauss hasta obtener una forma escalonada y luego continúe con el paso II.

II. Comenzando con el último renglón no cero, se avanza hacia arriba del siguiente modo:

   i. Vaya al último renglón no cero y, si el elemento delantero no es un 1, multiplique dicho renglón por un apropiado escalar para obtener un 1-delantero.

   ii. Obtenga ceros arriba del 1−delantero sumando múltiplos adecuados de este último renglón a los renglones arriba de él.

   iii. Cubra el último renglón no cero y aplique el mismo procedimiento comenzando con el paso i. a la matriz restante.

Ejemplo 1.21. Para obtener una matriz en forma escalonada reducida equivalente a la matriz A del ejemplo anterior, se continúa con el algoritmo de Gauss-Jordan a partir de la matriz E:

$$\begin{bmatrix} \boxed{1} & 1 & 2 & -1 & 1 \\ 0 & \boxed{2} & 3 & 3 & 0 \\ 0 & 0 & 0 & \boxed{5} & 10 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[\frac{1}{5}R_3 \to R_3]{} \begin{bmatrix} \boxed{1} & 1 & 2 & -1 & 1 \\ 0 & \boxed{2} & 3 & 3 & 0 \\ 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[\substack{R_2 - 3R_3 \to R_2 \\ R_1 + R_3 \to R_1}]{} \begin{bmatrix} \boxed{1} & 1 & 2 & 0 & 3 \\ 0 & \boxed{2} & 3 & 0 & -6 \\ 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \to$$

$$\xrightarrow[\frac{1}{2}R_2 \to R_2]{} \begin{bmatrix} \boxed{1} & 1 & 2 & 0 & 3 \\ 0 & \boxed{1} & \frac{3}{2} & 0 & -3 \\ 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[R_1 - R_2 \to R_1]{} \begin{bmatrix} \boxed{1} & 0 & \frac{1}{2} & 0 & 6 \\ 0 & \boxed{1} & \frac{3}{2} & 0 & -3 \\ 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

La matriz obtenida

$$R = \begin{bmatrix} 1 & 0 & \frac{1}{2} & 0 & 6 \\ 0 & 1 & \frac{3}{2} & 0 & -3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

está en forma escalonada reducida y es equivalente a la matriz A.◀

Observación 1.1. Pueden obtenerse 1-delanteros en el paso 3 del algoritmo de Gauss aunque esta práctica tendería a introducir fracciones en una etapa temprana del cálculo si los elementos son enteros.

Observación 1.2. Debe tenerse en cuenta que una vez que se hayan hecho ceros abajo (o arriba) del elemento delantero de un renglón debe comenzarse otra vez con el primer paso aplicado a la matriz que resulta de haber eliminado dicho renglón. Esto es así para no "arruinar" los ceros que ya se habían hecho. Por ejemplo: el paso siguiente en la conversión de la matriz

$$\begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

es hacer cero el elemento en la posición $(3, 2)$. El algoritmo de Gauss indica cubrir el renglón 1 (no se tiene en cuenta el renglón 1) y hacer $R_3 - R_2 \to R_3$ para obtener:

$$\begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Nótese que se hubiese logrado un cero en la posición $(3, 2)$ con la operación $R_3 + R_1 \to R_3$, pero esto hubiese conducido a la matriz

$$\begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ 2 & 0 & 1 \end{bmatrix}$$

"arruinando" el cero que ya se había obtenido en la posición $(3, 1)$.

Observación 1.3. En este texto se hará uso de expresiones tales como "llevar la matriz A a una forma escalonada", que significa obtener una matriz en forma escalonada equivalente a la matriz A.

Nota: Una matriz no cero es equivalente a infinitas matrices en forma escalonada pero solo a una en forma escalonada reducida. Se aceptará sin demostrar el siguiente teorema:

Teorema 1.6 (Unicidad de la forma escalonada reducida)**.** Toda matriz es equivalente a una y solo una matriz escalonada reducida.

Si A, E y R son tres matrices equivalentes tales que E está en forma escalonada y R en forma escalonada reducida, se dice que E es **una** forma escalonada de A y que R es **la** forma escalonada reducida de A.

En cualquier forma escalonada de una matriz A los elementos delanteros se encuentran en las mismas posiciones, a dichos elementos delanteros se los llama **pivotes**. Se llaman **posiciones pivote** (o de pivoteo) de una matriz A a las posiciones de los pivotes de una (cualquiera) forma escalonada de A.

Las columnas de A que contienen una posición pivote se llaman **columnas pivote** de A. Debe notarse que para determinar cuáles son las columnas pivote de una matriz A que no está en forma escalonada debe primero llevársela a una forma escalonada para conocer cuáles son las posiciones pivote.

Ejemplo 1.22. Las matrices

$$B = \begin{bmatrix} 2 & 2 & 3 \\ 0 & 4 & 2 \\ 0 & 0 & -5 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & -5 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 2 & 3 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & -1 \end{bmatrix} \quad \text{y} \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

son formas escalondas de la matriz

$$A = \begin{bmatrix} 2 & 2 & 3 \\ 4 & 8 & 8 \\ 6 & 6 & 4 \end{bmatrix}.$$

La matriz I (esta matriz es la **matriz identidad** de orden 3) es la forma escalonada reducida de A (también es la forma escalonada reducida de B, C y D).

Las posiciones pivote de las matrices A, B, C, D y I son $(1,1)$, $(2,2)$ y $(3,3)$, y sus columnas pivote son las columnas 1, 2 y 3. Los pivotes de B son 2, 4 y -5, los de C son 2, 2 y -5, los de D son 2, 1 y -1, y los pivotes de I son 1, 1 y 1.◄

Actividad 1.7. Considere las matrices A, B, C, D y I del Ejemplo 1.22. Verifique que las matrices B, C, D y I son formas escalonadas de la matriz A y que la matriz I es su forma escalonada reducida.

Actividad 1.8. Dada la matriz $A = \begin{bmatrix} 1 & -1 & 2 & 3 \\ -2 & 2 & -4 & -6 \\ 2 & -1 & 6 & 8 \\ 1 & 0 & 4 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, halle su forma escalonada reducida y exhiba sus columnas pivote.

Actividad 1.9. ¿Cuántas matrices cuadradas de orden 3 escalonadas reducidas tienen exactamente 2 pivotes?, ¿cuántas tienen exactamente 3 pivotes?

Ejercicios 1.3

$\boxed{1.}$ Escriba la matriz de coeficientes y la matriz aumentada asociada al sistema lineal

$$\left\{\begin{array}{rrrrcr} x & +y & +z & -w & = & -1 \\ 2x & +y & -z & & = & 0 \\ & y & +z & -w & = & 1 \\ x & +y & +z & +w & = & 4 \\ 2x & & -z & +w & = & 0 \end{array}\right. .$$

2. Teniendo en cuenta la Actividad 1.6 demuestre que si A, B y C son matrices $m \times n$ entonces se verifican las siguientes propiedades:

   a) $A \sim B \Rightarrow B \sim A$

   b) Si $A \sim B$ y $B \sim C$, entonces $A \sim C$.

$\boxed{3.}$ Indique cuáles de las siguientes matrices están en forma escalonada y cuáles de éstas en forma escalonada reducida. Recuadre los pivotes de cada matriz que están en forma escalonada e indique sus columnas pivote.

$$A = \begin{bmatrix} 1 & -2 & 1 & 1 \\ 0 & 5 & -2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 5 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 & 9 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$E = \begin{bmatrix} 7 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad F = \begin{bmatrix} 1 & 6 & 0 & 0 & 1 \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 & 0 & 5 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

4. Sean A y B dos matrices de $m \times n$, y A' y B' sus respectivas formas escalonadas reducidas. Demuestre que: $A \sim B$ si y solo si $A' = B'$.

   Nota: Dos matrices son **iguales** si tienen el mismo tamaño y los elementos que ocupan las mismas posiciones son iguales.

$\boxed{5.}$ Determine en cada caso si las matrices A y B son equivalentes.

a) $A = \begin{bmatrix} 1 & 2 & -1 & 2 \\ -1 & 2 & 2 & 1 \\ 0 & 4 & 1 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & -\frac{3}{2} & \frac{1}{2} \\ 0 & 4 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

b) $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix}$

c) $A = \begin{bmatrix} 2 & 1 & -1 & 3 \\ 4 & 0 & 1 & -1 \\ 0 & 2 & 0 & 2 \\ 1 & 1 & -1 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

$\boxed{6.}$ Demuestre que si $a_{11}a_{22} - a_{12}a_{21} \neq 0$ entonces

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$\boxed{7.}$ En cada caso obtenga la forma escalonada reducida para cada matriz A. Indique las columnas pivote de A.

$$a)\ A = \begin{bmatrix} 1 & -1 & -1 & 3 \\ 1 & -1 & 0 & 1 \\ 0 & 2 & 1 & -1 \end{bmatrix} \quad b)\ A = \begin{bmatrix} 2 & -1 & 5 & 0 \\ -3 & 0 & 2 & 1 \\ 1 & -2 & 12 & 1 \end{bmatrix} \quad c)\ A = \begin{bmatrix} 1 & 3 & 2 & -2 \\ 2 & 6 & 4 & -4 \\ 3 & 9 & 6 & -6 \end{bmatrix}$$

8. Sea A una matriz de 15×18. ¿Es posible que el número de posiciones pivote de A sea 19?, ¿y que sea 16?. ¿Cuál es el número máximo de posiciones pivote de A?.

$\boxed{9.}$ Justifique que si A es una matriz de $m \times n$ entonces el número de posiciones pivote de A es menor o igual al mínimo entre m y n.

1.4. Resolución de sistemas lineales

Dada la identificación entre un sistema lineal y su matriz aumentada, y habiéndose demostrado que toda matriz es equivalente a alguna matriz en forma escalonada y a una (y solo una) en forma escalonada reducida, el mismo resultado se aplica a los sistemas de ecuaciones lineales. Esto es, todo sistema de ecuaciones lineales es equivalente a alguno escalonado.

En esta sección se verá cómo emplear la eliminación gaussiana y la de Gauss-Jordan para resolver un sistema lineal. Cualquiera de estos procedimientos se aplica a la matriz aumentada del sistema para producir una matriz en forma escalonada equivalente. El sistema correspondiente a esta última matriz es equivalente al original y como es escalonado es fácil de resolver, o es evidente su inconsistencia.

Eliminación gaussiana con sustitución hacia atrás

El procedimiento es el siguiente:

1. Escriba la matriz aumentada del sistema lineal.
2. Utilice el algoritmo de eliminación gaussiana para llevar la matriz a una forma escalonada. Si durante cualquier etapa de este proceso se encuentra una matriz con un renglón de la forma $\begin{bmatrix} 0 & 0 & ... & 0 & : & b \end{bmatrix}$ con $b \neq 0$, deténgase, en este caso el sistema es inconsistente. De lo contrario continúe con 3).
3. Escriba el sistema de ecuaciones lineales escalonado correspondiente a la matriz obtenida en 2) sin tener en cuenta los renglones cero (si los tiene).
4. Para el sistema escalonado obtenido, distinga las variables delanteras de las libres y aplique el método de sustitución hacia atrás para encontrar la solución del sistema.

Con respecto al paso 3, cuando se llega a una matriz en forma escalonada, si ésta tiene renglones cero, las ecuaciones lineales correspondientes son de la forma $0x_1 + 0x_2 + ... + 0x_n = 0$. Este tipo de ecuaciones se satisfacen para cualquier n-upla de números reales, por lo tanto no aportan soluciones al sistema y se descartan del mismo. En estos casos el sistema escalonado equivalente tendrá menos ecuaciones que el sistema original.

Ejemplo 1.23. La matriz

$$\begin{bmatrix} 0 & 2 & 3 & 3 & 0 \\ 1 & 1 & 2 & -1 & 1 \\ 2 & 0 & 1 & 0 & 12 \\ 3 & 1 & 3 & -1 & 13 \end{bmatrix}$$

del Ejemplo 1.20 es la matriz aumentada del sistema lineal

$$\begin{cases} & 2x_2 & +3x_3 & +3x_4 & = & 0 \\ x_1 & +x_2 & +2x_3 & -x_4 & = & 1 \\ 2x_1 & & +x_3 & & = & 12 \\ 3x_1 & +x_2 & +3x_3 & -x_4 & = & 13 \end{cases}$$

Aplicando eliminación gaussiana se obtuvo la siguiente forma escalonada de su matriz aumentada

$$\begin{bmatrix} \boxed{1} & 1 & 2 & -1 & 1 \\ 0 & \boxed{2} & 3 & 3 & 0 \\ 0 & 0 & 0 & \boxed{5} & 10 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Como esta matriz tiene un renglón cero, se descarta la ecuación correspondiente. Además no hay un renglón de la forma $\begin{bmatrix} 0 & 0 & ... & 0 & : & b \end{bmatrix}$ con $b \neq 0$, por lo tanto el sistema escalonado correspondiente es consistente. Este sistema es

$$\begin{cases} x_1 & +x_2 & +2x_3 & -x_4 & = & 1 \\ & 2x_2 & +3x_3 & +3x_4 & = & 0 \\ & & & 5x_4 & = & 10 \end{cases}$$

Las variables delanteras son x_1, x_2 y x_4, que son las que corresponden a las columnas pivote 1°, 2° y 4° de la matriz de coeficientes. La variable x_3 (que corresponde a una columna no pivote) es libre, por lo que toma cualquier valor. Sea

$$x_3 = t, \quad t \in \mathbb{R}.$$

Se aplica ahora el método de sustitución hacia atrás. De la última ecuación

$$x_4 = 2.$$

Sustituyendo x_4 y x_3 en la ecuación anterior se tiene $2x_2 + 3t + 3(2) = 0$ y despejando x_2

$$x_2 = -\frac{3}{2}t - 3.$$

Reemplazando ahora x_4, x_3 y x_2 en la primera ecuación se tiene $x_1 + \left(-\frac{3}{2}t - 1\right) + 2t - 2 = 1$ y depejando x_1 se obtiene

$$x_1 = -\frac{1}{2}t + 6.$$

El sistema lineal tiene infinitas soluciones y su solución general es

$$x_1 = -\frac{1}{2}t + 6, \quad x_2 = -\frac{3}{2}t - 3, \quad x_3 = t, \quad x_4 = 2 \quad \text{para todo} \quad t \in \mathbb{R}. \blacktriangleleft$$

La eliminación de Gauss-Jordan para la resolución de sistemas, encuentra un sistema lineal equivalente al dado que es **escalonado reducido**, esto es, su matriz aumentada está en forma escalonada reducida.

Eliminación de Gauss-Jordan

Los pasos a seguir son los siguientes:

1. Escriba la matriz aumentada del sistema lineal.
2. Emplee el método de eliminación de Gauss-Jordan para obtener la forma escalonada reducida de la matriz. Si en alguna etapa (temprana) del proceso se encuentra una matriz con un renglón de la forma $\left[\begin{array}{ccccc} 0 & 0 & \ldots & 0 & : & b \end{array}\right]$ con $b \neq 0$, deténgase. El sistema es inconsistente. De lo contrario continúe con el paso 3).
3. Escriba el sistema de ecuaciones lineales escalonado "reducido" correspondiente a la matriz obtenida en 2) sin tener en cuenta los renglones cero.
4. Distinga las variables delanteras de las libres (si las hay) y despeje la variable delantera de cada ecuación en función de las libres, de las constantes o de ambas.

Observación 1.4. En un sistema escalonado reducido la variable delantera de cada ecuación "no aparece" en otra ecuación, en particular no aparece en las ecuaciones anteriores, y por lo tanto no se requiere aplicar sustitución hacia atrás para resolverlo.

Ejemplo 1.24. Retomando el ejemplo anterior, se puede resolver el sistema

$$\left\{\begin{array}{rrrrcr} & 2x_2 & +3x_3 & +3x_4 & = & 0 \\ x_1 & +x_2 & +2x_3 & -x_4 & = & 1 \\ 2x_1 & & +x_3 & & = & 12 \\ 3x_1 & +x_2 & +3x_3 & -x_4 & = & 13 \end{array}\right.$$

llevándolo al escalonado reducido.

Aplicando eliminación gaussiana se obtuvo la siguiente forma escalonada de su matriz aumentada:

$$\begin{bmatrix} \boxed{1} & 1 & 2 & -1 & 1 \\ 0 & \boxed{2} & 3 & 3 & 0 \\ 0 & 0 & 0 & \boxed{5} & 10 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Ahora, en lugar de usar sustitución hacia atrás para resolver el sistema escalonado correspondiente, se continúa con la eliminación de Gauss-Jordan hasta obtener la forma escalonada reducida. El Ejemplo 1.21 muestra la matriz escalonada reducida que se obtuvo:

$$\begin{bmatrix} \boxed{1} & 0 & \frac{1}{2} & 0 & 6 \\ 0 & \boxed{1} & \frac{3}{2} & 0 & -3 \\ 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

El sistema escalonado reducido correpondiente es

$$\left\{\begin{array}{rrrrcr} x_1 & & +\frac{1}{2}x_3 & & = & 6 \\ & x_2 & +\frac{3}{2}x_3 & & = & -3 \\ & & & x_4 & = & 2 \end{array}\right.$$

La variable x_3 es libre, puede tomar cualquier valor. Si $x_3 = t$, despejando la variable delantera de cada ecuación en función de t y las constantes se obtiene directamente la solución general del sistema, a saber:

$$\left\{\begin{array}{rcr} x_1 & = & 6 - \frac{1}{2}t \\ x_2 & = & -3 - \frac{3}{2}t \\ x_3 & = & t \\ x_4 & = & 2 \end{array}\right. \quad \text{para todo} \quad t \in \mathbb{R}.$$

Observe que el trabajo que se agrega al completar el proceso para obtener la forma escalonada reducida de la matriz aumentada, se compensa con la resolución directa del sistema reducido equivalente.◀

Ejemplo 1.25 (Sistema consistente con única solución). Para resolver el sistema lineal

$$\left\{\begin{array}{rrrcr} x_1 & -x_2 & -x_3 & = & 0 \\ -x_1 & +x_2 & -x_3 & = & -6 \\ 2x_1 & -x_2 & -x_3 & = & 1 \end{array}\right.$$

aplicando el algoritmo de eliminación de Gauss-Jordan, procedemos así:

$$\left[\begin{array}{rrrcr} \boxed{1} & -1 & -1 & : & 0 \\ -1 & 1 & -1 & : & -6 \\ 2 & -1 & -1 & : & 1 \end{array}\right] \xrightarrow[\substack{R_2 + R_1 \to R_2 \\ R_3 - 2R_1 \to R_3}]{} \left[\begin{array}{rrrcr} \boxed{1} & -1 & -1 & : & 0 \\ 0 & 0 & -2 & : & -6 \\ 0 & 1 & 1 & : & 1 \end{array}\right] \xrightarrow[R_3 \leftrightarrow R_2]{} \left[\begin{array}{rrrcr} \boxed{1} & -1 & -1 & : & 0 \\ 0 & \boxed{1} & 1 & : & 1 \\ 0 & 0 & -2 & : & -6 \end{array}\right] \to$$

$$\xrightarrow[-\frac{1}{2}R_3 \to R_3]{} \left[\begin{array}{rrrcr} \boxed{1} & -1 & -1 & : & 0 \\ 0 & \boxed{1} & 1 & : & 1 \\ 0 & 0 & \boxed{1} & : & 3 \end{array}\right] \xrightarrow[\substack{R_1 + R_3 \to R_1 \\ R_2 - R_3 \to R_2}]{} \left[\begin{array}{rrrcr} \boxed{1} & -1 & 0 & : & 3 \\ 0 & \boxed{1} & 0 & : & -2 \\ 0 & 0 & \boxed{1} & : & 3 \end{array}\right] \to$$

$$\xrightarrow[R_1 + R_2 \to R_1]{} \left[\begin{array}{rrrcr} \boxed{1} & 0 & 0 & : & 1 \\ 0 & \boxed{1} & 0 & : & -2 \\ 0 & 0 & \boxed{1} & : & 3 \end{array}\right].$$

El sistema escalonado reducido equivalente al dado es:

$$\left\{\begin{array}{rrrcr} x_1 & & & = & 1 \\ & x_2 & & = & -2 \\ & & x_3 & = & 3 \end{array}\right.$$

Todas las variables del sistema son delanteras (no posee variables libres). La solución es única y se lee directamente: $x_1 = 1$, $x_2 = -2$ y $x_3 = 3$. Su conjunto solución es $\{(1, -2, 3)\}$. ◀

Ejemplo 1.26 (Sistema consistente con infinitas soluciones). Considere el sistema lineal

$$\left\{\begin{array}{rrrrcr} 2x_1 & -x_2 & +x_3 & -x_4 & = & 0 \\ x_1 & +3x_2 & +x_3 & -x_4 & = & 1 \\ -3x_1 & +5x_2 & -x_3 & +x_4 & = & 1 \end{array}\right.$$

Para resolver el sistema aplicando eliminación gaussiana, procedemos como sigue:

$$\left[\begin{array}{rrrrcr} \boxed{2} & -1 & 1 & -1 & : & 0 \\ 1 & 3 & 1 & -1 & : & 1 \\ -3 & 5 & -1 & 1 & : & 1 \end{array}\right] \xrightarrow[R_1 \leftrightarrow R_2]{} \left[\begin{array}{rrrrcr} \boxed{1} & 3 & 1 & -1 & : & 1 \\ 2 & -1 & 1 & -1 & : & 0 \\ -3 & 5 & -1 & 1 & : & 1 \end{array}\right] \to$$

$$\xrightarrow[R_3+3R_1\to R_3]{R_2-2R_1\to R_2}\left[\begin{array}{cccccc}\boxed{1} & 3 & 1 & -1 & : & 1\\ 0 & \boxed{-7} & -1 & 1 & : & -2\\ 0 & 14 & 2 & -2 & : & 4\end{array}\right]\xrightarrow{R_3+2R_2\to R_3}\left[\begin{array}{cccccc}\boxed{1} & 3 & 1 & -1 & : & 1\\ 0 & \boxed{-7} & -1 & 1 & : & -2\\ 0 & 0 & 0 & 0 & : & 0\end{array}\right]$$

(observe que en el primer paso se intercambiaron los renglones R_1 con R_2 para usar como pivote el 1, ya que de esta forma se evita introducir fracciones en una etapa temprana del proceso).

Un sistema escalonado equivalente al dado es

$$\begin{cases} x_1 +3x_2 +x_3 -x_4 = 1\\ \quad -7x_2 -x_3 +x_4 = -2\end{cases}$$

Las variables x_1 y x_2 son delanteras, x_3 y x_4 son libres pudiendo tomar cualquier valor real. Así

$$x_4 = t,\ t\in\mathbb{R},$$

y

$$x_3 = s,\ s\in\mathbb{R}.$$

Se aplica ahora el método de sustitución hacia atrás. Despejando x_2 en función de s y t se obtiene

$$x_2=\frac{2}{7}-\frac{1}{7}s+\frac{1}{7}t.$$

Sustituyendo ahora x_2 en la primera ecuación y despejando x_1 en función de s y t se tiene

$$x_1=\frac{1}{7}-\frac{4}{7}s+\frac{4}{7}t.$$

El sistema es consistente con infinitas soluciones y su conjunto solución es

$$\left\{\left(\frac{1}{7}-\frac{4}{7}s+\frac{4}{7}t,\frac{2}{7}-\frac{1}{7}s+\frac{1}{7}t,s,t\right):t,s\in\mathbb{R}\right\}.\ \blacktriangleleft$$

Ejemplo 1.27 (Sistema inconsistente). Sea el sistema lineal

$$\begin{cases} x -2y +z = 1\\ 2x +y = 3\\ x +3y -z = 0\end{cases}.$$

Si se aplica eliminación gaussiana a su matriz aumentada

$$\left[\begin{array}{ccccc}\boxed{1} & -2 & 1 & : & 1\\ 2 & 1 & 0 & : & 3\\ 1 & 3 & -1 & : & 0\end{array}\right]\xrightarrow[R_3-R_1\to R_3]{R_2-2R_1\to R_2}\left[\begin{array}{ccccc}\boxed{1} & -2 & 1 & : & 1\\ 0 & \boxed{5} & -2 & : & 1\\ 0 & 5 & -2 & : & -1\end{array}\right]\xrightarrow{R_3-R_2\to R_3}\left[\begin{array}{ccccc}\boxed{1} & -2 & 1 & : & 1\\ 0 & \boxed{5} & -2 & : & 1\\ 0 & 0 & 0 & : & \boxed{-2}\end{array}\right]$$

se llega a una matriz escalonada equivalente cuyo último renglón, $\left[\begin{array}{ccccc}0 & 0 & 0 & : & -2\end{array}\right]$ corresponde a la ecuación inconsistente $0z=-2$ por lo que el sistema no tiene solución. $\blacktriangleleft$

Soluciones de un sistema lineal

Se han visto ejemplos de sistemas lineales que tienen una única solución, infinitas soluciones o son inconsistentes. También se han establecido criterios para determinar la consistencia o inconsistencia de los mismos.

Se obtendrán ahora resultados importantes que permiten determinar la “cantidad” de soluciones de un sistema lineal.

El siguiente resultado ya ha sido demostrado (Teorema 1.3). Ahora se agrega el enunciado equivalente referido a su matriz aumentada:

Teorema 1.7. (**Consistencia de un sistema lineal**).

Un sistema lineal es consistente si y solo si cualquier sistema escalonado equivalente no tiene una ecuación de la forma

$$0x_1 + 0x_2 + ... + 0x_n = b$$

con $b \neq 0$.

En término de matrices:

Un sistema lineal es consistente si y solo si cualquier forma escalonada de la matriz aumentada no tiene un renglón de la forma

$$\begin{bmatrix} 0 & 0 & ... & 0 & : & b \end{bmatrix}$$

con $b \neq 0$, es decir, la última columna de la matriz aumentada no es una columna pivote.

Con respecto a los sistemas consistentes se tiene el teorema que sigue. Dado que las variables delanteras corresponden a columnas pivote de la matriz aumentada y las variables libres corresponden a columnas no pivote, se agrega el enunciado equivalente referido a su matriz aumentada:

Teorema 1.8. (**Sistemas consistentes: número de soluciones**).

Un sistema lineal consistente tiene única solución si y solo si no tiene variables libres, es decir, todas sus variables son delanteras.

Un sistema lineal consistente tiene infinitas soluciones si y solo tiene variables libres.

En término de matrices:

Un sistema lineal consistente tiene una única solución si y solo si cada columna de la matriz aumentada, excepto la última, es una columna pivote (la última columna no debe ser columna pivote ya que el sistema es consistente).

Un sistema lineal consistente tiene infinitas soluciones si y solo si su matriz aumentada tiene otras columnas no pivote además de la última (la última columna no debe ser columna pivote ya que el sistema es consistente).

Demostración. Para resolver un sistema lineal consistente las variables delanteras de un sistema escalonado equivalente se despejan en función de las libres (si las tiene) y de las constantes. Si el sistema tiene variables libres entonces no tiene una única solución, más aún tiene infinitas soluciones ya que las variables libres asumen cualquier valor real. Si no tiene variables libres entonces cada variable asume un único valor y por lo tanto el sistema tendrá solución única. □

Dado que un sistema de ecuaciones lineales tiene variables libres o no las tiene, los teoremas 1.7 y 1.8 implican el siguiente:

Teorema 1.9 (Cantidad de soluciones de un sistema lineal). Para cualquier sistema lineal una y solo una de las tres propiedades siguientes es válida:

- El sistema no tiene soluciones (inconsistente).
- El sistema tiene solo una solución.
- El sistema tiene infinitas soluciones.

Actividad 1.10. ¿Qué puede asegurar respecto a las soluciones de los sitemas lineales cuyas matrices aumentadas son equivalentes a las siguientes formas escalonadas?

$$A = \begin{bmatrix} 1 & a & b & d & g \\ 0 & 2 & c & e & h \\ 0 & 0 & 5 & f & i \\ 0 & 0 & 0 & 3 & j \end{bmatrix}, \quad B = \begin{bmatrix} 1 & a & b & c & e \\ 0 & 0 & 2 & d & f \\ 0 & 0 & 0 & 3 & 0 \end{bmatrix}; \quad C = \begin{bmatrix} 1 & a & b & d & f \\ 0 & 2 & c & e & g \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

donde a, b, c, d, e, f, g, h, i y j son constantes reales.

Para los sistemas lineales homogéneos se tienen los siguientes resultados:

Teorema 1.10. (**Soluciones de un sistema lineal homogéneo**).

1. Un sistema lineal homogéneo tiene solo la solución trivial o bien un número infinito de soluciones.
2. Si un sistema lineal homogéneo tiene más incógnitas que ecuaciones entonces tiene infinitas soluciones.

Demostración. 1. Todo sistema lineal homogéneo es consistente, ya que tiene al menos la solución trivial, luego, de acuerdo al Teorema 1.9 tiene una única solución o infinitas soluciones. Si tiene solo una solución, ésta será la solución trivial y si no, tendrá un número infinito de soluciones.

2. Si un sistema tiene más incógnitas (variables) que ecuaciones entonces tiene variables libres, ya que el número de variables delanteras no puede exceder al número de ecuaciones. Como el sistema es consistente (dado que es homogéneo) y tiene variables libres, entonces tiene infinitas soluciones de acuerdo al Teorema 1.8.

□

Observación 1.5. Cuando se dice que un sistema lineal homogéneo tiene soluciones no triviales debe entenderse que tiene otras (infinitas) soluciones, además de la trivial.

Actividad 1.11. Determine si los siguientes sistemas homogéneos tienen soluciones no triviales:

$$\left\{ \begin{array}{rrrcl} x_1 & +x_2 & +x_3 & = & 0 \\ 2x_1 & -x_2 & -x_3 & = & 0 \\ x_1 & +2x_2 & -x_3 & = & 0 \\ 4x_1 & -x_2 & +x_3 & = & 0 \\ 2x_1 & +3x_2 & & = & 0 \end{array} \right., \quad \left\{ \begin{array}{rcl} x_1 + x_2 + x_3 & = & 0 \\ 2x_1 - x_2 - x_3 & = & 0 \\ 5x_1 + x_2 - x_3 & = & 0 \\ 4x_1 - x_2 + x_3 & = & 0 \end{array} \right., \quad \left\{ \begin{array}{rrrrcl} x_1 & +x_2 & +x_3 & +5x_4 & = & 0 \\ 2x_1 & -x_2 & -x_3 & +20x_4 & = & 0 \\ 77x_1 & +6x_2 & +45x_3 & -x_4 & = & 0 \end{array} \right..$$

Se terminará la sección considerando el siguiente ejemplo:

Ejemplo 1.28. Sean los sistemas lineales

$$\left\{\begin{array}{rcl} x+2y+3z & = & 2 \\ 2x+4y+6z & = & 4 \end{array}\right. \quad \text{y} \quad \left\{\begin{array}{rcl} x+2y+3z & = & 2 \\ 2x+4y+6z & = & 5 \end{array}\right.$$

que tienen la misma matriz de coeficientes.

El primer sistema es consistente mientras que el segundo no lo es. Las correspondientes matrices aumentadas se reducen a las siguientes formas escalonadas:

$$\begin{bmatrix} 1 & 2 & 3 & : & 2 \\ 0 & 0 & 0 & : & 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 2 & 3 & : & 2 \\ 0 & 0 & 0 & : & 1 \end{bmatrix},$$

respectivamente.

La matriz de coeficientes de ambos sistemas, $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$, no tiene una posición pivote en el segundo renglón, entonces cualquiera sea el vector de constantes $\mathbf{b}$ el último renglón de cualquier forma escalonada de $[A : \mathbf{b}]$ será de la forma $\begin{bmatrix} 0 & 0 & 0 & : & c \end{bmatrix}$ para algún $c \in \mathbb{R}$. El sistema lineal será consistente o inconsistente dependiendo del valor de c: si $c = 0$ el sistema es consistente mientras que es inconsistente si $c \neq 0$.◀

El ejemplo anterior despierta una inquietud: dada una matriz A ¿podrá suceder que un sistema lineal con dicha matriz de coeficientes sea siempre consistente?. Es claro que para que esto ocurra las formas escalonadas de A no deben tener renglones ceros, es decir A debe tener un pivote por renglón. Se tiene así el siguiente teorema:

Teorema 1.11. Sea A una matriz de $m \times n$. El sistema lineal de matriz aumentada $[A : \mathbf{b}]$ es consistente para todo $\mathbf{b} \in \mathbb{R}^m$ si y solo si ninguna forma escalonada de A tiene un renglón cero; esto es A tiene m posiciones pivote (un pivote por cada renglón).

Obsérvese que si una matriz A de $m \times n$ tiene m posiciones pivote necesariamente el número de renglones, m, debe ser menor o igual al número de columnas, n. Es decir, el número de ecuaciones del sistema no debe exceder al número de incógnitas.

Actividad 1.12. ¿Cuáles de las siguientes matrices son matrices de coeficientes de sistemas lineales que siempre resultan consistentes?

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ -1 & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 2 & 1 & -2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 & 0 \\ 3 & 4 & -1 & 1 \\ 0 & -1 & 2 & 2 \\ 4 & 6 & 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 \end{bmatrix}.$$

Ejercicios 1.4

1. Resuelva los siguientes sistemas de ecuaciones lineales:

$$a)\ \begin{cases} x_1 & & +x_3 & +x_4 & = & -5 \\ x_1 & & -x_3 & +x_4 & = & -1 \\ x_1 & +x_2 & +x_3 & +x_4 & = & -3 \\ 2x_1 & & +2x_3 & & = & -2 \end{cases}$$

$$b)\ \begin{cases} & y & & -z & +t & = & 5 \\ x & & & & -t & = & -2 \\ & -y & & +z & -t & = & -5 \\ & y & & -z & +t & = & 5 \\ & -y & -w & & & = & -1 \end{cases}$$

$$c)\ \begin{cases} & 2y & -3z & = & 9 \\ -x & +3y & +4z & = & -2 \\ 3x & -5y & -18z & = & 0 \end{cases}$$

$$d)\ \begin{cases} x_1 & +2x_2 & +3x_3 & +4x_4 & = & 0 \\ 2x_1 & +2x_2 & +3x_3 & +4x_4 & = & 0 \\ 3x_1 & +3x_2 & +3x_3 & +4x_4 & = & 0 \\ 4x_1 & +4x_2 & +4x_3 & +4x_4 & = & 0 \\ x_1 & +x_2 & +x_3 & -2x_4 & = & 0 \end{cases}$$

$$e)\ \begin{cases} 2x_1 & -x_2 & & -2x_4 & = & 3 \\ x_1 & -2x_2 & +x_3 & -x_4 & = & 1 \\ 2x_1 & -x_2 & -x_3 & +x_4 & = & 0 \\ 3x_1 & -3x_2 & & & = & 1 \\ x_1 & +x_2 & -x_3 & -x_4 & = & 2 \end{cases}$$

$$f)\ \begin{cases} -x & +y & -3z & +w & = & 9 \\ 2x & -5y & +z & +9w & = & 0 \\ -4x & +7y & -7z & -7w & = & 18 \end{cases}$$

2. **El problema de los celulares.** Una fábrica de celulares elabora cuatro modelos de teléfonos móviles: AA1, AA2, AA3 y B1. Una vez fabricados, se requiere de la inspección de los mismos en algunos de los sectores S1, S2, S3 y S4. El modelo AA1 requiere de 5 horas de inspección en el sector S1, 4 horas en S2, 1 hora en S3 y 3 horas en S4. El modelo AA2 requiere de 2 horas de inspección en cada uno de los cuatro sectores. El modelo AA3 requiere de 3 horas en el sector S2, 1 hora en S3 y 5 horas en S4. Por último, el modelo más nuevo B1 requiere de 1 hora en S1, 7 horas en S2 y 10 horas en S3. Si el sector S1 dispone de 230 horas semanales de trabajo, S2 de 369 horas, S3 de 299 horas y S4 de 231 horas, ¿cuántos celulares de cada modelo se pueden inspeccionar por semana?

3. **Elaboración de una dieta.** Por lo general, el empaque de un cereal indica el número de calorías y las cantidades de proteínas, carbohidratos y grasa contenidas en una porción del producto. Para dos marcas de cereales, se dan las cantidades en el Cuadro 1.1.

Nutriente	Cereal X	Cereal Y
Calorías	110	130
Proteínas (g)	4	3
Carbohidratos (g)	20	18
Grasa (g)	2	5

Cuadro 1.1: Cantidades de nutrientes de cada cereal

Suponga que se preparará una mezcla de esos cerales, de manera que contenga exactamente 295 calorías, 9g de proteínas, 48g de carbohidratos y 8g de grasa. Determine si es posible preparar la mezcla deseada de los dos cereales.

4. **Flujo de redes.** Los sitemas de ecuaciones lineales son utilizados también para estudiar el flujo de alguna cantidad a través de una red. Por ejemplo, los planificadores urbanos y los ingenieros de tráfico monitorizan el patrón de flujo de tránsito en una rejilla de las calles de una ciudad.

Una red consiste en un conjunto de puntos llamados nodos, con líneas o arcos llamados ramas, que conectan algunos o todos los nodos. Se indica la dirección y el sentido de flujo en cada rama, y la cantidad de flujo (o tasa) se denota con una variable.

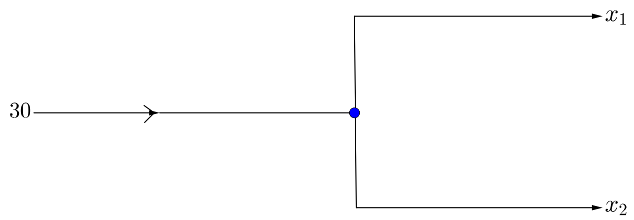


Figura 1.2: Nodo y ramas

La suposición básica en el flujo de red es que el flujo total en la red es igual al flujo total de salida de la red, y que el flujo total en un nodo es igual al flujo total de salida en dicho nodo. Por ejemplo, la Figura 1.2 muestra 30 unidades que fluyen por una rama hacia una unión; x_1 y x_2 denotan los flujos de salida del nodo a través de otras ramas.

Como el flujo se "conserva" en cada unión, entonces $x_1 + x_2 = 30$. En forma similar, el flujo en cada nodo se describe mediante una ecuación lineal. El problema de análisis de redes es determinar el flujo en cada rama cuando se tiene información parcial (como el flujo de entrada y salida de la red).

a) La red de la Figura 1.3 representa el flujo de tránsito (en vehículos por hora) en varias calles de un solo sentido en el centro de Baltimore en un dáa común, poco después del mediodía.

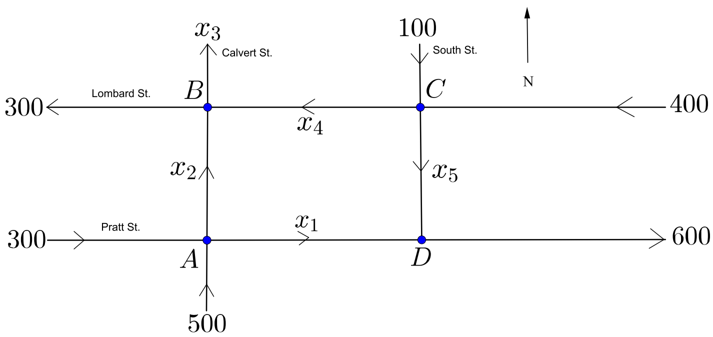


Figura 1.3: Red de tránsito (Baltimore)

Para determinar el patrón de flujo general para la red, se escriben las ecuaciones que describen el flujo. En cada intersección de las calles (nodos) se iguala el flujo de entrada al de salida como se muestra en el Cuadro 1.2.

Intersección	Flujo de entrada		Flujo de salida
A	$300 + 500$	$=$	$x_1 + x_2$
B	$x_2 + x_4$	$=$	$300 + x_3$
C	$100 + 400$	$=$	$x_4 + x_5$
D	$x_1 + x_5$	$=$	600

Cuadro 1.2: Flujos en intersecciones

También, el flujo total de entrada en la red $(500 + 300 + 100 + 400)$ es igual al flujo total de salida $(300 + x_3 + 600)$, que se simplifica a $x_3 = 400$.

Escriba el sistema de ecuaciones lineales que modeliza el problema y encuentre su solución para determinar el patrón de flujo general para la red. Determine el posible rango de las variables del problema.

Nota: Un flujo negativo en una rama de la red corresponde a un flujo en el sentido opuesto al indicado en el modelo. Como las calles en este problema son de un solo sentido, ninguna de las variables puede ser negativa.

b) Encuentre el patrón de flujo general de la red que se ilustra en la Figura 1.4.

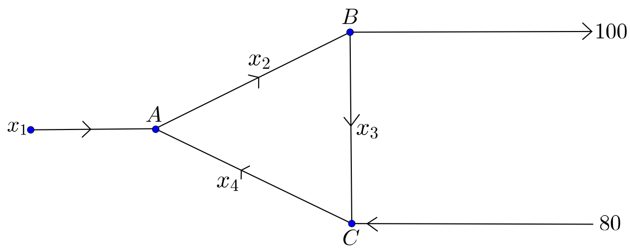


Figura 1.4: Red de tránsito

Suponiendo que todos los flujos son no negativos, ¿cuál es el valor más pequeño posible para x_4?

5. ¿Para qué valores de a, b y c el sistema lineal

$$\begin{cases} -x & +2y & -z & = & a \\ 3x & +y & +2z & = & b \\ 5x & -3y & +4z & = & c \end{cases}$$

admite al menos una solución?

6. Determine para qué valores de λ y μ el sistema lineal

$$\begin{cases} 2x & -\lambda y & +\mu z & = & 4 \\ x & & +z & = & 2 \\ x & +y & +z & = & 2 \end{cases}$$

es (a) inconsistente, (b) consistente con infinitas soluciones, (c) consistente con única solución.

7. Halle el o los valores de α para los cuales el sistema de ecuaciones lineales

$$\begin{cases} x & +3y & +z & = & \alpha \\ -x & +y & +\alpha z & = & -1 \\ 2x & +\alpha y & -z & = & \alpha+1 \end{cases}$$

admite (a) única solución, (b) infinitas soluciones, (c) ninguna solución. Si existe un único valor de α para el cual el sistema tiene única solución, resuelva el sistema. Idem en el caso de infinitas soluciones.

8. Analice la consistencia del sistema lineal según los valores de a y b.

$$\left\{\begin{array}{ccccccc} 2x_1 & + & x_2 & + & 2x_3 & = & 1 \\ x_1 & + & x_2 & + & x_3 & = & a \\ 2x_1 & + & 3x_2 & + & bx_3 & = & -1 \end{array}\right.$$

9. ¿Para qué valores de $\lambda \in \mathbb{R}$ los sistemas homogéneos:

$a)$ $\left\{\begin{array}{rcl} (\lambda-5)\,x - y & = & 0 \\ -x + (\lambda-5)\,y & = & 0 \end{array}\right.$

$\boxed{b)}$ $\left\{\begin{array}{ccc} (\lambda-1)\,x_1 - 3x_3 & = & 0 \\ -x_1 + (\lambda+1)\,x_2 - 2x_3 & = & 0 \\ x_1 - x_2 + (\lambda+2)\,x_3 & = & 0 \end{array}\right.$

tienen soluciones no triviales?

$\boxed{10.}$ **Resolución simultánea de sistemas lineales que tienen la misma matriz de coeficientes.** Es frecuente en la práctica tener que resolver sistemas lineales que tienen la misma matriz de coeficientes pero distintos términos independientes. Por ejemplo, los sistemas:

$$\left\{\begin{array}{rrrcr} x & -2y & +3z & = & -2 \\ -x & +4y & +5z & = & 7 \end{array}\right. \tag{1.4}$$

y

$$\left\{\begin{array}{rrrcr} u & -2v & +3w & = & 1 \\ -u & +4v & +5w & = & -4 \end{array}\right. \tag{1.5}$$

tienen la misma matriz de coeficientes $\begin{bmatrix} 1 & -2 & 3 \\ -1 & 4 & 5 \end{bmatrix}$ y vectores de términos independientes $\begin{bmatrix} -2 \\ 7 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ -4 \end{bmatrix}$.

En lugar de resolver cada sistema por separado, se los puede resolver simultáneamente aplicando el algoritmo de eliminación gaussiana, o el de Gauss-Jordan, a la matriz ampliada que tiene en el lado derecho las dos columnas que corresponden a los términos independientes. Si se aplica Gauss-Jordan a esta matriz se obtiene su forma escalonada reducida y en consecuencia los dos sistemas escalonados reducidos equivalentes a los originales.

$$\left[\begin{array}{rrrcrr} 1 & -2 & 3 & : & -2 & 1 \\ -1 & 4 & 5 & : & 7 & -4 \end{array}\right] \xrightarrow[R_2+R_1 \to R_2]{} \left[\begin{array}{rrrcrr} 1 & -2 & 3 & : & -2 & 1 \\ 0 & 2 & 8 & : & 5 & -3 \end{array}\right] \xrightarrow[\frac{1}{2}R_2 \to R_2]{}$$

$$\to \left[\begin{array}{rrrcrr} 1 & -2 & 3 & : & -2 & 1 \\ 0 & 1 & 4 & : & \frac{5}{2} & -\frac{3}{2} \end{array}\right] \xrightarrow[R_1+2R_2 \to R_1]{} \left[\begin{array}{rrrcrr} 1 & 0 & 11 & : & 3 & -2 \\ 0 & 1 & 4 & : & \frac{5}{2} & -\frac{3}{2} \end{array}\right].$$

Rápidamente se obtienen las soluciones:

Para el sistema (1.4)

$$\left\{\begin{array}{rcr} x & = & 3-11s \\ y & = & \frac{5}{2}-4s \\ z & = & s \end{array}\right., \; s \in \mathbb{R},$$

y para el sistema (1.5)

$$\left\{\begin{array}{rcr} u & = & -2-11t \\ v & = & -\frac{3}{2}-4t \\ w & = & t \end{array}\right. ,\ t\in\mathbb{R}.$$

En el caso de k sistemas lineales de n ecuaciones con n incógnitas cuya matriz de coeficientes tiene n pivotes, los k sistemas tienen solución única y la solución del i-ésimo sistema se lee directamente en la $(n+i)$-ésima columna de la matriz ampliada reducida equivalente.

Resuelva simultáneamente los sistemas:

$$a)\ \left\{\begin{array}{rrcr} 2x & +y & = & 5 \\ -x & +3y & = & 1 \\ x & +4y & = & 6 \end{array}\right. ,\quad \left\{\begin{array}{rrcr} 2x & +y & = & 5 \\ -x & +3y & = & 1 \\ x & +4y & = & 0 \end{array}\right. \quad \text{y} \quad \left\{\begin{array}{rrcr} 2x & +y & = & -1 \\ -x & +3y & = & 4 \\ x & +4y & = & 3 \end{array}\right.$$

$$b)\ \left\{\begin{array}{rrrcr} x & +2y & +z & = & 1 \\ x & +y & +3z & = & 2 \\ 2x & +y & +3z & = & 3 \end{array}\right. ,\quad \left\{\begin{array}{rrrcr} x & +2y & +z & = & 1 \\ x & +y & +3z & = & 0 \\ 2x & +y & +3z & = & 1 \end{array}\right. \quad \text{y} \quad \left\{\begin{array}{rrrcr} x & +2y & +z & = & 0 \\ x & +y & +3z & = & 1 \\ 2x & +y & +3z & = & -1 \end{array}\right.$$

$$c)\ \left\{\begin{array}{rrrcr} x_1 & -2x_2 & +x_3 & = & 1 \\ x_1 & +x_2 & -x_3 & = & 0 \\ 2x_1 & -x_2 & & = & 1 \end{array}\right. \quad \text{y} \quad \left\{\begin{array}{rrrcr} x_1 & -2x_2 & +x_3 & = & 1 \\ x_1 & +x_2 & -x_3 & = & 1 \\ 2x_1 & -x_2 & & = & 0 \end{array}\right.$$

11. Demuestre el Teorema 1.11.

1.5. Introducción a Scilab

SCILAB es un software libre que puede ser descargado de su página oficial http://www.scilab.org/. Para introducir matrices, los elementos de un renglón se separan por espacios y/o comas, y las columas se separan con un punto y coma.

Por ejemplo,

$$A=[1\ \ 2\ \ -5;3\ \ 2\ \ 0;-9\ \ 1\ \ 8;0\ \ 0\ \ -3] \quad \text{y} \quad b=[1;3;0;0]$$

produce la matriz

$$A=\begin{bmatrix} 1 & 2 & -5 \\ 3 & 2 & 0 \\ -9 & 1 & 8 \\ 0 & 0 & -3 \end{bmatrix},$$

y el vector columna

$$\mathbf{b}=\begin{bmatrix} 1 \\ 3 \\ 0 \\ 1 \end{bmatrix},$$

respectivamente. Hay que tener en cuenta que SCILAB distingue mayúsculas de minúsculas.

La notación $A(2,3)$ produce

$$0$$

que es el elemento en la posición $(2, 3)$ de la matriz A. La notación $A(2, :)$ produce

$$3 \quad 2 \quad 0$$

que son los elementos del segundo renglón de la matriz A. Y la notación $A(:, 3)$ produce

$$\begin{array}{r} -5 \\ 0 \\ 8 \\ -3 \end{array}$$

que son los elementos de la tercera columna de la matriz A.

La notación $B = [A\ b]$ produce la matriz aumentada

$$B = \begin{bmatrix} 1 & 2 & -5 & 1 \\ 3 & 2 & 0 & 3 \\ -9 & 1 & 8 & 0 \\ 0 & 0 & -3 & 1 \end{bmatrix}.$$

Las operaciones elementales entre renglones de una matriz A se indican de la siguiente manera:

- Intercambio: para intercambiar los renglones i y j se utiliza la notación

$$A([i\ j], :) = A([j\ i], :).$$

- Escalamiento: para multiplicar el renglón i por una constante no nula c se utiliza la notación

$$A(i, :) = c * A(i, :).$$

- Eliminación: para sumarle al renglón i el renglón j multiplicado por la constante c se indica

$$A(i, :) = A(i, :) + c * A(j, :).$$

El comando "rref" produce la forma escalonada reducida de una matriz. Por ejemplo, si quisiéramos determinar la forma escalonada reducida de la matriz A de nuestro ejemplo, lo indicamos con rref(A) y produce

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Ejercicios 1.5

1. Sean

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 \\ -1 & -2 & -1 & 3 \\ -2 & -3 & 0 & 3 \\ 1 & 4 & 5 & -9 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 9 \\ 1 \\ -1 \\ -7 \end{bmatrix} \quad \text{y} \quad \mathbf{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \\ 3 \end{bmatrix}$$

a) Ingrese la matriz A y los vectores columna $\mathbf{b}$ y $\mathbf{c}$.

b) Obtenga la forma escalonada reducida de cada una de las matrices aumentadas $[A : \mathbf{b}]$ y $[A : \mathbf{c}]$. Luego resuelva los sistemas lineales correspondientes que tienen a dichas matrices como matrices aumentadas.

c) Verifique los resultados obtenidos en (b) haciendo uso del comando "rref".

2. Use el comando "rref" para verificar las soluciones de los sistemas lineales de los ejercicios 1 y 10 de la sección 1.4.

3. **Ajuste polinomial de curvas.** Suponga que n puntos del plano xy

$$(x_1, y_1), (x_2, y_2), (x_3, y_3), ..., (x_n, y_n),$$

representan un conjunto de datos y se desea hallar una función polinómica de grado $n-1$:

$$p(x) = a_0 + a_1 x + a_2 x^2 + ... + a_{n-1} x^{n-1}$$

cuya gráfica contenga todos los puntos dados. El conjunto de datos permite calcular los coeficientes de la función polinómica p.

Halle la curva que contiene a los puntos dados en la siguiente tabla:

x	-1	$-0,5$	0	0,5	1	1,5	2	2,5	3
y	-5	$-0,8$	1	1,3	1	0,6	1	2,8	7

Capítulo 2

MATRICES

2.1. Vectores de $\mathbb{R}^n$

En el capítulo anterior se describió una solución de un sistema lineal con n incógnitas como una n-upla ordenada de números reales. Estas n-uplas son llamadas *vectores*.

> **Definición (Vector de n componentes).** Un *vector* de n componentes es una n-upla ordenada de n números. Puede representarse como un *vector renglón* o un *vector columna*.
>
> Un *vector renglón* de n componentes se escribe de la siguiente manera:
>
> $$(c_1, c_2, ..., c_n),$$
>
> mientras que un *vector columna* se expresa:
>
> $$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

En este texto las componentes de un vector serán números reales. En ambas expresiones c_1 es la primer componente, c_2 la segunda,..., c_k la k-ésima componente. Cualquier vector cuyas componentes son todas cero, se denomina **vector cero**.

Ejemplo 2.1. $(1, 0, 3, -7)$ es un vector renglón de 4 componentes, $\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ es un vector columna de 3 componentes y $(0, 0, 0, 0, 0, 0)$ es el vector renglón cero de 6 componentes. ◄

A lo largo del texto los vectores se representarán con letras minúsculas negritas como $\mathbf{u}$, $\mathbf{v}$, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{x}$, etcétera. Un vector cero se denota por $\mathbf{0}$.

Los vectores de n componentes que tienen un 1 en la $i-$ésima componente y 0 en el resto de las componentes los simbolizamos $\mathbf{e}_i$, es decir

$$\mathbf{e}_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \leftarrow i\text{-ésima componente.}$$

Por ejemplo, los vectores de 3 componentes $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$ son, respectivamente:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Definición (El símbolo $\mathbb{R}^n$). Al conjunto de todos los vectores de n componentes reales se lo representa con $\mathbb{R}^n$.

Los vectores de $\mathbb{R}^2$ y $\mathbb{R}^3$ pueden interpretarse geométricamente como puntos en el plano y como puntos en el espacio, respectivamente.

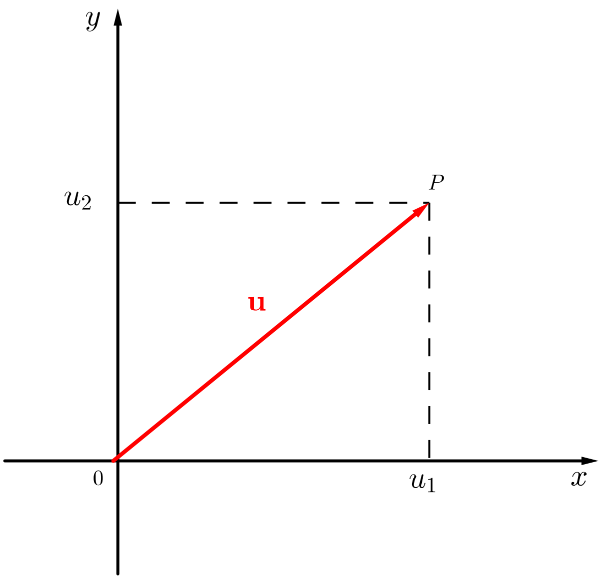


Figura 2.1: Representación geométrica de un vector de dos componentes

Cualquier vector de dos componentes $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ ó $\mathbf{u} = (u_1, u_2)$ puede representarse gráficamente con el punto P del plano cuyas coordenadas son (u_1, u_2). También al vector $\mathbf{u}$ se lo considera como la flecha que comienza en el punto origen $(0, 0)$ y cuya punta es el punto P, como se observa en la Figura 2.1. El vector cero de $\mathbb{R}^2$ geométricamente es el punto origen de coordenadas $(0, 0)$.

Similarmente se interpretan geométricamente a los vectores de $\mathbb{R}^3$.

Definición (Igualdad de vectores). Dos vectores $\mathbf{u}$ y $\mathbf{v}$ con el mismo número de componentes son *iguales*, y se expresa $\mathbf{u} = \mathbf{v}$, si sus componentes respectivas son iguales. Los vectores que tienen distinto número de componentes nunca son iguales.

Ejemplo 2.2. La igualdad

$$\begin{bmatrix} 1 \\ a+b \\ c \\ 3 \end{bmatrix} = \begin{bmatrix} a \\ -1 \\ 8 \\ c+d \end{bmatrix}$$

es válida si y solo si $a = 1$, $b = -2$, $c = 8$ y $d = -5$. ◀

Suma de vectores y multiplicación por escalar

Definición (Suma de vectores de $\mathbb{R}^n$). La *suma* de dos vectores $\mathbf{u}$ y $\mathbf{v}$ de $\mathbb{R}^n$ es el vector de $\mathbb{R}^n$, expresado por $\mathbf{u} + \mathbf{v}$, que se obtiene de sumar las correspondientes componentes.

Así, si $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, entonces $\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$. Análogamente:

Si $\mathbf{u} = (u_1, u_2, ..., u_n)$ y $\mathbf{v} = (v_1, v_2, ..., v_n)$, entonces $\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, ..., u_n + v_n)$.

Por ejemplo, $\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 7 \end{bmatrix}$.

La suma $\mathbf{u} + \mathbf{v}$ de dos vectores de $\mathbb{R}^2$ o $\mathbb{R}^3$ se representa en forma geométrica como la flecha diagonal del paralelogramo determinado por los vectores $\mathbf{u}$ y $\mathbf{v}$ como se observa en la Figura 2.2.

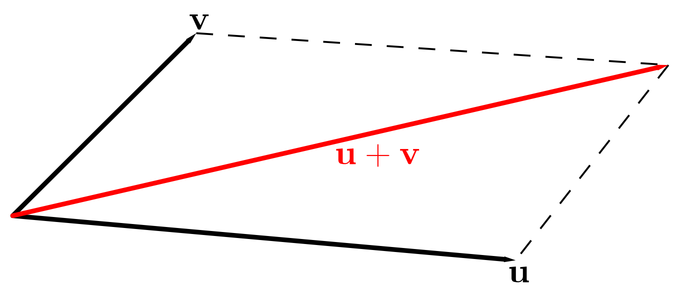


Figura 2.2: Suma de vectores

Definición (Producto de un vector por un escalar). El *producto de un vector* $\mathbf{u}$ de $\mathbb{R}^n$ *por un escalar* (número real) k es el vector de $\mathbb{R}^n$, expresado por $k\mathbf{u}$, cuyas componentes se obtienen multiplicando por k las correspondientes componentes de $\mathbf{u}$.

El vector $k\mathbf{u}$ se dice que es un **múltiplo escalar** del vector $\mathbf{u}$.

Asi, si $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ la multiplicación por el escalar k es el vector

$$k\mathbf{u} = \begin{bmatrix} ku_1 \\ ku_2 \\ \vdots \\ ku_n \end{bmatrix}.$$

Análogamente si $\mathbf{u} = (u_1, u_2, ..., u_n)$ la multiplicación por el escalar k es el vector

$$k\mathbf{u} = (ku_1, ku_2, ..., ku_n).$$

Por ejemplo, $(-3)\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}$ y $2\begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \\ -2 \end{bmatrix}$.

El vector $(-1)\mathbf{v}$ se llama **opuesto de v** y se representa $-\mathbf{v}$. Entonces si $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, el vector opuesto de $\mathbf{v}$ es

$$-\mathbf{v} = (-1)\mathbf{v} = \begin{bmatrix} (-1)v_1 \\ (-1)v_2 \\ \vdots \\ (-1)v_n \end{bmatrix} = \begin{bmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{bmatrix}.$$

Por ejemplo, si $\mathbf{u} = \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix}$ entonces $-\mathbf{u} = \begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix}$.

Dados $\mathbf{u}$ y $\mathbf{v}$ vectores de $\mathbb{R}^n$, se define la **diferencia** $\mathbf{u} - \mathbf{v}$ como la operación $\mathbf{u} + (-\mathbf{v})$. Entonces si $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, la diferencia $\mathbf{u} - \mathbf{v}$ es

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v}) = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{bmatrix} = \begin{bmatrix} u_1 + (-v_1) \\ u_2 + (-v_2) \\ \vdots \\ u_n + (-v_n) \end{bmatrix} = \begin{bmatrix} u_1 - v_1 \\ u_2 - v_2 \\ \vdots \\ u_n - v_n \end{bmatrix}.$$

Por ejemplo $\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} - \begin{bmatrix} 3 \\ -5 \\ -2 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ 2 \end{bmatrix}$.

Geométricamente la multiplicación $k\mathbf{u}$ de un vector de $\mathbb{R}^2$ o $\mathbb{R}^3$ por el escalar k es la flecha escalada por el factor k. Si $k > 0$ entonces la flecha $k\mathbf{u}$ apunta en el mismo sentido que la flecha $\mathbf{u}$, en cambio si $k < 0$ la flecha apunta en sentido opuesto a la flecha $\mathbf{u}$. Si $k = 0$ el vector $k\mathbf{u}$ es el vector cero. Si $|k| > 1$ entonces la flecha $k\mathbf{u}$ es una expansión de la flecha $\mathbf{u}$. Si $0 < |k| < 1$ entonces la flecha $k\mathbf{u}$ es una contracción de la flecha $\mathbf{u}$. Observe la Figura 2.3.

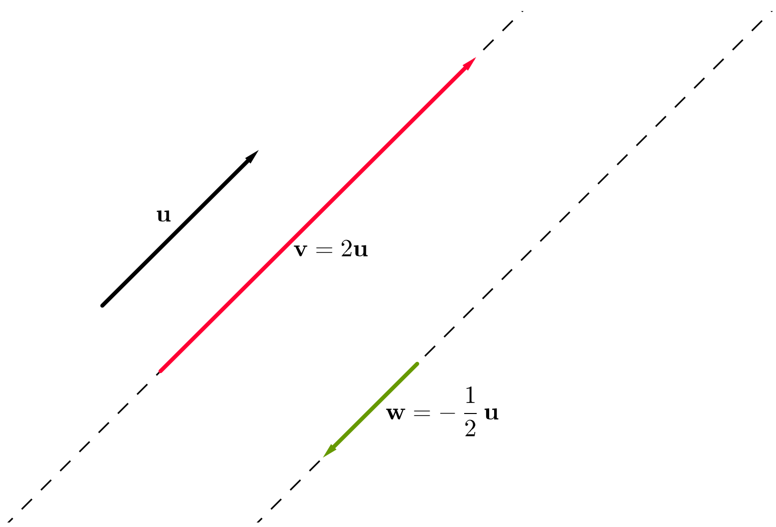


Figura 2.3: Multiplicación de un vector por un escalar

La suma de vectores y la multiplicación por un escalar, satisfacen algunas propiedades fundamentales que se resumen en el siguiente teorema:

Teorema 2.1. Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ cualesquiera vectores de $\mathbb{R}^n$ y sean a y b dos escalares cualesquiera. Entonces se satisfacen las siguientes propiedades:

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ [Conmutativa]
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ [Asociativa]
3. $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$ [Elemento neutro o identidad]
4. $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$ [Existencia de elementos opuestos]
5. $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ [Distributiva]
6. $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$ [Distributiva]
7. $a(b\mathbf{u}) = (ab)\mathbf{u} = b(a\mathbf{u})$
8. $1\mathbf{u} = \mathbf{u}$

Demostración. solo comprobaremos la propiedad 6. Se dejan las demás demostraciones como ejercicio.

Sean el vector $\mathbf{u} = (u_1, u_2, ..., u_n)$ y los escalares a y b. Entonces:

$$\begin{aligned} (a+b)\,\mathbf{u} &= ((a+b)\,u_1, (a+b)\,u_2, ..., (a+b)\,u_n) = (au_1 + bu_1, au_2 + bu_2, ..., au_n + bu_n) = \\ &= (au_1, au_2, ..., au_n) + (bu_1, bu_2, ..., bu_n) = a\mathbf{u} + b\mathbf{u}. \end{aligned}$$

□

Actividad 2.1. Demuestre las siguientes propiedades:

1. $0\mathbf{u} = \mathbf{0}$ para todo vector $\mathbf{u}$ de $\mathbb{R}^n$.

2. $-(a\mathbf{u}) = (-a)\,\mathbf{u}$ para todo vector $\mathbf{u}$ de $\mathbb{R}^n$ y escalar a.

El teorema anterior pueda aplicarse para resolver ciertas ecuaciones que involucran vectores. Por ejemplo, sean $\mathbf{v}$ y $\mathbf{u}$ dos vectores de $\mathbb{R}^n$ y se desea determinar el vector $\mathbf{x}$ que satisface la ecuación $2\mathbf{x} + 4\mathbf{v} = 3\mathbf{u}$. Entonces:

$$\begin{aligned} 2\mathbf{x} + 4\mathbf{v} &= 3\mathbf{u}, \\ (2\mathbf{x} + 4\mathbf{v}) + (-(4\mathbf{v})) &= 3\mathbf{u} + (-(4\mathbf{v})), \\ 2\mathbf{x} + (4\mathbf{v} + (-(4\mathbf{v})) &= 3\mathbf{u} + (-(4\mathbf{v})), \\ 2\mathbf{x} + \mathbf{0} &= 3\mathbf{u} + (-4)\,\mathbf{v}, \\ 2\mathbf{x} &= 3\mathbf{u} + (-4)\,\mathbf{v}, \\ \frac{1}{2}(2\mathbf{x}) &= \frac{1}{2}(3\mathbf{u} - 4\mathbf{v}), \\ \left(\frac{1}{2}2\right)\mathbf{x} &= \frac{1}{2}(3\mathbf{u} - 4\mathbf{v}), \\ 1\mathbf{x} &= \frac{1}{2}(3\mathbf{u} - 4\mathbf{v}), \\ \mathbf{x} &= \frac{1}{2}(3\mathbf{u} - 4\mathbf{v}). \end{aligned}$$

Actividad 2.2. Indique qué propiedades se aplicaron sucesivamente para resolver la ecuación anterior.

Matriz como sucesión de vectores columna

En algunas situaciones es de gran utilidad representar a una matriz como una sucesión de vectores. Por ejemplo, la matriz A puede escribirse

$$A = \begin{bmatrix} -3 & 0 & 2 & -5 \\ 0 & -1 & 0 & 0 \\ 1 & 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{bmatrix},$$

donde $\mathbf{a}_1 = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}$, $\mathbf{a}_3 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$ y $\mathbf{a}_4 = \begin{bmatrix} -5 \\ 0 \\ 0 \end{bmatrix}$.

En general, cualquier matriz A de tamaño $m \times n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix},$$

puede expresarse como la sucesión de los n vectores columna de m componentes $\mathbf{a}_1$, $\mathbf{a}_2$,...$\mathbf{a}_n$:

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix},$$

donde $\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$ para todo $j = 1, 2, ..., n$.

En Álgebra y Geometría Analítica se estudió que cualquier vector del plano puede escribirse como combinación lineal de dos vectores no nulos ni paralelos. También, se probó que todo vector del espacio se puede escribir como combinación lineal de tres vectores no nulos ni coplanares. Definiremos la noción de *combinación lineal* para cualquier conjunto de vectores de $\mathbb{R}^n$.

Definición (Combinación lineal)**.** Una *combinación lineal* de los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_k$ de $\mathbb{R}^n$ es un vector de $\mathbb{R}^n$ de la forma

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k,$$

donde c_1, c_2,..., c_k son escalares cualesquiera que son llamados *coeficientes* de la combinación lineal.

Así, un vector $\mathbf{u}$ de $\mathbb{R}^n$ es una combinación lineal de ciertos vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_k$ de $\mathbb{R}^n$ si y solo si existen escalares c_1, c_2,..., c_k tal que

$$\mathbf{u} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k.$$

Ejemplo 2.3. El vector $\mathbf{u} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ es combinación lineal de los vectores $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. En efecto,

$$\begin{bmatrix} -2 \\ 3 \end{bmatrix} = -2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Trivialmente $\mathbf{u}$ no es combinación lineal de los vectores $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -3 \\ 0 \end{bmatrix}$ dado que no existen escalares c_1 y c_2 tales que

$$c_1 \begin{bmatrix} 2 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -3 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}. \blacktriangleleft$$

Ejemplo 2.4. Para determinar si $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ es combinación lineal de los vectores $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ -3 \\ -5 \end{bmatrix}$ y $\mathbf{v}_4 = \begin{bmatrix} -4 \\ 1 \\ 4 \end{bmatrix}$ debe analizarse si existen escalares c_1, c_2, c_3 y c_4 tales que satisfagan la siguiente ecuación en las incógnitas x_1, x_2, x_3 y x_4:

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + x_4\mathbf{v}_4 = \mathbf{u}, \tag{2.1}$$

esto es:

$$x_1 \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ -3 \\ -5 \end{bmatrix} + x_4 \begin{bmatrix} -4 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

Operando en el primer miembro se obtiene:

$$\begin{bmatrix} x_1 - 2x_2 + x_3 - 4x_4 \\ x_1 + x_2 - 3x_3 + x_4 \\ -x_1 + 5x_2 - 5x_3 + 4x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix},$$

y, por igualdad de vectores, la ecuación (2.4) es equivalente al sistema lineal

$$\left\{ \begin{array}{rrrrcl} x_1 & -x_2 & +x_3 & -4x_4 & = & 1 \\ x_1 & +x_2 & -3x_3 & +x_4 & = & 2 \\ -x_1 & +5x_2 & -5x_3 & +4x_4 & = & 0 \end{array} \right.$$

con lo cual el problema se reduce a estudiar la consistencia de este sistema. Aplicando eliminación gaussiana a la matriz aumentada se obtiene una forma escalonada que no tiene renglones de la forma $\begin{bmatrix} 0 & 0 & ... & 0 & : & b \end{bmatrix}$ con $b \neq 0$:

$$\begin{bmatrix} 1 & -2 & 1 & -4 & : & 1 \\ 1 & 1 & -3 & 1 & : & 2 \\ -1 & 5 & -5 & 4 & : & 0 \end{bmatrix} \xrightarrow[\substack{R_2 - R_1 \to R_2 \\ R_3 + R_1 \to R_3}]{} \begin{bmatrix} 1 & -2 & 1 & -4 & : & 1 \\ 0 & 3 & -4 & 5 & : & 1 \\ 0 & 3 & -4 & 0 & : & 1 \end{bmatrix} \to$$

$$\xrightarrow[R_3 - R_2 \to R_3]{} \begin{bmatrix} \boxed{1} & -2 & 1 & -4 & : & 1 \\ 0 & \boxed{3} & -4 & 5 & : & 1 \\ 0 & 0 & 0 & \boxed{-5} & : & 0 \end{bmatrix}.$$

Luego, el sistema lineal es consistente por lo que el vector $\mathbf{u}$ resulta ser combinación lineal de los vectores $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{v}_4$. Más aún, el sistema tiene infinitas soluciones y por lo tanto hay una infinidad de maneras de expresar al vector $\mathbf{u}$ como combinación lineal de dichos vectores.

La solución general del sistema es

$$\left\{ \begin{array}{l} x_1 = \frac{5}{3} + \frac{5}{3}t \\ x_2 = \frac{1}{3} + \frac{4}{3}t \\ x_3 = t \\ x_4 = 0 \end{array} \right., \ t \in \mathbb{R}.$$

Así, por ejemplo, si $t = 0$ se obtiene $c_1 = \frac{5}{3}$, $c_2 = \frac{1}{3}$, $c_3 = 0$ y $c_4 = 0$:

$$\frac{5}{3} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ -3 \\ -5 \end{bmatrix} + 0 \begin{bmatrix} -4 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

y si $t = 1$ se obtiene $c_1 = \frac{10}{3}$, $c_2 = \frac{5}{3}$, $c_3 = 1$ y $c_4 = 0$:

$$\frac{10}{3}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + \frac{5}{3}\begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} + 1\begin{bmatrix} 1 \\ -3 \\ -5 \end{bmatrix} + 0\begin{bmatrix} -4 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

En general, las infinitas maneras de escribir al vector $\mathbf{u}$ como combinación lineal de los vectores $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$ y $\boldsymbol{v}_4$ está dada por

$$\left(\frac{5}{3} + \frac{5}{3}t\right)\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + \left(\frac{1}{3} + \frac{4}{3}t\right)\begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} + t\begin{bmatrix} 1 \\ -3 \\ -5 \end{bmatrix} + 0\begin{bmatrix} -4 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad t \in \mathbb{R}. \blacktriangleleft$$

Ejercicios 2.1

$\boxed{1.}$ Indique para qué valores de x, y y z los vectores

$$\mathbf{u} = \begin{bmatrix} -x - y + z \\ 2x + y \\ -y - 3z \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$$

son iguales.

$\boxed{2.}$ Efectúe las siguientes operaciones, en el caso en que sea posible:

a) $\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} + \begin{bmatrix} -3 \\ 5 \\ 0 \end{bmatrix}$

b) $(-1, 0, 3, 4) - (3, -4, 5, 0)$

c) $-3\begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}$

d) $-\frac{1}{2}(-2, 3, 1, 5) + 2(0, 0, 3, 4) - (-1, 2, 4, 0)$

e) $-4\begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix} + 5\begin{bmatrix} 0 \\ -2 \\ 4 \end{bmatrix}$

f) $(-1, 3, 4) + 2(0, 3)$

g) $-3(0, 0, 1, -1, 5) + 0(2, 3, 1, 8, -9)$

3. Demuestre las restantes propiedades del Teorema 2.1.

4. Sean $\mathbf{a} = \begin{bmatrix} 1 \\ -1 \\ 3 \\ 0 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \\ -1 \\ 2 \end{bmatrix}$ y $\mathbf{c} = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$. Obtenga, si es posible, el vector $\mathbf{x}$ tal que:

a) $\mathbf{x} = -2\mathbf{a} + \frac{1}{2}\mathbf{b} - \mathbf{c}$.

$\boxed{b)}$ $\mathbf{x} + 2\mathbf{b} - \mathbf{c} = 2\mathbf{a}$

c) $4\mathbf{x} - 2\mathbf{b} - \mathbf{c} = \mathbf{0}$

$\boxed{d)}$ $2\mathbf{a} - 3(-\mathbf{b} + \mathbf{c}) + \mathbf{x} = \mathbf{x} - \mathbf{c}$

5. Determine si el vector $\mathbf{b}$ es combinación lineal de los restantes vectores. En caso afirmativo, halle coeficientes de la combinación lineal.

$\boxed{a)}$ $\mathbf{b} = \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix}$; $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ y $\mathbf{a}_3 = \begin{bmatrix} 5 \\ -6 \\ -8 \end{bmatrix}$.

b) $\mathbf{b} = \begin{bmatrix} 4 \\ 0 \\ 5 \end{bmatrix}$; $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ y $\mathbf{a}_3 = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$.

c) $\mathbf{b} = (4, 0, 4)$; $\mathbf{a}_1 = (1, -1, 0)$, $\mathbf{a}_2 = (2, 1, 3)$, $\mathbf{a}_3 = (3, 2, 5)$ y $\mathbf{a}_4 = (1, 1, 2)$.

$\boxed{d)}$ $\mathbf{b} = \begin{bmatrix} -10 \\ -5 \\ 17 \\ -14 \end{bmatrix}$; $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -1 \\ 3 \\ 0 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$ y $\mathbf{a}_3 = \begin{bmatrix} 4 \\ 1 \\ -2 \\ -3 \end{bmatrix}$.

$\boxed{6.}$ Calcule el o los valores de a de manera que el vector $\mathbf{u} = \begin{bmatrix} a \\ -2 \\ 0 \end{bmatrix}$ sea una combinación lineal de los vectores $\mathbf{v} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 1 \\ 0 \\ a \end{bmatrix}$.

7. Sean $\mathbf{v}_1$, $\mathbf{v}_2$,...,$\mathbf{v}_k$ vectores de $\mathbb{R}^3$ y suponga que para cada $j = 1, 2, ..., k$ un objeto con masa m_j está posicionado en el extremo del vector $\mathbf{v}_j$. En Física llaman *masas puntuales* a esos objetos. La *masa total* del sistema de masas puntuales es $m = m_1 + m_2 + ... + m_k$. El *centro de gravedad* o (*centro de masa*) del sistema es

$$\mathbf{v} = \frac{1}{m} \left(m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2 + ... + m_k \mathbf{v}_k \right).$$

a) Calcule el centro de gravedad del sistema de cuatro masas puntuales cuyas posiciones y masas son: $\mathbf{v}_1 = (2, -2, 4)$ y $m_1 = 4$g; $\mathbf{v}_2 = (-4, 2, 3)$ y $m_2 = 2$g; $\mathbf{v}_3 = (4, 0, -2)$ y $m_3 = 3$g; $\mathbf{v}_4 = (1, -6, 0)$ y $m_4 = 5$g.

b) Una delgada placa triangular de densidad y grosor uniformes, tiene vértices en $\mathbf{v}_1 = (0, 1)$, $\mathbf{v}_2 = (8, 1)$ y $\mathbf{v}_3 = (2, 4)$. Si la masa total de la placa es de 6g, determine cómo distribuirla en los tres vértices de manera que el centro de gravedad sea $\mathbf{v} = (2, 2)$.

2.2. Operaciones matriciales

Recordemos que una matriz A de orden o tamaño $m \times n$ es un arreglo rectangular de mn números dispuestos en m filas o renglones y n columnas. Los números del arreglo son los elementos de la matriz.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = [a_{ij}].$$

Si todos los mn números son números reales, se dice que A es una **matriz real**. En este texto, salvo indicación expresa, todas las matrices serán matrices reales.

La matriz de $m \times n$ que tiene todos sus elementos ceros se denomina **matriz cero** y se simboliza $\mathbf{O}_{m\times n}$. Por ejemplo la matriz cero de 2×3 es

$$\mathbf{O}_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

y la matriz cero cuadrada de orden 4 se denota $\mathbf{O}_4$ y es

$$\mathbf{O}_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Cuando se sobreentiende el tamaño de la matriz, las matrices cero pueden denotarse simplemente por $\mathbf{O}$.

La matriz cuadrada de orden n que tiene solo unos en su diagonal principal y los restantes elementos son ceros se llama **matriz identidad** y se simboliza I_n. Por ejemplo la matriz identidad de orden 3 es

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Cuando se sobreentiende el tamaño de la matriz, una matriz identidad puede denotarse simplemente por I.

Observe que la primera, segunda y tercera columnas de I_3 son respectivamente los vectores $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$ de $\mathbb{R}^3$. En general I_n expresada en términos de sus columnas es $I_n = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \ldots & \mathbf{e}_n \end{bmatrix}$, donde la j-ésima columna de I_n es el vector $\mathbf{e}_j$ de $\mathbb{R}^n$ con $j = 1, 2, ..., n$.

Definición (Igualdad de matrices)**.** Dos matrices A y B son *iguales* si tienen el mismo tamaño y sus elementos correspondientes son iguales.

Ejemplo 2.5. $A = \begin{bmatrix} 1 & -2 \\ 1 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & -2 \\ a & b-5 \end{bmatrix}$ son iguales si y solo si $a = 1$ y $b = 5$.◄

Suma de matrices y multiplicación de una matriz por un escalar

Definición (Suma de matrices)**.** Sean $A = [a_{ij}]$ y $B = [b_{ij}]$ dos matrices de $m \times n$. La *suma* de A y B es la matriz de $m \times n$, denotada por $A + B$, dada por

$$A + B = [a_{ij} + b_{ij}].$$

Es decir, $A + B$ es la matriz de $m \times n$ que se obtiene de sumar los elementos correspondientes de A y B.

Ejemplo 2.6. $\begin{bmatrix} -2 & 3 & 0 \\ 1 & 4 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 1 \\ -3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 3 & 1 \\ -2 & 5 & 0 \end{bmatrix}$.◄

Definición (Producto de una matriz por un escalar)**.** Sean $A = [a_{ij}]$ una matriz de $m \times n$ y k un escalar. El *producto de la matriz A por el escalar k* es la matriz de $m \times n$, denotada por kA, dada por
$$kA = [ka_{ij}] .$$
Es decir, kA es la matriz de $m \times n$ que se obtiene de multiplicar todos los elementos de A por k.

La matriz kA se dice que es un **múltiplo escalar** de la matriz A.

Ejemplo 2.7. $3\begin{bmatrix} -1 & 3 \\ -2 & 5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -3 & 9 \\ -6 & 15 \\ 0 & 3 \end{bmatrix}$.◄

La matriz $(-1)A$ se llama **opuesta** de A y se representa por $-A$. Si $A = [a_{ij}]$ entonces
$$-A = (-1)A = [(-1)a_{ij}] = [-a_{ij}] .$$
Por ejemplo, si $A = \begin{bmatrix} -1 & 2 & 3 \\ 0 & -4 & 1 \\ 2 & -1 & 0 \end{bmatrix}$ entonces $-A = \begin{bmatrix} 1 & -2 & -3 \\ 0 & 4 & -1 \\ -2 & 1 & 0 \end{bmatrix}$.

Se define la **diferencia** $A - B$ como la operación $A + (-B)$. Si $A = [a_{ij}]$ y $B = [b_{ij}]$ son matrices de $m \times n$ entonces
$$A - B = A + (-B) = [a_{ij} + (-b_{ij})] = [a_{ij} - b_{ij}] ,$$
es la matriz de $m \times n$ que se obtiene de restar a cada elemento de A el correspondiente elemento de B. Por ejemplo:
$$\begin{bmatrix} -1 & 2 & 3 \\ 0 & -4 & 1 \\ 2 & -1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 1 \\ 3 & 5 & -1 \\ -4 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 2 \\ -3 & -9 & 2 \\ 6 & -1 & -2 \end{bmatrix}.$$

A continuación se presentan propiedades de la suma de matrices y de la multiplicación de un escalar por una matriz:

Teorema 2.2. Sean A, B y C matrices $m \times n$ y b y c escalares. Entonces se verifican las siguientes propiedades:

1. $A + B = B + A$ [Conmutativa]
2. $A + (B + C) = (A + B) + C$ [Asociativa]
3. $A + \mathbf{O} = \mathbf{O} + A = A$ [Elemento neutro o identidad]
4. $A + (-A) = (-A) + A = \mathbf{O}$ [Existencia de elementos opuestos]
5. $c(A + B) = cA + cB$ [Distributiva]
6. $(b + c)A = bA + cA$ [Distributiva]
7. $b\,(cA) = (bc)\,A = c\,(bA)$
8. $1A = A$

Demostración. Se probarán aquí las propiedades 1. y 5. Se dejan las demás demostraciones como ejercicio.

Sean las matrices $A = [a_{ij}]$ y $B = [b_{ij}]$ de $m \times n$. Entonces:

1. Por definición de suma de matrices y la conmutatividad de la suma en $\mathbb{R}$:

$$A + B = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = B + A.$$

5. Por definición de suma de matrices, de producto de una matriz por un escalar y por la propiedad distributiva del producto respecto a la suma en $\mathbb{R}$:

$$c\,(A + B) = [c\,(a_{ij} + b_{ij})] = [ca_{ij} + cb_{ij}] = [ca_{ij}] + [cb_{ij}] = cA + cB.$$

□

Actividad 2.3. Sean A una matriz de $m \times n$ y k un escalar. Demuestre:

1. $kA = \boldsymbol{O} \Leftrightarrow k = 0$ ó $A = \boldsymbol{O}$.
2. $-\,(kA) = (-k)\,A$.

El teorema pueda aplicarse para resolver ciertas ecuaciones que involucran matrices. Por ejemplo, sean A y B dos matrices de $m \times n$ y se desea determinar la matriz que satisface la ecuación $2X + 4B = 3A$.

$$\begin{aligned}
2X + 4B &= 3A, \\
(2X + 4B) + (-\,(4B)) &= 3A + (-\,(4B))\,, \\
2X + (4B + (-\,(4B)) &= 3A - 4B, \\
2X + \mathbf{O} &= 3A - 4B, \\
2X &= 3A - 4B, \\
\frac{1}{2}\,(2X) &= \frac{1}{2}\,(3A - 4B)\,, \\
\left(\frac{1}{2}2\right) X &= \frac{1}{2}\,(3A - 4B)\,, \\
1X &= \frac{1}{2}\,(3A - 4B)\,, \\
X &= \frac{1}{2}\,(3A - 4B)\,.
\end{aligned}$$

Actividad 2.4. Indique qué propiedades se aplicaron sucesivamente para resolver la ecuación anterior.

Producto punto y multiplicación de matrices

Definición (Producto punto)**.** El *producto punto* o *producto escalar* de dos vectores de n componentes $\mathbf{u}$ y $\mathbf{v}$ es el número real, denotado por $\mathbf{u}\cdot\mathbf{v}$, que se obtiene sumando los productos de las componentes correspondientes de dichos vectores.

Así, si $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, o bien, si $\mathbf{u} = (u_1, u_2, ..., u_n)$ y $\mathbf{v} = (v_1, v_2, ..., v_n)$, entonces

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + ... + u_n v_n = \sum_{i=1}^{n} u_i v_i.$$

Similarmente si uno de ellos es un vector columna y el otro un vector renglón:

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i.$$

Ejemplo 2.8. Si $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \\ 3 \\ 5 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ -1 \\ 0 \end{bmatrix}$, el producto punto de $\mathbf{u}$ y $\mathbf{v}$ es:

$$\mathbf{u} \cdot \mathbf{v} = 1(2) + (-2)(3) + 3(-1) + 5(0) = -7. \blacktriangleleft$$

Nota: El producto punto es una operación importante que se retomará en el Capítulo 5 y ha sido introducido en este momento por conveniencia ya que interviene en la multiplicación entre matrices.

Definición (Producto de matrices)**.** Si $A = [a_{ik}]$ es una matriz de $m \times p$ y $B = [b_{kj}]$ es una matriz de $p \times n$, el *producto* A por B, es la matriz de de $m \times n$ denotada por $AB = [c_{ij}]$ donde

$$c_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + ... + a_{ip} b_{pj} = \sum_{k=1}^{p} a_{ik} b_{kj}. \tag{2.2}$$

La ecuación (2.2) establece que el (i, j)-ésimo elemento de la matriz producto AB, es el producto punto del vector cuyas componentes son las del renglón i-ésimo de A por el vector cuyas componentes son las de la columna j-ésima de B.

$$\begin{array}{ccc} & & \text{col}_j(B)\downarrow \\ \text{ren}_i(A) \rightarrow & \begin{bmatrix} a_{11} & a_{12} & ... & a_{1p} \\ \vdots & \vdots & ... & \vdots \\ a_{i1} & a_{i2} & ... & a_{ip} \\ \vdots & \vdots & ... & \vdots \\ a_{m1} & a_{m2} & ... & a_{mp} \end{bmatrix} & \begin{bmatrix} b_{11} & b_{12} & ... & b_{1j} & ... & b_{1n} \\ b_{21} & b_{22} & ... & b_{2j} & ... & b_{2n} \\ \vdots & \vdots & ... & \vdots & \vdots & \vdots \\ b_{p1} & b_{p2} & ... & b_{pj} & ... & b_{pn} \end{bmatrix} = \end{array}$$

$$= \begin{bmatrix} c_{11} & c_{12} & ... & c_{1j} & ... & c_{1n} \\ \vdots & \vdots & ... & \vdots & \vdots & \vdots \\ c_{i1} & c_{i2} & ... & \boxed{c_{ij}} & ... & c_{in} \\ \vdots & \vdots & ... & \vdots & \vdots & \vdots \\ c_{m1} & c_{m2} & ... & c_{mj} & ... & c_{mn} \end{bmatrix}$$

$$c_{ij} = \text{ren}_i(A) \cdot \text{col}_j(B) = a_{i1} b_{1j} + a_{i2} b_{2j} + ... + a_{ip} b_{pj} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$

Observe que el producto AB está definido solo si el número de columnas de la matriz A es igual al número de renglones de la matriz B:

$$A_{m\times \boxed{p}} B_{\boxed{p}\times n} = AB_{m\times n}$$

Ejemplo 2.9. Dadas $A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 5 \end{bmatrix}$ y $B = \begin{bmatrix} -2 & 5 \\ 3 & -4 \\ 1 & 2 \end{bmatrix}$. Como A es una matriz de 2×3 y B es de 3×2 es posible efectuar AB y resultará una matriz de 2×2:

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 5 \end{bmatrix} \begin{bmatrix} -2 & 5 \\ 3 & -4 \\ 1 & 2 \end{bmatrix} = \\ &= \begin{bmatrix} 1(-2)+2(3)+(-1)(1) & 1(5)+2(-4)+(-1)(2) \\ 3(-2)+1(3)+5(1) & 3(5)+1(-4)+5(2) \end{bmatrix} = \\ &= \begin{bmatrix} 3 & -5 \\ 2 & 21 \end{bmatrix}. \blacktriangleleft \end{aligned}$$

Ejemplo 2.10. Sean las matrices $A = \begin{bmatrix} 1 & -2 & 3 \\ 3 & 0 & 2 \\ 2 & -1 & 3 \end{bmatrix}$ y $B = \begin{bmatrix} 2 & 3 \\ -1 & 0 \\ 3 & 2 \end{bmatrix}$.

Para calcular el elemento $(3,2)$ de la matriz AB se debe efectuar el producto punto del vector renglón 3 de la matriz A por el el vector columna 2 de la matriz B. Luego el elemento $(3,2)$ de AB es $2(3)+(-1)(0)+3(2) = 12.\blacktriangleleft$

Una disposición práctica para efectuar el producto AB se visualiza en este ejemplo:

Ejemplo 2.11. Sean las matrices $A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 5 \end{bmatrix}$ y $B = \begin{bmatrix} -2 & 5 \\ 3 & -4 \\ 1 & 2 \end{bmatrix}$.

En el lado inferior izquierdo se ubican los elementos de la matriz A y en el lado superior derecho los elementos de la matriz B. Haciendo el producto punto de cada vector renglón de A por la columna $j-$ésima de B se obtiene la columna $j-$ésima de AB. La matriz AB es la obtenida en el lado inferior derecho.

$$\begin{array}{ccc|cc} & & & -2 & 5 \\ & AB & & 3 & -4 \\ & & & 1 & 2 \\ \hline 1 & 2 & -1 & 3 & -5 \\ 3 & 1 & 5 & 2 & 21 \end{array}$$

Luego, $AB = \begin{bmatrix} 3 & -5 \\ 2 & 21 \end{bmatrix}. \blacktriangleleft$

Nota: La suma de matrices se define de manera natural, sumando elemento a elemento. Esto hace que la suma de matrices verifique las mismas propiedades que la suma de números reales.

La multiplicación de matrices requiere más cuidado que la suma, no está definida de la manera natural que se hubiese esperado (multiplicar elemento a elemento), por lo cual ciertas propiedades del producto de números reales no se verifican en el producto matricial. Para los números reales a y b siempre se tiene que $ab = ba$, sin embargo en el caso de las matrices AB y BA no son por lo general iguales.

Por definición, el producto AB se puede realizar solo si el número de columnas de A es igual al número de renglones de B. Si A es una matriz de $m \times p$ y B es una matriz de $p \times n$, AB es una matriz de $m \times n$ y con respecto a BA pueden darse distintas situaciones:

- Puede ocurrir que BA no esté definido. Por ejemplo, si A es de 2×3 y B es de 3×4, AB es una matriz de 2×4 mientras que el producto BA no es posible.
- En el caso en que BA también esté definido, pueden ocurrir dos situaciones:
  - AB y BA son de distinto tamaño y por lo tanto son matrices distintas. Por ejemplo, si A es de 2×3 y B es de 3×2 entonces AB es de orden 2 mientras que BA es de orden 3. Trivialmente $AB \neq BA$.
  - AB y BA son del mismo orden (esto ocurre cuando A y B son matrices cuadradas de igual orden). Sin embargo, en general, no resultan iguales.

    Por ejemplo si $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix}$ entonces

    $$AB = \begin{bmatrix} 0 & 2 \\ 3 & 3 \end{bmatrix} \quad \text{y} \quad BA = \begin{bmatrix} 2 & 2 \\ 4 & 1 \end{bmatrix}.$$

Se ha visto que aunque ambos productos AB y BA estén definidos no tienen por qué ser iguales. Puede concluirse entonces que **el producto de matrices no es conmutativo**.

Otras peculiaridades del producto de matrices son las siguientes:

Supongamos que A, B y C son matrices de tamaños tales que los productos matriciales que se indican están definidos:

- $AB = \mathbf{O}$ no implica $A = \mathbf{O}$ ó $B = \mathbf{O}$ (claramente, si $A = \mathbf{O}$ ó $B = \mathbf{O}$ entonces $AB = \mathbf{O}$).

  Por ejemplo, si $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ y $B = \begin{bmatrix} -2 & -4 & 2 \\ 1 & 2 & -1 \end{bmatrix}$ resulta $AB = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, siendo A y B matrices no cero.

- $CA = CB$ no implica $A = B$.

  Por ejemplo, para las matrices $C = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, $A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & -3 \end{bmatrix}$ y $B = \begin{bmatrix} 3 & 2 & -4 \\ 0 & 0 & 0 \end{bmatrix}$, resulta $CA = CB = \begin{bmatrix} 3 & 2 & -4 \\ 6 & 4 & -8 \end{bmatrix}$ siendo las matrices A y B distintas.

- $AC = BC$ no implica $A = B$.

Actividad 2.5. Proponga un ejemplo de tres matrices A, B y C tales que $AC = BC$ y $A \neq B$.

Multiplicación de una matriz por un vector: el producto $A\mathbf{c}$

Sea $A = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & ... & \vdots \\ a_{m1} & a_{m2} & ... & a_{mn} \end{bmatrix}$ una matriz de $m \times n$ y sea $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$ un vector columna de n componentes.

Como A es una matriz de $m \times n$ y $\mathbf{c}$ es una matriz de $n \times 1$, por definición de multiplicación de matrices, el producto $A\mathbf{c}$ es la siguiente matriz de $m \times 1$:

$$A\mathbf{c} = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & ... & \vdots \\ a_{m1} & a_{m2} & ... & a_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \mathrm{ren}_1(A) \cdot \mathbf{c} \\ \mathrm{ren}_2(A) \cdot \mathbf{c} \\ \vdots \\ \mathrm{ren}_m(A) \cdot \mathbf{c} \end{bmatrix} = \begin{bmatrix} a_{11}c_1 + a_{12}c_2 + ... + a_{1n}c_n \\ a_{21}c_1 + a_{22}c_2 + ... + a_{2n}c_n \\ \vdots \\ a_{m1}c_1 + a_{m2}c_2 + ... + a_{mn}c_n \end{bmatrix}.$$

La última expresión encontrada de $A\mathbf{c}$ puede escribirse como

$$c_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + c_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + ... + c_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Es decir

$$A\mathbf{c} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + ... + c_n\mathbf{a}_n,$$

donde $\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$ es la j-ésima columna de la matriz A. Entonces el producto de una matriz A de $m \times n$ por un vector columna, $\mathbf{c}$, de n componentes (matriz de $n \times 1$) es una combinación lineal de las columnas de A, cuyos coeficientes son las componentes (elementos) de $\mathbf{c}$.

Ejemplo 2.12. Si $A = \begin{bmatrix} -1 & 3 \\ -2 & 5 \\ 0 & 1 \end{bmatrix}$ y $\boldsymbol{c} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$, el producto $A\boldsymbol{c}$ dado como una combinación lineal de las columnas de A es

$$A\boldsymbol{c} = 3 \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix} + (-1) \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -6 \\ 0 \end{bmatrix} + \begin{bmatrix} -3 \\ -5 \\ -1 \end{bmatrix} = \begin{bmatrix} -6 \\ -11 \\ -1 \end{bmatrix}. \blacktriangleleft$$

Ejemplo 2.13 (Producto de una matriz por un vector $\mathbf{e}_k$). Si se multiplica una matriz A de $m \times n$ por el vector (columna) $\mathbf{e}_k$ de $\mathbb{R}^n$ se obtiene la k-ésima columna de A. Esto es:

$$A\mathbf{e}_k = \mathbf{a}_k,$$

donde $\mathbf{a}_k$ es la k-ésima columna de A. ◀

Actividad 2.6. Pruebe el resultado del Ejemplo 2.13. Sugerencia: expresa $A\mathbf{e}_k$ como una combinación lineal de las columnas de A.

Ejemplo 2.14 (Producto de la matriz identidad I_n por un vector c de $\mathbb{R}^n$). Si I es la matriz identidad de orden n y $\mathbf{c}$ es un vector columna de n componentes, entonces

$$I\mathbf{c} = \mathbf{c}. \blacktriangleleft$$

Actividad 2.7. Pruebe el resultado del Ejemplo 2.14.

El producto AB como sucesión de productos de una matriz por un vector

Si A es una matriz de $m \times p$ y B es una matriz de $p \times n$ cuyas columnas son $\mathbf{b}_1$, $\mathbf{b}_2$,..., $\mathbf{b}_n$, por definición de multiplicación de matrices, el producto AB es

$$AB = \begin{bmatrix} \text{ren}_1(A)\cdot\mathbf{b}_1 & \text{ren}_1(A)\cdot\mathbf{b}_2 & ... & \text{ren}_1(A)\cdot\mathbf{b}_n \\ \text{ren}_2(A)\cdot\mathbf{b}_1 & \text{ren}_2(A)\cdot\mathbf{b}_2 & ... & \text{ren}_2(A)\cdot\mathbf{b}_n \\ \vdots & \vdots & ... & \vdots \\ \text{ren}_m(A)\cdot\mathbf{b}_1 & \text{ren}_m(A)\cdot\mathbf{b}_2 & ... & \text{ren}_m(A)\cdot\mathbf{b}_n \end{bmatrix}.$$

Se observa que la primer columna de la matriz AB es el producto $A\mathbf{b}_1$, la segunda columna es $A\mathbf{b}_2$, etc. Por lo tanto la matriz AB puede describirse en término de sus columnas como

$$AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & ... & A\mathbf{b}_n \end{bmatrix}.$$

Forma matricial de un sistema lineal

El sistema lineal de m ecuaciones con n incógnitas

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m \end{cases}$$

puede escribirse en **forma matricial** como

$$A\mathbf{x} = \mathbf{b},$$

donde $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$ es la matriz de coeficientes del sistema, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ es el vector de incógnitas y $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ el vector de términos constantes.

En efecto,

$$A\mathbf{x}=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=\begin{bmatrix} a_{11}x_1+a_{12}x_2+...+a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+...+a_{2n}x_n \\ \vdots \\ a_{m1}x_1+a_{m2}x_2+...+a_{mn}x_n \end{bmatrix}.$$

Luego, por definición de igualdad de vectores:

$$A\mathbf{x}=\mathbf{b} \Leftrightarrow \begin{cases} a_{11}x_1+a_{12}x_2+...+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+...+a_{2n}x_n=b_2 \\ \vdots \\ a_{m1}x_1+a_{m2}x_2+...+a_{mn}x_n=b_m \end{cases}$$

Esto muestra que resolver la ecuación matricial $A\mathbf{x}=\mathbf{b}$ equivale a resolver el sistema lineal de matriz aumentada $[A:\mathbf{b}]$.

Una solución de la ecuación $A\mathbf{x}=\mathbf{b}$ es un vector $\mathbf{c}=\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$ si y solo si la $n-$upla $(c_1,c_2,...,c_n)$ es una solución del sistema lineal de matriz aumentada $[A:\mathbf{b}]$.

Ejemplo 2.15. El sistema lineal $\begin{cases} 2x & +y & -3z & = & 5 \\ 4x & & -5z & = & 8 \end{cases}$ puede escribirse en forma matricial como

$$\underbrace{\begin{bmatrix} 2 & 1 & -3 \\ 4 & 0 & -5 \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}_{\mathbf{x}}=\underbrace{\begin{bmatrix} 5 \\ 8 \end{bmatrix}}_{\mathbf{b}}. \blacktriangleleft$$

Sistemas lineales y combinaciones lineales

El producto $A\mathbf{x}$ es la siguiente combinación lineal:

$$A\mathbf{x}=x_1\mathbf{a}_1+x_2\mathbf{a}_2+...+x_n\mathbf{a}_n,$$

donde $\mathbf{a}_1$, $\mathbf{a}_2$,..., $\mathbf{a}_n$ son las sucesivas columnas de A y x_1, x_2,..., x_n son las componentes del vector $\mathbf{x}$.

Luego,

$$A\mathbf{x}=\mathbf{b} \Leftrightarrow x_1\mathbf{a}_1+x_2\mathbf{a}_2+...+x_n\mathbf{a}_n=\mathbf{b}.$$

Así, un sistema lineal de matriz aumentada $[A:\mathbf{b}]$ también puede escribirse como la ecuación

$$x_1\mathbf{a}_1+x_2\mathbf{a}_2+...+x_n\mathbf{a}_n=\mathbf{b},$$

en las incógnitas x_1, x_2,...,x_n

Resumiendo: si $A=\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & ... & \mathbf{a}_n \end{bmatrix}$, resolver la ecuación $A\mathbf{x}=\mathbf{b}$ o bien la ecuación $x_1\mathbf{a}_1+x_2\mathbf{a}_2+...+x_n\mathbf{a}_n=\mathbf{b}$, equivale a resolver un sistema lineal de matriz aumentada $[A:\mathbf{b}]$.

Ejemplo 2.16. El sistema lineal del Ejemplo 2.15 se puede expresar como la ecuación:

$$x\begin{bmatrix} 2 \\ 4 \end{bmatrix} + y\begin{bmatrix} 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} -3 \\ -5 \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}. \blacktriangleleft$$

Como consecuencia de las expresiones equivalentes para un sistema lineal, resulta el siguiente teorema:

Teorema 2.3. Sean A una matriz de $m \times n$ y $\mathbf{b}$ un vector columna de n componentes. Entonces las siguientes afirmaciones son equivalentes:

1. El sistema lineal de matriz aumentada $[A : \mathbf{b}]$ es consistente.
2. La ecuación matricial $A\mathbf{x} = \mathbf{b}$ tiene solución.
3. Existen escalares c_1, c_2,..., c_n tales que $\mathbf{b} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + ... + c_n\mathbf{a}_n$, donde $\mathbf{a}_j$ es la j-ésima columna de A, $j = 1, 2, ..., n$.
4. $\mathbf{b}$ es una combinación lineal de las columnas de A.

Ejemplo 2.17. De acuerdo al teorema anterior, para determinar si el vector

$$\mathbf{b} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$$

es una combinación lineal de los vectores $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{v}_3 = \begin{bmatrix} -3 \\ -5 \end{bmatrix}$, se requiere estudiar la consistencia del sistema lineal de matriz aumentada

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & : & \mathbf{b} \end{bmatrix}.$$

Aplicando eliminación gaussiana

$$\begin{bmatrix} 2 & 1 & -3 & : & 5 \\ 4 & 0 & -5 & : & 8 \end{bmatrix} \sim \begin{bmatrix} \boxed{2} & 1 & -3 & : & 5 \\ 0 & \boxed{-2} & 1 & : & -2 \end{bmatrix}.$$

Como se observa, el sistema lineal es consistente por lo tanto $\mathbf{b}$ es una combinación lineal de las columnas de A. Más aún, el sistema tiene infinitas soluciones por lo tanto hay infinitas maneras de escribir al vector $\mathbf{b}$ como combinación lineal de los vectores $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$.

La solución general del sistema lineal viene dada por $\left(2 + \frac{5}{4}t, 1 + \frac{1}{2}t, t\right)$ entonces

$$\begin{bmatrix} 5 \\ 8 \end{bmatrix} = \left(2 + \frac{5}{4}t\right)\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \left(1 + \frac{1}{2}t\right)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -3 \\ -5 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Por ejemplo, si $t = 1$:

$$\begin{bmatrix} 5 \\ 8 \end{bmatrix} = \frac{13}{4}\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \frac{3}{2}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} -3 \\ -5 \end{bmatrix}. \blacktriangleleft$$

Propiedades de la multiplicación de matrices

La multiplicación de matrices satisface algunas propiedades fundamentales que se resumen en el siguiente teorema:

Teorema 2.4. En cada ítem sea A una matriz de $m \times n$ y B, C y la matriz cero, $\mathbf{O}$, de tamaños tales que estén bien definidas las operaciones que se indican. Entonces se verifican las siguientes propiedades:

1. $A(BC) = (AB)C$ [Asociativa]
2. $A(B + C) = AB + AC$ [Distributiva a izquierda]
3. $(B + C)A = BA + CA$ [Distributiva a derecha]
4. $a(BC) = (aB)C = B(aC)$, para todo $a \in \mathbb{R}$
5. $I_m A = AI_n = A$
6. $\mathbf{O}A = \mathbf{O}$ y $A\mathbf{O} = \mathbf{O}$

Demostración. Se probará aquí solo la propiedad 5.

Considere la matriz A y la matriz identidad I_n expresadas en término de sus columnas:

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \text{ y } I_n = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \end{bmatrix}.$$

Escribiendo los productos AI_n y $I_m A$ en términos de sus columnas y teniendo en cuenta los Ejemplos 2.13 y 2.14, se obtienen

$$AI_n = \begin{bmatrix} A\mathbf{e}_1 & A\mathbf{e}_2 & \dots & A\mathbf{e}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} = A,$$

y

$$I_m A = \begin{bmatrix} I_m\mathbf{a}_1 & I_m\mathbf{a}_2 & \dots & I_m\mathbf{a}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} = A.$$

□

Nota: La propiedad asociativa permite eliminar paréntesis para la multiplicación de tres o más matrices. Por ejemplo:

$$(AB)(CD) = A((BC)D) = A(B(CD)) = ABCD.$$

Actividad 2.8. Verifique la propiedad 1. del teorema anterior para las matrices:

$$A = \begin{bmatrix} 5 & -3 \\ 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & -2 \\ 1 & 2 & -3 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 0 & -3 & 0 & 1 \\ 1 & 2 & 5 & -1 \\ 2 & 0 & -1 & 3 \end{bmatrix}.$$

Observación 2.1. El producto de una matriz de $m \times n$ por un vector columna de n componentes es un caso particular de la multiplicación de matrices. Por lo tanto:

Si A es una matriz de $m \times n$, $\mathbf{x}$ y $\mathbf{y}$ vectores columna de n componentes y k un escalar, entonces

$$A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y},$$

$$A(k\mathbf{x}) = k(A\mathbf{x}).$$

En general si A es una matriz de $m \times n$ y $\mathbf{x}_1$, $\mathbf{x}_2$,..., $\mathbf{x}_k$ son k vectores columna de n componentes entonces

$$A(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + ... + c_k\mathbf{x}_k) = c_1(A\mathbf{x}_1) + c_2(A\mathbf{x}_2) + ... + c_k(A\mathbf{x}_k),$$

para cualesquiera escalares c_1, c_2,..., c_k.

Definición (Matriz traspuesta)**.** Sea A una matriz de $m \times n$. La *matriz traspuesta* de A, que se denota A^t, es la matriz de $n \times m$ cuyas columnas son los renglones de A en el mismo orden.

Ejemplo 2.18. Dadas las matrices $A = \begin{bmatrix} 5 & -3 & 2 \\ 8 & 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 0 & 0 \\ -1 & 2 & 1 \end{bmatrix}$ y $C = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$, entonces

$$A^t = \begin{bmatrix} 5 & 8 \\ -3 & 0 \\ 2 & -1 \end{bmatrix}, \; B^t = \begin{bmatrix} 1 & 1 & -1 \\ 3 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix} \text{ y } C^t = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix}. \blacktriangleleft$$

Para la trasposición de matrices se tienen las siguientes propiedades que se enuncian sin demostración:

Teorema 2.5. En cada ítem, sea A una matriz de $m \times n$, B una matriz de tamaño tal que estén bien definidas las operaciones que se indican y c un escalar. Entonces se satisfacen las siguientes propiedades:

1. $\left(A^t\right)^t = A$
2. $(A + B)^t = A^t + B^t$
3. $(cA)^t = cA^t$
4. $(AB)^t = B^tA^t$

Observación 2.2 (Transpuesta de un producto)**.** La propiedad 4. del teorema anterior se puede extender a un número finito de matrices. Si A_1, A_2,..., A_k son matrices de tamaños tales que esté bien definido el producto $A_1A_2...A_{k-1}A_k$ entonces

$$(A_1A_2...A_{k-1}A_k)^t = A_k^tA_{k-1}^t...A_2^tA_1^t.$$

Actividad 2.9. Verifique la propiedad 4. del teorema anterior para las matrices

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 1 & 2 & -3 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & -3 & 0 & 1 \\ 1 & 2 & 5 & -1 \\ 2 & 0 & -1 & 3 \end{bmatrix}.$$

Definición (Matriz simétrica)**.** Una matriz A es *simétrica* si $A^t = A$.

De la definición anterior se deduce que necesariamente una matriz simétrica es cuadrada ya que si A es de $m \times n$ entonces A^t es de $n \times m$ y por lo tanto si $A^t = A$ debe ser $n = m$.

Ejemplo 2.19. La matriz $\begin{bmatrix} 9 & 0 & -1 \\ 0 & 2 & 5 \\ -1 & 5 & 3 \end{bmatrix}$ es simétrica mientras que $\begin{bmatrix} 5 & 8 \\ -3 & 5 \end{bmatrix}$ no lo es.◄

Actividad 2.10. Pruebe que para toda matriz A las matrices AA^t y A^tA están bien definidas y son matrices simétricas.

Definición (Potenciación de matrices)**.** Sea A una matriz cuadrada de orden n y $k \in \mathbb{N}$. La *potencia $k-$ésima de A*, que se denota por A^k, es la matriz cuadrada de orden n dada por

$$A^k = \overbrace{AAA...A}^{k\text{-veces}}.$$

Por conveniencia si $A \neq \mathbf{O}$ se define $A^0 = I$.

Ejemplo 2.20. Sea la matriz $A = \begin{bmatrix} -1 & 2 & 0 \\ 3 & 0 & -1 \\ 0 & -1 & 4 \end{bmatrix}$. Para efectuar el cálculo de A^3 primero se puede obtener A^2

$$A^2 = AA = \begin{bmatrix} -1 & 2 & 0 \\ 3 & 0 & -1 \\ 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} -1 & 2 & 0 \\ 3 & 0 & -1 \\ 0 & -1 & 4 \end{bmatrix} = \begin{bmatrix} 7 & -2 & -2 \\ -3 & 7 & -4 \\ -3 & -4 & 17 \end{bmatrix},$$

para luego efectuar el producto A^2A:

$$A^3 = AAA = (AA)\,A = A^2A = \begin{bmatrix} 7 & -2 & -2 \\ -3 & 7 & -4 \\ -3 & -4 & 17 \end{bmatrix} \begin{bmatrix} -1 & 2 & 0 \\ 3 & 0 & -1 \\ 0 & -1 & 4 \end{bmatrix} = \begin{bmatrix} -13 & 16 & -6 \\ 24 & -2 & -23 \\ -9 & -23 & 72 \end{bmatrix}. ◄$$

Para la potenciación de matrices se tienen las siguientes propiedades:

Teorema 2.6. Si A es una matriz cuadrada y p y q son cualesquiera enteros positivos entonces

1. $A^pA^q = A^{p+q}$.
2. $(A^p)^q = A^{pq}$.
3. $(cA)^p = c^pA^p$, para todo escalar c.

Actividad 2.11. Demuestre el resultado del teorema anterior para $p = 2$ y $q = 3$.

Ejercicios 2.2

$\boxed{1.}$ Obtenga los valores x, y y z tales que las matrices A y B resulten iguales:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -2 & -5 \end{bmatrix}, \quad B = \begin{bmatrix} x & 0 & y-z \\ x+y+2z & -2 & -5 \end{bmatrix}.$$

2. Calcule, si es posible, las siguientes operaciones. Si no es posible efectuar el cálculo, explique por qué.

$\boxed{a)}$ $\begin{bmatrix} 1 & 2 \\ -3 & -2 \end{bmatrix} + \begin{bmatrix} -4 & 3 \\ 1 & 0 \end{bmatrix}$

b) $-3 \begin{bmatrix} 1 & 2 & 9 \\ 3 & 2 & -5 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 2 \\ 3 & 0 \\ 1 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 9 \\ 3 & 2 & 0 \end{bmatrix}$

$\boxed{d)}$ $2 \begin{bmatrix} 1 & 2 & 0 \\ -3 & -2 & 1 \\ 1 & -1 & 4 \end{bmatrix} - 4 \begin{bmatrix} 3 & -2 & -1 \\ 1 & 0 & 1 \\ 0 & -1 & 2 \end{bmatrix}$

3. Demuestre las restantes propiedades del Teorema 2.2.

$\boxed{4.}$ Dadas las matrices

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 3 & 4 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 0 & -2 & -1 \\ 2 & 0 & 1 \end{bmatrix},$$

determine la matriz X que satisfaga las siguientes ecuaciones:

a) $2X - A + B + 3C = \mathbf{O}$

b) $-X + \frac{1}{2}A - 3C = 4X + B$

c) $2(A - 2B) + 2X = 0C$

d) $X - \frac{1}{2}B = C + 4X$

5. Considere los vectores $\mathbf{u} = (-1, 2, -2, 1)$, $\mathbf{v} = (4, -2, 5, 3)$, $\mathbf{a} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix}$. Calcule los siguientes productos punto:

a) $\mathbf{u} \cdot \mathbf{v}$

b) $\mathbf{u} \cdot (2\mathbf{u} - \mathbf{v})$

c) $\mathbf{a} \cdot \mathbf{b}$

d) $-\mathbf{b} \cdot 2\mathbf{a}$

$\boxed{6.}$ Sean $\mathbf{u} = \begin{bmatrix} a \\ 0 \\ -2 \\ 3 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} -2 \\ 1 \\ 3 \\ -1 \end{bmatrix}$. Halle el valor de a de manera que $\mathbf{u} \cdot \mathbf{v} = 3$.

$\boxed{7.}$ Calcule, si están definidos, los siguientes productos de matrices:

a) $\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} \begin{bmatrix} 4 & 0 & 2 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

c) $\begin{bmatrix} 1 & -2 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} 4 & 0 & 2 \\ 1 & -1 & 5 \end{bmatrix}$

d) $\begin{bmatrix} 4 & 0 & 2 \\ 1 & -1 & 5 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -3 & 0 \end{bmatrix}$

e) $\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 4 \\ -3 & 3 \end{bmatrix}$

f) $\begin{bmatrix} 1 & 2 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 2 & -2 & 0 \\ 1 & 4 & 0 \end{bmatrix}$

8. Obtenga el elemento $(2,2)$ de $\begin{bmatrix} 2 & -1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 4 & -5 \end{bmatrix}$.

$\boxed{9.}$ Encuentre la 3*er* columna de AB siendo

$$A = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 2 & 0 & 2 \\ -1 & 0 & -2 & -5 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \\ 2 & 0 & -2 \\ 4 & -2 & 3 \end{bmatrix}.$$

$\boxed{10.}$ Una matriz cuadrada $A = [a_{ij}]$ es una *matriz diagonal* si $a_{ij} = 0$ si $i \neq j$. Demuestre que el producto de dos matrices diagonales de orden n es también una matriz diagonal de orden n y enuncie una regla para multiplicar dos matrices diagonales.

11. Efectúe el producto $A\mathbf{c}$ siendo

a) $A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \\ 2 & 0 & -2 \end{bmatrix}$ y $\mathbf{c} = \begin{bmatrix} 2 \\ 2 \\ -3 \end{bmatrix}$.

$\boxed{b)}$ $A = \begin{bmatrix} -1 & 1 & 1 & 2 \\ 1 & -1 & 1 & 3 \\ 3 & 1 & -2 & 0 \\ 1 & -1 & 0 & 2 \\ 0 & 0 & -3 & 4 \end{bmatrix}$ y $\mathbf{c} = \begin{bmatrix} -2 \\ 3 \\ -1 \\ 1 \end{bmatrix}$.

12. Escriba los siguientes sistemas en la forma $A\mathbf{x} = \mathbf{b}$:

a) $\begin{cases} 2x & -4y & = & 1 \\ x & +y & = & 0 \\ -x & -3y & = & 4 \end{cases}$

$\boxed{b)}$ $\begin{cases} x_1 & -4x_2 & -x_3 & = & 0 \\ 2x_1 & +x_2 & 3x_3 & = & 1 \\ -x_1 & +x_2 & +2x_3 & = & 2 \end{cases}$

$\boxed{13.}$ Sean los vectores $\mathbf{a}_1 = \begin{bmatrix} -2 \\ 3 \\ -1 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{a}_3 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$. Exprese la ecuación

$$x\mathbf{a}_1 + y\mathbf{a}_2 + z\mathbf{a}_3 = \mathbf{b},$$

como un sistema de ecuaciones lineales y resuélvalo.

14. Determine si el vector $\mathbf{b}$ es combinación lineal de las columnas de la matriz A siendo:

$\boxed{a)}$ $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & -1 & 0 & 0 \\ 2 & 1 & 3 & 4 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix}$.

b) $A = \begin{bmatrix} 1 & 3 \\ -2 & -1 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

c) $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 5 \\ 3 & -2 & 7 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 3 \\ 3 \\ 0 \end{bmatrix}$.

15. Demuestre que si $\mathbf{x}_1$ y $\mathbf{x}_2$ son soluciones del sistema $A\mathbf{x} = \mathbf{b}$, entonces $\mathbf{x}_1 - \mathbf{x}_2$ es una solución del sistema homogéneo asociado $A\mathbf{x} = \mathbf{0}$.

$\boxed{16.}$ Obtenga la matriz X que satisfaga la ecuación $(2X)^t - 2AB^t = \left(BA^t\right)^t$ siendo

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 3 \\ 2 & -1 & 2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & 2 \end{bmatrix}.$$

$\boxed{17.}$ Indique cuáles de las siguientes matrices son simétricas, justifique:

$$A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 0 & 5 \\ 1 & 5 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 1 & 4 \\ -1 & 0 & 5 & 0 \\ 2 & 3 & 7 & -2 \\ 4 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{O}_4, \quad \text{y} \quad C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

18. Una matriz A es *antisimétrica* si $A^t = -A$.

a) Indique cuáles de las siguientes matrices son antisimétricas:

$$A = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & 5 \\ 2 & -5 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 1 & 4 \\ -1 & 0 & 5 & 0 \\ 2 & 3 & 7 & -2 \\ 4 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{O}_4, \quad \text{y} \quad C = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$$

b) Demuestre que si A es una matriz tanto simétrica como antisimétrica, entonces $A = \mathbf{O}$.

19. Dadas las matrices $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & -1 & 1 \\ 2 & 3 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. Halle las matrices:

a) $(2A)^3$

b) $(A+B)^2 - I^4$

c) $(-AB)^3$

d) $(0A)^3 + \mathbf{O}^{15} + 2I$

20. Halle, si existe, una matriz A cuadrada de orden 2 tal que:

a) $A^2 = \mathbf{O}$ y $A \neq \mathbf{O}$.

b) $A^2 = I$ siendo $A \neq I$ y $A \neq -I$.

$\boxed{21.}$ Dadas las matrices $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ y $B = \begin{bmatrix} 2 & -1 \\ -3 & -2 \end{bmatrix}$. ¿Es cierto que $(A+B)^2 = A^2 + 2AB + B^2$?. Justifique su respuesta.

22. Sean A y B matrices cuadradas de orden n. Determine si las siguientes proposiciones son verdaderas o falsas. Justifique.

a) $A^2 - B^2 = (A - B)(A + B)$.

b) $(AB)^2 = A^2B^2$.

c) $(A + B)^2 = A^2 + 2AB + B^2 \Leftrightarrow AB = BA$

d) Si A y B conmutan, es decir $AB = BA$, entonces $(AB)^2 = A^2B^2$.

e) Si A es simétrica entonces la matriz $2A^2BB^t$ es simétrica.

2.3. Matriz inversa

Definición (Matriz invertible)**.** Una matriz cuadrada A de $n \times n$ es es *invertible*, o *no singular*, si existe una matriz B de $n \times n$, llamada *inversa* de A, tal que

$$AB = I \quad \text{y} \quad BA = I.$$

Si no existe tal matriz B, la matriz A es una matriz **singular** o **no invertible**.

Ejemplo 2.21. Sean $A = \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix}$ y $B = \begin{bmatrix} 3 & -1 \\ -1 & \frac{1}{2} \end{bmatrix}$. Puede verificarse que $AB = BA = I$, por lo tanto A es invertible y B es una inversa de A.◀

¿Existirá otra matriz C, distinta de B, tal que $AC = CA = I$?. El teorema que sigue muestra que esto no puede suceder.

Teorema 2.7. Una matriz invertible tiene una única inversa.

Demostración. Sean A una matriz invertible y B y C dos matrices inversas de A, es decir

$$AB = BA = I \quad \text{y} \quad AC = CA = I.$$

Entonces:

$$B = BI = B(AC) = (BA)C = IC = C.$$

Dado que $B = C$ queda probado que una matriz invertible no puede admitir más que una matriz inversa.

□

Observación 2.3. A la matriz inversa de una matriz A invertible se la denota A^{-1}. Así, para una matriz A invertible:

$$AA^{-1} = A^{-1}A = I.$$

Debido a la unicidad de la matriz inversa, si A y B son dos matrices de $n \times n$ tales que $AB = BA = I$ se puede concluir que A es invertible y que B es su inversa, es decir $A^{-1} = B$.

Ejemplo 2.22. Retomando el Ejemplo 2.21 puede concluirse que la matriz B es *la* inversa de A, esto es: $A^{-1} = B$.◀

Ejemplo 2.23 (Inversa de la matriz identidad). Como $II = I\,(= II)$, la matriz identidad I es invertible y $I^{-1} = I$.◀

Los siguientes son dos ejemplos de matrices no invertibles (singulares):

Ejemplo 2.24. La matriz cero, $\mathbf{O}_n$, es singular. ¿Por qué?.◀

Ejemplo 2.25. La matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ es singular.

En efecto, si A fuera invertible, debería existir una matriz B tal que $AB = I$ (y $BA = I$). Por lo tanto la ecuación $AX = I$ debería tener solución. Siendo $X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$:

$$AX = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} x_{11} + 2x_{21} & x_{12} + 2x_{22} \\ 3x_{11} + 6x_{21} & 3x_{12} + 6x_{22} \end{bmatrix},$$

luego

$$AX = I \Leftrightarrow \begin{bmatrix} x_{11} + 2x_{21} & x_{12} + 2x_{22} \\ 3x_{11} + 6x_{21} & 3x_{12} + 6x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Por igualdad de matrices, se obtienen los siguientes sistemas lineales

$$\left\{ \begin{array}{rcl} x_{11} \; +2x_{21} & = & 1 \\ 3x_{11} \; +6x_{21} & = & 0 \end{array} \right. \quad \text{y} \quad \left\{ \begin{array}{rcl} x_{12} \; +2x_{22} & = & 0 \\ 3x_{12} \; +6x_{22} & = & 1 \end{array} \right. .$$

Es fácilmente comprobable que estos sistemas son inconsistentes, luego $AX = I$ no tiene solución y por lo tanto A no es invertible.◀

Algunas propiedades de las matrices invertibles son las siguientes:

Teorema 2.8. Sean A y B matrices invertibles de $n \times n$ y c un escalar distinto de cero. Entonces A^{-1}, cA, A^t y el producto AB son matrices invertibles. Además:

1. $\left(A^{-1}\right)^{-1} = A$.
2. $(cA)^{-1} = \frac{1}{c}A^{-1}$.
3. $\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$.
4. $(AB)^{-1} = B^{-1}A^{-1}$.

Demostración. Probaremos aquí las propiedades 3. y 4. Las matrices A y B son invertibles, entonces $AA^{-1} = A^{-1}A = I$ y $BB^{-1} = B^{-1}B = I$.

3. Las matrices A^t y $\left(A^{-1}\right)^t$ son ambas cuadradas de orden n,

$$A^t \left(A^{-1}\right)^t = \left(A^{-1}A\right)^t = I^t = I \quad \text{y} \quad \left(A^{-1}\right)^t A^t = \left(AA^{-1}\right)^t = I^t = I.$$

De la Observación 2.3, A^t es invertible y su inversa es $\left(A^{-1}\right)^t$ es decir $\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$.

4. Las matrices AB y $B^{-1}A^{-1}$ son ambas cuadradas de orden n. Dado que:

$$(AB)\left(B^{-1}A^{-1}\right) = A\left(B\left(B^{-1}A^{-1}\right)\right) = A\left(\left(BB^{-1}\right)A^{-1}\right) = A\left(IA^{-1}\right) = AA^{-1} = I, \quad \text{y}$$

$$\left(B^{-1}A^{-1}\right)(AB) = B^{-1}\left(A^{-1}(AB)\right) = B^{-1}\left(\left(A^{-1}A\right)B\right) = B^{-1}(IB) = B^{-1}B = I,$$

se deduce, de la Observación 2.3, que AB es invertible y que su inversa es el prodcuto $B^{-1}A^{-1}$ es decir

$$(AB)^{-1} = B^{-1}A^{-1}.$$

□

Inversa del producto

La propiedad 4. del teorema anterior puede generalizarse para k matrices invertibles. Si A_1, A_2,..., A_k son matrices invertibles de igual tamaño, también lo es el producto $A_1A_2...A_{k-1}A_k$. Su inversa es

$$(A_1A_2...A_{k-1}A_k)^{-1} = A_k^{-1}A_{k-1}^{-1}...A_2^{-1}A_1^{-1}.$$

Observación 2.4. La suma de matrices invertibles no es, en general, una matriz invertible. No es cierto que $(A+B)^{-1}$ sea igual que $A^{-1}+B^{-1}$ (Véase Ejercicio 2.3.a).

Actividad 2.12. Pruebe las propiedades 1. y 2. del Teorema 2.8.

Actividad 2.13. Pruebe que si A es invertible, A^3 es invertible y $\left(A^3\right)^{-1} = \left(A^{-1}\right)^3$.

El resultado de la actividad anterior puede generalizarse: Si $k \in \mathbb{N}$ y A es una matriz invertible entonces A^k es invertible y

$$\left(A^k\right)^{-1} = \left(A^{-1}\right)^k.$$

Es más, este resultado permite definir un nuevo concepto: potenciación de matrices de exponente negativo:

Definición (Potencia de matrices de exponente negativo)**.** Sea A una matriz invertible y $k \in \mathbb{N}$. La *potencia* $-k$ de A, que se denota A^{-k}, se define por $A^{-k} = \left(A^{-1}\right)^k$.

Se puede generalizar el Teorema 2.2 para la potenciación de matrices con cualquier exponente entero y se tiene el siguiente resultado:

Teorema 2.9. Si A es una matriz invertible, p y q enteros cualesquiera y c un escalar, entonces

1. $(A^p)^{-1} = \left(A^{-1}\right)^p$.
2. $A^pA^q = A^{p+q}$.
3. $(A^p)^q = A^{pq}$.
4. $(cA)^p = c^pA^p$ (si $p \leq 0$ debe ser $c \neq 0$).

Actividad 2.14. Verifique las propiedades anteriores para $p=-2$ y $q=3$.

Recuerde que, en general, si $AC = BC$ no es válida la cancelación de la matriz C en ambos miembros. Sin embargo, esto es posible si la matriz C es invertible:

Teorema 2.10. Si C es una matriz invertible y en cada caso A y B son matrices de tamaños apropiados entonces

1. $CA = CB \Rightarrow A = B$.
2. $AC = BC \Rightarrow A = B$.

Demostración. Se probará la proposición 1. La 2. se demuestra en forma análoga.

1. Dada la existencia de C^{-1} se tiene

$$CA = CB \Rightarrow C^{-1}(CA) = C^{-1}(CB) \Rightarrow \left(C^{-1}C\right)A = \left(C^{-1}C\right)B \Rightarrow IA = IB \Rightarrow A = B.$$

□

Ecuaciones con productos de matrices

Las propiedades de las operaciones matriciales nos permiten resolver algunas ecuaciones que involucran matrices (ecuaciones matriciales).

Es importante tener en cuenta que el producto de matrices no es conmutativo. Por lo tanto cuando se multiplica a ambos miembros de una ecuación por una matriz se debe usar la multiplicación por la izquierda, o por la derecha, en ambos miembros. Así, para matrices C y D de tamaños apropiados,

$$A = B \Rightarrow CA = CB,$$

y

$$A = B \Rightarrow AD = BD,$$

mientras que $A = B$ no implica que $CA = BC$.

Por ejemplo, considere la ecuación matricial $CXA - B = \mathbf{O}$ donde A y C son invertibles y todas las matrices son de tamaños compatibles.

Se puede despejar X (matriz incógnita) de la siguiente manera:

$$\begin{aligned}
CXA - B &= \mathbf{O}, \\
CXA &= B, \\
(CXA)\,A^{-1} &= BA^{-1}, \\
CX\left(AA^{-1}\right) &= BA^{-1}, \\
CXI &= BA^{-1}, \\
CX &= BA^{-1}, \\
C^{-1}(CX) &= C^{-1}\left(BA^{-1}\right), \\
\left(C^{-1}C\right)X &= C^{-1}BA^{-1}, \\
IX &= C^{-1}BA^{-1}, \\
X &= C^{-1}BA^{-1}.
\end{aligned}$$

En lo que sigue se proveen algunos resultados que permitirán hallar la inversa de una matriz invertible o probar que es singular.

De acuerdo a la Observación 2.3, una matriz B es la inversa de una matriz A si $AB = I$ y $BA = I$. En general para dos matrices A y B de $n \times n$, AB no es igual a BA pero en el caso en que $AB = I$ el siguiente teorema, cuya demostración se omitirá en este texto, muestra que también $BA = I$.

Teorema 2.11. Si A y B son matrices de $n \times n$ y $AB = I$, entonces $BA = I$.

Una consecuencia inmediata de este teorema es el siguiente importante resultado:

Corolario 2.1. Sea A una matriz de $n \times n$. Entonces:

A es invertible si y solo si la ecuación matricial $AX = I$ tiene única solución.

Demostración. En un sentido es trivial ya que si A es invertible entonces $AA^{-1} = I$, y por lo tanto A^{-1} es solución de $AX = I$.

Recíprocamente, si $AX = I$ tiene solución entonces existe una matriz B tal que $AB = I$, por el Teorema 2.11 también $BA = I$ por lo cual A es una matriz invertible y $B = A^{-1}$. □

Ejemplo 2.26. De acuerdo al corolario anterior, para determinar si la matriz $A = \begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}$ es invertible, debe analizarse si la ecuación $AX = I$ tiene solución. En caso afirmativo, la (única) solución de dicha ecuación es la matriz inversa de A.

Siendo $X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$ la matriz incógnita,

$$\begin{aligned} AX = I \quad &\Leftrightarrow \quad \begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\ &\Leftrightarrow \quad \begin{bmatrix} x_{11} + 4x_{21} & x_{12} + 4x_{22} \\ -x_{11} - 3x_{21} & -x_{12} - 3x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Por igualdad de matrices, se obtienen los dos sistemas lineales de matrices aumentadas $[A : \mathbf{e}_1]$ y $[A : \mathbf{e}_2]$:

$$\left\{ \begin{array}{rrcl} x_{11} & +4x_{21} & = & 1 \\ -x_{11} & -3x_{21} & = & 0 \end{array} \right. \quad \text{y} \quad \left\{ \begin{array}{rrcl} x_{12} & +4x_{22} & = & 0 \\ -x_{12} & -3x_{22} & = & 1 \end{array} \right. .$$

Puede verificarse que ambos sistemas tienen solución única. La solución del primer sistema es $x_{11} = -3$ y $x_{21} = 1$ que son los elementos de la primer columna de la matriz X. De manera similar, resolviendo el segundo sistema, se obtienen los elementos de la segunda columna de la matriz incógnita, $x_{12} = -4$ y $x_{22} = 1$. Se obtiene así la matriz

$$B = \begin{bmatrix} -3 & -4 \\ 1 & 1 \end{bmatrix}$$

tal que $AB = I$ y por lo tanto (Corolario 2.1) A es invertible y $B = A^{-1}$.◀

Actividad 2.15. Verifique, en el ejemplo anterior, que también $BA = I$.

Cálculo de la inversa de una matriz: método de Gauss-Jordan

Se puede generalizar el procedimiento empleado en el Ejemplo 2.26 para encontrar la inversa de una matriz o probar que no es invertible.

Sabemos que una matriz cuadrada, A, de orden n es invertible si y solo si la ecuación matricial $AX = I$ tiene solución (en cuyo caso la única solución es A^{-1}). Si las sucesivas columnas de la matriz incógnita X son $\mathbf{x}_1$, $\mathbf{x}_2$,..., $\mathbf{x}_n$:

$$AX = \left[\begin{array}{cccc} A\mathbf{x}_1 & A\mathbf{x}_2 & ... & A\mathbf{x}_n \end{array}\right].$$

Dado que $I = \left[\begin{array}{cccc} \mathbf{e}_1 & \mathbf{e}_2 & ... & \mathbf{e}_n \end{array}\right]$, por definición de igualdad de matrices:

$$AX = I \;\Leftrightarrow\; A\mathbf{x}_1 = \mathbf{e}_1,\; A\mathbf{x}_2 = \mathbf{e}_2, ...,\; A\mathbf{x}_n = \mathbf{e}_n.$$

Luego, la ecuación $AX = I$ equivale a resolver los n sistemas lineales $A\mathbf{x}_j = \mathbf{e}_j$, $j = 1, 2, ..., n$. Estos n sistemas lineales tienen la misma matriz, A, de coeficientes y por lo tanto pueden resolverse en forma simultánea (véase Ejercicio 10, Sección 1.4) aplicando el método de eliminación de Gauss-Jordan a la matriz aumentada:

$$\left[\begin{array}{ccccc} A & : & \mathbf{e}_1 & \mathbf{e}_2 & ... & \mathbf{e}_n \end{array}\right] = [A : I].$$

Si en algún momento del proceso se llega a una matriz escalonada $[H : J]$ (H es una forma escalonada de A) donde H no tiene renglones cero, entonces todos los sistemas lineales $A\mathbf{x}_j = \mathbf{e}_j$ tienen solución y por lo tanto $AX = I$ tiene solución y A es invertible.

Dado que A es de $n \times n$ y tiene n posiciones pivote, el proceso de Gauss-Jordan conduce a la forma escalonada reducida $[I : B]$ siendo la j-ésima columna de B la solución del sistema $A\mathbf{x}_j = \mathbf{e}_j$. Así, la matriz B es la solución de $AX = I$ y por lo tanto $B = A^{-1}$.

Supongamos que en algún momento del proceso se llega a una matriz $[C : D]$ donde C tiene un renglón cero. Debe tenerse en cuenta que en cualquier etapa del proceso, la matriz de la derecha no tiene renglones cero ya que es equivalente por renglones a la matriz original I identidad. Esto significa que alguno de los sistemas $A\mathbf{x}_j = \mathbf{e}_j$ es inconsistente (¿ por qué?) y por lo tanto $AX = I$ no tiene solución de lo que se deduce que A no es invertible.

Todo lo anterior se puede resumir en el siguiente algoritmo:

Algoritmo para invertir una matriz A

1. Escriba la matriz de $n \times 2n$, $[A : I]$.
2. Aplique eliminación de Gauss-Jordan a la matriz del paso 1 para encontrar su forma escalonada reducida. Si en algún momento del proceso se encuentra una matriz $[C : D]$ donde C tiene un renglón cero, deténgase: la matriz A no es invertible. De lo contrario A es invertible y la forma escalonada reducida de $[A : I]$ es $\left[I : A^{-1}\right]$.

Ejemplo 2.27. Para determinar la inversa de $A = \begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & -2 \\ 2 & -1 & 0 \end{bmatrix}$, si existe, utilizamos el algoritmo anterior. Se comienza aplicando el proceso de eliminación de Gauss-Jordan a la matriz $[A : I]$:

$$[A : I] = \left[\begin{array}{ccccccc} -1 & 1 & 0 & : & 1 & 0 & 0 \\ 1 & 0 & -2 & : & 0 & 1 & 0 \\ 2 & -1 & 0 & : & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_3 + 2R_1 \to R_3]{R_2 + R_1 \to R_2} \left[\begin{array}{ccccccc} -1 & 1 & 0 & : & 1 & 0 & 0 \\ 0 & 1 & -2 & : & 1 & 1 & 0 \\ 0 & 1 & 0 & : & 2 & 0 & 1 \end{array}\right] \xrightarrow{R_3 - R_2 \to R_3}$$

$$\rightarrow\left[\begin{array}{rrrcrrr} -1 & 1 & 0 & : & 1 & 0 & 0 \\ 0 & 1 & -2 & : & 1 & 1 & 0 \\ 0 & 0 & 2 & : & 1 & -1 & 1 \end{array}\right] \xrightarrow[\frac{1}{2}R_3 \to R_3]{} \left[\begin{array}{rrrcrrr} -1 & 1 & 0 & : & 1 & 0 & 0 \\ 0 & 1 & -2 & : & 1 & 1 & 0 \\ 0 & 0 & 1 & : & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array}\right] \xrightarrow[R_2+2R_3 \to R_2]{}$$

$$\rightarrow\left[\begin{array}{rrrcrrr} -1 & 1 & 0 & : & 1 & 0 & 0 \\ 0 & 1 & 0 & : & 2 & 0 & 1 \\ 0 & 0 & 1 & : & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array}\right] \xrightarrow[R_1-R_2 \to R_1]{} \left[\begin{array}{rrrcrrr} -1 & 0 & 0 & : & -1 & 0 & -1 \\ 0 & 1 & 0 & : & 2 & 0 & 1 \\ 0 & 0 & 1 & : & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array}\right] \xrightarrow[-R_1 \to R_1]{}$$

$$\rightarrow\left[\begin{array}{rrrcrrr} 1 & 0 & 0 & : & 1 & 0 & 1 \\ 0 & 1 & 0 & : & 2 & 0 & 1 \\ 0 & 0 & 1 & : & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array}\right] = [I : B]\,.$$

Se completó el proceso de eliminación de Gauss-Jordan. La forma escalonada reducida de $[A : I]$ es $[I : B]$ por lo tanto A es invertible y $A^{-1} = B = \left[\begin{array}{rrr} 1 & 0 & 1 \\ 2 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array}\right]$.◀

Ejemplo 2.28. Sea la matriz $A = \left[\begin{array}{rrrr} 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & -1 \\ 3 & 1 & 2 & 1 \\ -1 & 5 & 0 & 0 \end{array}\right]$. Como en el ejemplo anterior, para determinar A^{-1}, si ésta existe, se comienza aplicando el proceso de eliminación de Gauss-Jordan a la matriz $[A : I]$:

$$[A : I] = \left[\begin{array}{rrrrcrrrr} 1 & 0 & 1 & 2 & : & 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & -1 & : & 0 & 1 & 0 & 0 \\ 3 & 1 & 2 & 1 & : & 0 & 0 & 1 & 0 \\ -1 & 5 & 0 & 0 & : & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{R_2-2R_1 \to R_2 \\ R_3-3R_1 \to R_3 \\ R_4+R_1 \to R_4}]{} \left[\begin{array}{rrrrcrrrr} 1 & 0 & 1 & 2 & : & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -5 & : & -2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -5 & : & -3 & 0 & 1 & 0 \\ 0 & 5 & 1 & 2 & : & 1 & 0 & 0 & 1 \end{array}\right] \rightarrow$$

$$\xrightarrow[\substack{R_3-R_2 \to R_3 \\ R_4-5R_2 \to R_4}]{} \left[\begin{array}{rrrrcrrrr} 1 & 0 & 1 & 2 & : & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -5 & : & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & : & -1 & -1 & 1 & 0 \\ 0 & 0 & 5 & 27 & : & 11 & -5 & 0 & 1 \end{array}\right].$$

En este momento se detiene el procedimiento ya que se ha llegado a una matriz equivalente $[C : D]$ donde C tiene un renglón cero. Luego, A no es invertible.◀

Debe observarse que para determinar A^{-1} no es necesrio saber de antemano si existe o no. Simplemente se inicia el procedimiento indicado y se obtiene A^{-1} o se concluye que A es singular.

El siguiente teorema da a conocer diferentes caracterizaciones de matrices invertibles.

Teorema 2.12. Sea A una matriz de $n \times n$. Entonces las siguientes proposiciones son equivalentes:

1. A es invertible.
2. $A\mathbf{x} = \mathbf{b}$ es consistente con única solución para todo $\mathbf{b} \in \mathbb{R}^n$.
3. $A\mathbf{x} = \mathbf{0}$ tiene única solución la trivial.
4. A tiene n posiciones pivote.
5. A es equivalente (por renglones) a la matriz identidad I.
   Es decir, la forma escalonada reducida de A es la matriz identidad, I.

Demostración. Se seguirá el siguiente orden en las implicancias (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5). Probando luego que (5) $\Rightarrow$ (1) se tendrán demostradas todas las equivalencias.

(1) $\Rightarrow$ (2) Sea $\mathbf{b}$ un vector columna de $\mathbb{R}^n$. El vector $A^{-1}\mathbf{b}$ es solución de $A\mathbf{x} = \mathbf{b}$ ya que

$$A\left(A^{-1}\mathbf{b}\right) = \left(AA^{-1}\right)\mathbf{b} = I\mathbf{b} = \mathbf{b}.$$

Por lo tanto $A\mathbf{x} = \mathbf{b}$ es consistente para todo $\mathbf{b} \in \mathbb{R}^n$. Además $A^{-1}\mathbf{b}$ es la única solución de $A\mathbf{x} = \mathbf{b}$ puesto que si $\mathbf{y}$ es una solución de $A\mathbf{x} = \mathbf{b}$ resulta:

$$A\mathbf{y} = \mathbf{b} \Rightarrow A^{-1}(A\mathbf{y}) = A^{-1}\mathbf{b} \Rightarrow \left(A^{-1}A\right)\mathbf{y} = A^{-1}\mathbf{b} \Rightarrow I\mathbf{y} = A^{-1}\mathbf{b} \Rightarrow \mathbf{y} = A^{-1}\mathbf{b}.$$

(2) $\Rightarrow$ (3) Si $A\mathbf{x} = \mathbf{b}$ tiene única solución para todo $\mathbf{b} \in \mathbb{R}^n$, esto vale, en particular, si $\mathbf{b}$ es el vector cero. Por lo tanto $A\mathbf{x} = \mathbf{0}$ tiene única solución, la trivial.

(3) $\Rightarrow$ (4) Si $A\mathbf{x} = \mathbf{0}$ tiene única solución, entonces cada columna de la matriz A es una columna pivote y por lo tanto A tiene n posiciones pivote.

(4) $\Rightarrow$ (5) Sea Q la forma escalonada reducida de A. Como A tiene n posiciones pivote, Q tiene un pivote en cada renglón, las posiciones pivote de Q serán $(1, j_1)$, $(2, j_2)$,...., (n, j_n) donde $1 \leq j_1 < j_2 < ... < j_n \leq n$. Luego, necesariamente $j_1 = 1$, $j_2 = 2$,..., $j_n = n$ es decir las posiciones pivote de Q son $(1,1)$, $(2,2)$,...., (n,n). Dado que Q es escalonada reducida los pivotes son unos y en cada columna donde hay un pivote el resto de los elementos son ceros. Así, las columnas de Q son los vectores $\mathbf{e}_1$, $\mathbf{e}_2$,..., $\mathbf{e}_n$ y por lo tanto $Q = I$.

(5) $\Rightarrow$ (1) Si la forma escalonada reducida de A es I entonces ninguna forma escalonada de A tiene un renglón cero y por lo tanto todo sistema que tiene a A como matriz de coeficientes es consistente. En particular, los sistemas $A\mathbf{x}_j = \mathbf{e}_j$, $j = 1, 2, ..., n$ son consistentes y por lo tanto la ecuación $AX = I$ tiene solución lo que prueba que A es invertible.

$\square$

Observación 2.5. La equivalencia (1) $\Leftrightarrow$ (2) es el siguiente resultado importante en relación a los sistemas lineales cuadrados:

El sistema lineal cuadrado $A\mathbf{x} = \mathbf{b}$ tiene única solución si y solo si A es invertible. Dicha solución es $\mathbf{x} = A^{-1}\mathbf{b}$.

Ejemplo 2.29. Considere el sistema lineal cuadrado

$$\begin{cases} x + 4y = 2 \\ -x - 3y = 1 \end{cases}$$

cuya forma matricial es

$$\begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Su matriz de coeficientes $A = \begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}$ es invertible y su inversa es $A^{-1} = \begin{bmatrix} -3 & -4 \\ 1 & 3 \end{bmatrix}$ (véase Ejemplo 2.26). Luego el vector solución es

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}^{-1}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 & -4 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -10 \\ 3 \end{bmatrix},$$

por consiguiente la única solución del sistema es $x = -10$, $y = 3$.◄

Ejercicios 2.3

$\boxed{1.}$ Halle la inversa de A o bien determine que es singular.

a) $A = \begin{bmatrix} -2 & 3 \\ 4 & 1 \end{bmatrix}$

b) $A = \begin{bmatrix} 3 & -1 \\ -6 & 2 \end{bmatrix}$

c) $A = \begin{bmatrix} 1 & -2 & 3 \\ 1 & 2 & 0 \\ 0 & 2 & -3 \end{bmatrix}$

d) $A = \begin{bmatrix} 2 & -2 & 4 \\ 3 & 2 & 3 \\ 7 & -2 & 11 \end{bmatrix}$

e) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$

f) $A = \begin{bmatrix} 1 & -2 & 4 & 5 \\ 1 & 2 & 3 & 0 \\ 1 & 1 & 3 & 0 \\ -1 & 0 & 1 & 2 \end{bmatrix}$

g) $A = \begin{bmatrix} 1 & 2 & -1 \\ 4 & 5 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

h) $A = \begin{bmatrix} 1 & -1 & 0 & 1 \\ 1 & 0 & 3 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$

2. Halle la inversa de $10A$ si $A^{-1} = \begin{bmatrix} 2 & 2 \\ -3 & 3 \end{bmatrix}$.

3. Halle las matrices A^2, A^3 y, si es posible, A^{-2} y A^{-3} siendo $A = \begin{bmatrix} -2 & 0 & 1 \\ 1 & 3 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

$\boxed{4.}$ Calcule todos los valores de a de modo que la siguiente matriz sea no singular:

$$\begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ -1 & 2 & a \end{bmatrix}.$$

5. En caso de ser posible resuelva los sistemas lineales calculando primero la inversa de la matriz de coeficientes.

$\boxed{a)}$ $\left\{ \begin{array}{rcrcrcl} x & - & y & - & z & = & 1 \\ x & & & & -z & = & 4 \\ x & - & 2y & - & z & = & 0 \end{array} \right.$

b) $\left\{ \begin{array}{rcrcrcl} -x & + & y & + & z & = & b \\ -x & & & + & z & = & b^2 \\ & & y & - & z & = & b^3 \end{array} \right.$

6. Demuestre la propiedad 2. del Teorema 2.10.

$\boxed{7.}$ En caso de ser posible, obtenga la matriz X de las siguientes ecuaciones matriciales:

a) $XA + B^t = X$ siendo $A = \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 3 & -1 \\ 1 & 0 \end{bmatrix}$.

b) $2X - A^t = BX$ siendo $A = \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 2 & 3 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$.

c) $(X+2F)\,G = X$ siendo $F = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & -2 \\ 0 & 1 & 3 \end{bmatrix}$ y $G = \begin{bmatrix} 3 & 3 & -2 \\ 0 & 0 & -3 \\ 0 & 0 & 2 \end{bmatrix}$.

d) $XB^{-1} = C^t$ siendo $B^{-1} = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ 3 & -2 & 1 \end{bmatrix}$ y $C = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 2 & -1 \end{bmatrix}$.

e) $N^t - NX = 2X$ siendo $N = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$.

8. Demuestre que una matriz cuadrada que tiene una columna o una fila de ceros es una matriz singular.

9. Sea $A = [a_{ij}]$ una matriz diagonal. Demuestre que si los elementos de la diagonal a_{11}, a_{22},..., a_{nn} son distintos de cero, entonces A es invertible y A^{-1} es una matriz diagonal cuyos elementos de la diagonal son $\frac{1}{a_{11}}$, $\frac{1}{a_{22}}$,..., $\frac{1}{a_{nn}}$.

10. La ecuación real $a^2 = 1$ tiene exactamente dos soluciones. Encuentre al menos cuatro matrices de 2×2 que son autoinversas, es decir, $A^2 = I$ (busque entre las matrices diagonales).

11. Sean A y B matrices cuadradas de orden n. Determine si las siguientes proposiciones son verdaderas o falsas. Justifique su respuesta:

a) Si A y B son invertibles, entonces $(A+B)^{-1} = A^{-1} + B^{-1}$.

b) Si A es invertible y $AB = \mathbf{O}$, entonces $B = \mathbf{O}$.

c) Si $A^2 = \mathbf{O}$ entonces $I - A$ es invertible y su inversa es $A + I$.

d) Si A y B son invertibles entonces $(2AB)^{-1} = \frac{1}{2}A^{-1}B^{-1}$.

e) Si $A^2 + 2A - I = \mathbf{O}$ entonces A es invertible y $A^{-1} = A + 2I$.

12. **Scilab.** Las operaciones suma, multiplicación y potencia se indican con "+", "∗" y "∧", respectivamente. La transpuesta de A se obtiene mediante el comando "A' ". Para calcular la inversa de una matriz A es posible utilizar el comando "inv(A)" o bien simplemente ingresando "A∧ − 1". En el caso en que A no es invertible, devuelve "!–error 19 Problem is singular". También es posible determinar si una matriz A es invertible usando eliminación de Gauss-Jordan aplicando el comando "rref" a la matriz ampliada $[A : I]$ que devuelve su forma escalonada reducida $[C : D]$, si $C = I$ entonces $D = A^{-1}$ mientras que si $C \neq I$ entonces A no es invertible.

Sean las matrices

$$A = \begin{bmatrix} 0{,}1 & 2 & 3 & 4 & -1 & 2 & 2 \\ 1 & -0{,}2 & 2 & -1 & 5 & 4 & -1 \\ 0 & 0 & -1{,}2 & 4 & 2 & 0 & -1 \\ 0{,}1 & 3 & 2 & -0{,}1 & 4 & 2 & 1{,}3 \\ 1 & 0{,}1 & 2 & 0{,}1 & 0 & 0 & 2 \\ 1 & 0{,}4 & 3 & 0{,}1 & 5 & 6 & 7 \\ -0{,}1 & 1 & 0{,}2 & 4 & 1 & 6 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 0{,}3 & 1 & 0 & 0{,}2 & -1 & 1 & 1 \\ -0{,}1 & 0{,}2 & 0 & 1 & -2 & 0{,}2 & 1 \\ 1 & 1 & 1{,}2 & -0{,}3 & 2 & 0 & 1 \\ 0{,}1 & 3 & 2 & -0{,}1 & 0{,}4 & 2 & 1 \\ 2 & 0{,}3 & 0{,}22 & 0{,}1 & 0 & 0 & 2 \\ 1 & 0{,}4 & 3 & 0{,}1 & 1{,}5 & -0{,}2 & -2{,}1 \\ 1 & 1 & 0{,}2 & 4 & 1 & 6 & 0 \end{bmatrix},$$

y el vector $\mathbf{b} = \begin{bmatrix} 0{,}1 \\ 0{,}2 \\ 1 \\ 1{,}1 \\ 2 \\ 0 \\ -1 \end{bmatrix}$. Calcule si es posible:

a) $2A^t - AB$

b) A^{-1}

c) $(10A)^{-1} B$

d) $A\mathbf{b}$

e) A^{-3}

f) $\left(2A^{-1}B^2 - B\right)^3$

g) La solución del sistema lineal $A\mathbf{x} = \mathbf{b}$, hallando previamente la inversa de A.

h) ¿Admite soluciones no triviales el sistema lineal $B\mathbf{x} = \mathbf{0}$?

13. **Análisis de insumos-producción.** El ganador del Premio Nobel, W. Leontief (1906-1999), es famoso por su modelo de una economía basada en insumos y producción. Una de las suposiciones del modelo es que todo cuanto se produce se consumirá. La demanda de los productos de una industria puede provenir de dos fuentes: (1) demanda de industrias diferentes y (2) demanda de otras fuentes que no sean las industrias. Por ejemplo, las compañías de electricidad generan energía que (a) se necesita para operar sus propias plantas, (b) se requiere para satisfacer las necesidades eléctricas de otras empresas, y (c) se necesita para satisfacer las exigencias de los consumidores en general.
Los dos primeros (a) y (b) son ejemplos de la *demanda interindustrial* y el tercero (c) de la *demanda no industrial*. El objeto del análisis de insumos-producción consiste en determinar cuánto producirá una industria, para que ambos tipos de la demanda queden exactamente satisfechos. Es decir determinar cuánto se deberá producir a fin de alcanzar un equilibrio entre oferta y demanda.

La demanda interindustrial suele sintetizarse en una *matriz tecnológica* o de *insumos-producción*. Por ejemplo, suponga que la producción de una industria se mide en dólares y la matriz

$$A = \begin{bmatrix} 0{,}3 & 0{,}3 & 0{,}2 \\ 0{,}1 & 0{,}2 & 0{,}3 \\ 0{,}2 & 0{,}1 & 0{,}4 \end{bmatrix}$$

es la matriz de insumos-producción. El elemento general a_{ij} representará la cantidad de producción requerida de la industria i para generar un dólar en la producción de la industria j. Esta

matriz representa una situación de tres industrias. Por ejemplo, el elemento $a_{12} = 0{,}3$ indica que por cada dólar de producción en la industria 2, el 30 % es aportado por la industria 1.
Sea x_j la producción de la industria j (en dólares) y sea d_j la demanda no industrial (en dólares) de la producción de la indsutria j. En nuestro ejemplo,$j = 1, 2, 3$. Puede formularse un conjunto de ecuaciones simultáneas que, al resueltas, determinen los niveles de producción de x_j en los cuales la oferta y la demanda totales estarán en equilibrio. Las ecuaciones de este sistema presentarán la forma general:

$$\text{Producción de la industria} = \text{demanda interindustrial} + \text{demanda no industrial.}$$

En este ejemplo de tres industrias, el sistema es:

$$\begin{array}{ccccc} x_1 & = & \overbrace{0{,}3x_1 + 0{,}3x_2 + 0{,}2x_3}^{\text{demanda interindustrial}} & + & \overbrace{d_1}^{\text{demanda no industrial}} \\ x_2 & = & 0{,}1x_1 + 0{,}2x_2 + 0{,}3x_3 & + & d_2 \\ x_3 & = & 0{,}2x_1 + 0{,}1x_2 + 0{,}4x_3 & + & d_3 \end{array}$$

Llamando $\mathbf{x}$ al vector columna:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

cuyos elementos son los niveles de equilibrio de producción, y $\mathbf{d}$ al vector columna:

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

cuyos elementos son los niveles de las demandas no industriales, el sistema anterior puede expresarse matricialmente de la forma

$$\mathbf{x} = A\mathbf{x} + \mathbf{d}.$$

Si la matriz de insumos-producción es cuadrada, como en nuestro ejemplo, y la matriz $I - A$ es invertible, es posible despejar $\mathbf{x}$ y obtener el vector de niveles de equilibrio de producción:

$$\mathbf{x} = (I - A)^{-1}\,\mathbf{d}$$

conociendo el vector demanda no industrial $\mathbf{d}$. En nuestro ejemplo de las tres industrias, verifique usando Scilab que:

$$(I - A)^{-1} = \begin{bmatrix} 1{,}837 & 0{,}816 & 1{,}020 \\ 0{,}490 & 1{,}551 & 0{,}939 \\ 0{,}694 & 0{,}531 & 2{,}163 \end{bmatrix}.$$

Supongamos además que los niveles de las demandas no industriales son:

$$d_1 = \$50000000,$$

$$d_2 = \$30000000,$$

y

$$d_3 = \$60000000,$$

entonces verifique usando Scilab que en nuestro ejemplo, los niveles de equilibrio son:

$$\mathbf{x} = \begin{bmatrix} 177530000 \\ 127370000 \\ 180410000 \end{bmatrix},$$

es decir la industria 1 debería generar una producción con un valor de 177530000 dólares, la industria 2 una producción con un valor de 127370000 dólares y la industria 3 una producción con un valor de 180410000 dólares.

a) (**Scilab**) La matriz tecnológica para un modelo de insumos-producción de tres industrias es

$$\begin{bmatrix} 0{,}5 & 0 & 0{,}2 \\ 0{,}2 & 0{,}8 & 0{,}12 \\ 1 & 0{,}4 & 0 \end{bmatrix}.$$

Si la demanda no industrial de la producción de las industrias mencionadas es $d_1 = 5$ millones de dólares, $d_2 = 3$ millones de dólares y $d_3 = 4$ millones de dólares, determine los niveles de producción de las tres industrias.

b) (**Scilab**) La siguiente matriz de insumos-producción para una economía de seis industrias es:

$$\begin{bmatrix} 0{,}180 & 0{,}005 & 0 & 0{,}003 & 0 & 0{,}010 \\ 0{,}005 & 0{,}290 & 0{,}020 & 0{,}002 & 0{,}004 & 0{,}015 \\ 0{,}030 & 0{,}170 & 0{,}450 & 0{,}008 & 0{,}010 & 0{,}006 \\ 0{,}035 & 0{,}040 & 0{,}020 & 0{,}040 & 0{,}010 & 0{,}050 \\ 0{,}010 & 0{,}001 & 0{,}040 & 0{,}250 & 0{,}360 & 0{,}250 \\ 0{,}120 & 0{,}080 & 0{,}100 & 0{,}120 & 0{,}180 & 0{,}230 \end{bmatrix}.$$

Si el vector de las demandas no industriales está dada por:

$$\mathbf{d} = \begin{bmatrix} 20000 \\ 10000 \\ 30000 \\ 5000 \\ 12000 \\ 30000 \end{bmatrix},$$

i) determine los niveles de la producción para las seis industrias.

ii) las demandas interindustriales en las seis industrias.

Capítulo 3

DETERMINANTES

En este capítulo se estudiará uno de los conceptos más importantes del Álgebra Lineal. A cada matriz cuadrada se le puede asignar cierto número real llamado *determinante* de la misma, el cual tiene diversas aplicaciones en esta asignatura, como por ejemplo el cálculo del polinomio característico para encontrar los valores y vectores propios de una matriz, la regla de Cramer para resolver ciertos sistemas lineales, el cálculo de una matriz inversa, etc.

3.1. Introducción

Se utilizará un método de recurrencia para la definición de determinante, esto es, el determinante de una matriz de cuadrada de orden n se define en términos de determinantes de matrices de orden menor.

Se debe tener en cuenta, como en el caso de la inversa de una matriz, que el concepto de determinante es aplicable solo a matrices cuadradas.

Definición (Determinante)**.** Dada una matriz $A = [a_{ij}]$ cuadrada de orden n, el *determinante* de A, que simbolizamos $\det(A)$ o $|A|$, es el número que se obtiene, recursivamente, de la siguiente manera:

i. Si $n = 1$, es decir $A = [a_{11}]$, $\det(A) = a_{11}$.

ii. Si $n \geq 2$,

$$\begin{aligned} \det(A) &= (-1)^{1+1}a_{11}\det(A_{1,1}) + (-1)^{1+2}a_{12}\det(A_{1,2}) + ... + (-1)^{1+n}a_{1n}\det(A_{1,n}) \\ &= \sum_{j=1}^{n}(-1)^{1+j}\ a_{1j}\det(A_{1,j}), \end{aligned}$$

donde $A_{1,j}$ es la matriz cuadrada de orden $n-1$ que se obtiene al eliminar la fila 1 y la columna j de la matriz A.

Ejemplo 3.1. Si $A = [-3]$, una matriz cuadrada de orden $n = 1$, por definición (parte i.) es $\det(A) = -3$. ◀

Determinante de una matriz 2×2

Para cualquier matriz cuadrada de orden 2

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

por definición (de acuerdo a ii.) es:

$$\begin{aligned} \det(A) &= (-1)^{1+1} a_{11} \det(A_{1,1}) + (-1)^{1+2} a_{12} \det(A_{1,2}) = \\ &= (-1)^2 a_{11} \det([a_{22}]) + (-1)^3 a_{12} \det([a_{21}]) = \\ &= a_{11}a_{22} - a_{12}a_{21}. \end{aligned}$$

Luego:

$$\boxed{\text{Si } A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ entonces } \det(A) = |A| = a_{11}a_{22} - a_{12}a_{21}.}$$

Nota: En este texto $\det(A)$ y $|A|$ se usan indistintamente para representar el determinante de una matriz cuadrada A. Es práctica común eliminar los corchetes de una matriz y escribir por ejemplo

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad \text{en lugar de} \quad \left| \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \right|.$$

Ejemplo 3.2. Sean la matrices $A = \begin{bmatrix} -2 & 4 \\ -3 & \frac{1}{2} \end{bmatrix}$ y $B = \begin{bmatrix} 3 & -2 \\ -6 & 4 \end{bmatrix}$, entonces:

$$|A| = \begin{vmatrix} -2 & 4 \\ -3 & \frac{1}{2} \end{vmatrix} = (-2)\left(\frac{1}{2}\right) - 4(-3) = 11,$$

y

$$|B| = \begin{vmatrix} 3 & -2 \\ -6 & 4 \end{vmatrix} = 3\,(4) - (-2)(-6) = 0. \blacktriangleleft$$

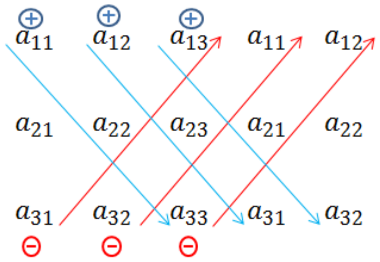
Determinante de una matriz 3×3

Para cualquier matriz cuadrada de orden 3

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

$$\begin{aligned} \det(A) &= (-1)^{1+1} a_{11} \det(A_{1,1}) + (-1)^{1+2} a_{12} \det(A_{1,2}) + (-1)^{1+3} a_{13} \det(A_{1,3}) = \\ &= (-1)^2 a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^3 a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^4 a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = \\ &= a_{11}\,(a_{22}a_{33} - a_{23}a_{32}) - a_{12}\,(a_{21}a_{33} - a_{23}a_{31}) + a_{13}\,(a_{21}a_{32} - a_{31}a_{22}) = \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}. \end{aligned}$$

Hay un artificio para memorizar esta fórmula, llamado **regla de Sarrus**, que se ilustra a continuación:



Se agregan las dos primeras columnas a la derecha de la matriz A. Se suman los productos de los elementos que indican las flechas que con el signo $\oplus$ y se restan los productos de los elementos que indican las flechas con el signo $\ominus$.

Desarrollo de un determinante por cofactores

En la definición de determinante de una matriz A de orden n intervienen los determinantes de n matrices de un orden menor, las matrices $A_{1,j}$, $j = 1, 2, ..., n$, que son aquellas que resultan de eliminar el renglón 1 y la columna j de la matriz A. En lo que sigue de esta sección se verá que el determinante de A se puede calcular también tomando en cuenta otro renglón o cualquier otra columna. Para ello se dará la siguiente definición:

Definición (Menores y Cofactores)**.** Sea $A = [a_{ij}]$ una matriz cuadrada de orden n, $n \geq 2$.

i. El *menor del elemento* a_{ij}, o simplemente el *menor* (i, j), M_{ij}, es el determinante de la matriz cuadrada de orden $n-1$, $A_{i,j}$, que resulta de eliminar el renglón i y la columna j de la matriz A. Esto es: $M_{ij} = \det(A_{i,j})$.

ii. El *cofactor* del elemento a_{ij} de A, o simplemente el *cofactor* (i, j), es el número denotado C_{ij} y dado por $C_{ij} = (-1)^{i+j}\, M_{ij}$.

Ejemplo 3.3. Si $A = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & -1 & 1 \end{bmatrix}$ el menor $(2, 3)$ es el determinante de la matriz que se obtiene de eliminar al renglón 2 y la columna 3 de A. Así, $M_{23} = \det(A_{2,3}) = \begin{vmatrix} -1 & 0 \\ 1 & -1 \end{vmatrix} = 1$. El cofactor $(2, 3)$ es $C_{23} = (-1)^{2+3} M_{23} = (-1)^5(1) = -1$. ◀

Dado que $\det(A_{1,j}) = M_{1j}$, en términos de menores y cofactores, la definición dada de determinante de una matriz A cuadrada de orden $n \geq 2$ se reescribe de la siguiente manera

$$\det(A) = |A| = \sum_{j=1}^{n} (-1)^{1+j}\, a_{1j} M_{1j}.$$

O bien:

$$\det(A) = |A| = \sum_{j=1}^{n} a_{1j} C_{1j}.$$

Sea $A = [a_{ij}]$ una matriz cuadrada de orden n ($n \geq 2$).

Para cada renglón i puede calcularse el número:

$$\Delta_{\mathbf{ren}_i} = \sum_{j=1}^{n} a_{ij}C_{ij} = a_{i1}C_{i1} + a_{i2}C_{i2} + ... + a_{in}C_{in},$$

y para cada columna j el número:

$$\Delta_{\mathbf{col}_j} = \sum_{i=1}^{n} a_{ij}C_{ij} = a_{1j}C_{1j} + a_{2j}C_{2j} + ... + a_{nj}C_{nj},$$

donde C_{ij} es el cofactor del elemento a_{ij}.

Nótese que el número $\Delta_{\mathbf{ren}_1} = \sum\limits_{j=1}^{n} a_{1j}C_{1j}$ es el determinante de la matriz A.

Por ejemplo, si $A = \begin{bmatrix} 2 & 0 & -1 \\ -1 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}$ entonces:

$$\begin{aligned}
\Delta_{\mathbf{ren}_1} &= \det(A) = \sum_{j=1}^{3} a_{1j}C_{1j} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = 2(-1)^{1+1}M_{11} + 0(-1)^{1+2}M_{12} + \\
&+(-1)(-1)^{1+3}M_{13} = 2\begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} + 0 - \begin{vmatrix} -1 & 1 \\ 3 & 2 \end{vmatrix} = 2(1) - (-5) = 7. \\
\Delta_{\mathbf{ren}_2} &= \sum_{j=1}^{3} a_{2j}C_{2j} = a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} = (-1)(-1)^{2+1}M_{21} + 1(-1)^{2+2}M_{22} + \\
&+0(-1)^{2+3}M_{23} = 1\begin{vmatrix} 0 & -1 \\ 2 & 1 \end{vmatrix} + 1\begin{vmatrix} 2 & -1 \\ 3 & 1 \end{vmatrix} + 0 = 1(2) + 1(5) = 7. \\
\Delta_{\mathbf{ren}_3} &= \sum_{j=1}^{3} a_{3j}C_{3j} = a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} = 3(-1)^{3+1}M_{31} + 2(-1)^{3+2}M_{32} + \\
&+1(-1)^{3+3}M_{33} = 3\begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} + (-2)\begin{vmatrix} 2 & -1 \\ -1 & 0 \end{vmatrix} + 1\begin{vmatrix} 2 & 0 \\ -1 & 1 \end{vmatrix} = 3(1) + \\
&+(-2)(-1) + 1(2) = 7. \\
\Delta_{\mathbf{col}_1} &= \sum_{i=1}^{3} a_{i1}C_{i1} = a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31} = 2(-1)^{1+1}M_{11} + (-1)(-1)^{2+1}M_{21} + \\
&+3(-1)^{3+1}M_{31} = 2\begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} + (-1)(-1)\begin{vmatrix} 0 & -1 \\ 2 & 1 \end{vmatrix} + 3\begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} = 2(1) + \\
&+1(2) + 3(1) = 7. \\
\Delta_{\mathbf{col}_2} &= \sum_{i=1}^{3} a_{i2}C_{i2} = a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32} = 0(-1)^{1+2}M_{12} + 1(-1)^{2+2}M_{22} + \\
&+2(-1)^{3+2}M_{32} = 0 + 1\begin{vmatrix} 2 & -1 \\ 3 & 1 \end{vmatrix} + 2(-1)\begin{vmatrix} 2 & -1 \\ -1 & 0 \end{vmatrix} = 1(5) + (-2)(-1) = 7.
\end{aligned}$$

$$\begin{aligned}\Delta_{\mathbf{col}_3} &= \sum_{i=1}^{3} a_{i3}C_{i3} = a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33} = (-1)(-1)^{1+3}M_{13} + 0(-1)^{2+3}M_{23} + \\ &\quad +1(-1)^{3+3}M_{33} = (-1)\begin{vmatrix} -1 & 1 \\ 3 & 2 \end{vmatrix} + 0 + 1\begin{vmatrix} 2 & 0 \\ -1 & 1 \end{vmatrix} = (-1)(-5) + 1(2) = 7.\end{aligned}$$

Obsérvese que todos estos números son iguales e iguales al determinante de A. Esto no es casualidad sino el resultado de un teorema más general que se enuncia a continuación y que no se demostrará en este texto:

Teorema 3.1. Sea A una matriz cuadrada de orden de n ($n \geq 2$). Entonces:

1. $\Delta_{\mathbf{ren}_i} = \Delta_{\mathbf{ren}_k}$ para todo $i = 1, 2, ..., n$ y $k = 1, 2, ..., n$.
2. $\Delta_{\mathbf{col}_j} = \Delta_{\mathbf{col}_h}$ para todo $j = 1, 2, ..., n$ y $h = 1, 2, ..., n$.
3. $\Delta_{\mathbf{ren}_i} = \Delta_{\mathbf{col}_j}$ para todo $i = 1, 2, ..., n$ y $j = 1, 2, ..., n$.

Dado que $\det(A) = \Delta_{\mathbf{ren}_1}$ y, por el teorema anterior, $\Delta_{\mathbf{ren}_1} = \Delta_{\mathbf{ren}_i} = \Delta_{\mathbf{col}_j}$ para todo $i = 1, 2, ..., n$ y $j = 1, 2, ..., n$, se tiene el siguiente resultado que muestra que el determinante de una matriz puede calcularse empleando cualquier renglón o cualquier columna. Más precisamente:

Corolario 3.1. Sea $A = [a_{ij}]$ una matriz cuadrada de orden de n ($n \geq 2$). Entonces:

1. $\det(A) = \sum_{j=1}^{n} a_{ij}C_{ij} = a_{i1}C_{i1} + a_{i2}C_{i2} + ... + a_{in}C_{in}$, para cualquier $i = 1, 2, ..., n$.
2. $\det(A) = \sum_{i=1}^{n} a_{ij}C_{ij} = a_{1j}C_{1j} + a_{2j}C_{2j} + ... + a_{nj}C_{nj}$, para cualquier $j = 1, 2, ..., n$.

En el caso 1. se dice que el determinante se ha desarrollado por cofactores respecto al renglón i y en el caso 2. que se ha desarrollado por cofactores respecto a la columna j.

Es evidente que el cálculo de un determinante puede ser muy tedioso. Por ejemplo, para calcular un determinante de orden 4 deben calcularse 4 determinantes de orden 3 y para calcular un determinante de orden 5, debe calcularse 5 determinantes de orden 4.

Sabiendo ahora que $\det(A)$ puede calcularse desarrollándolo por cualquier renglón o cualquier columna, es claro que para simplificar el cálculo es conveniente elegir una fila o columna que tenga el máximo número de ceros. Así se evita el cálculo de algunos de los menores.

Ejemplo 3.4. Si $A = \begin{bmatrix} -1 & 3 & 0 & 1 \\ 4 & 2 & 2 & 5 \\ 0 & 6 & 3 & 4 \\ 2 & 1 & 0 & 1 \end{bmatrix}$ para calcular el determinante de A conviene desarrollarlo por la tercera columna. Luego:

$$\begin{aligned} |A| &= \sum_{i=1}^{4} a_{i3}C_{i3} = \overbrace{a_{13}}^{0} C_{13} + a_{23}C_{23} + a_{33}C_{33} + \overbrace{a_{43}}^{0} C_{43} = \\ &= 2(-1)^5 M_{23} + 3(-1)^6 M_{33} = -2 \begin{vmatrix} -1 & 3 & 1 \\ 0 & 6 & 4 \\ 2 & 1 & 1 \end{vmatrix} + 3 \begin{vmatrix} -1 & 3 & 1 \\ 4 & 2 & 5 \\ 2 & 1 & 1 \end{vmatrix}, \end{aligned}$$

y calculando los determinantes de orden 3 indicados se obtiene (verifíquelo): $|A| = 43$. ◀

Ya se ha dicho que obtener el determinante de una matriz puede convertirse en un trabajo que requiere una gran cantidad de cálculos a menos que la matriz sea pequeña o que tenga muchos ceros. En la sección siguiente se verán procedimientos que ayudan a minimizar dichos cálculos. Pero además existen algunas matrices que aunque sean grandes sus determinantes se pueden evaluar fácilmente. Se dará la siguiente definición:

Definición (Matrices triangulares)**.** Una matriz cuadrada es *triangular superior* si todos sus elementos debajo de la diagonal principal son ceros. Es *triangular inferior* si todos sus elementos por encima de la diagonal principal son ceros. Una matriz cuadrada es *diagonal* si todos sus elementos fuera de la diagonal principal son ceros.

Es decir, la matriz cuadrada $A = [a_{ij}]$ es triangular superior si $a_{ij} = 0$ para todo $i > j$, es triangular inferior si $a_{ij} = 0$ para todo $i < j$ y es diagonal si $a_{ij} = 0$ para todo $i \neq j$.

Ejemplo 3.5. Las matrices $A = \begin{bmatrix} -1 & 3 & 0 & 0 \\ 0 & 1 & 2 & 5 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 5 \\ 0 & 0 & 3 \end{bmatrix}$ son triangulares superiores, $C = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 8 \end{bmatrix}$ y $D = \begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix}$ son triangulares inferiores y $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ es diagonal.◀

Ejemplo 3.6 (Determinante de una matriz triangular)**.** La matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

es triangular superior. Desarrollando $\det(A)$ por cofactores con respecto a la primer columna:

$$\begin{aligned} \det(A) &= a_{11}C_{11} + 0C_{21} + 0C_{31} + 0C_{41} = a_{11}C_{11} = a_{11}(-1)^2 M_{11} = \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ 0 & a_{33} & a_{34} \\ 0 & 0 & a_{44} \end{vmatrix} = a_{11}a_{22} \begin{vmatrix} a_{33} & a_{34} \\ 0 & a_{44} \end{vmatrix} = a_{11}a_{22}a_{33}a_{44}. \end{aligned}$$

Para una matriz A triangular inferior, se obtendría el mismo resultado: $\det(A) = a_{11}a_{22}a_{33}a_{44}$ (en este caso convendría elegir la última columna o el primer renglón para desarrollarlo). ◀

El ejemplo dado puede generalizarse:

Teorema 3.2. Si $A = [a_{ij}]$ es una matriz triangular de orden n entonces

$$\det(A) = a_{11}a_{22}...a_{nn}.$$

Es decir, el determinante de una matriz triangular es el producto de los elementos de la diagonal principal.

Claramente el determinante de una matriz triangular es cero si y solo si en la diagonal principal hay al menos un elemento cero.

Nótese que dado que una matriz diagonal es una matriz triangular (superior e inferior), entonces su determinante es el producto de los elementos de la diagonal principal.

Ejemplo 3.7. Para las matrices A, B, C, D y E del Ejemplo 3.5 se tiene:

$$\begin{aligned} \det(A) &= (-1)(1)(3)(1) = -3, \\ \det(B) &= (1)(0)(3) = 0, \\ \det(C) &= (1)(1)(8) = 8, \\ \det(D) &= (2)(1) = 2, \\ \det(E) &= (1)(-4)(3) = -12. \blacktriangleleft \end{aligned}$$

Ejercicios 3.1

1. Dada la matriz $A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 4 & 2 \\ -2 & 1 & -1 \end{bmatrix}$.

   a) Calcule todos los cofactores de A.

   b) Obtenga $\det(A)$ desarrollando por cofactores respecto a cada una de las columnas y cada uno de los renglones de A.

$\boxed{2.}$ Calcule los determinantes siguientes:

a) $\begin{vmatrix} 4 & -3 \\ 2 & 5 \end{vmatrix}$

b) $\begin{vmatrix} \sqrt{2} & -\sqrt{2} \\ 1 & -3 \end{vmatrix}$

c) $\begin{vmatrix} 3 & -5 \\ -6 & 10 \end{vmatrix}$

d) $\begin{vmatrix} 0 & 2 & -1 \\ 1 & 1 & 3 \\ 4 & -2 & 3 \end{vmatrix}$

e) $\begin{vmatrix} 2 & 3 & 1 \\ -1 & 2 & 4 \\ 0 & 0 & -2 \end{vmatrix}$

f) $\begin{vmatrix} 0 & -2 & 1 \\ 1 & 4 & 2 \\ 0 & 3 & -2 \end{vmatrix}$

g) $\begin{vmatrix} 4 & -2 & 5 \\ 1 & 4 & 4 \\ -2 & 1 & -2 \end{vmatrix}$

$h)$ $\begin{vmatrix} 1 & 2 & 0 & -1 \\ 2 & 1 & 2 & 2 \\ 0 & 1 & 0 & 1 \\ 2 & -2 & 1 & -3 \end{vmatrix}$

$i)$ $\begin{vmatrix} 2 & 0 & 0 & 1 \\ 3 & -3 & 0 & 0 \\ 2 & 1 & 5 & 0 \\ 3 & 2 & 1 & -1 \end{vmatrix}$

$j)$ $\begin{vmatrix} 2 & 3 & -1 & 8 & 2 \\ 0 & 3 & 1 & -1 & 4 \\ 0 & 0 & 5 & 0 & -9 \\ 0 & 0 & 0 & -1 & 12 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix}$

$k)$ $\begin{vmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{vmatrix}$

$l)$ $\begin{vmatrix} -3 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & -3 \end{vmatrix}$

3. Obtenga el o los valores de x, si existen, que satisfacen las siguientes ecuaciones:

$\boxed{a)}$ $\begin{vmatrix} 2x & -3 \\ 4 & -3 \end{vmatrix} = 0$

$b)$ $\begin{vmatrix} 2-x & 0 & 0 \\ 1 & -3+x & 0 \\ 2 & 1 & 4-6x \end{vmatrix} = 0$

$\boxed{c)}$ $\begin{vmatrix} x & 0 & 2 \\ 2 & 0 & -6+x \\ 0 & x-3 & -1 \end{vmatrix} = \begin{vmatrix} x-1 & 1 & 0 \\ 3 & 0 & x-2 \\ -2 & x-7 & 0 \end{vmatrix}$

3.2. Propiedades de los determinantes

Los determinantes tienen muchas propiedades las cuales se pueden emplear con el fin de reducir el trabajo en el cálculo de los mismos.

Teorema 3.3. Si cualquier renglón o columna de una matriz A es el vector cero entonces $|A| = 0$.

Demostración. Basta con desarrollar el determinante por cofactores con respecto al renglón o columna que contiene solo ceros. $\square$

Ejemplo 3.8. $\begin{vmatrix} 1 & 0 & 2 \\ -3 & 0 & 5 \\ 8 & 0 & 3 \end{vmatrix} = 0$, dado que la segunda columna contiene solo ceros.◀.

Teorema 3.4. Si el $i-$ésimo renglón o la $j-$ésima columna de una matriz $A = [a_{ij}]$ se multiplica por la constante c entonces el determinante de la nueva matriz B es el producto de c por el determinante de A. Esto es:

1. $$|B| = \begin{vmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & ... & \vdots \\ ca_{i1} & ca_{i2} & ... & ca_{in} \\ \vdots & \vdots & ... & \vdots \\ a_{n1} & a_{n2} & ... & a_{nn} \end{vmatrix} = c \begin{vmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & ... & \vdots \\ a_{i1} & a_{i2} & ... & a_{in} \\ \vdots & \vdots & ... & \vdots \\ a_{n1} & a_{n2} & ... & a_{nn} \end{vmatrix} = c\,|A|.$$

2. $$|B| = \begin{vmatrix} a_{11} & a_{12} & ... & ca_{1j} & ... & a_{1n} \\ a_{21} & a_{22} & ... & ca_{2j} & ... & a_{2n} \\ \vdots & \vdots & ... & ... & \vdots & \\ a_{n1} & a_{n2} & ... & ca_{nj} & ... & a_{nn} \end{vmatrix} = c \begin{vmatrix} a_{11} & a_{12} & ... & a_{1j} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2j} & ... & a_{2n} \\ \vdots & \vdots & ... & ... & \vdots & \\ a_{n1} & a_{n2} & ... & a_{nj} & ... & a_{nn} \end{vmatrix} = c\,|A|.$$

Demostración. Se probará aquí 1. Dado que las matrices A y B difieren solo en el i-ésimo renglón, entonces los cofactores de los elementos del renglón i de A, C_{i1}, C_{i2},..., C_{in}, son respectivamente iguales a los cofactores correspondientes al renglón i de B. Considerando el desarrollo de los determinantes de ambas matrices, A y B, respecto al i-ésimo renglón es:

$$|B| = (ca_{i1})C_{i1} + (ca_{i2})C_{i2} + ... + (ca_{in})C_{in} = c\,(a_{i1}C_{i1} + a_{i2}C_{i2} + ... + a_{in}C_{in}) = c\,|A|\,.$$

La demostración de 2. es análoga (se considera el desarrollo de los determinantes de A y B con respecto a la j-ésima columna). □

Observación 3.1. Esta propiedad establece que en un determinante se permite "extraer" un factor común de un renglón o columna del mismo.

Ejemplo 3.9. Supongamos que se quiere calcular el determinante de la matriz $B = \begin{bmatrix} 2 & 0 & -1 \\ -1 & 1 & 0 \\ 15 & 10 & 5 \end{bmatrix}$.

Sea $A = \begin{bmatrix} 2 & 0 & -1 \\ -1 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}$ la matriz que se obtiene de dividir entre 5 (multiplicar por $\frac{1}{5}$) el último renglón de B, es decir, la matriz A se obtiene de realizar una operación elemental de renglón a la matriz B $\left(\frac{1}{5}R_3 \to R_3\right)$. Por el teorema anterior la relación entre $|B|$ y $|A|$ es:

$$|A| = \frac{1}{5}\,|B| \quad \text{o} \quad |B| = 5\,|A|\,.$$

De esta manera se puede simplificar el cálculo de $|B|$ calculando un determinante más fácil de resolver. Dado que $|A| = 7$ entonces

$$|B| = 5\,|A| = 5(7) = 35.$$

En la práctica se procede "sacando factor común" 5 en el último renglón:

$$\begin{vmatrix} 2 & 0 & -1 \\ -1 & 1 & 0 \\ 15 & 10 & 5 \end{vmatrix} = 5 \begin{vmatrix} 2 & 0 & -1 \\ -1 & 1 & 0 \\ 3 & 2 & 1 \end{vmatrix} = 5(7) = 35. \blacktriangleleft$$

El siguiente es un ejemplo de un resultado general muy interesante:

Ejemplo 3.10. Sea la matriz $5A$ donde A es la matriz del ejemplo anterior. Extrayendo sucesivamente el factor común 5 en cada columna:

$$\begin{aligned} |5A| &= \begin{vmatrix} 10 & 0 & -5 \\ -5 & 5 & 0 \\ 15 & 10 & 5 \end{vmatrix} = \begin{vmatrix} 5(2) & 5(0) & 5(-1) \\ 5(-1) & 5(1) & 5(0) \\ 5(3) & 5(2) & 5(1) \end{vmatrix} = 5 \begin{vmatrix} 2 & 5(0) & 5(-1) \\ -1 & 5(1) & 5(0) \\ 3 & 5(2) & 5(1) \end{vmatrix} = \\ &= 5(5) \begin{vmatrix} 2 & 0 & 5(-1) \\ -1 & 1 & 5(0) \\ 3 & 2 & 5(1) \end{vmatrix} = 5^2(5) \begin{vmatrix} 2 & 0 & -1 \\ -1 & 1 & 0 \\ 3 & 2 & 1 \end{vmatrix} = \\ &= 5^3 |A| = 5^3(7) = 875. \end{aligned}$$

El mismo resultado se hubiese obtenido extrayendo sucesivamente factor común 5 en cada renglón.◀

La aplicación reiterada del resultado del Teorema 3.4 permite demostrar el siguiente:

Corolario 3.2. Si A es una matriz cuadrada de orden n y c es un escalar, entonces

$$|cA| = c^n |A|.$$

Otra propiedad se demuestra en el siguiente teorema:

Teorema 3.5. Al intercambiar dos renglones o dos columnas cualesquiera de una matriz A, el determinante de la matriz B obtenida, es el opuesto del determinante de A. Esto es: $|B| = -|A|$.

Demostración. La prueba se hará para el intercambio de dos renglones (análogamente se demuestra para el intercambio de dos columnas).

Es fácil verificar este resultado para matrices cuadradas de orden 2 ó 3.

Supongamos que A es de orden 2 y se intercambian sus renglones. Esto es: $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ y $B = \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix}$ entonces

$$|B| = a_{21}a_{12} - a_{11}a_{22} = -(a_{11}a_{22} - a_{21}a_{12}) = -|A|.$$

Supongamos ahora que A es de orden 3 y se intercambian dos renglones obteniéndose la matriz B. Las matrices A y B coinciden en el renglón l que no fue intercambiado. Los menores (de orden 2) correspondientes a los elementos del renglón l de ambas matrices son opuestos (¿por qué?). Los términos del desarrollo del determinante de B por cofactores correspondiente al renglón l resultan entonces opuestos a los términos del desarrollo del determinante de A por dicho renglón. De este modo, $|B| = -|A|$.

El mismo razonamiento permite generalizar este resultado para una matriz A de orden mayor. □

Actividad 3.1. Dadas las matrices $A = \begin{bmatrix} 1 & 3 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 2 & 1 & 9 & 6 \\ 3 & 2 & 4 & 8 \end{bmatrix}$ y $B = \begin{bmatrix} 5 & 3 & 1 & 2 \\ 3 & -1 & 0 & 4 \\ 9 & 1 & 2 & 6 \\ 4 & 2 & 3 & 8 \end{bmatrix}$. Compruebe, sin hacer cálculos, que $|B| = -|A|$. Halle luego $|A|$ y por consiguiente $|B|$.

Corolario 3.3. Si A tiene dos renglones o columnas iguales entonces $|A| = 0$.

Demostración. Se probará para el caso de dos renglones iguales (análogamente se demuestra para dos columnas iguales).

Supongamos que los renglones h y k de la matriz A son iguales. La matriz B que se obtiene de intercambiar los renglones h y k es la misma matriz A, por lo tanto

$$|B| = |A| . \tag{3.1}$$

Por otro lado, dado que B se obtiene de A por un intercambio de renglones entonces del teorema anterior

$$|B| = -|A| . \tag{3.2}$$

De (3.1) y (3.2) resulta $|A| = -|A|$ y así $|A| = 0$. □

Ejemplo 3.11. $\begin{vmatrix} 2 & 2 & 2 & 9 \\ -1 & 1 & -1 & 8 \\ 3 & 3 & 3 & 6 \\ 4 & 2 & 4 & 8 \end{vmatrix} = 0$ dado que la primer y tercer columnas son iguales.◀

Teorema 3.6. Si a un renglón, o columna, de una matriz A se le suma un múltiplo de otro renglón, o columna, de A, entonces el determinante de la nueva matriz B es igual al determinante de A.

Demostración. Sean R_h y R_k dos renglones arbitrarios de la matriz $A = [a_{ij}]$.

Sea B la matriz que es igual a A exceptuando el renglón $k-$ésimo el cual es la suma del R_k más c veces R_h es decir,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{h1} & a_{h2} & \cdots & a_{hn} \\ \vdots & \vdots & \ldots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \vdots & \vdots & \ldots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{h1} & a_{h2} & \cdots & a_{hn} \\ \vdots & \vdots & \ldots & \vdots \\ a_{k1}+ca_{h1} & a_{k2}+ca_{h2} & \cdots & a_{kn}+ca_{hn} \\ \vdots & \vdots & \ldots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Dado que A y B difieren solo en el k-ésimo renglón entonces los cofactores de los elementos de dicho renglón de A, C_{k1}, C_{k2},..., C_{kn} son iguales a los cofactores del k-ésimo renglón de B.

Desarollando el determinante de B por el k-ésimo renglón:

$$\begin{aligned} |B| &= (a_{k1}+ca_{h1})\,C_{k1} + (a_{k2}+ca_{h2})\,C_{k2} + \ldots + (a_{kn}+ca_{hn})\,C_{kn} = \\ &= (a_{k1}C_{k1} + a_{k2}C_{k2} + \ldots + a_{kn}C_{kn}) + c\,(a_{h1}C_{k1} + a_{h2}C_{k2} + \ldots + a_{hn}C_{kn})\,. \end{aligned}$$

El primer término de la suma es el determinante de A desarrollado por el k-ésimo renglón. El segundo término es 0 puesto que $a_{h1}C_{k1} + a_{h2}C_{k2} + \ldots + a_{hn}C_{kn}$ es el determinante, desarrollado por el k-ésimo renglón, de la matriz

$$A' = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{h1} & a_{h2} & \cdots & a_{hn} \\ \vdots & \vdots & \ldots & \vdots \\ a_{h1} & a_{h2} & \cdots & a_{hn} \\ \vdots & \vdots & \ldots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{array}{l} \\ \\ \\ \\ \leftarrow k\text{-ésimo renglón.} \\ \\ \\ \end{array}$$

que se obtiene de la matriz A al reemplazar el renglón k por el renglón h. Dado que A' tiene dos renglones iguales entonces $|A'| = 0$. Así,

$$|B| = |A| + 0 = |A|\,.$$

Análogamente se demuestra que si a una columna se le suma un múltiplo escalar de otra columna, el determinante no varía. □

Ejemplo 3.12. Sea la matriz $A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & -1 & 0 \\ 6 & 4 & 3 \end{bmatrix}$. Se observa rápidamente que si el 3° renglón se le resta dos veces el 1° se obtiene la matriz $B = \begin{bmatrix} 3 & 2 & 1 \\ 2 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ cuyo determinante es -7.

Así por propiedad enunciada en el teorema anterior es $|A| = |B| = -7$.◄

Operaciones elementales de renglón y determinantes

De acuerado al Teorema 3.4 si se aplica una operación de escalamiento a un renglón de una matriz ($cR_i \to R_i$), el determinante de la matriz original es $\frac{1}{c}$ por el determinante de la matriz nueva. De los teoremas 3.5 y 3.6 se sabe que una operación de intercambio ($R_i \leftrightarrow R_j$) cambia el signo del determinante y que una operación de eliminación ($R_i + cR_j \to R_i$) no cambia el determinante.

El siguiente ejemplo muestra cómo aplicar operaciones elementales de renglón a una matriz A de modo que el cálculo de su determinante involucre el determinante de una matriz triangular el cual, como se sabe, es de cálculo directo.

$$|A| = \begin{vmatrix} 2 & 3 & 4 & 4 \\ -1 & 5 & 0 & -1 \\ 1 & -2 & 3 & 4 \\ 1 & -1 & 1 & 0 \end{vmatrix} \underset{R_1 \leftrightarrow R_3}{=} - \begin{vmatrix} 1 & -2 & 3 & 4 \\ -1 & 5 & 0 & -1 \\ 2 & 3 & 4 & 4 \\ 1 & -1 & 1 & 0 \end{vmatrix} \underset{\substack{R_2 + R_1 \to R_2 \\ R_3 - 2R_1 \to R_3 \\ R_4 - R_1 \to R_4}}{=} - \begin{vmatrix} 1 & -2 & 3 & 4 \\ 0 & 3 & 3 & 3 \\ 0 & 7 & -2 & -4 \\ 0 & 1 & -2 & -4 \end{vmatrix} =$$

$$\underset{\frac{1}{3}R_2 \to R_2}{=} -3 \begin{vmatrix} 1 & -2 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 7 & -2 & -4 \\ 0 & 1 & -2 & -4 \end{vmatrix} \underset{\substack{R_3 - 7R_2 \to R_3 \\ R_4 - R_2 \to R_4}}{=} -3 \begin{vmatrix} 1 & -2 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -9 & -11 \\ 0 & 0 & -3 & -5 \end{vmatrix} =$$

$$\underset{-\frac{1}{9}R_3 \to R_3}{=} -3(-9) \begin{vmatrix} 1 & -2 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & \frac{11}{9} \\ 0 & 0 & -3 & -5 \end{vmatrix} \underset{R_4 + 3R_3 \to R_4}{=} 27 \begin{vmatrix} 1 & -2 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & \frac{11}{9} \\ 0 & 0 & 0 & -\frac{4}{3} \end{vmatrix} =$$

$$= 27\left(-\frac{4}{3}\right) = -36.$$

En este ejemplo, se llegó a una matriz triangular equivalente haciendo uso de las tres operaciones elementales de renglón (intercambio, escalamiento y eliminación), pero siempre es posible llevar una matriz cuadrada a una forma triangular usando únicamente operaciones de eliminación y de intercambio. Por lo tanto si B es una matriz triangular equivalente a A obtenida sin escalamientos, es $\det(A) = \det(B)$ o $\det(A) = -\det(B)$ y, como el determinante de una matriz triangular es el producto de los elementos de la diagonal principal, se tiene entonces el siguiente resultado:

Teorema 3.7. Sea A una matriz cuadrada de orden n y $B = [b_{ij}]$ una matriz triangular obtenida de A a partir de operaciones elementales de renglón sin escalamientos. Entonces

$$|A| = (-1)^k |B| = (-1)^k b_{11} b_{22} ... b_{nn},$$

donde k es el número de operaciones de intercambios entre renglones realizadas.

Actividad 3.2. Aplique el Teorema 3.7, es decir use solo operaciones de eliminación e intercambio, para llevar la matriz A del ejemplo anterior a una matriz triangular equivalente, B. Muestre que $|A| = (-1)^k |B|$ siendo k es el número de intercambios realizados.

El teorema anterior permite concluir una importante caracterización de matrices invertibles.

Corolario 3.4. Una matriz A es invertible si y solo si $\det(A) \neq 0$

Demostración. Sea A una matriz cuadrada de orden n y $B = [b_{ij}]$ una forma escalonada de A (por lo tanto triangular) obtenida mediante operaciones de renglón sin escalamientos. De acuerdo con el teorema anterior:

$$\det(A) = (-1)^k b_{11} b_{22} ... b_{nn}, \tag{3.3}$$

donde k es el número de intercambios de renglones realizados.

Según el Teorema 2.12, A es invertible si y solo si tiene n posiciones pivote. Esto es, A es invertible si y solo si en cualquier (toda) forma escalonada de A los elementos de la diagonal principal son distintos de cero. Luego, A es invertible si y solo si $b_{ii} \neq 0$ para todo $i = 1, 2, ..., n$ que, por (3.3), equivale a $\det(A) \neq 0$. □

Ejemplo 3.13. La matriz $A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 0 & -2 & 5 \end{bmatrix}$ es invertible ya que $\det(A) = 16 \neq 0$.◀

Otras propiedades de los determinantes

Teorema 3.8 (Determinante de la transpuesta)**.** Sea A una matriz cuadrada de orden n. Entonces $|A| = |A^t|$.

Demostración. Es fácil verificar este resultado para el caso $n = 2$:

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

y

$$|A^t| = \begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12} = a_{11}a_{22} - a_{12}a_{21} = |A|.$$

Para el caso $n = 3$, dado que el primer renglón de A es igual a la primera columna de A^t, desarrollando el determinante de A^t por la primera columna:

$$|A^t| = a_{11}M'_{11} - a_{12}M'_{21} + a_{13}M'_{31},$$

donde M'_{i1} es el menor $(i, 1)$ de la matriz A^t, $i = 1, 2, 3$. Desarrollando el determinante de A por el primer renglón:

$$|A| = a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13},$$

donde M_{1i} es el menor $(1, i)$ de la matriz A, $i = 1, 2, 3$.

Para probar que $|A| = |A^t|$ basta verificar que $M_{1i} = M'_{i1}$ para todo $i = 1, 2, 3$.

El menor M_{1i} es el determinante de la matriz $A_{1,i}$ que resulta de suprimir el renglón 1 y la columna i de A, mientras que el menor M'_{i1} es el determinante de la matriz $A'_{i,1}$ que resulta de suprimir el renglón

i y la columna 1 de A^t. Es fácil ver que $A'_{i,1} = (A_{1,i})^t$ y, dado que son matrices cuadradas de orden 2, es $|A_{1,i}| = \left|(A_{1,i})^t\right|$, es decir $M_{1i} = M'_{i1}$ para todo $i = 1, 2, 3$.

Este mismo razonamiento se puede generalizar para una matriz de orden $n > 3$.

□

Actividad 3.3. Sea $A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 0 & -2 & 5 \end{bmatrix}$. Verifique que $\det\left(A^t\right) = \det(A) = 16$.

A continuación enunciaremos, sin probar, el siguiente importante resultado:

Teorema 3.9 (Determinante del producto)**.** Sean A y B matrices cuadrada de orden n. Entonces

$$|AB| = |A|\,|B|\,.$$

En general, si A_1, A_2,..., A_k son k matrices cuadradas de orden n, entonces

$$|A_1A_2...A_k| = |A_1|\,|A_2|\,...\,|A_k|\,.$$

Actividad 3.4. Sean $A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 0 & -2 & 5 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 2 & 0 \\ -5 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Compruebe que $|AB| = |A|\,|B|$.

Nota: No debe inferirse un resultado similar al del teorema anterior para la suma de matrices. No es cierto que determinante de una suma es la suma de los determinantes.

En efecto, sean las matrices $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$.

$$|A + B| = \begin{vmatrix} 2 & 3 \\ 3 & 5 \end{vmatrix} = 1,$$

mientras que

$$|A| + |B| = \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} + \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = -1 + 1 = 0.$$

Así,

$$|A + B| \neq |A| + |B|\,.$$

El Teorema 3.9 proporciona la relación que existe entre el determinante de una matriz invertible y el de su inversa.

Corolario 3.5. Si A es una matriz invertible entonces $|A| \neq 0$ y $\left|A^{-1}\right| = \dfrac{1}{|A|}$.

Demostración. Como la matriz A es invertible entonces $AA^{-1} = I$. Luego,

$$\left|AA^{-1}\right| = |I| = 1,$$

y por el Teorema 3.9,

$$\left|AA^{-1}\right| = |A|\left|A^{-1}\right|,$$

de donde

$$|A|\left|A^{-1}\right| = 1.$$

De la última igualdad se deduce que $|A| \neq 0$ y $\left|A^{-1}\right| = \dfrac{1}{|A|}$. □

Ejemplo 3.14. La matriz $A = \begin{bmatrix} 0 & 2 & 0 \\ -5 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ es invertible. Su inversa es $A^{-1} = \begin{bmatrix} 0 & -\frac{1}{5} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (verifíquelo). Puede comprobarse que $|A| = 10\,(\neq 0)$ y que $\left|A^{-1}\right| = \dfrac{1}{10} = \left(\dfrac{1}{|A|}\right)$. ◄

El Corolario 3.4 agrega un resultado a los del Teorema 2.12 visto en el capítulo 2:

Teorema 3.10. Sea A una matriz A de $n \times n$. Entonces las siguientes proposiciones son equivalentes:

1. $|A| \neq 0$.
2. A es invertible.
3. $A\mathbf{x} = \mathbf{b}$ es consistente con única solución para todo $\mathbf{b} \in \mathbb{R}^n$.
4. $A\mathbf{x} = \mathbf{0}$ tiene única solución la trivial.
5. A tiene n posiciones pivote.
6. A es equivalente (por renglones) a la matriz identidad I.

Demostración. Del Teorema 2.12 se sabe que 2., 3., 4., 5. y 6. son equivalentes. El Teorema 3.4 agrega la equivalencia entre 1. y 2.

Por lo tanto, 1., 2., 3., 4., 5. y 6. son equivalentes.

□

De la equivalencia (1) $\Leftrightarrow$ (4) se deduce el siguiente resultado importante en relación a los sistemas lineales cuadrados homogéneos:

El sistema lineal homogéneo cuadrado $A\mathbf{x} = \mathbf{0}$ tiene soluciones no triviales si y solo si $|A| = 0$.

Ejemplo 3.15. Sean los sistemas homogéneos

$$\left\{\begin{array}{rcrcrcl} 2x & - & 3y & + & z & = & 0 \\ 4x & & & - & 2z & = & 0 \\ 3x & - & y & - & 3z & = & 0 \end{array}\right. \quad \text{y} \quad \left\{\begin{array}{rcrcrcl} 2x_1 & + & 3x_2 & + & x_3 & = & 0 \\ x_1 & - & x_2 & + & 2x_3 & = & 0 \\ x_1 & + & 4x_2 & - & x_3 & = & 0 \end{array}\right.$$

Dado que $\begin{vmatrix} 2 & -3 & 1 \\ 4 & 0 & -2 \\ 3 & -1 & -3 \end{vmatrix} = -26 \neq 0$ entonces el primer sistema tiene única solución $x = y = z = 0$ (solución trivial).

En cambio $\begin{vmatrix} 2 & 3 & 1 \\ 1 & -1 & 2 \\ 1 & 4 & -1 \end{vmatrix} = 0$ por lo cual el segundo sistema tiene soluciones no triviales.◀

Ejercicios 3.2

$\boxed{1.}$ Sin efectuar cálculos, explique por qué los siguientes determinantes son cero:

$$a)\ \begin{vmatrix} -1 & 3 & 0 & 1 \\ 2 & 9 & 0 & 5 \\ 1 & 2 & 0 & 4 \\ 2 & 3 & 0 & -1 \end{vmatrix} \qquad b)\ \begin{vmatrix} -1 & 3 & 1 & 4 \\ 2 & 9 & 1 & -1 \\ -1 & 3 & 1 & 4 \\ 0 & 2 & 5 & -1 \end{vmatrix} \qquad c)\ \begin{vmatrix} 1 & 2 & 1 & -4 \\ 3 & -1 & 1 & 2 \\ 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & -4 \end{vmatrix}$$

$\boxed{2.}$ Calcule los siguientes determinantes de las siguientes matrices aplicando operaciones elementales de renglón:

$$a)\ \begin{bmatrix} 1 & 3 & 1 & 1 \\ 2 & 2 & 1 & 0 \\ -1 & 1 & 1 & 1 \\ 0 & 2 & 2 & -1 \end{bmatrix} \qquad b)\ \begin{bmatrix} 2 & 1 & 0 & 1 \\ 0 & -2 & 1 & 1 \\ 1 & 1 & 3 & -1 \\ 1 & 3 & 0 & -2 \end{bmatrix} \qquad c)\ \begin{bmatrix} 1 & -1 & 0 & 0 \\ 2 & 3 & 4 & 0 \\ 1 & 1 & 1 & 1 \\ -2 & 2 & 1 & -1 \end{bmatrix}$$

3. Sea A una matriz cuadrada de orden n y k un número entero positivo. Teniendo en cuenta el Teorema 3.9, pruebe que
$$\left|A^k\right| = |A|^k .$$
Además, demuestre que si A es invertible entonces
$$\left|A^{-k}\right| = \left|A^{-1}\right|^k .$$

$\boxed{4.}$ Sea A una matriz cuadrada de orden 4 tal que $\det(A) = -3$, y sean las matrices:

- B, obtenida de intercambiar el 1° y 3° renglón de A.
- C, obtenida de sumarle a la 3° columna de A, la 2° columna de A multiplicada por 5.
- D, obtenida multiplicando por -2 el 4° renglón de A.
- $E = 2A$.

Calcule los determinantes de las matrices: B, C, D, E, $F = (-A)^5$, $G = 2A^{-3}$ y $H = 3A^tB^{-1}$.

$\boxed{5.}$ Sea $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = -2$. Calcule los siguientes determinantes:

a) $\begin{vmatrix} a & d & g \\ b & e & h \\ c & f & i \end{vmatrix}$

b) $\begin{vmatrix} a & b & c \\ d & e & f \\ -2g & -2h & -2i \end{vmatrix}$

c) $\begin{vmatrix} 3a & 3b & 3c \\ 3d & 3e & 3f \\ 3g & 3h & 3i \end{vmatrix}$

d) $\begin{vmatrix} g & h & i \\ d & e & f \\ -a & -b & -c \end{vmatrix}$

e) $\begin{vmatrix} g & 2a & d \\ h & 2b & e \\ i & 2c & f \end{vmatrix}$

f) $\begin{vmatrix} a & b & c-2a \\ d & e & f-2d \\ g & h & i-2g \end{vmatrix}$

$\boxed{6.}$ Analice si las siguientes matrices son invertibles.

a) $A = \begin{bmatrix} 2 & 3 \\ -4 & 5 \end{bmatrix}$

b) $B = \begin{bmatrix} 2 & 3 & -1 \\ 1 & 4 & 0 \\ 0 & 2 & 1 \end{bmatrix}$

c) $C = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ 0 & 3 & 4 & 0 \\ 1 & 2 & 5 & 10 \end{bmatrix}$

d) $D = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -3 \\ 4 & -2 & -1 \end{bmatrix}$

7. Determine si existe algún valor de a de manera que la matriz A resulte invertible.

$\boxed{a)}$ $A = \begin{bmatrix} a & -1 & 3 \\ 2 & 1 & a \\ -2a & 1 & 2 \end{bmatrix}$

b) $A = \begin{bmatrix} a & -1 & a \\ 1 & -3 & 4 \\ 2 & -a & 0 \end{bmatrix}$

$\boxed{8.}$ Sean A, B, C y D matrices cuadradas de orden 4 tal que $\det(A) = -2$, $\det(C) = 3$, $\det(D) = 0$ y

$$AB^tC^{-1} = \begin{bmatrix} 2 & -1 & 3 & 5 \\ 1 & 3 & 1 & 0 \\ 2 & 1 & -1 & 4 \\ 1 & 0 & -2 & 1 \end{bmatrix}.$$

a) Determine si la matriz B es invertible.

b) Calcule, si es posible, los determinantes de las matrices: $2B^t$, ABC, $DA + DB$, $-A^4$ y $(-3A)^2$.

9. Analice para qué valores de λ el sistema lineal homogéneo

$$\begin{cases} (\lambda-2)x-y &= 0 \\ x+\lambda y-z &= 0 \\ -x-3y+(\lambda-1)z &= 0 \end{cases}$$

admite soluciones no triviales.

10. Sean A y B matrices cuadradas de orden n. Demuestre que:

   a) $\det(AB)=\det(BA)$.

   b) $\left|AA^t\right|=|A|^2$.

   c) Si B es invertible entonces $\det\left(B^{-1}AB\right)=\det(A)$.

   d) $||A|\,A|=|A|^{n+1}$.

3.3. La inversa a través de la adjunta. La regla de Cramer

En esta sección se verá cómo calcular la inversa de una matriz empleando determinantes. También se desarrollará un método de resolución de sistemas lineales cuadrados con matrices de coeficientes invertibles.

Definición (Matriz de cofactores y matriz adjunta)**.** Dada una matriz $A=[a_{ij}]$ cuadrada de orden $n\geq 2$. La *matriz de cofactores* de A, que se simboliza $\mathrm{Cof}(A)$, es la matriz cuadrada de orden n cuyo elemento (i,j)-ésimo es el cofactor C_{ij} de la matriz A.

La matriz transpuesta de la matriz de cofactores de A es la *matriz adjunta* de A y se simboliza $\mathrm{Adj}(A)$. Es decir:

$$\mathrm{Adj}(A)=(\mathrm{Cof}(A))^t=\begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & \dots & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix}.$$

Ejemplo 3.16. Sea la matriz $A=\begin{bmatrix} 2 & -3 & 1 \\ 4 & 0 & -2 \\ 3 & -1 & -3 \end{bmatrix}$. Los cofactores de A son: $C_{11}=-2$, $C_{12}=6$, $C_{13}=-4$, $C_{21}=-10$, $C_{22}=-9$, $C_{23}=-7$, $C_{31}=6$, $C_{32}=8$ y $C_{33}=12$, por lo cual

$$\mathrm{Cof}(A)=\begin{bmatrix} -2 & 6 & -4 \\ -10 & -9 & -7 \\ 6 & 8 & 12 \end{bmatrix} \quad \text{y} \quad \mathrm{Adj}(A)=\begin{bmatrix} -2 & -10 & 6 \\ 6 & -9 & 8 \\ -4 & -7 & 12 \end{bmatrix}. \blacktriangleleft$$

A continuación se establece un resultado que nos permitirá deducir un teorema aún más importante para obtener la inversa de una matriz:

Lema 3.1. Sea $A = [a_{ij}]$ una matriz cuadrada de orden n.

1. Si $h \neq k$ entonces $\sum_{j=1}^{n} a_{hj} C_{kj} = 0$.
2. Si $l \neq t$ entonces $\sum_{i=1}^{n} a_{il} C_{it} = 0$.

Esto es, la suma de los productos de los elementos de un renglón (o columna) de A por los cofactores correspondientes a los elementos de otro renglón (o columna) de A es igual a cero.

Demostración. Se probará 1. (similar demostración para 2.).

Considere la matriz, B, que se obtiene de la matriz A al reemplazar el renglón k por el renglón h:

$$B = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \dots & \vdots \\ a_{h1} & a_{h2} & \dots & a_{hn} \\ \vdots & \vdots & \dots & \vdots \\ a_{h1} & a_{h2} & \dots & a_{hn} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \leftarrow \text{renglón } k \\ \\ \\ \end{matrix}$$

Dado que las matrices A y B difieren solo en el k-ésimo renglón, los cofactores correspondientes a dicho renglón en la matriz A, C_{k1}, C_{k2},..., C_{kn}, son iguales a los cofactores del k-ésimo renglón en la matriz B. Luego la suma $\sum_{j=1}^{n} a_{hj} C_{kj}$ es el determinante de B (desarrollado por el k-ésimo renglón). Como B tiene dos renglones iguales entonces su determinante es cero. Por lo tanto

$$\sum_{j=1}^{n} a_{hj} C_{kj} = 0.$$

□

Este lema es utilizado en la demostración del teorema que sigue. Primero, veamos el siguiente ejemplo:

Ejemplo 3.17. Sea la matriz A del ejemplo anterior. Observemos la matriz que se obtiene de hacer el producto de la matriz A por su adjunta:

$$\begin{aligned} A \operatorname{Adj}(A) &= \begin{bmatrix} 2 & -3 & 1 \\ 4 & 0 & -2 \\ 3 & -1 & -3 \end{bmatrix} \begin{bmatrix} -2 & -10 & 6 \\ 6 & -9 & 8 \\ -4 & -7 & 12 \end{bmatrix} = \begin{bmatrix} -26 & 0 & 0 \\ 0 & -26 & 0 \\ 0 & 0 & -26 \end{bmatrix} = (-26) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \\ &= (-26)I. \end{aligned}$$

Puede verificarse fácilmente que -26 es el determinante de la matriz A. Por lo tanto

$$A \operatorname{Adj}(A) = |A|\, I. \blacktriangleleft$$

Lo obtenido en el ejemplo anterior no es casual ya que se verifica para cualquier matriz cuadrada:

Teorema 3.11. Sea $A = [a_{ij}]$ una matriz cuadrada de orden n. Entonces

$$A \ \mathrm{Adj}(A) = |A| \, I.$$

Demostración. Sea $D = [d_{ij}]$ la matriz producto de A por su adjunta:

$$D = A \ \mathrm{Adj}(A) = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & ... & \vdots \\ a_{n1} & a_{n2} & ... & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} & ... & C_{n1} \\ C_{12} & C_{22} & ... & C_{n2} \\ \vdots & \vdots & ... & \vdots \\ C_{1n} & C_{2n} & ... & C_{nn} \end{bmatrix}.$$

Por definición de producto de matrices:

$$d_{ij} = a_{i1}C_{j1} + a_{i2}C_{j2} + ... + a_{in}C_{jn}. \tag{3.4}$$

Si $i = j$ la suma en (3.4) es $|A|$ (desarrollado por el i-ésimo renglón) y si $i \neq j$, por Lema 3.1, la suma en (3.4) es igual a 0, es decir

$$d_{ij} = \begin{cases} |A| & \text{si} \quad i = j \\ 0 & \text{si} \quad i \neq j \end{cases},$$

luego

$$D = A \ \mathrm{Adj}(A) = \begin{bmatrix} |A| & 0 & ... & 0 \\ 0 & |A| & ... & 0 \\ \vdots & \vdots & ... & \vdots \\ 0 & 0 & ... & |A| \end{bmatrix} = |A| \, I.$$

□

El teorema anterior nos permite probar el siguiente importante resultado:

Teorema 3.12. Si A es una matriz invertible entonces $A^{-1} = \dfrac{1}{|A|}\mathrm{Adj}(A)$.

Demostración. De acuerdo al teorema anterior:

$$A\mathrm{Adj}(A) = |A| \, I.$$

Premultiplicando por A^{-1},

$$\mathrm{Adj}(A) = |A| \, A^{-1}.$$

Dado que A es invertible, $|A| \neq 0$ (Corolario 3.5). Por lo tanto:

$$A^{-1} = \frac{1}{|A|}\mathrm{Adj}(A).$$

□

Ejemplo 3.18. Considere la matriz $A = \begin{bmatrix} 2 & -3 & 1 \\ 4 & 0 & -2 \\ 3 & -1 & -3 \end{bmatrix}$ del Ejemplo 3.16. Su adjunta es

$$\text{Adj}(A) = \begin{bmatrix} -2 & -10 & 6 \\ 6 & -9 & 8 \\ -4 & -7 & 12 \end{bmatrix}.$$

Dado que $|A| = -26 \neq 0$, A es invertible. Su inversa es:

$$A^{-1} = \frac{1}{|A|}\text{Adj}(A) = \frac{1}{-26}\begin{bmatrix} -2 & -10 & 6 \\ 6 & -9 & 8 \\ -4 & -7 & 12 \end{bmatrix} = \begin{bmatrix} 1/13 & 5/13 & -3/13 \\ -3/13 & 9/26 & -4/13 \\ 2/13 & 7/26 & -6/13 \end{bmatrix}. \blacktriangleleft$$

Nota: Para obtener la matriz $\text{Adj}(A)$ se requiere calcular n^2 cofactores, es decir n^2 determinantes de $(n-1)\times(n-1)$. Por ejemplo, si $n = 10$ hay que hallar 100 determinantes de orden 9×9 y esto lleva al cálculo de 8100 determinantes de 8×8, y así siguiendo. Debido a esta cantidad de cómputos casi no se usa el Teorema 3.12 para el cálculo de A^{-1}, la reducción de la matriz aumentada $[A : I]$ sigue siendo el método a elegir. Sin embargo, el teorema es muy útil para demostrar propiedades de las inversas. También para comprobar una fórmula para la solución de un sistema lineal cuadrado.

El siguiente teorema, denominado Regla de Cramer, es uno de los resultados más conocidos en la historia de la Matemática. Durante muchos años fue fundamental en la enseñanza del Álgebra y de la teoría de ecuaciones. Aunque se usa muy poco en la actualidad, debido al gran número de cálculos requeridos, sigue siendo un resultado muy importante a los fines teóricos.

Teorema 3.13 (La regla de Cramer)**.** Sea A una matriz cuadrada de orden n con determinante distinto de cero. Entonces la única solución del sistema lineal $A\mathbf{x} = \mathbf{b}$ está dada por

$$x_1 = \frac{|A_1|}{|A|}, \quad x_2 = \frac{|A_2|}{|A|}, \quad ... \quad , \quad x_n = \frac{|A_n|}{|A|}$$

donde A_i es la matriz que se obtiene al sustituir la i-ésima columna de A por el vector $\mathbf{b}$. Es decir, si en términos de sus columnas es $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & ... & \mathbf{a}_n \end{bmatrix}$ entonces

$$A_1 = \begin{bmatrix} \mathbf{b} & \mathbf{a}_2 & ... & \mathbf{a}_n \end{bmatrix}, \quad A_2 = \begin{bmatrix} \mathbf{a}_1 & \mathbf{b} & ... & \mathbf{a}_n \end{bmatrix}, \quad ... \quad , \quad A_n = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & ... & \mathbf{b} \end{bmatrix}.$$

Demostración. Como $|A| \neq 0$ la matriz A es invertible. Por consiguiente $A\mathbf{x} = \mathbf{b}$ tiene la solución única: $\mathbf{x} = A^{-1}\mathbf{b}$. Por Teorema 3.12, $A^{-1} = \frac{1}{|A|}\text{Adj}(A)$. Así:

$$\mathbf{x} = A^{-1}\mathbf{b} = \left(\frac{1}{|A|}\text{Adj}(A)\right)\mathbf{b} = \frac{1}{|A|}\begin{bmatrix} C_{11} & C_{21} & ... & C_{n1} \\ \vdots & \vdots & ... & \vdots \\ C_{1i} & C_{2i} & ... & C_{ni} \\ \vdots & \vdots & ... & \vdots \\ C_{1n} & C_{2n} & ... & C_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} \sum_{j=1}^{n} C_{j1}b_j \\ \vdots \\ \sum_{j=1}^{n} C_{ji}b_j \\ \vdots \\ \sum_{j=1}^{n} C_{jn}b_j \end{bmatrix}.$$

Dado que las matrices A y A_i difieren solo en la i-ésima columna, los cofactores correspondientes a esta columna en la matriz A, C_{1i}, C_{2i},..., C_{ni}, son iguales a los cofactores de la i-ésima columna en la matriz A_i. Por lo tanto, desarrollado por la i-ésima columna, es:

$$|A_i| = \sum_{j=1}^{n} b_j C_{ji} = \sum_{j=1}^{n} C_{ji} b_j.$$

Así, el vector solución es:

$$\mathbf{x} = \frac{1}{|A|} \begin{bmatrix} |A_1| \\ \vdots \\ |A_i| \\ \vdots \\ |A_n| \end{bmatrix},$$

y en consecuencia la única solución del sistema está dada por $x_1 = \dfrac{|A_1|}{|A|}$, $x_2 = \dfrac{|A_2|}{|A|}$,..., $x_n = \dfrac{|A_n|}{|A|}$. □

Ejemplo 3.19. Sea el sistema lineal $\begin{cases} 2x - 3y + z = 1 \\ \quad 4x - 2z = 2 \\ 3x - y - 3z = 0 \end{cases}$.

La matriz de coeficientes del sistema es $A = \begin{bmatrix} 2 & -3 & 1 \\ 4 & 0 & -2 \\ 3 & -1 & -3 \end{bmatrix}$ y como $|A| = -26 \neq 0$, el sistema puede resolverse mediante la regla de Cramer. Se pueden verificar los siguientes resultados:

$$|A_1| = \begin{vmatrix} 1 & -3 & 1 \\ 2 & 0 & -2 \\ 0 & -1 & -3 \end{vmatrix} = -22, \quad |A_2| = \begin{vmatrix} 2 & 1 & 1 \\ 4 & 2 & -2 \\ 3 & 0 & -3 \end{vmatrix} = -12 \quad \text{y} \quad |A_3| = \begin{vmatrix} 2 & -3 & 1 \\ 4 & 0 & 2 \\ 3 & -1 & 0 \end{vmatrix} = -6.$$

Entonces la única solución del sistema es

$$x = \frac{|A_1|}{|A|} = \frac{-22}{-26} = \frac{11}{13}, \qquad y = \frac{|A_2|}{|A|} = \frac{-12}{-26} = \frac{6}{13}, \qquad z = \frac{|A_3|}{|A|} = \frac{-6}{-26} = \frac{3}{13}. \blacktriangleleft$$

Ejercicios 3.3

1. Utilice la matriz adjunta para calcular la inversa, si existe, de cada una de las siguientes matrices:

   a) $A = \begin{bmatrix} 2 & 3 \\ 1 & -9 \end{bmatrix}$ *b*) $B = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$

   c) $C = \begin{bmatrix} 1 & 1 & 1 \\ -2 & -1 & 1 \\ 1 & 4 & 0 \end{bmatrix}$ *d*) $D = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & -1 \\ 1 & -1 & 2 \end{bmatrix}$

2. Cuando sea posible, aplique la regla de Cramer para resolver los siguientes sistemas de ecuaciones lineales:

a) $\begin{cases} 2x - 9y & = & 6 \\ x + 3y & = & 7 \end{cases}$

b) $\begin{cases} 2x_1 & -x_2 & & = & 0 \\ x_1 & +x_2 & -x_3 & = & 1 \\ 3x_1 & -4x_2 & +3x_3 & = & 2 \end{cases}$

c) $\begin{cases} 2x & -y & +3z & = & 1 \\ x & -y & +2z & = & 0 \\ 3x & -2y & +5z & = & -3 \end{cases}$

d) $\begin{cases} x_1 & +x_2 & +x_3 & = & -5 \\ x_1 & -x_2 & +x_3 & = & -1 \\ 2x_1 & -x_2 & +3x_3 & = & -3 \end{cases}$

3. Aplique la regla de Cramer para hallar la solución del sistema lineal:

$$\begin{cases} \cos\theta\ x - \operatorname{sen}\theta\ y & = & a \\ \operatorname{sen}\theta\ x + \cos\theta\ y & = & b \end{cases}$$

$\boxed{4.}$ Detemine el o los valores de a de manera que el sistema lineal

$$\begin{cases} ax_1 & +x_2 & -x_3 & = & 1 \\ & 2x_2 & +2x_3 & = & -1 \\ ax_1 & +3x_2 & -x_3 & = & 0 \end{cases}$$

se pueda resolver mediante la regla de Cramer.

5. Sea A una matriz cuadrada de orden $n \geq 2$. Pruebe los siguientes resutados:

   a) A es invertible si y solo si $\mathrm{Adj}(A)$ es invertible.

   b) $|\mathrm{Adj}(A)| = |A|^{n-1}$.

6. **Scilab.** Sea A una matriz cuadrada. Para calcular el determinante de A se usa el comando "det(A)" una vez ingresada la matriz A.

   Sean las matrices

$$A = \begin{bmatrix} 0{,}1 & -1 & 2{,}3 & 3 \\ 1 & 9 & -1 & 0{,}2 \\ 1{,}1 & 8 & 1{,}3 & 3{,}2 \\ 0 & 5 & -1 & 9 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 0 & -1 & 2 & 3 \\ 1 & 9 & -1 & 0{,}2 \\ 1{,}1 & 8 & 0{,}5 & 3{,}2 \\ 0 & 5 & -1 & 9 \end{bmatrix}$$

$$C = \begin{bmatrix} \frac{1}{3} & 2 & -\frac{1}{4} & 0 \\ 1 & -1 & 1 & \frac{1}{3} \\ \frac{1}{2} & 0 & -\frac{3}{4} & 0 \\ 1 & -2 & \frac{1}{2} & -1 \end{bmatrix} \quad , \quad D = \begin{bmatrix} 1 & 1 & -1 & 0 \\ 3 & 2 & 0{,}2 & 0{,}1 \\ 5 & 0{,}1 & 2 & 2 \\ 1 & -0{,}1 & 2 & -0{,}5 \end{bmatrix}$$

   a) Calcule los determinantes de cada una de las matrices dadas. ¿Cuáles de esas matrices son invertibles?.

   b) Calcule: $\left|A^2\right|$, $\left|-3AB^t\right|$, $|AB| - |CD|$, $|ABC|$ y $\left|D^{-2}\right|$.

   c) Aplique la regla de Cramer para resolver el sistema lineal $A\mathbf{x} = \mathbf{b}$ siendo $\mathbf{b} = \begin{bmatrix} 0 \\ -1 \\ 0{,}2 \\ 0{,}1 \end{bmatrix}$.

Capítulo 4

ESPACIOS VECTORIALES

4.1. Espacios vectoriales

En la sección 2.1, se definió el conjunto $\mathbb{R}^n$, las operaciones suma y multiplicación por escalares y en el Teorema 2.1 se establecieron las propiedades fundamentales de dichas operaciones. Estas mismas propiedades se cumplen en muchos otros contextos. Por ejemplo la suma y la multiplicación por escalares en el conjunto de las matrices de $m \times n$ (véase Teorema 2.2).

Se generalizará esta situación para definir un nuevo concepto. Cualquier conjunto de objetos juntamente con dos operaciones que satisfacen las mismas propiedades que la suma y multiplicación por escalares definidas en $\mathbb{R}^n$ se denomina *espacio vectorial*.

Los espacios vectoriales son el corazón del Álgebra Lineal y en este concepto se han sintetizado varias aspectos fundamentales de la Matemática y otras disciplinas. La idea de dotar a ciertos conjuntos con un tipo particular de estructura hizo posible la creación de esta noción matemática.

Definición (Espacio vectorial real). Sea $\mathbb{V}$ un conjunto y dos operaciones llamadas suma y multiplicación por escalar. La suma es una regla que asocia a cada par de elementos $\mathbf{u}$, $\mathbf{v}$ de $\mathbb{V}$ un objeto, $\mathbf{u}+\mathbf{v}$, llamado *suma* de $\mathbf{u}$ y $\mathbf{v}$. La multiplicación por un escalar es una regla que asocia a cada par compuesto por un escalar real k y un elemento $\mathbf{u}$ de $\mathbb{V}$ un objeto, $k\mathbf{u}$, llamado producto de k por $\mathbf{u}$.

El conjunto $\mathbb{V}$ juntamente con las dos operaciones conforman una estructura matemática denominada *espacio vectorial real* si se verifican las diez propiedades siguientes llamadas axiomas de un espacio vectorial:

(S_1) $\mathbf{u}+\mathbf{v}$ está en $\mathbb{V}$, para todo $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{V}$.

(S_2) $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$, para todo $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{V}$.

(S_3) $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$, para todo $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ en $\mathbb{V}$.

(S_4) Existe un elemento distinguido, $\mathbf{0}$, en $\mathbb{V}$ tal que para cualquier $\mathbf{u}$ en $\mathbb{V}$,

$$\mathbf{u}+\mathbf{0}=\mathbf{0}+\mathbf{u}=\mathbf{u}.$$

(S_5) Para cada $\mathbf{u}$ en $\mathbb{V}$, existe un elemento $-\mathbf{u}$ en $\mathbb{V}$ tal que

$$\mathbf{u}+(-\mathbf{u})=(-\mathbf{u})+\mathbf{u}=\mathbf{0}.$$

(M_1) $a\mathbf{u}$ está en $\mathbb{V}$, para todo $\mathbf{u}$ en $\mathbb{V}$ y todo a en $\mathbb{R}$.

(M_2) $a(\mathbf{u}+\mathbf{v})=a\mathbf{u}+a\mathbf{v}$, para todo $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{V}$ y todo a en $\mathbb{R}$.

(M_3) $(a+b)\,\mathbf{u}=a\mathbf{u}+b\mathbf{u}$, para todo $\mathbf{u}$ en $\mathbb{V}$ y todo a, b en $\mathbb{R}$.

(M_4) $a(b\mathbf{u})=(ab)\,\mathbf{u}$, para todo $\mathbf{u}$ en $\mathbb{V}$ y todo a, b en $\mathbb{R}$.

(M_5) $1\mathbf{u}=\mathbf{u}$, para todo $\mathbf{u}$ en $\mathbb{V}$.

A los elementos del conjunto $\mathbb{V}$ se los llama **vectores**. El elemento $\mathbf{0}$ se denomina **vector cero**, o simplemente **el cero del espacio vectorial**, y $-\mathbf{u}$ se llama **vector opuesto** de $\mathbf{u}$. Se puede demostrar que en un espacio vectorial hay un único $\mathbf{0}$ y que cada vector $\mathbf{u}$ de $\mathbb{V}$, tiene un único opuesto (véase Ejercicio 4).

(S_1) y (M_1) son las **propiedades de cerradura** para la suma y multiplicación por escalares, respectivamente. También se dice que el **conjunto** $\mathbb{V}$ es **cerrado** bajo dichas operaciones.

(S_2) y (S_3) son las **propiedades conmutativa** y **asociativa** de la suma, respectivamente.

(M_2) y (M_3) son **propiedades distributivas**.

Debe observarse que en un espacio vectorial el conjunto de vectores $\mathbb{V}$ es no vacío ya que de acuerdo a (S_4) $\mathbb{V}$ debe contener, al menos, un elemento (el vector cero).

Si en la definición anterior los escalares son números complejos se tiene la definición de **espacio vectorial complejo**. En este texto todos los escalares que se utilizan son números reales, por lo tanto en lo que sigue un espacio vectorial será siempre un espacio vectorial real.

Debe tenerse presente que en la definición no se especifica la naturaleza de los llamados vectores ni la de las operaciones. Cualquier clase de objetos puede "servir" como vector. Para comprobar que un determinado conjunto $\mathbb{V}$ de objetos con dos operaciones, una entre elementos del conjunto (suma) y otra entre un escalar y un elemento del conjunto (multiplicación por escalar) es un espacio vectorial, lo único que se requiere es que se satisfagan los 10 axiomas.

Para la verificación de estos axiomas puede ser conveniente seguir el siguiente orden:

1. Verificar que ambas operaciones son cerradas en $\mathbb{V}$: propiedades de cerradura (S_1) y (M_1).
2. Determinar la existencia de un elemento que actúe como cero: axioma (S_4).
3. Verificar la existencia de un elemento opuesto para cada elemento de $\mathbb{V}$: axioma (S_5).
4. Comprobar la vigencia de los restantes axiomas.

Los ejemplos que siguen dan una idea de la diversidad de espacios vectoriales posibles. Los dos primeros ejemplos son los modelos usados para tomar como axiomas, en la definición de espacio vectorial, las propiedades que verifican las operaciones habituales de suma y multiplicación por un escalar en dichos conjuntos.

Ejemplo 4.1. El espacio vectorial $\mathbb{R}^n$ (Casos especiales: $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3$).

El conjunto $\mathbb{V} = \mathbb{R}^n$ con las operaciones habituales o estándar de suma y multiplicación por escalar es un espacio vectorial. En efecto:

1. La suma de vectores y la multiplicación por escalar habituales en $\mathbb{R}^n$ son, por definición, operaciones cerradas en $\mathbb{R}^n$ (véase la sección 2.1). Por lo tanto se satisfacen (S_1) y (M_1).
2. El vector cero de $\mathbb{R}^n$ es el cero de este espacio vectorial. Luego, se satisface (S_4).
3. Para cada $\mathbf{u}$ de $\mathbb{R}^n$, el vector $-\mathbf{u}$ es su vector opuesto. Por lo tanto se satisface (S_5).
4. Todos los restantes axiomas se cumplen de acuerdo al Teorema 2.1.

Este espacio vectorial es el denominado espacio vectorial $\mathbb{R}^n$ o simplemente el espacio $\mathbb{R}^n$. Debe notarse que las operaciones estándar de suma y multiplicación por escalar en el espacio vectorial $\mathbb{R}$, son las habituales suma y multiplicación entre números reales.◄

Ejemplo 4.2. El espacio vectorial $\mathbb{M}_{mn}$.

El conjunto $\mathbb{V} = \mathbb{M}_{mn}$ de todas las matrices reales de $m \times n$, con las operaciones estándar de suma y multiplicación por escalar es un espacio vectorial. En efecto:

1. La suma de matrices y la multiplicación por escalar habituales en $\mathbb{M}_{mn}$ son, por definición, operaciones cerradas en $\mathbb{M}_{mn}$ (véase la sección 2.2). Por lo tanto se satisfacen (S_1) y (M_1).
2. La matriz cero, $\mathbf{O}_{m\times n}$, es el cero de este espacio vectorial. Luego, se satisface (S_4).
3. Para cada matriz A de $\mathbb{M}_{mn}$, la matriz $-A$ es su matriz opuesta. Por lo tanto se satisface (S_5).
4. Todos los restantes axiomas se cumplen de acuerdo al Teorema 2.2.

Este espacio vectorial es el denominado espacio vectorial $\mathbb{M}_{mn}$ o simplemente el espacio $\mathbb{M}_{mn}$.

En este contexto, un "vector" es una matriz de $m \times n$. En particular si $m = n$ se denotará simplemente $\mathbb{M}_n$.◄

Ejemplo 4.3. El espacio vectorial $\mathbb{P}_n$.

El conjunto $\mathbb{V} = \mathbb{P}_n$ formado por el polinomio cero y todos los polinomios a coeficientes reales de grado menor o igual a n, con las operaciones estándar de suma de polinomios y multiplicación por escalar es un espacio vectorial.

Un polinomio genérico en $\mathbb{P}_n$ es de la forma

$$p = a_0 + a_1 x + ... + a_{n-1}x^{n-1} + a_n x^n,$$

donde $a_0, a_1, ..., a_n$ son números reales.

Las operaciones estándar de suma y multiplicación por escalar están definidas por:

Si $p = a_0 + a_1 x + ... + a_{n-1}x^{n-1} + a_n x^n$ y $q = b_0 + b_1 x + ... + b_{n-1}x^{n-1} + b_n x^n$, entonces:

La suma de p y q es

$$p + q = (a_0 + b_0) + (a_1 + b_1)\, x + ... + (a_{n-1} + b_{n-1})\, x^{n-1} + (a_n + b_n)\, x^n.$$

El producto de un escalar c por el polinomio p es

$$cp = (ca_0) + (ca_1)\, x + ... + (ca_{n-1})\, x^{n-1} + (ca_n)\, x^n$$

1. La suma de dos polinomios de $\mathbb{P}_n$ es un polinomio de $\mathbb{P}_n$ ya que $p + q$ es de la forma $c_0 + c_1 x + c_2 x^2 + ... + c_n x^n$ donde $c_i = a_i + b_i$ es un número real, para todo $i = 0, 1, ..., n$.

   También cp es un polinomio de $\mathbb{P}_n$ pues cp es de la forma $d_0 + d_1 x + d_2 x^2 + ... + d_n x^n$ donde $d_i = ca_i$ es un número real, para todo $i = 0, 1, ..., n$.

   Por lo tanto se verifican (S_1) y (M_1).

2. El polinomio cero, $\mathbf{0} = 0 + 0x + ... + 0x^n$ es tal que $p + \mathbf{0} = \mathbf{0} + p$ para todo $p \in \mathbb{P}_n$. En efecto, $p + \mathbf{0} = (a_0 + 0) + (a_1 + 0)\, x + ... + (a_n + 0)\, x^n = a_0 + a_1 x + ... + a_n x^n = p$. Similarmente, $\mathbf{0} + p = p$. Luego, se satisface el axioma (S_4).

3. Es fácilmente comprobable que para cada polinomio $p = a_0 + a_1 x + ... + a_n x^n$ en $\mathbb{P}_n$ el polinomio $-p = (-a_0) + (-a_1)\, x + ... + (-a_n)\, x^n$ es un polinomio en $\mathbb{P}_n$ tal que $p + (-p) = \mathbf{0} = (-p) + p$. Por lo tanto se verifica el axioma (S_5).

4. Probaremos el axioma (M_3), dejando como ejercicio la prueba de los cinco restantes axiomas.

   Sean $a,\ b \in \mathbb{R}$ y $p = a_0 + a_1 x + ... + a_n x^n$ un polinomio en $\mathbb{P}_n$. Debe probarse que es válida la siguiente igualdad: $(a + b)\, p = ap + bp$. En efecto:

$$\begin{aligned}
(a+b)\,p &= (a+b)\,(a_0 + a_1 x + ... + a_n x^n) = && \text{[Def. multiplicación por escalar]}\\
&= ((a+b)\,a_0) + ((a+b)\,a_1)\,x + ... + ((a+b)\,a_n)\,x^n = && \text{[Prop. distributiva en } \mathbb{R}]\\
&= (aa_0 + ba_0) + (aa_1 + ba_1)\,x + ... + (aa_n + ba_n)\,x^n = && \text{[Def. suma de polinomios]}\\
&= (aa_0 + aa_1 x + ... + aa_n x^n) + (ba_0 + ba_1 x + ... + ba_n x^n) = && \text{[Def. multiplicación por escalar]}\\
&= a\,(a_0 + a_1 x + ... + a_n x^n) + b\,(a_0 + a_1 x + ... + a_n x^n) = &&\\
&= ap + bp. &&
\end{aligned}$$

Este espacio vectorial es el denominado espacio vectorial $\mathbb{P}_n$ o simplemente el espacio $\mathbb{P}_n$.
En este contexto, un polinomio $p = a_0 + a_1 x + ... + a_n x^n$ es un vector de este espacio vectorial.◀

Actividad 4.1. Demuestre los axiomas (S_2), (S_3), (M_2), (M_4) y (M_5) del Ejemplo 4.3.

Ejemplo 4.4. El espacio vectorial $\mathbb{P}$.

El conjunto $\mathbb{V} = \mathbb{P}$ formado por todos los polinomios a coeficientes reales, con las operaciones estándar de suma de polinomios y multiplicación por escalar, es un espacio vectorial.

Las operaciones estándar de suma y multiplicación por escalar están definidas de la siguiente manera: si $p = a_0 + a_1x + ... + a_nx^n$ y $q = b_0 + b_1x + ... + b_mx^m$ son dos polinomios de $\mathbb{P}$ entonces la suma de p y q es:

$$p + q = (a_0 + b_0) + (a_1 + b_1)\,x + ... + (a_n + b_n)\,x^n \quad \text{si} \quad m = n,$$
$$p + q = p' + q \quad \text{si} \quad m \neq n,$$

donde $p' = a_0 + a_1x + ... + a_nx^n + ... + 0x^{m-1} + 0x^m$, suponiendo $m > n$. El producto del escalar c por el polinomio p es:

$$cp = (ca_0) + (ca_1)\,x + ... + (ca_n)\,x^n.$$

Este espacio vectorial es el denominado espacio vectorial $\mathbb{P}$ o simplemente el espacio $\mathbb{P}$. ◄

Ejemplo 4.5. El espacio vectorial $\mathbb{F}(D)$.

El conjunto $\mathbb{V} = \mathbb{F}(D)$ de todas las funciones reales definidas en D donde D es un subconjunto de $\mathbb{R}$ (por ejemplo, un intervalo $[a, b]$ o el propio $\mathbb{R}$) con las operaciones estándar de suma de funciones y multiplicación por escalar, es un espacio vectorial.

Las operaciones estándar de suma y multiplicación por escalar están definidas del siguiente modo: si f y g son funciones de $\mathbb{F}(D)$:

La suma de f y g es la función $f + g$ tal que

$$(f + g)\,(x) = f(x) + g(x), \quad \text{para todo} \quad x \in D.$$

El producto de un escalar c por la función f es la función cf tal que

$$(cf)\,(x) = cf(x), \quad \text{para todo} \quad x \in D.$$

1. Las funciones $f + g$ y cf son funciones reales ya que $f(x) + g(x)$ y $cf(x)$ son números reales. Además como f y g están definidas para todo $x \in D$, también $f + g$ y cf están definidas en D. Así, ambas operaciones son cerradas en $\mathbb{F}(D)$ y por lo tanto se verifican los axiomas (S_1) y (M_1).

2. La función constante cero, O, cuya ley es $O(x) = 0$ para todo $x \in D$, es el vector cero de este espacio vectorial. Puede verificarse fácilmente que $f + O = O + f = f$ para todo f en $\mathbb{F}(D)$. En efecto para todo $x \in D$,

$$(f + O)\,(x) = f(x) + O(x) = f(x) + 0 = f(x),$$

y similarmente

$$(O + f)\,(x) = f(x).$$

Por lo tanto, se satisface el axioma (S_4).

3. Es fácilmente comprobable que si f es una función en $\mathbb{F}(D)$ entonces la función $-f$ cuya ley es

$$(-f)(x) = -f(x),$$

es una función en $\mathbb{F}(D)$ tal que $f + (-f) = (-f) + f = O$. Luego, se satisface el axioma (S_5).

La verificación de los restantes axiomas se deja como ejercicio.

Este espacio vectorial es el denominado espacio vectorial $\mathbb{F}(D)$ o simplemente el espacio $\mathbb{F}(D)$. En este contexto, una función real definida en D es un vector de este espacio vectorial.◀

Ejemplo 4.6. El espacio vectorial cero $\{\mathbf{0}\}$.

Sea $\mathbb{V} = \{\alpha\}$, un conjunto que consta del único objeto α. Las únicas operaciones suma y multiplicación por escalar que se pueden dar de modo que sean cerradas en $\mathbb{V}$ son las definidas por:

$$\alpha + \alpha = \alpha,$$

$$c\alpha = \alpha, \quad \text{para todo} \quad c \in \mathbb{R}.$$

Claramente α es el vector cero y el opuesto de α es el mismo α. Los demás axiomas se verifican de manera trivial. Por ejemplo: para probar el axioma (M_4), dado que $c\alpha = \alpha$ para todo c, se tiene que

$$a(b\alpha) = a\alpha = \alpha,$$

y también

$$(ab)\alpha = \alpha.$$

Luego, $a(b\alpha) = (ab)\alpha$ para todo a y b en $\mathbb{R}$.

Los espacios vectoriales unitarios difieren solo en el único objeto que constituye el conjunto $\mathbb{V}$. Este objeto es trivialmente el cero de dicho espacio. Por esta razón se denomina *espacio vectorial cero* a todo espacio vectorial unitario y a su único elemento se lo simboliza $\mathbf{0}$.◀

Ejemplo 4.7. El conjunto formado por el polinomio cero y todos los polinomios a coeficientes reales de grado 2, con las operaciones estándar de suma de polinomios y multiplicación por escalar, no es un espacio vectorial.

En efecto, la suma de dos polinomios de grado 2 no es necesariamente un polinomio de grado 2. Por ejemplo: si $p = 2 + 3x - 5x^2$ y $q = x + 5x^2$ entonces $p + q = 2 + 4x$ que no es un polinomio de grado 2. Por lo tanto no se satisface el axioma de cerradura de la suma.◀

Ejemplo 4.8. El conjunto $\mathbb{R}^2$ con la suma habitual pero con la multiplicación por escalar definida por

$$c(x, y) = (cx, 0),$$

no es un espacio vectorial.

Si bien se satisfacen los primeros seis axiomas: (S_1), (S_2), (S_3), (S_4), (S_5) y (M_1), ya que la suma es la habitual en $\mathbb{R}^2$ y la multiplicación por escalar así definida es cerrada, es fácilmente comprobable que no se satisface el axioma (M_5) ya que

$$1(x, y) = (1x, 0) = (x, 0),$$

y por lo tanto si $y \neq 0$ resulta $1(x, y) \neq (x, y)$.◀

Nota: Es importante observar, como la definición de espacio vectorial establece, que éste se compone de un conjunto $\mathbb{V}$ de "vectores", un conjunto de escalares y dos operaciones con ciertas propiedades especiales.

Un mismo conjunto $\mathbb{V}$ puede ser el conjunto de objetos de dos espacios vectoriales distintos, es decir, bajo operaciones diferentes. Cuando no hay posibilidad de confusión, en todo lo que sigue, se denotará simplemente con $\mathbb{V}$ a un espacio vectorial arbitrario.

De los axiomas de la definición de espacio vectorial, se deducen las siguientes propiedades:

Teorema 4.1. Sean $\mathbb{V}$ un espacio vectorial, $\mathbf{u}$ un vector de $\mathbb{V}$ y c un escalar. Entonces:

1. $0\mathbf{u} = \mathbf{0}$
2. $c\mathbf{0} = \mathbf{0}$
3. $(-1)\mathbf{u} = -\mathbf{u}$
4. Si $c\mathbf{u} = \mathbf{0}$ entonces $c = 0$ ó $\mathbf{u} = \mathbf{0}$

Demostración. Se probará la propiedad 1. y quedará como ejercicio la prueba de las demás propiedades. Sea $\mathbf{u}$ un vector de $\mathbb{V}$. Como $0 + 0 = 0$ entonces

$$0\mathbf{u} = (0+0)\,\mathbf{u}, \tag{4.1}$$

y por el axioma (M_3)

$$(0+0)\,\mathbf{u} = 0\mathbf{u} + 0\mathbf{u}. \tag{4.2}$$

Luego, de (4.1) y (4.2) se tiene:

$$0\mathbf{u} + 0\mathbf{u} = 0\mathbf{u}.$$

Por el axioma (S_5), el vector $0\mathbf{u}$ tiene un opuesto, $-0\mathbf{u}$. Al sumar este vector a los dos miembros de la igualdad anterior se obtiene:

$$(0\mathbf{u} + 0\mathbf{u}) + (-0\mathbf{u}) = 0\mathbf{u} + (-0\mathbf{u})\,,$$

o bien por el axioma (S_3):

$$0\mathbf{u} + (0\mathbf{u} + (-0\mathbf{u})) = 0\mathbf{u} + (-0\mathbf{u})\,,$$

por el axioma (S_5):

$$0\mathbf{u} + \mathbf{0} = \mathbf{0},$$

y por el axioma (S_4)

$$0\mathbf{u} = \mathbf{0}.$$

□

Nota: Cabe la siguiente importante observación: *En un espacio vectorial real el conjunto* $\mathbb{V}$ *consta de un número infinito de vectores o bien de un único vector* (véase Ejercicio 8).

Ejercicios 4.1

1. Demuestre lo enunciado en el Ejemplo 4.4.
2. Demuestre los restantes axiomas del Ejemplo 4.5.
3. Determine si el conjunto dado V con las operaciones indicadas es un espacio vectorial.
    $\boxed{a)}$ V es el conjunto formado por todos los vectores de $\mathbb{R}^3$ que tienen la tercera componente cero (plano xy) con las operaciones habituales de suma y multiplicación por escalar en $\mathbb{R}^3$.
    $\boxed{b)}$ V es el conjunto de las matrices de 2×2 invertibles, con las operaciones habituales de suma y multiplicación por escalar definidas en $\mathbb{M}_2$.
    c) V es el conjunto de todos los polinomios a coeficientes reales con términos constantes no negativos, con las operaciones de suma y multiplicación por escalar habituales en $\mathbb{P}$.
    d) V es el conjunto $\mathbb{R}^2$, con la suma habitual pero con la mutiplicación por escalar definida por $c\,(x, y) = (0, 0)$.
    e) V es el conjunto de todas las funciones reales derivables en $\mathbb{R}$, con las operaciones de suma y multiplicación por escalar habituales en $\mathbb{F}\,(\mathbb{R})$.
4. Demuestre que en un espacio vectorial $\mathbb{V}$, hay un único vector cero y para cada vector $\mathbf{u}$ de $\mathbb{V}$, hay un único opuesto.
5. Demuestre las restantes propiedades del Teorema 4.1.
6. Sea $\mathbb{V}$ un espacio vectorial y $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores de $\mathbb{V}$. Demuestre que si $\mathbf{u}+\mathbf{w} = \mathbf{v}+\mathbf{w}$ entonces $\mathbf{u} = \mathbf{v}$.
7. Sea $\mathbb{V}$ un espacio vectorial, $\mathbf{u}$ y $\mathbf{v}$ vectores de $\mathbb{V}$ y a un escalar. Demuestre que si $a\mathbf{u} = a\mathbf{v}$ entonces $a = 0$, o bien $\mathbf{u} = \mathbf{v}$.
8. Demuestre que el único espacio vectorial que tiene un número finito de vectores es el espacio vectorial cero. Es decir, todo espacio vectorial (con escalares reales) contiene infinitos vectores o solo el vector cero.

4.2. Subespacios vectoriales

Puede verificarse (Ejercicio 3.*a*) de la sección anterior) que el conjunto formado por todos los vectores de $\mathbb{R}^3$ que tienen la tercera componente cero (plano xy) con la suma y multiplicación por escalar habituales, satisface los diez axiomas de un espacio vectorial. Así el plano xy es un subconjunto de $\mathbb{R}^3$ que es, a su vez, un espacio vectorial respecto a las mismas operaciones del espacio vectorial $\mathbb{R}^3$. Se dirá que el plano xy es un *subespacio* del espacio $\mathbb{R}^3$.

Se tiene entonces la siguiente definición:

Definición (Subespacio vectorial). Sea $\mathbb{V}$ un espacio vectorial. Un *subespacio vectorial*, o simplemente un *subespacio de* $\mathbb{V}$ es un subconjunto no vacío, $\mathbb{W}$, de $\mathbb{V}$ que, con las mismas operaciones de suma y multiplicación por escalar definidas en $\mathbb{V}$ es, también, un espacio vectorial.

Ejemplo 4.9. Claramente $\mathbb{P}_n$ es un subespacio del espacio $\mathbb{P}$ ya que $\mathbb{P}_n$ está contenido en $\mathbb{P}$ y es un espacio vectorial con las operaciones habituales de suma de polinomios y multiplicación por escalar.◀

Para probar que un conjunto $\mathbb{W}$ con dos operaciones, suma y multiplicación por escalar, es un espacio vectorial, se deben verificar los 10 axiomas de la definición. Sin embargo, si $\mathbb{W}$ es parte de un conjunto mayor $\mathbb{V}$ del que ya se sabe que es un espacio vectorial con esas mismas operaciones, entonces no es necesario verificar ciertos axiomas para $\mathbb{W}$ porque se "heredan" de $\mathbb{V}$. Por ejemplo no es necesario verificar que $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$ (axioma (S_2)) en $\mathbb{W}$ porque esto se cumple para todos los vectores en $\mathbb{V}$ y, en consecuencia, para todos los vectores en $\mathbb{W}$. Es fácil ver que los otros axiomas "heredados" por $\mathbb{W}$ son: (S_3), (M_2), (M_3), (M_4) y (M_5). Por lo tanto para demostrar que un conjunto $\mathbb{W}$ es un subespacio de un espacio vectorial $\mathbb{V}$ solo es necesario verificar los axiomas (S_1), (S_4), (S_5) y (M_1), que son las dos cerraduras, la existencia de un vector cero y la de los opuestos. El siguiente teorema hace ver que, más aún, se puede omitir la verificación de los axiomas (S_4) y (S_5).

Teorema 4.2 (Caracterización de un subespacio)**.** Si $\mathbb{W}$ es un subconjunto no vacío de un espacio vectorial $\mathbb{V}$, entonces $\mathbb{W}$ es un subespacio de $\mathbb{V}$ si y solo si $\mathbb{W}$ es cerrado bajo la suma y bajo la multiplicación por escalar definidas en $\mathbb{V}$. Es decir:

(a) Si $\mathbf{u}$ y $\mathbf{v}$ son vectores de $\mathbb{W}$, entonces $\mathbf{u}+\mathbf{v}$ está en $\mathbb{W}$.

(b) Si $\mathbf{u}$ es un vector de $\mathbb{W}$ y k es un número real, entonces $k\mathbf{u}$ está en $\mathbb{W}$.

Demostración. Si $\mathbb{W}$ es un subespacio de $\mathbb{V}$ entonces, por definición de subespacio, se satisfacen todos los axiomas de un espacio vectorial; en particular se cumplen los axiomas (S_1) y (M_1), que son las condiciones (a) y (b), respectivamente.

Recíprocamente: debe probarse que $\mathbb{W}$ es un subespacio de $\mathbb{V}$ sabiendo que se cumplen las condiciones (a) y (b). Para ello, debe demostrarse que $\mathbb{W}$ satisface los 10 axiomas de espacio vectorial con las operaciones suma y multiplicación por escalar definidas en $\mathbb{V}$.

Los axiomas (S_1) y (M_1) se satisfacen por hipótesis ya que son las condiciones (a) y (b), respectivamente. Los axiomas (S_2), (S_3), (M_2), (M_3), (M_4) y (M_5) se satisfacen inmediatamente en $\mathbb{W}$ dado que "se heredan" de $\mathbb{V}$. Por lo tanto, para completar la demostración basta con verificar que los axiomas (S_4) y (S_5) se satisfacen en $\mathbb{W}$.

Sea $\mathbf{u}$ un vector cualquiera en $\mathbb{W}$ ($\mathbb{W}$ es no vacío). Por la condición (b), tomando $k=0$, el vector $0\mathbf{u}$ está en $\mathbb{W}$. Dado que $\mathbf{u}$ está en $\mathbb{V}$ y $\mathbb{V}$ es un espacio vectorial, por la propiedad 1 del Teorema 4.1, resulta $0\mathbf{u}=\mathbf{0}$ (el cero de $\mathbb{V}$). Luego, el $\mathbf{0}$ de $\mathbb{V}$ es un vector de $\mathbb{W}$ y como $\mathbf{v}+\mathbf{0}=\mathbf{v}=\mathbf{0}+\mathbf{v}$ para todo $\mathbf{v}$ en $\mathbb{W}$, resulta que el $\mathbf{0}$ de $\mathbb{V}$ es el vector cero de $\mathbb{W}$. Por lo tanto el axioma (S_4) se verifica en $\mathbb{W}$.

Dado cualquier vector $\mathbf{u}$ en $\mathbb{W}$, por la condición (b), $(-1)\,\mathbf{u}$ está en $\mathbb{W}$. Dado que $\mathbf{u}$ está en $\mathbb{V}$ y $\mathbb{V}$ es un espacio vectorial, por la propiedad 3 del Teorema 4.1, $(-1)\,\mathbf{u}=-\mathbf{u}$. Luego, para todo $\mathbf{u}$ en $\mathbb{W}$, $-\mathbf{u}$ está en $\mathbb{W}$ y como $\mathbf{u}+(-\mathbf{u})=\mathbf{0}=(-\mathbf{u})+\mathbf{u}$, resulta que $-\mathbf{u}$ es el opuesto de $\mathbf{u}$ en $\mathbb{W}$. De esta forma, se ha probado que se satiface el axioma (S_5) en $\mathbb{W}$. □

La Figura 4.1 representa lo que establece el teorema de caracterización de un subespacio:

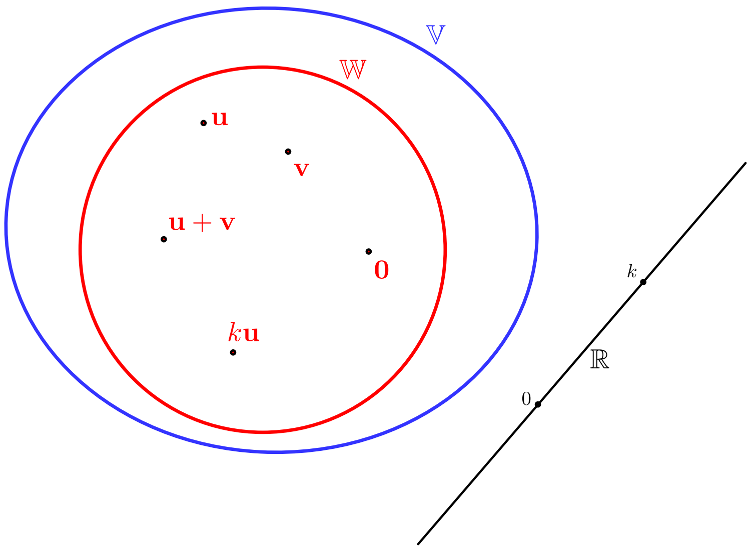


Figura 4.1: $\mathbb{W}$ es un subespacio de $\mathbb{V}$

Un resultado inmediato de la definición de subespacio vectorial es el siguiente:

> Todo subespacio de un espacio vectorial $\mathbb{V}$ contiene al vector cero de $\mathbb{V}$.

En efecto, si $\mathbb{W}$ es un subespacio de $\mathbb{V}$ entonces $\mathbb{W}$ es no vacío y es cerrado bajo la multiplicación por escalares, por lo tanto (como se probó en el teorema anterior), el vector cero de $\mathbb{V}$ está en $\mathbb{W}$.

Este hecho con frecuencia facilita ver que un cierto subconjunto de $\mathbb{V}$ no es un subespacio. Es decir, un subconjunto de un espacio vectorial $\mathbb{V}$ que no contiene al vector cero no es un subespacio.

Por ejemplo, el subconjunto W de $\mathbb{R}^3$ formado por todos los puntos (x, y, z) de $\mathbb{R}^3$ que satisfacen la ecuación (un plano en el espacio):

$$x + 2y - 3z + 1 = 0,$$

no es un subespacio de $\mathbb{R}^3$. En efecto, el vector cero de $\mathbb{R}^3$, $\mathbf{0} = (0, 0, 0)$, no satisface la ecuación.

Es claro que la existencia del vector cero en el conjunto W no es condición suficiente para que W sea un subespacio de $\mathbb{V}$: ¡deben verificarse las dos cerraduras en W!

Sea, por ejemplo, el subconjunto W de $\mathbb{R}^2$ formado por todos los puntos (x, y) de $\mathbb{R}^2$ tales que

$$x^2 - y^2 = 0.$$

Es fácil ver que el vector cero $\mathbf{0} = (0, 0)$ de $\mathbb{R}^2$ satisface la ecuación, por lo que $\mathbf{0}$ está en W. Pero W no es cerrado bajo la suma. Si se consideran los elementos $\mathbf{u} = (1, -1)$ y $\mathbf{v} = (1, 1)$ de W, su suma $\mathbf{u} + \mathbf{v} = (2, 0)$ no es un elemento de W. Luego, W no es un subespacio de $\mathbb{R}^2$.

Ejemplo 4.10. Subespacios triviales y subespacios propios

En todo espacio vectorial $\mathbb{V}$ el subconjunto unitario $\{\mathbf{0}\}$, es un subespacio de $\mathbb{V}$, ya que en él se verifican las dos cerraduras (a) y (b) del Teorema 4.2: $\mathbf{0} + \mathbf{0} = \mathbf{0}$ y $k\mathbf{0} = \mathbf{0}$ para todo k real.

Es inmediato, por definición de subespacio, que el propio $\mathbb{V}$ es un subespacio de sí mismo.

Así, todo espacio vectorial $\mathbb{V}$ contiene los subespacios $\{\mathbf{0}\}$ y $\mathbb{V}$ (que coinciden si $\mathbb{V} = \{\mathbf{0}\}$). Dichos subespacios son los llamados *subespacios triviales* de $\mathbb{V}$.

Si $\mathbb{V} \neq \{\mathbf{0}\}$, además de los subespacios triviales, hay otros subespacios más interesantes que son llamados *subespacios propios* de $\mathbb{V}$.◀

Nota: A menos que se indiquen otras operaciones, cuando se mencionen a los espacios vectoriales $\mathbb{R}^n$, $\mathbb{P}$, $\mathbb{P}_n$, $\mathbb{M}_{mn}$ o $\mathbb{F}(D)$ se entenderá que son los espacios vectoriales con respecto a las operaciones habituales, como se definieron en la Sección 4.1.

Ejemplo 4.11. Todo plano que contiene al origen es un subespacio de $\mathbb{R}^3$.

Sea $\mathbb{W}$ el conjunto de los puntos (x, y, z) de $\mathbb{R}^3$ que satisfacen la ecuación

$$ax + by + cz = 0,$$

donde a, b y c son constantes reales no todas cero.

Claramente $\mathbb{W}$ es un subconjunto no vacío de $\mathbb{R}^3$ dado que por ejemplo $\mathbf{0} = (0, 0, 0)$ está en $\mathbb{W}$.

Por otro lado, sea un escalar k y dos puntos $\mathbf{u} = (x_1, y_1, z_1)$ y $\mathbf{v} = (x_2, y_2, z_2)$ del conjunto $\mathbb{W}$, es decir puntos que satisfacen la ecuación $ax + by + cz = 0$. Para demostrar que $\mathbb{W}$ es cerrado bajo la suma, se debe probar que $\mathbf{u} + \mathbf{v} = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$ está en $\mathbb{W}$. Dado que:

$$\begin{aligned} a(x_1 + x_2) + b(y_1 + y_2) + c(z_1 + z_2) &= ax_1 + ax_2 + by_1 + by_2 + cz_1 + cz_2 = \\ &= (ax_1 + by_1 + cz_1) + (ax_2 + by_2 + cz_2) = \\ &= 0 + 0 = 0, \end{aligned}$$

entonces $\mathbf{u} + \mathbf{v}$ es un punto de $\mathbb{W}$ puesto que satisface la ecuación del plano.

De manera similar para probar que $\mathbb{W}$ es cerrado bajo la multiplicación por escalares, debe demostrarse que $k\mathbf{u} = (kx_1, ky_1, kz_1)$ está en $\mathbb{W}$. Como:

$$a(kx_1) + b(ky_1) + c(kz_1) = k(ax_1) + k(by_1) + k(cz_1) = k(ax_1 + by_1 + cz_1) = k(0) = 0,$$

entonces $k\mathbf{u}$ satisface la ecuación del plano y por lo tanto es un punto de $\mathbb{W}$.

Así, $\mathbb{W}$ es cerrado bajo la suma y la multiplicación por escalares y por lo tanto es un subespacio de $\mathbb{R}^3$.◀

Actividad 4.2. Pruebe que toda recta en el espacio que contiene al origen de coordenadas es un subespacio de $\mathbb{R}^3$.

Actividad 4.3. Pruebe que toda recta en el plano que contiene al origen de coordenadas es un subespacio de $\mathbb{R}^2$.

Ejemplo 4.12. El conjunto solución de un sistema lineal homogéneo de m ecuaciones con n incógnitas es un subespacio de $\mathbb{R}^n$.

Se facilitará la demostración utilizando la forma matricial del sistema, $A\mathbf{x} = \mathbf{0}$, donde A es la matriz de $m \times n$ de coeficientes y $\mathbf{x}$ es el vector de incógnitas.

Claramente, el conjunto solución de $A\mathbf{x} = \mathbf{0}$ es no vacío ya que el vector $\mathbf{0}$ de $\mathbb{R}^n$ es solución del sistema. Por otro lado, sean $\mathbf{x}_1$ y $\mathbf{x}_2$ soluciones de $A\mathbf{x} = \mathbf{0}$ y c un escalar, es decir $A\mathbf{x}_1 = \mathbf{0}$ y $A\mathbf{x}_2 = \mathbf{0}$. Entonces:

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

lo que prueba que $\mathbf{x}_1 + \mathbf{x}_2$ es también solución del sistema. Además:

$$A(c\mathbf{x}_1) = c(A\mathbf{x}_1) = c\mathbf{0} = \mathbf{0},$$

por lo que $c\mathbf{x}_1$ es solución del sistema. Se ha probado que el conjunto solución de $A\mathbf{x} = \mathbf{0}$ es cerrado bajo la suma y la multiplicación por escalares y por lo tanto es un subespacio de $\mathbb{R}^n$.◀

Actividad 4.4. ¿Es el conjunto de soluciones de un sistema lineal consistente de m ecuaciones con n incógnitas, $A\mathbf{x} = \mathbf{b}$ con $\mathbf{b} \neq \mathbf{0}$, un subespacio de $\mathbb{R}^n$?

Ejemplo 4.13. Sea $\mathbb{V}$ un espacio vectorial y $\mathbf{u}$ un vector fijo de $\mathbb{V}$. Se mostrará que el conjunto $\mathbb{W}$ de todos los múltiplos escalares de $\mathbf{u}$:

$$\mathbb{W} = \{c\mathbf{u} : c \in \mathbb{R}\}$$

es un subespacio de $\mathbb{V}$.

En efecto, $\mathbb{W}$ no es vacío puesto que si $c = 0$ entonces $0\mathbf{u} = \mathbf{0}$ es un elemento de $\mathbb{W}$.

Por otro lado, sea k un escalar y dos vectores $\mathbf{v}$ y $\mathbf{w}$ del conjunto $\mathbb{W}$ es decir $\mathbf{v} = a\mathbf{u}$ y $\mathbf{w} = b\mathbf{u}$ con a y b en $\mathbb{R}$. Para demostrar que $\mathbb{W}$ es cerrado bajo la suma, se debe probar que $\mathbf{v} + \mathbf{w}$ está en $\mathbb{W}$. Dado que:

$$\mathbf{v} + \mathbf{w} = a\mathbf{u} + b\mathbf{u} = (a + b)\,\mathbf{u},$$

entonces $\mathbf{v} + \mathbf{w}$ es un múltiplo escalar de $\mathbf{u}$. De manera similar para probar que $\mathbb{W}$ es cerrado bajo la multiplicación por escalar, debe demostrarse que $k\mathbf{v}$ está en $\mathbb{W}$. Como:

$$k\mathbf{v} = k(a\mathbf{u}) = (ka)\,\mathbf{u},$$

entonces $k\mathbf{v}$ es un múltiplo escalar de $\mathbf{u}$. Así, $\mathbb{W}$ es cerrado bajo la suma y la multiplicación por escalares, por lo cual es un subespacio de $\mathbb{V}$.

Nótese que las rectas en el plano o en el espacio que contienen al origen de coordenadas, constituyen un caso particular de este ejemplo para $\mathbb{V} = \mathbb{R}^2$ o $\mathbb{R}^3$, respectivamente. ◀

Ejemplo 4.14. Sea W el conjunto de los polinomios $p = a + bx + cx^2$ de $\mathbb{P}_2$ tales que $a + 2b = c$. Se mostrará que W es un subespacio de $\mathbb{P}_2$.

Es fácil ver que W es un subconjunto no vacío de $\mathbb{P}_2$. Por ejemplo, el polinomio cero de $\mathbb{P}_2$, $\mathbf{O} = 0 + 0x + 0x^2$, pertenece a W ya que satisface la condición requerida: $0 + 2(0) = 0$. Otros polinomios de $\mathbb{W}$ son, por ejemplo: $2 - 5x - 8x^2$ y $1 + x^2$.

Por otro lado, sea k un escalar, y dos polinomios $p = a_1 + b_1 x + c_1 x^2$ y $q = a_2 + b_2 x + c_2 x^2$ de W, es decir polinomios tales que $a_1 + 2b_1 = c_1$ y $a_2 + 2b_2 = c_2$. Para verificar si W es cerrado bajo la suma, debe analizarse si el polinomio $p+q = (a_1 + a_2) + (b_1 + b_2)\,x + (c_1 + c_2)\,x^2$ satisface la condición requerida. Dado que:

$$(a_1 + a_2) + 2\,(b_1 + b_2) = a_1 + a_2 + 2b_1 + 2b_2 = (a_1 + 2b_1) + (a_2 + 2b_2) = c_1 + c_2,$$

se concluye que $p + q$ es un polinomio de W.

Además, debe verificarse que W es cerrado bajo la multiplicación por escalares. Para ello, se debe analizar si $kp = (ka_1) + (kb_1)\,x + (kc_1)\,x^2$, es un polinomio de W. Como:

$$ka_1 + 2\,(kb_1) = k\,(a_1 + 2b_1) = kc_1,$$

entonces kp está en W.

Se ha probado que el conjunto W es un subespacio de $\mathbb{P}_2$.◀

Ejemplo 4.15. $\mathbb{W} = \left\{ \begin{bmatrix} a & -b & 0 \\ b & 0 & 2a \end{bmatrix} : a, b \in \mathbb{R} \right\}$ es un subespacio de $\mathbb{M}_{23}$.

En efecto, la matriz cero $\mathbf{O}$ de $\mathbb{M}_{23}$ está en $\mathbb{W}$ ($a = b = 0$) por lo que $\mathbb{W}$ es un subconjunto no vacío de $\mathbb{M}_{23}$. Otras matrices de $\mathbb{W}$ son, por ejemplo:

$$\begin{bmatrix} 2 & -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 4 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} -\sqrt{3} & -2 & 0 \\ 2 & 0 & -2\sqrt{3} \end{bmatrix}.$$

Sean dos matrices $A = \begin{bmatrix} a_1 & -b_1 & 0 \\ b_1 & 0 & 2a_1 \end{bmatrix}$ y $B = \begin{bmatrix} a_2 & -b_2 & 0 \\ b_2 & 0 & 2a_2 \end{bmatrix}$ de $\mathbb{W}$ y un número real k. Entonces:

$$\begin{aligned} A + B &= \begin{bmatrix} a_1 & -b_1 & 0 \\ b_1 & 0 & 2a_1 \end{bmatrix} + \begin{bmatrix} a_2 & -b_2 & 0 \\ b_2 & 0 & 2a_2 \end{bmatrix} = \begin{bmatrix} a_1 + a_2 & -b_1 - b_2 & 0 + 0 \\ b_1 + b_2 & 0 + 0 & 2a_1 + 2a_2 \end{bmatrix} = \\ &= \begin{bmatrix} a_1 + a_2 & -(b_1 + b_2) & 0 \\ b_1 + b_2 & 0 & 2\,(a_1 + a_2) \end{bmatrix}, \end{aligned}$$

por lo que $A + B$ tiene la forma de las matrices de $\mathbb{W}$.

En cuanto a la multiplicación por escalares,

$$kA = k \begin{bmatrix} a_1 & -b_1 & 0 \\ b_1 & 0 & 2a_1 \end{bmatrix} = \begin{bmatrix} k\,(a_1) & k\,(-b_1) & k\,(0) \\ k\,(b_1) & k\,(0) & k\,(2a_1) \end{bmatrix} = \begin{bmatrix} ka_1 & -(kb_1) & 0 \\ kb_1 & 0 & 2\,(ka_1) \end{bmatrix},$$

y por lo tanto kA es una matriz de $\mathbb{W}$.

Luego, $\mathbb{W}$ es cerrado bajo la suma y la multiplicación por escalares por lo cual es un subespacio de $\mathbb{M}_{23}$.◀

Ejemplo 4.16. El conjunto $C[a, b]$ formado por todas las funciones reales continuas en el intervalo cerrado $[a, b]$, es un subespacio de $\mathbb{F}\,[a, b]$.

En efecto, la función constante cero, O, de $\mathbb{F}\,[a, b]$, $O(x) = 0$ para toda $x \in [a, b]$, es una función continua. Por lo tanto, O es una función de $C[a, b]$.

Por otro lado, sean f y g funciones de $C[a,b]$ y k un número real. Del Cálculo se sabe que la suma de funciones continuas en un conjunto es una función continua en dicho conjunto, por lo que $f+g$ está en $C[a,b]$.

Además, el producto de una constante k por una funnción continua f en un conjunto es una función continua en dicho conjunto, de lo que se tiene que kf está en $C[a,b]$.

Así, $C[a,b]$ es cerrado bajo la suma y la multiplicación por escalares por lo cual es un subespacio de $\mathbb{F}[a,b]$. ◄

Nota: Como se probó anteriormente, el conjunto

$$W = \left\{(x,y) \in \mathbb{R}^2 : x^2 - y^2 = 0\right\}$$

no es un subespacio de $\mathbb{R}^2$. Para ello, se mostró un par de elementos de W cuya suma no está en W.

En general, para probar que un subconjunto W de un espacio vectorial $\mathbb{V}$ no es un subespacio de $\mathbb{V}$ basta mostrar una de las siguientes:

- que el $\mathbf{0}$ de $\mathbb{V}$ no está en W,
- un par de elementos $\mathbf{u}$ y $\mathbf{v}$ de W cuya suma $\mathbf{u}+\mathbf{v}$ no esté en W (no se cumple la cerradura de W bajo la suma),
- un elemento $\mathbf{u}$ de W y un escalar k, cuyo producto $k\mathbf{u}$ no esté en W (no se cumple la cerradura de W bajo la multiplicación por escalares).

Los siguientes son ejemplos de subconjuntos de espacios vectoriales que no son subespacios:

Ejemplo 4.17. $W = \{A \in \mathbb{M}_2 : \det(A) = 0\}$ no es un subespacio de $\mathbb{M}_2$.

W no es cerrado bajo la suma. En efecto, para las matrices $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$ y $B = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$ de W es $A+B = \begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix}$ y $\det(A+B) = 10 \neq 0$, por lo que $A+B$ no está en W.◄

Ejemplo 4.18. $W = \left\{a + bx + cx^2 + dx^3 \in \mathbb{P}_3 : d \geq 0\right\}$ no es un subespacio de $\mathbb{P}_3$.

No se verifica la cerradura de W bajo la multiplicación por escalares. En efecto, si se considera un polinomio p con $d > 0$, por ejemplo, $p = 1 + x + 5x^3$ y un número $k < 0$, por ejemplo $k = -2$, entonces el producto $kp = (-2)p = -2 - 2x - 10x^3$ no está en W puesto que el coeficiente de x^3 es negativo.◄

Ejercicios 4.2

$\boxed{1.}$ Determine si el conjunto W es un subespacio del espacio $\mathbb{V}$ indicado:

a) W es el conjunto de todos los puntos (x,y) de $\mathbb{R}^2$ tal que $y = 0$; $\mathbb{V} = \mathbb{R}^2$.

b) W es el conjunto de todos los puntos (x,y) de $\mathbb{R}^2$ tal que $xy = 0$; $\mathbb{V} = \mathbb{R}^2$.

c) W es el conjunto de todos los puntos (x,y,z) de $\mathbb{R}^3$ tal que $x=y$; $\mathbb{V}=\mathbb{R}^3$.

d) W es el conjunto de todos los puntos (x,y,z) de $\mathbb{R}^3$ tal que $-x+2y=z$; $\mathbb{V}=\mathbb{R}^3$.

e) W es el conjunto de todos los puntos (x,y,z) de $\mathbb{R}^3$ tal que $x^2+y^2=z$; $\mathbb{V}=\mathbb{R}^3$.

f) W es el conjunto de todos los polinomios $a+bx+cx^2$ de $\mathbb{P}_2$ tal que $a+b+c=0$; $\mathbb{V}=\mathbb{P}_2$.

g) W es el conjunto de todos los polinomios $a+bx+cx^2$ de $\mathbb{P}_2$ tal que $abc=0$; $\mathbb{V}=\mathbb{P}_2$.

h) W es el conjunto de los polinomios de $\mathbb{P}_3$ de la forma $a-ax+bx^2-2bx^3$, $a,b\in\mathbb{R}$; $\mathbb{V}=\mathbb{P}_3$.

i) W es el conjunto de todas las matrices A de $\mathbb{M}_2$ tal que A es invertible; $\mathbb{V}=\mathbb{M}_2$.

j) W es el conjunto de todas las matrices de $\mathbb{M}_2$ de la forma $\begin{bmatrix} a & 0 \\ b & 2a \end{bmatrix}$ con $a,b\in\mathbb{R}$; $\mathbb{V}=\mathbb{M}_2$.

k) W es el conjunto de todas las matrices A de $\mathbb{M}_n$ tal que A es simétrica; $\mathbb{V}=\mathbb{M}_n$.

l) $W=\{X\in\mathbb{M}_2 : AX=XA\}$ donde A es una matriz fija de $\mathbb{M}_2$; $V=\mathbb{M}_2$.

m) $W=\mathbb{D}_n$ es el conjunto de las matrices diagonales de orden n; $\mathbb{V}=\mathbb{M}_n$.

n) W es el conjunto de todas las funciones polinómicas de $\mathbb{R}$ en $\mathbb{R}$; $\mathbb{V}=\mathbb{F}(\mathbb{R})$.

2. En el espacio $\mathbb{F}(\mathbb{R})$ considere los subconjuntos de *funciones pares* e *impares* de $\mathbb{F}(\mathbb{R})$:

$$\begin{aligned} W_p &= \{f : f(-x)=f(x)\}, \\ W_i &= \{f : f(-x)=-f(x)\}, \end{aligned}$$

respectivamente. Pruebe que W_p y W_i son subespacios de $\mathbb{F}(\mathbb{R})$.

3. Considere el espacio $C[a,b]$ de las funciones continuas en un intervalo cerrado $[a,b]$. Determine si el subconjunto W es un subespacio de $C[a,b]$ siendo:

a) W es el conjunto de las funciones f de $C[a,b]$ tal que $f(a)f(b)<0$.

$\boxed{b)}$ W es el conjunto de las funciones f de $C[a,b]$ tal que $f(a)=f(b)$.

$\boxed{c)}$ W es el conjunto formado por las funciones derivables en $[a,b]$.

d) W es el conjunto formado por las funciones integrables en $[a,b]$.

a) $W=\left\{f\in C[a,b] : \int_a^b f(x)\,dx=0\right\}$.

4. Sean $\mathbb{U}$ y $\mathbb{W}$ dos subespacios de un espacio vectorial $\mathbb{V}$. Demuestre que $\mathbb{U}\cap\mathbb{W}$ es un subespacio de $\mathbb{V}$. ¿$\mathbb{U}\cup\mathbb{W}$ es un subespacio de $\mathbb{V}$?.

4.3. Combinación lineal y espacio generado

En la Sección 2.1 se definió el concepto de *combinación lineal* de vectores de $\mathbb{R}^n$. Se estableció que una combinación lineal de los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_k$ de $\mathbb{R}^n$ es un vector de $\mathbb{R}^n$ de la forma

$$c_1\mathbf{u}_1+c_2\mathbf{u}_2+...+c_k\mathbf{u}_k,$$

donde c_1, c_2,..., c_k son escalares cualesquiera, llamados *coeficientes* de la combinación lineal. Es posible extender este concepto a un espacio vectorial cualquiera:

Definición (Combinación lineal). Sea $\mathbb{V}$ un espacio vectorial. Una combinación lineal de los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_k$ de $\mathbb{V}$ es un vector de $\mathbb{V}$ de la forma

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k,$$

donde c_1, c_2,..., c_k son escalares cualesquiera, llamados *coeficientes* de la combinación lineal.

Así, un vector $\mathbf{u}$ de $\mathbb{V}$ es una combinación lineal de ciertos vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_k$ de $\mathbb{V}$ si y solo si existen escalares c_1, c_2,..., c_k tal que

$$\mathbf{u} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k.$$

Ejemplo 4.19. En el espacio $\mathbb{R}^2$, el vector $\mathbf{u} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ es combinación lineal de los vectores $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ pero no es una combinación lineal de los vectores $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -3 \\ 0 \end{bmatrix}$.◄

Ejemplo 4.20. En el espacio $\mathbb{R}^3$, $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ es combinación lineal de los vectores $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ -3 \\ -5 \end{bmatrix}$ y $\mathbf{v}_4 = \begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix}$. En el Ejemplo 2.4 se mostró que hay una infinidad de maneras de expresar al vector $\mathbf{u}$ como combinación lineal de los vectores $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{v}_4$.◄

Ejemplo 4.21. En el espacio $\mathbb{P}_2$, sean los polinomios $p = 1 + 3x - x^2$, $q_1 = 1 + x + 2x^2$, $q_2 = 1 - x^2$ y $q_3 = 2 - x + 3x^2$. Para determinar si p es una combinación lineal de q_1, q_2 y q_3, debe analizarse si existen escalares c_1, c_2 y c_3 tales que

$$p = c_1q_1 + c_2q_2 + c_3q_3.$$

Esto es equivalente a estudiar la consistencia de la siguiente ecuación en las variables y_1, y_2 y y_3:

$$y_1q_1 + y_2q_2 + y_3q_3 = p.$$

es decir

$$y_1\left(1 + x + 2x^2\right) + y_2\left(1 - x^2\right) + y_3\left(2 - x + 3x^2\right) = 1 + 3x - x^2.$$

Operando en el primer miembro se obtiene:

$$(y_1 + y_2 + 2y_3) + (y_1 - y_3)\,x + (2y_1 - y_2 + 3y_3)\,x^2 = 1 + 3x - x^2,$$

y, por igualdad de polinomios, se llega al sistema lineal:

$$\begin{cases} y_1 & +y_2 & +2y_3 & = & 1 \\ y_1 & & -y_3 & = & 3 \\ 2y_1 & -y_2 & +3y_3 & = & -1 \end{cases}$$

El problema se reduce entonces a estudiar la consistencia de dicho sistema lineal. Aplicando eliminación gaussiana a la matriz aumentada:

$$\left[\begin{array}{rrrcr} 1 & 1 & 2 & : & 1 \\ 1 & 0 & -1 & : & 3 \\ 2 & -1 & 3 & : & -1 \end{array}\right] \xrightarrow[R_3 - 2R_1 \to R_3]{R_2 - R_1 \to R_2} \left[\begin{array}{rrrcr} 1 & 1 & 2 & : & 1 \\ 0 & -1 & -3 & : & 2 \\ 0 & -3 & -1 & : & -3 \end{array}\right] \xrightarrow[R_3 - 3R_2 \to R_3]{} \left[\begin{array}{rrrcr} \boxed{1} & 1 & 2 & : & 1 \\ 0 & \boxed{-1} & -3 & : & 2 \\ 0 & 0 & \boxed{8} & : & -9 \end{array}\right].$$

Puede observarse que el sistema es consistente. Luego el polinomio p es combinación lineal de los polinomios q_1, q_2 y q_3. Como el sistema tiene única solución, $y_1 = \frac{15}{8}$, $y_2 = \frac{11}{8}$ y $y_3 = -\frac{9}{8}$, hay una única manera de expresar al polinomio p como combinación lineal de q_1, q_2 y q_3. Los coeficientes de la combinación lineal son $c_1 = \frac{15}{8}$, $c_2 = \frac{11}{8}$ y $c_3 = -\frac{9}{8}$:

$$\frac{15}{8}\left(1 + x + 2x^2\right) + \frac{11}{8}\left(1 - x^2\right) - \frac{9}{8}\left(2 - x + 3x^2\right) = 1 + 3x - x^2. \blacktriangleleft$$

Ejemplo 4.22. En el espacio $\mathbb{F}(\mathbb{R})$, sean las funciones f, g y h definidas por $f(x) = \operatorname{sen}^2 x$, $g(x) = 3$ y $h(x) = \cos^2 x$, para todo $x \in \mathbb{R}$.

Puede comprobarse fácilmente que $g = 3f + 3h$. En efecto,

$$3f(x) + 3h(x) = 3\operatorname{sen}^2 x + 3\cos^2 x = 3\left(\operatorname{sen}^2 x + \cos^2 x\right) = 3 = g(x),$$

para todo $x \in \mathbb{R}$. Así, g es combinación lineal de f y h.$\blacktriangleleft$

La noción de combinación lineal permite definir el siguiente importante concepto:

Definición (Generado). Sea $\mathbb{V}$ un espacio vectorial y $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ un subconjunto de $\mathbb{V}$. El conjunto formado por todas las combinaciones lineales de los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_k$ es un subconjunto de $\mathbb{V}$ llamado *generado* por S, y se lo denota gen(S):

$$\operatorname{gen}(S) = \operatorname{gen}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\} = \{c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k : c_1, c_2, ..., c_k \in \mathbb{R}\}.$$

Así, un vector $\mathbf{v}$ de $\mathbb{V}$ está en gen(S) si y solo si existen escalares c_2,..., c_k tales que

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k.$$

Observación 4.1. A partir de esta definición puede notarse que:

1. El vector $\mathbf{0}$ está en gen(S). En efecto, el vector cero, $\mathbf{0}$, es combinación lineal de cualquier colección de vectores de $\mathbb{V}$ ya que una combinación lineal que tiene todos sus coeficientes cero dá como resultado el vector cero: $\mathbf{0} = 0\mathbf{u}_1 + 0\mathbf{u}_2 + ... + 0\mathbf{u}_k$.

2. $S \subset$ gen(S). En efecto, todo vector $\mathbf{u}_i$ de S puede escribirse como combinación lineal de los vectores de S de la siguiente manera:

$$\mathbf{u}_i = 0\mathbf{u}_1 + 0\mathbf{u}_2 + ... + 0\mathbf{u}_{i-1} + 1\mathbf{u}_i + 0\mathbf{u}_{i+1} + ... + 0\mathbf{u}_k.$$

Por lo tanto todo vector de S está en gen(S). La Figura 4.2 muestra un diagrama de la situación.

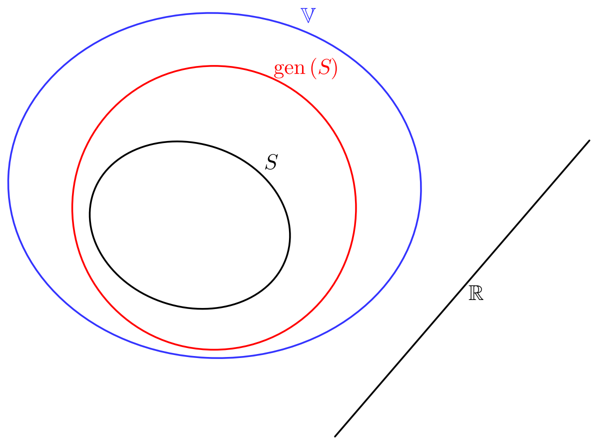


Figura 4.2: $S \subset \text{gen}(S) \subset \mathbb{V}$

Ejemplo 4.23. En el espacio $\mathbb{R}^2$, sea el conjunto unitario $S = \{(1,2)\}$. El generado por S es el conjunto

$$\text{gen}(S) = \{c(1,2) : c \in \mathbb{R}\} = \{(c, 2c) : c \in \mathbb{R}\}.$$

Nótese que éste es el conjunto de todos los múltiplos escalares del vector $\mathbf{u} = (1,2)$ que es un subespacio de $\mathbb{R}^2$ (Ejemplo 4.13). Geométricamente representa la recta en el plano que contiene al origen de coordenadas y tiene la dirección de $\mathbf{u}$ como se observa en la Figura 4.3.◀

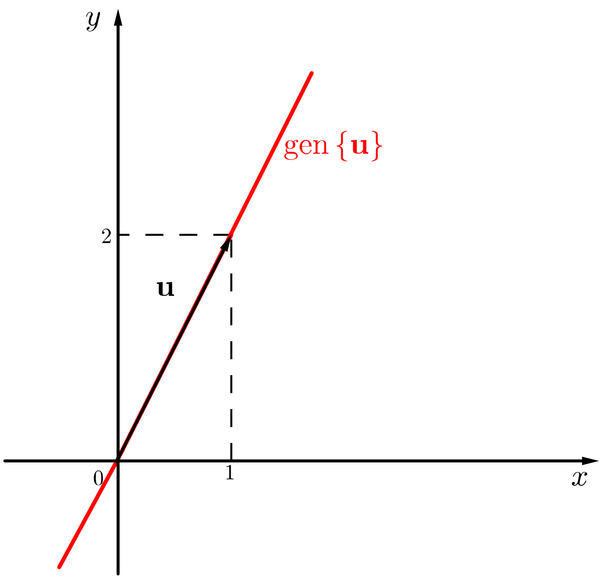


Figura 4.3: Recta generada por el vector $\mathbf{u}$

Ejemplo 4.24. En el espacio $\mathbb{P}_2$, sea el conjunto $S = \left\{1 - x^2, x\right\}$. El generado por S es el conjunto

$$\text{gen}(S) = \left\{c_1\left(1 - x^2\right) + c_2 x : c_1, c_2 \in \mathbb{R}\right\} = \left\{c_1 + c_2 x - c_1 x^2 : c_1, c_2 \in \mathbb{R}\right\},$$

es decir son los polinomios de $\mathbb{P}_2$ que tienen el coeficiente de x^2 opuesto al término constante (el coeficiente de x es cualquier número real). Por ejemplo, $2+x-2x^2$, $-5+3x+5x^2$, $-9x$ son polinomios de $\text{gen}(S)$ mientras que $3+x+3x^2$ no está en $\text{gen}(S)$.◀

Ejemplo 4.25. En el espacio $\mathbb{M}_2$, sean las matrices $A = \begin{bmatrix} 1 & 3 \\ 0 & -2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$ y $C = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$. Por definición,

$$\begin{aligned} \text{gen}\{A,B,C\} &= \left\{ c_1 \begin{bmatrix} 1 & 3 \\ 0 & -2 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} : c_1, c_2, c_3 \in \mathbb{R} \right\} = \\ &= \left\{ \begin{bmatrix} c_1 + c_2 & 3c_1 + 2c_2 + 2c_3 \\ -c_2 + c_3 & -2c_1 + c_3 \end{bmatrix} : c_1, c_2, c_3 \in \mathbb{R} \right\}. \end{aligned}$$

Entonces una matriz $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ está en $\text{gen}\{A,B,C\}$ si y solo si, existen escalares c_1, c_2 y c_3 tales que

$$\begin{bmatrix} c_1 + c_2 & 3c_1 + 2c_2 + 2c_3 \\ -c_2 + c_3 & -2c_1 + c_3 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Esto conduce a estudiar la consistencia del sistema lineal:

$$\left\{ \begin{array}{rrrcl} x_1 & +x_2 & & = & a \\ 3x_1 & +2x_2 & +2x_3 & = & b \\ & -x_2 & +x_3 & = & c \\ -2x_1 & & +x_3 & = & d \end{array} \right.$$

Aplicando eliminación gaussiana a la matriz aumentada:

$$\left[\begin{array}{rrrcc} 1 & 1 & 0 & : & a \\ 3 & 2 & 2 & : & b \\ 0 & -1 & 1 & : & c \\ -2 & 0 & 1 & : & d \end{array} \right] \to \left[\begin{array}{rrrcc} 1 & 1 & 0 & : & a \\ 0 & -1 & 2 & : & b-3a \\ 0 & -1 & 1 & : & c \\ 0 & 2 & 1 & : & d+2a \end{array} \right] \to \left[\begin{array}{rrrcl} 1 & 1 & 0 & : & a \\ 0 & -1 & 2 & : & b-3a \\ 0 & 0 & -1 & : & c-b+3a \\ 0 & 0 & 5 & : & d+2b-4a \end{array} \right] \to$$

$$\to \left[\begin{array}{rrrcl} 1 & 1 & 0 & : & a \\ 0 & -1 & 2 & : & b-3a \\ 0 & 0 & -1 & : & c-b+3a \\ 0 & 0 & 0 & : & d+5c-3b+11a \end{array} \right].$$

Este sistema es consistente si y solo si $d+5c-3b+11a=0$. Por lo tanto, el espacio generado por las matrices A, B y C está formado por todas las matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ de $\mathbb{M}_2$ cuyos elementos satisfacen la relación $d+5c-3b+11a=0$. Entonces,

$$\text{gen}\{A,B,C\} = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{M}_2 : d+5c-3b+11a=0 \right\}.$$

Al tener explicitada la forma de los elementos de $\text{gen}\{A,B,C\}$ se puede determinar rápidamente si una matriz de $\mathbb{M}_2$ está o no en dicho conjunto. Por ejemplo, la matriz $D = \begin{bmatrix} 3 & 2 \\ -4 & -7 \end{bmatrix}$ está en $\text{gen}\{A,B,C\}$ mientras que $E = \begin{bmatrix} 4 & 1 \\ 4 & -1 \end{bmatrix}$ no está en $\text{gen}\{A,B,C\}$.◀

Por definición, el conjunto $\text{gen}(S)$ es un subconjunto de $\mathbb{V}$. Si $\text{gen}(S) = \mathbb{V}$ entonces todo vector de $\mathbb{V}$ es una combinación lineal de los vectores de S. Al respecto se tiene:

Definición (Conjunto Generador). Se dice que un conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ de un espacio vectorial $\mathbb{V}$ *genera* al espacio $\mathbb{V}$ si todo vector de $\mathbb{V}$ es combinación lineal de los vectores de S. Si S genera a $\mathbb{V}$ se dice que S es un *generador* de $\mathbb{V}$, o que $\mathbb{V}$ está *generado* por S, o que los vectores $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k$ *generan* a $\mathbb{V}$.

La Figura 4.4 aclara este concepto:

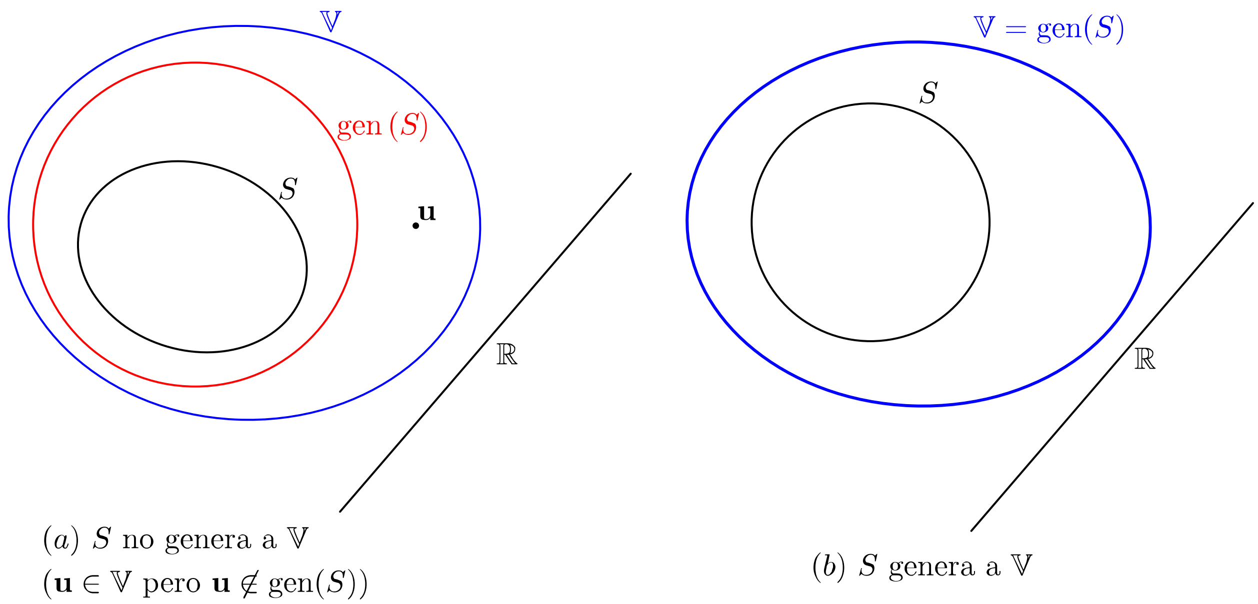


Figura 4.4: Conjunto generador

Ejemplo 4.26. Sea el conjunto $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 4 \\ 5 \end{bmatrix} \right\}$ del espacio $\mathbb{R}^3$.

Para determinar si S es un generador de $\mathbb{R}^3$, se debe analizar si todo vector $\mathbf{b}$ de $\mathbb{R}^3$ es una combinación lineal de los vectores de S. Es decir, debe analizarse si cualquiera sea el vector $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ de $\mathbb{R}^3$, la ecuación

$$x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 4 \\ 5 \end{bmatrix} = \mathbf{b},$$

tiene solución. Operando en el primer miembro y por igualdad de vectores se llega al sistema lineal

$$\left\{ \begin{array}{rrrrl} x_1 & +2x_2 & +3x_3 & -2x_4 = & b_1 \\ & x_2 & +2x_3 & +4x_4 = & b_2 \\ & & x_3 & +5x_4 = & b_3 \end{array} \right.$$

El problema se reduce a estudiar la consistencia de este sistema. Es inmediato que es consistente, y con infinitas soluciones, para todo b_1, b_2 y b_3. Luego, cualquier vector $\mathbf{b}$ de $\mathbb{R}^3$ puede escribirse como combinación lineal de los vectores de S (más aún, hay infinitas maneras de hacer ésto). Por lo tanto, S genera a $\mathbb{R}^3$.◄

Ejemplo 4.27. El conjunto $S = \left\{1 - x^2, x\right\}$ del Ejemplo 4.24 no genera a $\mathbb{P}_2$. En efecto, gen$(S) = \left\{c + bx + ax^2 \in \mathbb{P}_2 : a = -c\right\}$ que trivialmente no es todo $\mathbb{P}_2$.◄

Ejemplo 4.28. El conjunto $\{A, B, C\}$ de las matrices de $\mathbb{M}_2$ dadas en el Ejemplo 4.25, no genera a $\mathbb{M}_2$. Como se vió en el ejemplo, la matriz E de $\mathbb{M}_2$ no está en gen$\{A, B, C\}$.◄

Los siguientes son ejemplos de conjuntos generadores de los espacios vectoriales que se utilizan más a menudo:

Ejemplo 4.29. El conjunto $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ genera a $\mathbb{R}^3$.

En efecto, todo vector $\mathbf{x} = (x_1, x_2, x_3)$ se puede escribir como combinación lineal de los vectores $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ de la siguiente manera:

$$\mathbf{x} = (x_1, x_2, x_3) = x_1\,(1, 0, 0) + x_2\,(0, 1, 0) + x_3\,(0, 0, 1) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3. \blacktriangleleft$$

Generalizando el ejemplo anterior se tiene:

Ejemplo 4.30. El conjunto $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ genera a $\mathbb{R}^n$.◄

Ejemplo 4.31. El conjunto $\left\{1, x, x^2\right\}$ genera a $\mathbb{P}_2$.

En efecto, todo polinomio $p = a + bx + cx^2$ de $\mathbb{P}_2$ se puede escribir como combinación lineal de los polinomios 1, x, x^2 de la siguiente manera:

$$p = a + bx + cx^2 = a\,(1) + b\,(x) + c\left(x^2\right). \blacktriangleleft$$

Generalizando el ejemplo anterior:

Ejemplo 4.32. El conjunto $\left\{1, x, x^2, ..., x^n\right\}$ genera a $\mathbb{P}_n$.◄

Ejemplo 4.33. El conjunto $\{E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23}\}$ genera a $\mathbb{M}_{23}$, donde E_{ij} es la matriz de 2×3 que tiene un 1 en la posición (i, j) y 0 en las otras posiciones, por ejemplo: $E_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

En efecto, toda matriz $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$ se puede escribir como combinación lineal de las matrices E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23} de la siguiente manera:

$$\begin{aligned} A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} &= a_{11}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + a_{12}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + a_{13}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \\ &+ a_{21}\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} + a_{22}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + a_{23}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \\ &= a_{11}E_{11} + a_{12}E_{12} + a_{13}E_{13} + a_{21}E_{21} + a_{22}E_{22} + a_{23}E_{23}. \blacktriangleleft \end{aligned}$$

Generalizando este ejemplo:

Ejemplo 4.34. El conjunto $\{E_{11}, E_{12}, ..., E_{1n}; E_{21}, E_{22}, ..., E_{2n}; ...; E_{m1}, E_{m2}, ..., E_{mn}\}$ genera a $\mathbb{M}_{mn}$, donde E_{ij} es la matriz de $m \times n$ que tiene un 1 en la posición (i, j) y 0 en las otras posiciones.◀

Teorema 4.3 (Espacio generado)**.** Si $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ es un subconjunto de un espacio vectorial $\mathbb{V}$ entonces:

1. gen(S) es un subespacio de $\mathbb{V}$. Se lo llama *espacio generado* por S.
2. gen(S) es el subespacio más pequeño de $\mathbb{V}$ que contiene al conjunto S, en el sentido de que cualquier otro subespacio de $\mathbb{V}$ que contenga a S debe contener a gen(S).

Demostración. 1. Nótese primero que gen(S) es un subconjunto no vacío de $\mathbb{V}$ ya que $S \subset$ gen(S) y $S \neq \emptyset$. Se aprobará ahora que gen(S) es cerrado bajo la suma y la multiplicación por escalares.

Sean $\mathbf{v}$ y $\mathbf{w}$ vectores de gen(S). Luego, existen escalares $c_1, c_2, ..., c_k$ y $d_1, d_2, ..., d_k$ tal que

$$\begin{aligned}\mathbf{v} &= c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k,\\ \mathbf{w} &= d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + ... + d_k\mathbf{u}_k.\end{aligned}$$

Entonces:

$$\begin{aligned}\mathbf{v} + \mathbf{w} &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k) + (d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + ... + d_k\mathbf{u}_k) =\\ &= (c_1 + d_1)\,\mathbf{u}_1 + (c_2 + d_2)\,\mathbf{u}_2 + ... + (c_k + d_k)\,\mathbf{u}_k,\end{aligned}$$

por lo que $\mathbf{v} + \mathbf{w}$ es combinación lineal de los vectores de S y por lo tanto está en gen(S).

Por otro lado, la multiplicación de un escalar a por el vector $\mathbf{v}$ es

$$\begin{aligned}a\mathbf{v} &= a\,(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k) =\\ &= (ac_1)\,\mathbf{u}_1 + (ac_2)\,\mathbf{u}_2 + ... + (ac_k)\,\mathbf{u}_k,\end{aligned}$$

por lo que $a\mathbf{v}$ es combinación lineal de los vectores de S y luego está en gen(S).

Así, gen(S) es cerrado bajo la suma y multiplicación por escalares y por lo tanto es un subespacio de $\mathbb{V}$.

2. Se sabe que gen(S) es un subespacio de $\mathbb{V}$ que contiene a S.

Sea $\mathbb{W}$ un subespacio de $\mathbb{V}$ que contenga a los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_k$ de S. Dado que $\mathbb{W}$ es cerrado bajo la suma y multiplicación por escalares, entonces debe contener a todas las combinaciones lineales $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k$. Por lo tanto $\mathbb{W}$ contiene a todos los vectores de gen(S), es decir gen$(S) \subset \mathbb{W}$.

Luego, gen(S) es el subespacio "más pequeño" (en el sentido de inclusión) que contiene a S (véase Figura 4.5).

□

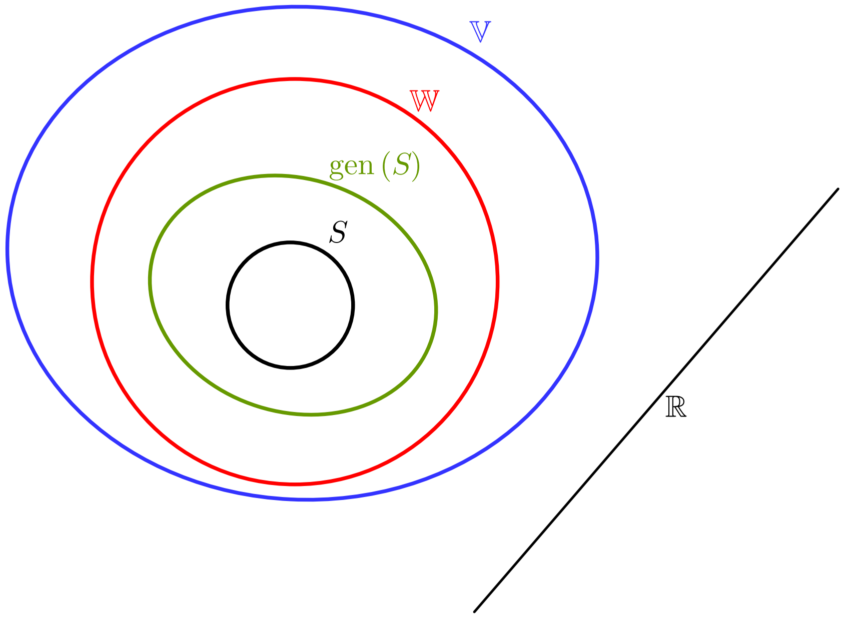


Figura 4.5: $S \subset \text{gen}(S) \subset \mathbb{W} \subset \mathbb{V}$

Con frecuencia se demuestra que un conjunto es un subespacio de un espacio vectorial $\mathbb{V}$ mostrando que es el espacio generado por algún subconjunto de $\mathbb{V}$. Los siguientes ejemplos muestran cómo utilizar ésto:

Ejemplo 4.35. Sea W el conjunto de todos los vectores de $\mathbb{R}^4$ de la forma $(a+2b, 2b, a-b, b)$ donde a y b son escalares arbritrarios, es decir

$$W = \{(a+2b, 2b, a-b, b) : a, b \in \mathbb{R}\}.$$

Por definición de suma y multiplicación por escalares, un vector genérico de W puede expresarse de la siguiente manera:

$$(a+2b, 2b, a-b, b) = (a, 0, a, 0) + (2b, 2b, -b, b) = a\,(1, 0, 1, 0) + b\,(2, 2, -1, 1).$$

Entonces todo vector de W se escribe como combinación lineal de los vectores $\mathbf{u}_1 = (1, 0, 1, 0)$ y $\mathbf{u}_2 = (2, 2, -1, 1)$. Luego,

$$W = \text{gen}\,\{\mathbf{u}_1, \mathbf{u}_2\},$$

y de acuerdo a la parte 1 del Teorema 4.3, W es un subespacio de $\mathbb{R}^4$.◄

Ejemplo 4.36. El conjunto

$$W = \left\{ \begin{bmatrix} a+b & c \\ -b & 2a \end{bmatrix} : a, b, c \in \mathbb{R} \right\}$$

es un subespacio de $\mathbb{M}_2$. En efecto, cualquier matriz de W puede expresarse como

$$\begin{bmatrix} a+b & c \\ -b & 2a \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} + b \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

para cualesquiera $a, b, c \in \mathbb{R}$. Luego,

$$W = \text{gen}\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}. \blacktriangleleft$$

Teorema 4.4 (Reducción de un conjunto generador)**.** Sea $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ un subconjunto de un espacio vectorial $\mathbb{V}$. Si un vector $\mathbf{u}_i$ de S es una combinación lineal de los restantes vectores de S, entonces $\text{gen}(S) = \text{gen}(S')$, donde S' es el conjunto que resulta de eliminar de S el vector $\mathbf{u}_i$.

Demostración. Sin perder generalidad, supóngase que $\mathbf{u}_k$ es combinación lineal de los restantes vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_{k-1}$ de S. Es decir, existen escalares c_1, c_2,..., c_{k-1} tales que

$$\mathbf{u}_k = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_{k-1}\mathbf{u}_{k-1}. \tag{4.3}$$

Se tiene que probar que $\text{gen}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{k-1}\} = \text{gen}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{k-1}, \mathbf{u}_k\}$. Para ello debe verificarse que todo vector de $\text{gen}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{k-1}\}$ es un vector de $\text{gen}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{k-1}, \mathbf{u}_k\}$ y recíprocamente.

Por un lado, cualquier vector de $\text{gen}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{k-1}\}$ es un vector de $\text{gen}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{k-1}, \mathbf{u}_k\}$ (¿por qué?).

Por otra parte, si $\mathbf{x}$ es un vector de $\text{gen}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{k-1}, \mathbf{u}_k\}$, entonces

$$\mathbf{x} = d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + ... + d_{k-1}\mathbf{u}_{k-1} + d_k\mathbf{u}_k,$$

para ciertos escalares d_1, d_2,..., d_{k-1}, d_k.

Sustituyendo en esta última igualdad, la expresión para $\mathbf{u}_k$ dada en (4.3):

$$\begin{aligned} \mathbf{x} &= d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + ... + d_{k-1}\mathbf{u}_{k-1} + d_k\left(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_{k-1}\mathbf{u}_{k-1}\right) = \\ &= \left(d_1 + d_kc_1\right)\mathbf{u}_1 + \left(d_2 + d_kc_2\right)\mathbf{u}_2 + ... + \left(d_{k-1} + d_kc_{k-1}\right)\mathbf{u}_{k-1}. \end{aligned}$$

Entonces, $\mathbf{x}$ es combinación lineal de los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_{k-1}$ y por lo tanto $\mathbf{x}$ está en $\text{gen}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{k-1}\}$. $\square$

Ejemplo 4.37. Los conjuntos $S = \{1 + x, -1 + x^2, x + x^2\}$ y $S' = \{1 + x, -1 + x^2\}$ generan el mismo subespacio de $\mathbb{P}_2$. En efecto, el último polinomio de S es la suma de los dos primeros.◄

Ejercicios 4.3

1. Indique si el vector $\mathbf{u}$ puede expresarse como una combinación lineal de los vectores $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$ y $\mathbf{u}_5$ de $\mathbb{R}^4$. En caso afirmativo, exhiba alguna de estas combinaciones.

   a) $\mathbf{u} = \begin{bmatrix} 2 \\ 9 \\ 7 \\ 3 \end{bmatrix}$; $\quad \mathbf{u}_1 = \begin{bmatrix} -1 \\ 0 \\ 4 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{u}_4 = \begin{bmatrix} 0 \\ 4 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_5 = \begin{bmatrix} 3 \\ 1 \\ -2 \\ 2 \end{bmatrix}.$

$\boxed{b)}$ $\mathbf{u} = \begin{bmatrix} -4 \\ 1 \\ 3 \\ 0 \end{bmatrix}$; $\quad \mathbf{u}_1 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ \frac{1}{2} \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_4 = \begin{bmatrix} 2 \\ 1 \\ -1 \\ 3 \end{bmatrix}, \mathbf{u}_5 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$

c) $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$; $\quad \mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_4 = \begin{bmatrix} 0 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_5 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}.$

$\boxed{2.}$ Sean los polinomios $p_1 = 1 + 3x - x^2$, $p_2 = -3x + 2x^2$, $p_3 = x^2$ y $p_4 = 4 - x^2$ de $\mathbb{P}_2$. ¿Es p_4 combinación lineal de los polinomios

a) p_1, p_2 y p_3? *b)* p_1 y p_2? *c)* p_2 y p_3?

$\boxed{3.}$ Determine valores de a, b y c de modo que la matriz

$$A = \begin{bmatrix} a+b & a+2c \\ -4 & 10 \\ 11 & 2a-b \end{bmatrix}$$

sea combinación lineal de las matrices $A_1 = \begin{bmatrix} 1 & 2 \\ -1 & 2 \\ 1 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & -1 \\ 0 & 1 \\ 2 & -1 \end{bmatrix}$ y $A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}$.

$\boxed{4.}$ Describa geométricamente el subespacio generado por los vectores $\mathbf{u}_1 = (-1, 0, 3)$, $\mathbf{u}_2 = (0, 1, -1)$ y $\mathbf{u}_3 = (-2, 3, 3)$ de $\mathbb{R}^3$.

$\boxed{5.}$ Explicite el espacio generado por las matrices $A = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 2 \\ 3 & -4 \end{bmatrix}$ de $\mathbb{M}_2$. Indique tres matrices de $\mathbb{M}_2$ que estén en gen $\{A, B\}$ y tres que no pertenezcan a dicho conjunto.

6. Sean los polinomios $p_1 = 1 - 2x + x^2$, $p_2 = x - x^2$, $p_3 = 2 - 3x + 2x^2$ y $p_4 = 2 - x - x^2$ de $\mathbb{P}_2$.

$\boxed{a)}$ ¿Está p_4 en el gen $\{p_1, p_2, p_3\}$?

b) ¿Está p_4 en el gen $\{p_1, p_2\}$?

c) Demuestre que gen $\{p_1, p_2, p_3\} = \mathbb{P}_2$.

d) Demuestre que gen $\{p_1, p_2, p_3, p_4\} = $ gen $\{p_1, p_2, p_3\}$.

$\boxed{7.}$ Indique si el conjunto:

a) $\{(1, 1, 2), (3, -1, 4), (0, 1, 3)\}$ genera al espacio $\mathbb{R}^3$.

b) $\left\{ \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \right\}$ genera a $\mathbb{R}^3$.

c) $\{1 + x, -3 + x^2\}$ genera a $\mathbb{P}_2$.

d) $\{1, 1 + x, 1 + x + x^2, 1 + x + x^2 + x^3, 3 - x^3\}$ genera a $\mathbb{P}_3$.

e) $\{E_{11}, E_{22}, E_{33}\}$ genera al espacio $\mathbb{D}_3$ de las matrices diagonales de orden 3.

f) $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\right\}$ genera a $\mathbb{M}_2$.

8. Halle, si existen, el o lo valores de a de modo que el conjunto $S = \left\{a, 1 + ax, -x + ax^2\right\}$, genere al espacio $\mathbb{P}_2$.

9. Complete la demostración del Teorema 4.4: “cualquier vector de gen $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{k-1}\}$ es un vector de gen $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{k-1}, \mathbf{u}_k\}$”

10. En cada ítem, pruebe que $\mathbb{W}$ es un subespacio del espacio vectorial $\mathbb{V}$ que se indica mostrando un conjunto generador de $\mathbb{W}$.

    a) $\mathbb{W} = \{(2a, 0, b, -a - 3b + c) : a, b, c \in \mathbb{R}\}$; $\mathbb{V} = \mathbb{R}^4$.
    b) $\mathbb{W}$ es el conjunto solución de la ecuación $x_1 - x_2 + x_3 - x_4 + 3x_5 = 0$; $\mathbb{V} = \mathbb{R}^5$.
    c) $\mathbb{W} = \left\{a - bx + (a + b)x^2 : a, b \in \mathbb{R}\right\}$; $\mathbb{V} = \mathbb{P}_2$.
    d) $\mathbb{W} = \{2a + b \cos x : a, b \in \mathbb{R}\}$; $\mathbb{V} = \mathbb{F}(\mathbb{R})$.
    e) $\mathbb{W} = \left\{\begin{bmatrix} 2a & b + c \\ -c & 0 \end{bmatrix} : a, b, c \in \mathbb{R}\right\}$; $\mathbb{V} = \mathbb{M}_2$.

11. Indique si las siguientes proposiciones son verdaderas o falsas. Justifique su respuesta:

    a) Dos vectores distintos de $\mathbb{R}^2$ generan $\mathbb{R}^2$.
    b) $\mathbb{M}_3$ puede ser generado por 10 matrices de 3×3.
    c) Tres vectores distintos de cero de $\mathbb{R}^3$ generan $\mathbb{R}^3$.
    d) gen $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} =$ gen $\{\mathbf{u}, \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} + \mathbf{w}\}$, para cualesquiera vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ de un espacio vectorial $\mathbb{V}$.

4.4. Independencia lineal

Se sabe de la sección anterior que un espacio vectorial $\mathbb{V}$ está generado por un conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ si cada vector de $\mathbb{V}$ es una combinación lineal de los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,...,$\mathbf{u}_k$.

Es posible estudiar un espacio vectorial $\mathbb{V}$ conociendo primero los vectores de un conjunto generador S y luego extendiendo los resultados al resto de los vectores de $\mathbb{V}$. Conviene entonces que el conjunto generador sea lo más pequeño posible. El problema de encontrar los conjunto generadores “más pequeños” depende de la noción de *independencia lineal*.

Si $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ es un conjunto de vectores de un espacio vectorial $\mathbb{V}$ entonces la ecuación

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + ... + x_k\mathbf{u}_k = \mathbf{0},$$

tiene siempre la solución trivial

$$x_1 = 0, x_2 = 0, ..., x_k = 0,$$

pero puede tener además otras soluciones.

Por ejemplo, en el conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ del espacio $\mathbb{R}^3$ donde $\mathbf{u}_1 = (1, 3, -1)$, $\mathbf{u}_2 = (0, 1, -2)$ y $\mathbf{u}_3 = (1, 0, 5)$, el vector $\mathbf{u}_1$ puede expresarse como una combinación lineal de los otros dos vectores de S de la siguiente manera:

$$\mathbf{u}_1 = 3\mathbf{u}_2 + \mathbf{u}_3,$$

o en forma equivalente

$$(1)\mathbf{u}_1 + (-3)\mathbf{u}_2 + (-1)\mathbf{u}_3 = \mathbf{0}.$$

Por lo tanto, la ecuación

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + x_3\mathbf{u}_3 = \mathbf{0},$$

admite como solución, entre otras, a $x_1 = 1$, $x_2 = -3$ y $x_3 = -1$.

En cambio, para el conjunto $H = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ del espacio $\mathbb{R}^3$ donde $\mathbf{v}_1 = (1,1,1)$, $\mathbf{v}_2 = (1,1,0)$ y $\mathbf{u}_3 = (1,0,0)$, si se resuelve la ecuación

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{0},$$

es decir

$$(x_1 + x_2 + x_3, x_2 + x_3, x_3) = (0,0,0)\,,$$

se encuentra una única solución: $x_1 = 0$, $x_2 = 0$ y $x_3 = 0$ (la trivial).

En general se tiene la siguiente, muy importante, definición:

Definición (**Dependencia e independencia lineal).** Sea $\mathbb{V}$ un espacio vectorial y $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ un subconjunto de $\mathbb{V}$. El conjunto S es *linealmente independiente* si la ecuación

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + ... + x_k\mathbf{u}_k = \mathbf{0} \tag{4.4}$$

tiene única solución: $x_1 = 0$, $x_2 = 0$, ..., $x_k = 0$ (la trivial).

Si por el contrario, la ecuación (4.4) admite otras (infinitas) soluciones, el conjunto S es *linealmente dependiente*.

De la definición anterior surge que:

El conjunto S es linealmente dependiente si y solo si existen escalares c_1, c_2,..., c_k, no todos ceros, tales que $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k = \mathbf{0}$.

En cambio, S es linealmente independiente si y solo si la única manera de obtener al vector cero como combinación lineal de los vectores de S, es la trivial. Es decir, si $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k = \mathbf{0}$ necesariamente $c_1 = 0$, $c_2 = 0$,..., $c_k = 0$ (todos los coeficientes cero).

Retomando los ejemplos introductorios, $S = \{(1,3,-1)\,,(0,1,-2)\,,(1,0,5)\}$ es un conjunto linealmente dependiente en $\mathbb{R}^3$, mientras que el conjunto $H = \{(1,1,1)\,,(1,1,0)\,,(1,0,0)\}$ es linealmente independiente.

Ejemplo 4.38. Sea el conjunto $S = \left\{ \begin{bmatrix} 1 \\ -2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ del espacio $\mathbb{R}^4$. Para determinar si S es linealmente dependiente o linealmente independiente, se debe analizar si la ecuación

$$x_1 \begin{bmatrix} 1 \\ -2 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

tiene única solución (la trivial) o infinitas soluciones, lo cual es equivalente a determinar si el sistema lineal homogéneo:

$$\begin{cases} x_1 & & +x_3 & = & 0 \\ -2x_1 & +x_2 & +x_3 & = & 0 \\ & -x_2 & +x_3 & = & 0 \\ x_1 & & +x_3 & = & 0 \end{cases}$$

tiene solo la solución trivial o tiene infinitas soluciones.

Es fácilmente comprobable que el sistema tiene solo la solución trivial $x_1 = x_2 = x_3 = 0$, y por lo tanto S es linealmente independiente.◀

Ejemplo 4.39. El conjunto $S = \left\{1 - x, 2 + x^2, 1 - 3x - x^2\right\}$ del espacio $\mathbb{P}_2$ es linealmente dependiente. En efecto, al plantear la ecuación en las incógnitas c_1, c_2 y c_3:

$$c_1 (1 - x) + c_2 \left(2 + x^2\right) + c_3 \left(1 - 3x - x^2\right) = 0,$$

operando en el primer miembro se obtiene

$$(c_1 + 2c_2 + c_3) + (-c_1 - 3c_3)\, x + (c_2 - c_3)\, x^2 = 0 + 0x + 0x^2.$$

Por igualdad de polinomios la ecuación es equivalente al siguiente sistema lineal homogéneo

$$\begin{cases} c_1 & +2c_2 & +c_3 & = & 0 \\ -c_1 & & -3c_3 & = & 0 \\ & c_2 & -c_3 & = & 0 \end{cases}$$

Como la matriz, A, de coeficientes es cuadrada puede determinarse si el sistema tiene única o infinitas soluciones por medio del determinante de la misma (Teorema 3.10). Rápidamente se obtiene $|A| = 0$, por lo tanto el sistema tiene soluciones no triviales lo que prueba que S es linealmente dependiente.◀

Los siguientes son ejemplos de conjuntos linealmente independientes de los espacios vectoriales que se encuentran más a menudo:

Ejemplo 4.40. El conjunto $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ de $\mathbb{R}^3$ es linealmente independiente.

En efecto, dada la ecuación

$$x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 = \mathbf{0},$$

es decir

$$x_1 (1, 0, 0) + x_2 (0, 1, 0) + x_3 (0, 0, 1) = (0, 0, 0)\,,$$

operando en el primer miembro se tiene

$$(x_1, x_2, x_3) = (0, 0, 0)\,,$$

y por igualdad de vectores, la única solución es $x_1 = x_2 = x_3 = 0$.◀

Generalizando el resultado anterior:

Ejemplo 4.41. El conjunto $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ de $\mathbb{R}^n$ es linealmente independiente.◀

Ejemplo 4.42. El conjunto $\left\{1, x, x^2\right\}$ de $\mathbb{P}_2$ es linealmente independiente.

En efecto,

$$c_1(1) + c_2(x) + c_2\left(x^2\right) = 0 = 0 + 0x + 0x^2,$$

y por igualdad de polinomios, la única solución es $c_1 = c_2 = c_3 = 0$.◀

Generalizando este resultado:

Ejemplo 4.43. El conjunto $\left\{1, x, x^2, ..., x^n\right\}$ de $\mathbb{P}_n$ es linealmente independiente.◀

Ejemplo 4.44. El conjunto $\{E_{11}, E_{12}, E_{21}, E_{22}, E_{31}, E_{32}\}$ de $\mathbb{M}_{32}$ es linealmente independiente.

En efecto, la ecuación

$$x_1 E_{11} + x_2 E_{12} + x_3 E_{21} + x_4 E_{22} + x_5 E_{31} + x_6 E_{32} = \mathbf{O},$$

queda

$$x_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} + x_6 \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

y operando en el primer miembro se obtiene

$$\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \\ x_5 & x_6 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Por igualdad de matrices, necesariamente $x_1 = x_2 = x_3 = x_4 = x_5 = x_6 = 0$.◀

Fácilmente puede generalizarse este resultado:

Ejemplo 4.45. El conjunto $\{E_{11}, E_{12}, ..., E_{1n}; E_{21}, E_{22}, ..., E_{2n}; ...; E_{m1}, E_{m2}, ..., E_{mn}\}$ de $\mathbb{M}_{mn}$ es linealmente independiente.◀

El término linealmente dependiente sugiere que los vectores "dependen" entre ellos de alguna manera. El conjunto $S = \{(1,3,-1), (0,1,-2), (1,0,5)\}$, dado en el ejemplo introductorio, es un conjunto linealmente dependiente en $\mathbb{R}^3$. Se mostró que, por ejemplo, el vector $\mathbf{u}_1 = (1,3,-1)$ es combinación lineal de los otros dos vectores de S.

Al respecto, vale el siguiente teorema:

Teorema 4.5. Un conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ $(k \geq 2)$ de un espacio vectorial $\mathbb{V}$ es linealmente dependiente si y solo si al menos uno de los vectores de S puede expresarse como combinación lineal de los restantes vectores de S.

Demostración. Si el conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{k-1}\mathbf{u}_k\}$ es linealmente dependiente, entonces existen escalares c_1, c_2,...,c_{k-1},c_k, no todos ceros, tales que

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_{k-1}\mathbf{u}_{k-1} + c_k\mathbf{u}_k = \mathbf{0}.$$

Sin pérdida de generalidad, supóngase $c_k \neq 0$, y despejando $\mathbf{u}_k$ de la expresión anterior se tiene

$$\mathbf{u}_k = \left(-\frac{c_1}{c_k}\right)\mathbf{u}_1 + \left(-\frac{c_2}{c_k}\right)\mathbf{u}_2 + ... + \left(-\frac{c_{k-1}}{c_k}\right)\mathbf{u}_{k-1},$$

por lo que $\mathbf{u}_k$ resulta combinación lineal de los restantes vectores, $\mathbf{u}_1$, $\mathbf{u}_2$,...,$\mathbf{u}_{k-1}$, de S.

Recíprocamente, supóngase que un vector de S es combinación lineal de los restantes vectores de S. Sin pérdida de generalidad, sea $\mathbf{u}_k$ combinación lineal de $\mathbf{u}_1$,$\mathbf{u}_2$,...,$\mathbf{u}_{k-1}$. Es decir, existen escalares d_1, d_2,..., d_{k-1} tal que

$$\mathbf{u}_k = d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + ... + d_{k-1}\mathbf{u}_{k-1},$$

lo que es equivalente a

$$d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + ... + d_{k-1}\mathbf{u}_{k-1} + (-1)\mathbf{u}_k = \mathbf{0}.$$

Esta última relación muestra que se puede obtener el vector cero como una combinación lineal de los vectores de S con no todos los coeficientes cero (el coeficiente que acompaña a $\mathbf{u}_k$ es -1). Luego, S es linealmente dependiente. □

Nota: Dado un conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ de un espacio vectorial $\mathbb{V}$ se puede decir indistintamente que S es linealmente dependiente (independiente) o que los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_k$ son linealmente dependientes (independientes).

Ejemplo 4.46. En el espacio $\mathbb{M}_2$, las matrices $A = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2 \\ 3 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 1 \\ 5 & 0 \end{bmatrix}$ y $D = \begin{bmatrix} 3 & 8 \\ 4 & 1 \end{bmatrix}$ son linealmente dependientes. Es fácilmente comprobable que la matriz C se puede escribir como combinación lineal de A, B y D. En efecto:

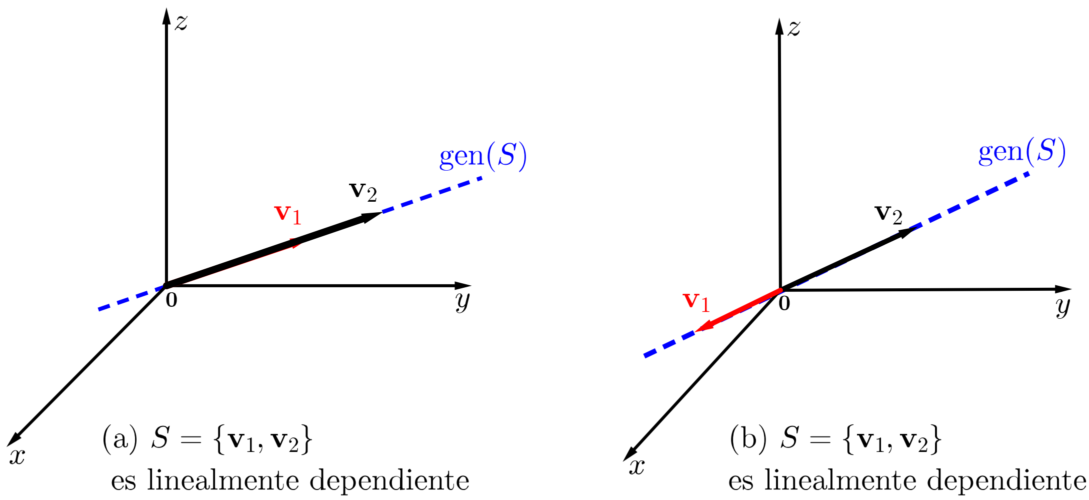
$$C = A + B = (1)A + (1)B + (0)D.$$

Observe que la matriz D no es combinación lineal de las restantes matrices, esto no contradice al teorema anterior (Teorema 4.5). Dicho teorema no dice que cada vector de un conjunto linealmente dependiente debe ser combinación lineal de los restantes sino que al menos uno de ellos debe serlo.◀

El Teorema 4.5 tiene un corolario que proporciona una prueba muy simple para determinar si dos vectores de un espacio vectorial son linealmente dependientes o no.

Corolario 4.1. Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores distintos de un espacio vectorial $\mathbb{V}$. Entonces $\mathbf{u}$ y $\mathbf{v}$ son linealmente dependientes si y solo uno de ellos es un múltiplo escalar del otro.

Demostración. De acuerdo al Teorema 4.5, el conjunto $S = \{\mathbf{u}, \mathbf{v}\}$ es linealmente dependiente si y solo si uno de los vectores de S es combinación lineal del otro. Esto es: $\mathbf{u} = c\mathbf{v}$ ó $\mathbf{v} = d\mathbf{u}$, para algunos escalares c y d. □



(a) $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente dependiente

(b) $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente dependiente

Figura 4.6: Dos vectores en $\mathbb{R}^3$ linealmente dependientes

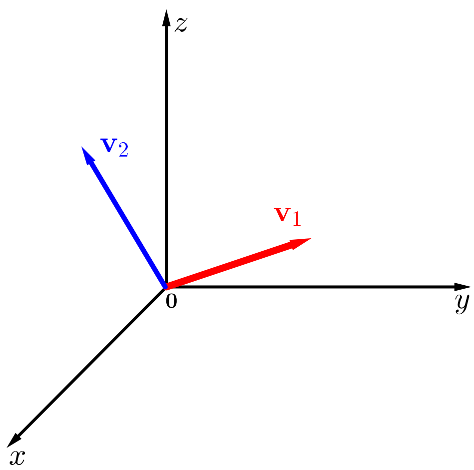


Figura 4.7: Dos vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ de $\mathbb{R}^3$ linealmente independientes

Ejemplo 4.47. Se deduce del corolario anterior que dos vectores en $\mathbb{R}^2$ o en $\mathbb{R}^3$ son linealmente dependientes si y solo si están sobre la misma recta que pasa por el origen. Observe las Figuras 4.6 y 4.7.◀

Actividad 4.5. Justifique que el espacio generado por dos vectores distintos cualesquiera en $\mathbb{R}^3$ es un plano que pasa por el origen o una recta que pasa por el origen.

Ejemplo 4.48. Los polinomios $p = 3-2x^2$ y $q = -6+4x^2$ del espacio $\mathbb{P}_2$ son linealmente dependientes dado que $q = -2p$ (ó $p = -\frac{1}{2}q$).◀

En el siguiente teorema se agregan otros importantes criterios para determinar la dependencia o independencia lineal de un subconjunto de un espacio vectorial. Queda como ejercicio la demostración del mismo:

Teorema 4.6. Sea S un subconjunto finito no vacío de un espacio vectorial $\mathbb{V}$.

1. Si S es un conjunto unitario, es decir $S = \{\mathbf{u}\}$, entonces S es linealmente independiente si y solo si $\mathbf{u} \neq \mathbf{0}$.
2. Si S contiene al vector cero, entonces S es linealmente dependiente.
3. Si S es linealmente independiente entonces todo subconjnunto no vacío de S es linealmente independiente.
4. Si S es linealmente dependiente entonces todo conjunto que lo contiene es linealmente dependiente.

Ejemplo 4.49. De acuerdo a la parte 4. del Teorema 4.6, el conjunto

$$\left\{ \begin{bmatrix} 3 & 1 \\ 0 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 5 \\ -9 & 8 \end{bmatrix}, \begin{bmatrix} 2 & 4 \\ 0 & 0 \end{bmatrix} \right\}$$

del espacio $\mathbb{M}_2$ es linealmente dependiente ya que contiene al conjunto $S = \left\{ \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 4 \\ 0 & 0 \end{bmatrix} \right\}$ que es linealmente dependiente puesto que una de las matrices es mútiplo escalar de la otra. ◀

Ejercicios 4.4

$\boxed{1.}$ Para cada conjunto S del espacio vectorial $\mathbb{V}$ indicado, determine si es linealmente independiente o linealmente dependiente:

a) $S = \{(1,-1),(2,0),(2,1)\}$, $\mathbb{V} = \mathbb{R}^2$.

b) $S = \{(1,-1,3),(2,0,-5),(0,1,1)\}$, $\mathbb{V} = \mathbb{R}^3$.

c) $S = \{(0,0,-1,2),(2,2,-5,1),(4,4,-21,5)\}$, $\mathbb{V} = \mathbb{R}^4$.

d) $S = \left\{1+x+x^2, 3x^2, -5+x\right\}$, $\mathbb{V} = \mathbb{P}_2$.

e) $S = \left\{x^3, 1-3x^2, 2+x\right\}$, $\mathbb{V} = \mathbb{P}$.

f) $S = \left\{ \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} -2 & -2 \\ 2 & 1 \end{bmatrix} \right\}$, $\mathbb{V} = \mathbb{M}_2$.

g) $S = \left\{ \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \right\}$, $\mathbb{V} = \mathbb{M}_{23}$.

h) $S = \{e^x, e^{-x}, \cosh x\}$, $\mathbb{V} = \mathbb{F}(\mathbb{R})$.

i) $S = \left\{e^{2x}, e^{-3x}\right\}$, $\mathbb{V} = \mathbb{F}(\mathbb{R})$.

$\boxed{2.}$ Determine para qué valores de a y b el subconjunto $\left\{1 + ax + ax^2, 1 + bx + bx^2, x^2\right\}$ del espacio $\mathbb{P}_2$ es linealmente independiente.

$\boxed{3.}$ Indique si las siguientes proposiciones son verdaderas o falsas. Justifique su respuesta:

a) Dos vectores distintos de un espacio vectorial son linealmente independientes.

b) Si $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto linealmente independiente en un espacio vectorial $\mathbb{V}$, entonces también lo es el conjunto $\{\mathbf{u}_1 + 2\mathbf{u}_2, \mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3, -\mathbf{u}_3\}$.

c) Si tanto $\{\mathbf{u}, \mathbf{v}\}$ como $\{\mathbf{v}, \mathbf{w}\}$ son conjuntos linealmente independientes de un espacio vectorial $\mathbb{V}$, entonces $\{\mathbf{u}, \mathbf{w}\}$ es también linealmente independiente.

$\boxed{4.}$ Demuestre el Teorema 4.6.

$\boxed{5.}$ En cada ítem, dé un ejemplo de un subconjunto S de $\mathbb{R}^4$ tal que

a) contenga exactamente dos elementos y sea linealmente dependiente.

b) contenga exactamente dos elementos y sea linealmente independiente.

c) contenga exactamente tres elementos y sea linealmente independiente.

d) contenga exactamente cuatro elementos y sea linealmente dependiente.

e) contenga exactamente cuatro elementos y sea linealmente independiente.

6. Repita el ejercicio anterior para subconjuntos S de $\mathbb{P}_4$.

4.5. Base y dimensión

Por lo común se concibe a una recta como un espacio unidimensional, a un plano como uno bidimensional y al espacio que lo rodea como a uno tridimensional.

En lo que sigue se precisará esta noción intuitiva de dimensión:

Definición (Base). Sea $\mathbb{V}$ un espacio vectorial distinto del espacio cero y S un subconjunto finito no vacío de $\mathbb{V}$. El conjunto S es una *base* de $\mathbb{V}$ si

1. S es linealmente independiente,
2. S genera a $\mathbb{V}$.

Ejemplo 4.50. En el Ejemplo 4.40 se demostró que $S = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ es un conjunto linealmente independiente en $\mathbb{R}^3$. Además, en el Ejemplo 4.29 se probó que S es un generador de $\mathbb{R}^3$. Por lo tanto, S es una base. Esta base se conoce como la **base estándar** de $\mathbb{R}^3$.◄

Ejemplo 4.51. Base estándar de $\mathbb{R}^n$

El conjunto $S = \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ es una base de $\mathbb{R}^n$. En los ejemplos 4.30 y 4.41 se mostró que S es un conjunto generador y linealmente independiente de $\mathbb{R}^n$. Esta base es conocida como la *base estándar*, o *canónica*, de $\mathbb{R}^n$.

Como caso especial, en el espacio vectorial $\mathbb{R}$, el conjunto unitario $S = \{1\}$ es trivialmente linealmente independiente y como para todo $x \in \mathbb{R}$, $x = x(1)$, entonces S es un conjunto generador de $\mathbb{R}$. Luego, S es una base de $\mathbb{R}$, la *base estándar o canónica* de $\mathbb{R}$.◀

Ejemplo 4.52. El conjunto $S = \{1, x, x^2\}$ es una base del espacio $\mathbb{P}_2$. Se mostró en los ejemplos 4.31 y 4.42 que S es un conjunto generador linealmente independiente en $\mathbb{P}_2$. Esta base es la **base estándar** del espacio vectorial $\mathbb{P}_2$.◀

Ejemplo 4.53. Base estándar de $\mathbb{P}_n$

El conjunto $S = \{1, x, x^2, ..., x^n\}$ es una base de $\mathbb{P}_n$. Se mostró que S es un conjunto generador y linealmente independiente en $\mathbb{P}_n$ en los ejemplos 4.43 y 4.31. Esta base es la llamada *base estándar* o *canónica* de $\mathbb{P}_n$.◀

Ejemplo 4.54. Base estándar de $\mathbb{M}_{mn}$

El conjunto $S = \{E_{11}, E_{12}, ..., E_{1n}; E_{21}, E_{22}, ..., E_{2n}; ...; E_{m1}, E_{m2}, ..., E_{mn}\}$ es una base del espacio $\mathbb{M}_{mn}$. Se vió en los ejemplos 4.34 y 4.45 que es un conjunto generador linealmente independiente de $\mathbb{M}_{mn}$. Esta base es la llamada *base estándar* o *canónica* del espacio $\mathbb{M}_{mn}$.◀

Ejemplo 4.55 (Una base de $\mathbb{R}^3$ que no es la estándar)**.** El conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ donde $\mathbf{u}_1 = (1, 2, 1)$, $\mathbf{u}_2 = (-2, 3, 0)$ y $\mathbf{u}_3 = (3, 3, 2)$ es una base de $\mathbb{R}^3$.

Debe comprobarse que S genera a $\mathbb{R}^3$ y que es linealmente independiente.

Para demostrar que *S genera a $\mathbb{R}^3$*, debe probarse que todo vector $\mathbf{b}$ de $\mathbb{R}^3$ es combinación lineal de los vectores de S. Es decir, debe mostrarse que cualquiera sea el vector $\mathbf{b} = (b_1, b_2, b_3)$ de $\mathbb{R}^3$, la ecuación

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + x_3\mathbf{u}_3 = \mathbf{b},$$

tiene solución.

Reemplazando, la ecuación anterior es

$$x_1 (1, 2, 1) + x_2 (-2, 3, 0) + x_3 (3, 3, 2) = (b_1, b_2, b_3),$$

y operando en el primer miembro e igualando las correspondientes componentes se obtiene el sistema lineal

$$\left\{\begin{array}{cccc} x_1 & -2x_2 & +3x_3 & = & b_1 \\ 2x_1 & +3x_2 & +3x_3 & = & b_2 \\ x_1 & & +2x_3 & = & b_3 \end{array}\right.$$

el cual debe ser consistente cualesquiera sean b_1, b_2 y b_3.

Para demostrar que S es *linealmente independiente* es necesario probar que la única solución de la ecuación

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + x_3\mathbf{u}_3 = \mathbf{0},$$

o bien

$$x_1\,(1,2,1) + x_2\,(-2,3,0) + x_3\,(3,3,2) = (0,0,0)\,,$$

es la trivial: $x_1 = x_2 = x_3 = 0$. La verificación de la independencia lineal se reduce a demostrar que el sistema homogéneo:

$$\left\{\begin{array}{rrrcl} x_1 & -2x_2 & +3x_3 & = & 0 \\ 2x_1 & +3x_2 & +3x_3 & = & 0 \\ x_1 & & +2x_3 & = & 0 \end{array}\right.$$

tiene únicamente la solución trivial.

Obsérvese que ambos sistemas tienen la misma matriz de coeficientes:

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 3 & 3 \\ 1 & 0 & 2 \end{bmatrix}.$$

Es posible probar simultáneamente que S genera a $\mathbb{R}^3$ y que es linealmente independiente mostrando que $\det(A) \neq 0$. De acuerdo al Teorema 3.10 y como $\det(A) = -1 \neq 0$, entonces el sistema $A\mathbf{x} = \mathbf{b}$ es consistente con única solución cualquiera sea el vector de términos constantes $\mathbf{b}$, y el sistema homogéneo $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial. ◄

Dimensión de un espacio vectorial

Sea $\mathbb{V}$ un espacio vectorial distinto del espacio cero. Si $\mathbb{V}$ contiene un conjunto finito no vacío de vectores que es una base de $\mathbb{V}$, se dice que $\mathbb{V}$ es un espacio vectorial **de dimensión finita** o **finito dimensional**. Si no existe un conjunto de este tipo se dice que $\mathbb{V}$ es **de dimensión infinita** o **infinito dimensional**.

Casi todos los espacios vectoriales considerados en este texto son de dimensión finita. Por ejemplo, los espacios $\mathbb{R}^n$, $\mathbb{P}_n$ y $\mathbb{M}_{mn}$ son espacios vectoriales de dimensión finita.

Sin embargo, es conveniente notar que muchos espacios vectoriales de importancia son de dimensión infinita. El espacio vectorial $\mathbb{P}$ es infinito dimensional ya que no es posible encontrar un conjunto finito que lo genere. Un conjunto generador de $\mathbb{P}$ es el conjunto infinito de la potencias de x, $S = \{x^n : n \in \mathbb{N}_0\} = \left\{1, x, x^2, x^3, ...\right\}$ ya que todo polinomio en $\mathbb{P}$ es una combinación lineal de polinomios de S.

Actividad 4.6. Explique por qué no existe un conjunto generador de $\mathbb{P}$ que contenga un número finito de elementos.

El siguiente teorema proporciona la clave hacia el concepto de **dimensión** de un espacio vectorial. Basándose en este teorema, se obtienen algunos de los resultados más importantes del Álgebra Lineal.

Teorema 4.7. Si un espacio vectorial $\mathbb{V}$ tiene una base que consta de n vectores entonces todo subconjunto de $\mathbb{V}$ que contiene más de n vectores es linealmente dependiente. Dicho de otro modo, todo subconjunto linealmente independiente de $\mathbb{V}$ contiene a lo sumo n vectores.

Demostración. Sea $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ una base del espacio vectorial $\mathbb{V}$ y $H = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m\}$ un subconjunto de $\mathbb{V}$ que consta de m vectores, donde $m > n$.

Se probará que H es linealmente dependiente mostrando que la ecuación

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + ... + x_m\mathbf{v}_m = \mathbf{0}, \tag{4.5}$$

tiene soluciones no triviales.

Como S es una base de $\mathbb{V}$ entonces S genera a $\mathbb{V}$. Así, cada vector de $\mathbb{V}$ puede expresarse como una combinación lineal de los vectores de S. En particular, cada $\mathbf{v}_i$ de H es una combinación lineal de vectores de S, por lo tanto:

$$\begin{array}{rcl} \mathbf{v}_1 & = & c_{11}\mathbf{u}_1 + c_{21}\mathbf{u}_2 + ... + c_{n1}\mathbf{u}_n \\ \mathbf{v}_2 & = & c_{12}\mathbf{u}_1 + c_{22}\mathbf{u}_2 + ... + c_{n2}\mathbf{u}_n \\ \vdots & & \vdots \\ \mathbf{v}_m & = & c_{1m}\mathbf{u}_1 + c_{2m}\mathbf{u}_2 + ... + c_{nm}\mathbf{u}_n \end{array}$$

Sustituyendo cada una de estas expresiones en la ecuación (4.5):

$$\begin{aligned} x_1\left(c_{11}\mathbf{u}_1 + c_{21}\mathbf{u}_2 + ... + c_{n1}\mathbf{u}_n\right) + x_2\left(c_{12}\mathbf{u}_1 + c_{22}\mathbf{u}_2 + ... + c_{n2}\mathbf{u}_n\right) + ... \\ ... + x_m\left(c_{1m}\mathbf{u}_1 + c_{2m}\mathbf{u}_2 + ... + c_{nm}\mathbf{u}_n\right) = \mathbf{0}, \end{aligned}$$

y operando y asociando en el primer miembro resulta

$$\begin{aligned} \left(c_{11}x_1 + c_{12}x_2 + ... + c_{1m}x_m\right)\mathbf{u}_1 + \left(c_{21}x_1 + c_{22}x_2 + ... + c_{2m}x_m\right)\mathbf{u}_2 + ... \\ ... + \left(c_{n1}x_1 + c_{n2}x_2 + ... + c_{nm}x_m\right)\mathbf{u}_n = \mathbf{0}. \end{aligned}$$

Dado que los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,...,$\mathbf{u}_n$ son linealmente independientes (S una base), entonces todos los coeficientes de la expresión anterior son iguales a cero. Así, resolver la ecuación (4.5) equivale a resolver el siguiente sistema lineal homogéneo en las incógnitas x_1, x_2,..., x_m:

$$\left\{ \begin{array}{c} c_{11}x_1 + c_{12}x_2 + ... + c_{1m}x_m = 0 \\ c_{21}x_1 + c_{22}x_2 + ... + c_{2m}x_m = 0 \\ \vdots \\ c_{n1}x_1 + c_{n2}x_2 + ... + c_{nm}x_m = 0 \end{array} \right.$$

Como este sistema es homogéneo y el número de incógnitas (m) es mayor que el número de ecuaciones (n), entonces tiene soluciones no triviales. Esto es, existen escalares d_1, d_2,...,d_m no todos ceros que satisfacen la ecuación (4.5).

Luego, el conjunto H es linealmente dependiente. $\square$

Ejemplo 4.56. Dado que $\mathbb{R}^3$ tiene una base que consta de 3 vectores (la base estándar), cualquier conjunto de $\mathbb{R}^3$ que consta de 4 ó más vectores es linealmente dependiente.

En general, en el espacio $\mathbb{R}^n$, todo conjunto que consta de más de n vectores es linealmente dependiente.◀

Ejemplo 4.57. En el espacio $\mathbb{P}_2$ todo conjunto que consta de 4 o más polinomios es linealmente dependiente puesto que $\mathbb{P}_2$ tiene una base que consta de 3 polinomios (la base estándar $\left\{1, x, x^2\right\}$).

En general en el espacio $\mathbb{P}_n$, todo conjunto que consta de más de $n+1$ polinomios es linealmente dependiente.◀

Ejemplo 4.58. En el espacio $\mathbb{M}_2$, el conjunto formado por las matrices $\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix},$ $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 4 & -1 \\ -2 & 1 \end{bmatrix}$ es linealmente dependiente. ¿Por qué?.◀

Debe observarse que si $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ es una base de un espacio vectorial $\mathbb{V}$, entonces $\{c\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ donde $c \neq 0$ es también una base de $\mathbb{V}$ (Ejercicio 5). Esto muestra que todo espacio vectorial real, distinto del espacio cero, tiene infinitas bases. A partir del teorema anterior, puede demostrarse que todas las (infinitas) bases de un espacio vectorial finito dimensional, tienen el mismo número de vectores.

Corolario 4.2. Si un espacio vectorial $\mathbb{V}$ tiene una base que consta de n vectores entonces toda otra base de $\mathbb{V}$ consta también de n vectores.

Demostración. Sea $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ una base del espacio vectorial $\mathbb{V}$ y $H = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m\}$ cualquier otra base de $\mathbb{V}$.

Por el Teorema 4.7, dado que S es una base de $\mathbb{V}$ y H es linealmente independiente (H es una base), entonces $m \leq n$. Haciendo el mismo razonamiento, y cambiando los roles, puesto que H es una base de $\mathbb{V}$ y S es linealmente independiente, entonces debe ser $n \leq m$.

En consecuencia, $m = n$ y por lo tanto ambas bases tienen el mismo número de vectores. □

A partir de este corolario, puede notarse que si se conoce una base de un espacio vectorial $\mathbb{V}$ de dimensión finita, una condición necesaria para que un conjunto S sea una base de $\mathbb{V}$ es que S tenga el mismo número de elementos que la base conocida.

Actividad 4.7. Indique, por simple inspección, por qué los siguientes conjuntos no son una base de los espacios vectoriales indicados:

1. $\{(1, -1, 3), (2, 0, 5)\}$ en $\mathbb{R}^3$.
2. $\left\{1 + x, 5 - x + x^2, 3x^2, x + x^2\right\}$ en $\mathbb{P}_2$.

El corolario anterior permite definir el concepto de **dimensión** de un espacio vectorial finito dimensional:

Definición (Dimensión). Si un espacio vectorial $\mathbb{V}$ tiene una base que consta de n vectores, a dicho número n se lo denomina *dimensión* de $\mathbb{V}$ y se denota $\dim(\mathbb{V}) = n$.

El espacio vectorial cero, $\mathbb{V} = \{\mathbf{0}\}$, está generado por el vector $\mathbf{0}$ pero $\{\mathbf{0}\}$ es un conjunto linealmente dependiente, y por lo tanto no puede ser una base. Ya que el espacio cero no contiene un conjunto, con al menos un elemento, que sea una base de él, se conviene en considerar que este espacio tiene dimensión 0 y que es finito dimensional.

Al contar el número de elementos de las bases estándar de los espacios $\mathbb{R}^n$, $\mathbb{P}_n$ y $\mathbb{M}_{mn}$, se obtiene la siguiente conclusión:

- $\dim(\mathbb{R}^n) = n$
- $\dim(\mathbb{P}_n) = n+1$
- $\dim(\mathbb{M}_{mn}) = mn$

El espacio vectorial $\mathbb{P}$ es infinito dimensional (no existe un conjunto finito de polinomios que lo genere). Similarmente, el espacio $\mathbb{F}(\mathbb{R})$ es de dimensión infinita ya que contiene un subespacio, el espacio de todas las funciones polinómicas, que es infinito dimensional (el Teorema 4.11 justifica esta afirmación).

Dado que un subespacio de un espacio vectorial es en sí mismo un espacio vectorial, tiene sentido hablar de la dimensión de un subespacio:

Ejemplo 4.59. Se vió en el Ejemplo 4.35 que el conjunto

$$W = \{(a+2b, 2b, a-b, -b) : a, b \in \mathbb{R}\}$$

es un subespacio de $\mathbb{R}^4$ mostrando que $W = \text{gen}\{\mathbf{u}_1, \mathbf{u}_2\}$ donde $\mathbf{u}_1 = (1,0,1,0)$ y $\mathbf{u}_2 = (2,2,-1,1)$.

El conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2\}$ es entonces un generador de W. S es linealmente independiente ya que consta de dos vectores tales que ninguno de ellos es múltiplo escalar del otro. Así, S es una base de W y $\dim(W) = 2$.◀

Ejemplo 4.60. En el Ejemplo 4.36 se probó que el conjunto

$$W = \left\{ \begin{bmatrix} a+b & c \\ -b & 2a \end{bmatrix} : a, b, c \in \mathbb{R} \right\}$$

es un subespacio de $\mathbb{M}_2$ mostrando que $W = \text{gen}\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}$.

Así, el conjunto $S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}$ es un generador de W. Puede probarse que S es linealmente independiente verificando que

$$a \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} + b \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

solo si $a = b = c = 0$. Luego, S es una base de W y $\dim(W) = 3$.◀

Ejemplo 4.61. Si S es un conjunto linealmente independiente en un espacio $\mathbb{V}$, trivialmente S es una base para el espacio $\text{gen}(S)$ (S es linealmente independiente y, por definición, genera a $\text{gen}(S)$).

Por ejemplo, el conjunto $S = \{(1,2,-1), (0,-1,3)\}$ de $\mathbb{R}^3$ es una base de $\text{gen}(S)$ puesto que S es linealmente independiente (¿por qué?). Geométricamente $\text{gen}(S)$ es un plano que pasa por el origen.◀

Ejemplo 4.62. Sea el subconjunto $S = \left\{1 + x^5, 2 + x - 4x^2\right\}$ del espacio $\mathbb{P}$. Entonces el generado por S es el siguiente conjunto:

$$\text{gen}(S) = \left\{a\left(1 + x^5\right) + b\left(2 + x - 4x^2\right) : a, b \in \mathbb{R}\right\} = \left\{(a + 2b) + bx - 4bx^2 + ax^5 : a, b \in \mathbb{R}\right\}.$$

El conjunto S es linealmente independiente y trivialmente genera a gen(S), por lo que S es una base de gen(S). Así, $\dim(\text{gen}(S)) = 2.$◄

El teorema siguiente muestra que la dimensión, n, de un espacio vectorial $\mathbb{V}$ finito dimensional es el *número máximo de vectores que puede tener un conjunto linealmente independiente* en $\mathbb{V}$ y el *número mínimo de vectores que debe tener un conjunto generador* de $\mathbb{V}$.

Teorema 4.8. Sea $\mathbb{V}$ un espacio vectorial no cero de dimensión n y S un subconjunto no vacío de $\mathbb{V}$ con m elementos.

1. Si S es linealmente independiente entonces $m \leq n$.
2. Si S genera a $\mathbb{V}$ entonces $m \geq n$.

Demostración.

1. Dado que la dimensión de $\mathbb{V}$ es n entonces $\mathbb{V}$ tiene una base que consta de n vectores. De acuerdo con el Teorema 4.7 cualquier conjunto linealmente independiente en $\mathbb{V}$ tiene a lo sumo n vectores. Así, $m \leq n$.
2. Si S es linealmente independiente y como, por hipótesis, S genera a $\mathbb{V}$, entonces S es una base de $\mathbb{V}$ y por lo tanto S consta también de n vectores. Así, $m = n$.

   Si en cambio S es linealmente dependiente entonces uno de los vectores de S es combinación lineal de los restantes, si se elimina este vector se obtiene un conjunto S_1 que consta de $m - 1$ vectores y que, de acuerdo al Teorema 4.4, también genera a $\mathbb{V}$. Si S_1 es linealmente independiente entonces S_1 es una base de $\mathbb{V}$ y dado que $\dim(\mathbb{V}) = n$ resulta $m - 1 = n$ y por lo tanto $m > n$. Si S_1 es linealmente dependiente, se repite este procedimiento. La eliminación reiterada de vectores de S se termina cuando se obtiene un subconjunto de S, S_k, que consta de $m - k$ elementos y que es linealmente independiente y generador de $\mathbb{V}$. Este conjunto S_k es entonces una base de $\mathbb{V}$ y por lo tanto debe ser $m - k = n$. Luego, $m > n$.

   Así, si S genera a $\mathbb{V}$ resulta $m \geq n$.

□

Actividad 4.8. Sea S un conjunto que consta de 10 vectores de $\mathbb{R}^n$. ¿Qué puede asegurarse respecto al número n en cada uno de los siguientes casos?

1. S es linealmente independiente.
2. S es generador de $\mathbb{R}^n$.
3. S es una base de $\mathbb{R}^n$.

En general, para demostrar que un conjunto de vectores $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ de un espacio vectorial $\mathbb{V}$ es una base de $\mathbb{V}$, se tiene que demostrar que los vectores son linealmente independientes y que generan a $\mathbb{V}$. Sin embargo, si se conoce de antemano que $\mathbb{V}$ tiene dimensión n, de modo que S tiene el número correcto de vectores para ser una base, el teorema siguiente (Teorema 4.9) asegura que basta comprobar solo uno de los dos requisitos. El siguiente lema es clave para la demostración de dicho teorema:

Lema 4.1. (Ampliación de un conjunto linealmente independiente) Si S es un subconjunto linealmente independiente de un espacio vectorial $\mathbb{V}$ y $\mathbf{w}$ es un vector de $\mathbb{V}$ que no pertenece al espacio generado por S, entonces el conjunto que se obtiene agregando el vector $\mathbf{w}$ a S es también linealmente independiente.

Demostración. Sea $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ un subconjunto linealmente independiente de $\mathbb{V}$ y $\mathbf{w}$ un vector de $\mathbb{V}$ que no pertenece al gen(S).

Para demostrar que $S' = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k, \mathbf{w}\}$ es linealmente independiente debe probarse que la única combinación lineal de vectores de S' que da como resultado el vector $\mathbf{0}$ es la trivial. Sean entonces c_1, c_2,..., c_k, c, escalares tales que

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k + c\mathbf{w} = \mathbf{0}. \tag{4.6}$$

Necesariamente $c = 0$ ya que de otro modo:

$$\mathbf{w} = \left(-\frac{c_1}{c}\right)\mathbf{u}_1 + \left(-\frac{c_2}{c}\right)\mathbf{u}_2 + ... + \left(-\frac{c_k}{c}\right)\mathbf{u}_k,$$

y $\mathbf{w}$ estaría en gen(S) contradiciendo la hipótesis.

Dado que $c = 0$ entonces $c\mathbf{w} = \mathbf{0}$ por lo cual la igualdad (4.6) se convierte en:

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k = \mathbf{0},$$

y como por hipótesis S es linealmente independiente, necesariamente $c_1 = c_2 = ... = c_k = 0$.

Así, la igualdad dada en (4.6) se verifica solo si $c_1 = c_2 = ... = c_k = c = 0$, que prueba que S' es linealmente independiente. □

Teorema 4.9. Sea $\mathbb{V}$ un espacio vectorial no cero de dimensión n y sea S un subconjunto de $\mathbb{V}$ con n vectores.

1. Si S es linealmente independiente entonces S es una base de $\mathbb{V}$.
2. Si S genera a $\mathbb{V}$ entonces S es una base de $\mathbb{V}$.

Demostración.

1. Sea $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ un subconjunto linealmente independiente de $\mathbb{V}$. Si S no genera a $\mathbb{V}$ entonces hay un vector $\mathbf{w}$ en $\mathbb{V}$ que no pertenece al gen(S). Por el lema anterior, el conjunto $S_1 = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n, \mathbf{w}\}$ es linealmente independiente. Dado que S_1 tiene $n+1$ elementos esto contradice la parte 1 del Teorema 4.8 (todo subconjunto linealmente independiente consta a lo sumo de n vectores). Por lo tanto, S genera a $\mathbb{V}$ y por ser linealmente independiente, S es una base de $\mathbb{V}$.

2. Sea $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ un conjunto generador de $\mathbb{V}$. Si S no es linealmente independiente entonces uno de los vectores de S es combinación lineal de los restantes. Si se elimina este vector del conjunto S, se tiene un subconjunto con $n - 1$ elementos y que, de acuerdo al Teorema 4.4, sigue generando a $\mathbb{V}$. Esto contradice la parte 2 del Teorema 4.8 (todo subconjunto generador de $\mathbb{V}$ consta al menos de n vectores). Luego, S es linealmente independiente y como genera a $\mathbb{V}$, entonces S es una base de $\mathbb{V}$.

□

Ejemplo 4.63. El conjunto $S = \{(1,1),(2,-4)\}$ de $\mathbb{R}^2$ es linealmente independiente (¿por qué?) y tiene 2 elementos. Como dim $\left(\mathbb{R}^2\right) = 2$ entonces S es una base de $\mathbb{R}^2$.◄

El teorema siguiente puede ser muy útil en la práctica. En él se afirma que en un espacio vectorial finito dimensional, procediendo adecuadamente, se puede obtener una base agregando vectores a un conjunto linealmente independiente o quitando vectores de un conjunto generador.

Teorema 4.10. Sea $\mathbb{V}$ un espacio vectorial no cero de dimensión n y S un subconjunto no vacío de $\mathbb{V}$ que consta de m elementos.

1. Si S es linealmente independiente y $m < n$, entonces S puede ampliarse a una base de $\mathbb{V}$.
2. Si S genera a $\mathbb{V}$ y $m > n$, entonces S contiene una base de $\mathbb{V}$.

Demostración. 1. Sea $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m\}$ un subconjunto linealmente independiente de $\mathbb{V}$ con $m < n$ elementos. S no puede generar a $\mathbb{V}$ puesto que se requieren al menos n vectores para generarlo (Teorema 4.8, parte 2). Por lo tanto hay un vector $\mathbf{u}$ en $\mathbb{V}$ que no pertenece al gen(S). El conjunto $S_1 = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m, \mathbf{u}\}$ consta de $m+1$ elementos y, de acuerdo al Lema 4.1, es linealmente independiente. Si $m+1 = n$ entonces, por Teorema 4.9 parte 1, S_1 es una base de $\mathbb{V}$. Si $m+1 < n$ entonces S_1 no genera a $\mathbb{V}$ y se repite este procedimiento con S_1 en lugar de S hasta llegar a un conjunto S_k que es linealmente independiente y consta de n vectores. Por el Teorema 4.9 parte 1, S_k es una base de $\mathbb{V}$ y contiene a S. Por lo tanto, S puede ampliarse a una base de $\mathbb{V}$.

2. Sea S un conjunto generador de $\mathbb{V}$ que consta de $m > n$ vectores. S es linealmente dependiente ya que n es el número máximo de vectores que puede tener un subconjunto linealmente independiente de $\mathbb{V}$ (Teorema 4.8, parte 1). Dado que $n \geq 1$ entonces S tiene al menos dos vectores distintos y, por ser linealmente dependiente, uno de los vectores de S es combinación lineal de los restantes. Si se elimina este vector, se obtiene un subconjunto S_1 de S que contiene $m-1$ elementos y que, de acuerdo al Teorema 4.4, sigue generando a $\mathbb{V}$. Si $m-1 = n$ entonces, por Teorema 4.9 parte 2, S_1 es una base de $\mathbb{V}$. Si $m-1 > n$ entonces S_1 es linealmente dependiente y se repite el procedimiento con S_1 en lugar de S hasta llegar a un subconjunto de S, S_k, que es generador de $\mathbb{V}$ y que consta de n elementos. Por el Teorema 4.9 parte 2, S_k es una base de $\mathbb{V}$. Por lo tanto todo conjunto generador de $\mathbb{V}$ contiene una base de $\mathbb{V}$.

□

Dado un conjunto finito, S, en un espacio vectorial $\mathbb{V}$ la parte 2 del teorema anterior provee un procedimiento para encontrar un subconjunto de S que es una base para el espacio gen(S). El ejemplo siguiente es ilustrativo de esta construcción. El método puede resultar tedioso puesto que cada vez que se elimina un vector se debe analizar la dependencia lineal de los vectores restantes. En la Sección 4.7 se proporcionará un método que permite encontrar de una manera más directa un subconjunto de S que sea una base para el gen(S).

Ejemplo 4.64. Sea el subconjunto $S = \{p_1, p_2, p_3, p_4, p_5\}$ del espacio vectorial $\mathbb{P}_3$ donde $p_1 = 1 + x + 2x^3$, $p_2 = x + x^2 - x^3$, $p_3 = 2 + 3x + x^2 + 3x^3$, $p_4 = -1 + x^2 - 3x^3$ y $p_5 = 1 + x + x^2 + x^3$, y se desea encontrar una base para el espacio $\mathbb{W} = \text{gen}(S)$.

Claramente S es linealmente dependiente ya que consta de 5 polinomios de $\mathbb{P}_3$ y $\dim(\mathbb{P}_3) = 4$. Se deduce entonces que uno de los polinomios de S es combinación lineal de los restantes. Si no se descubre a simple vista uno de tales polinomios, debe resolverse la ecuación:

$$c_1\left(1+x+2x^3\right)+c_2\left(x+x^2-x^3\right)+c_3\left(2+3x+x^2+3x^3\right)+c_4\left(-1+x^2-3x^3\right)+c_5\left(1+x+x^2+x^3\right)=0.$$

Operando se obtiene

$$(c_1+2c_3-c_4+c_5)+(c_1+c_2+3c_3+c_5)\,x+(c_2+c_3+c_4+c_5)\,x^2+(2c_1-c_2+3c_3-3c_4+c_5)\,x^3=0,$$

que conduce al sistema lineal homogéneo

$$\begin{cases} c_1+2c_3-c_4+c_5=0\\ c_1+c_2+3c_3+c_5=0\\ c_2+c_3+c_4+c_5=0\\ 2c_1-c_2+3c_3-3c_4+c_5=0\end{cases}$$

Aplicando Gauss-Jordan a la matriz aumentada del sistema, se llega a su forma escalonada reducida (verifíquelo):

$$\begin{bmatrix} \boxed{1} & 0 & 2 & -1 & 0 & : & 0\\ 0 & \boxed{1} & 1 & 1 & 0 & : & 0\\ 0 & 0 & 0 & 0 & \boxed{1} & : & 0\\ 0 & 0 & 0 & 0 & 0 & : & 0\end{bmatrix}$$

Las variables libres son c_3 y c_4 y la solución general del sistema esá dada por $c_1 = -2r+t$, $c_2 = -r-t$, $c_3 = r$, $c_4 = t$, $c_5 = 0$ con $t, r \in \mathbb{R}$.

Por lo tanto se mantiene la siguiente relación de dependencia lineal entre los polinomios p_1, p_2, p_3, p_4 y p_5:

$$(-2r+t)\,p_1+(-r-t)\,p_2+rp_3+tp_4+0p_5=0,\quad r,t\in\mathbb{R}. \tag{4.7}$$

El polinomio p_5 no es una combinación lineal de los polinomios restantes mientras que, tomando por ejemplo $t = 1$:

$$p_4=(2r-1)p_1+(r+1)p_2-rp_3+0p_5,\quad r\in\mathbb{R}.$$

Si se elimina p_4 de S se obtiene el conjunto $S_1 = \{p_1, p_2, p_3, p_5\}$ que genera el mismo espacio, $\mathbb{W}$, que S. Resolviendo la ecuación $d_1p_1 + d_2p_2 + d_3p_3 + d_4p_5 = 0$, o bien tomando $t = 0$ en la ecuación (4.7), puede comprobarse rápidamente que estos vectores verifican la siguiente relación de dependencia lineal:

$$(-2r)p_1+(-r)p_2+rp_3+0p_5=0,\quad r\in\mathbb{R},$$

que muestra que S_1 es linealmente dependiente y que p_3 puede expresarse como combinación lineal de los restantes polinomios de S_1. Considerando $r \neq 0$:

$$p_3=2p_1+p_2+0p_5.$$

Eliminando p_3 se obtiene el conjunto $S_2 = \{p_1, p_2, p_5\}$ que genera el mismo espacio, $\mathbb{W}$, que S.

Es fácil comprobar que S_2 es linealmente independiente sin resolver ninguna ecuación dado que p_1 y p_2 son linealmente independientes y p_5 no es combinación lineal de ellos. Así S_2 es una base para $\mathbb{W} = \text{gen}(S)$ que está contenida en S. ◀

Actividad 4.9. Considere el subconjunto $S = \{\mathbf{u}_1, \mathbf{u}_2\}$ del espacio $\mathbb{R}^4$ donde $\mathbf{u}_1 = (1, 2, 1, 0)$ y $\mathbf{u}_2 = (-1, 1, -1, 1)$. Siguiendo la demostración de la parte 1 del teorema anterior, amplíe el conjunto S hasta obtener una base de $\mathbb{R}^4$.

El siguiente teorema demuestra que todo subespacio de un espacio vectorial $\mathbb{V}$ finito dimensional es (como es de esperar) también finito dimensional y que su dimensión no supera a la dimensión de $\mathbb{V}$.

Teorema 4.11. Sea $\mathbb{V}$ un espacio vectorial no cero de dimensión n y $\mathbb{W}$ un subespacio del mismo. Entonces:

1. $\mathbb{W}$ es finito dimensional y $\dim(\mathbb{W}) \leq n$.
2. $\dim(\mathbb{W}) = n$ si y solo si $\mathbb{W} = \mathbb{V}$

Demostración. 1. Si $\mathbb{W} = \{\mathbf{0}\}$ entonces $\mathbb{W}$ es finito dimensional y $\dim(\mathbb{W}) = 0$. Dado que $n \geq 1$ entonces $\dim(\mathbb{W}) < n$.

Supóngase $\mathbb{W} \neq \{\mathbf{0}\}$. Debe probarse que $\mathbb{W}$ contiene un conjunto finito que es una base de $\mathbb{W}$ y que dicha base no contiene más de n vectores.

Dado que $\mathbb{W}$ no es el subespacio cero entonces $\mathbb{W}$ contiene subconjuntos linealmente independientes (para cualquier vector $\mathbf{w} \neq 0$ en $\mathbb{W}$, el conjunto unitario $\{\mathbf{w}\}$ es linealmente independiente).

Sea $S = \{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m\}$ un subconjunto de $\mathbb{W}$ que es linealmente independiente. Dado que $\dim(\mathbb{V}) = n$, S contiene a lo sumo n vectores, esto es: $m \leq n$. Si S genera a $\mathbb{W}$ entonces S es una base de $\mathbb{W}$. Así, $\mathbb{W}$ es finito dimensional y $\dim(\mathbb{W}) = m \leq n$.

Si S no genera a $\mathbb{W}$, hay un vector $\mathbf{w}$ en $\mathbb{W}$ que no pertence a $\text{gen}(S)$. Agregando el vector $\mathbf{w}$ al conjunto S se obtiene un subconjunto S_1 de $\mathbb{W}$ que, de acuerdo al Lema 4.1, es también linealmente independiente y por lo tanto tiene a lo sumo n vectores. Si S_1 genera a $\mathbb{W}$ entonces S_1 es una base de $\mathbb{W}$ y por lo tanto $\mathbb{W}$ es finito dimensional y $\dim(\mathbb{W}) \leq n$. Si S_1 no genera a $\mathbb{W}$, se repite el procedimiento con S_1 en lugar de S, y en no más de n pasos se llega a un subconjunto, S_k, de $\mathbb{W}$ que contiene a lo sumo n vectores y que es linealmente independiente y generador de $\mathbb{W}$. El conjunto S_k es una base de $\mathbb{W}$, así W es finito dimensional y $\dim(\mathbb{W}) \leq n$.

2. Trivialmente, si $\mathbb{W}$ es el propio espacio $\mathbb{V}$ entonces $\dim(\mathbb{W}) = n$.

Recíprocamente, supóngase que $\dim(\mathbb{W}) = n$. Si S es una base de $\mathbb{W}$ entonces S es un conjunto generador de $\mathbb{W}$ que consta de n vectores linealmente independientes. Dado que S es también un subconjunto de $\mathbb{V}$ que es linealmente independiente y $\dim(\mathbb{V}) = n$, de acuerdo al Teorema 4.9, S es una base de $\mathbb{V}$. Por lo tanto $\text{gen}(S) = \mathbb{V}$, y como $\text{gen}(S) = \mathbb{W}$, resulta $\mathbb{V} = \mathbb{W}$.

□

Ejemplo 4.65. Los subespacios de $\mathbb{R}^2$ son:

- Subespacios de dimensión cero: hay uno solo, el subespacio trivial $\{\mathbf{0}\}$.
- Subespacios unidimensionales: hay infinitos, todas las rectas que pasan por el origen. Son del tipo $\text{gen}\{\mathbf{u}\}$ con $\mathbf{u} \neq 0$.
- Subespacios bidimensionales: hay uno solo, el mismo espacio $\mathbb{R}^2$.

Actividad 4.10. Describa geométricamente todos los subespacios de $\mathbb{R}^3$.

Ejercicios 4.5

1. Determine si los conjuntos dados son bases de $\mathbb{R}^3$:

   a) $\{(1,-1,3),(0,4,-1),(1,1,1)\}$
   b) $\{(1,1,1),(0,-1,5)\}$
   c) $\{(0,0,1),(1,-1,2),(1,0,0),(0,1,-1)\}$
   d) $\{(1,-1,2),(2,0,4),(1,-2,2)\}$

2. Determine si los conjuntos dados son bases de $\mathbb{M}_2$:

   a) $\left\{\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 3 & 4 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 5 & 6 \\ 5 & 6 \end{bmatrix}, \begin{bmatrix} 7 & 8 \\ 7 & 8 \end{bmatrix}\right\}$
   b) $\left\{\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\}$
   c) $\left\{\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -2 & 3 \\ 1 & 0 \end{bmatrix}\right\}$

3. Determine si los conjuntos dados son bases de $\mathbb{P}_2$:

   a) $\left\{1+x+x^2, 2-x, x+3x^2\right\}$
   b) $\left\{1, 1+2x, 1+2x+3x^2, 1-x^2\right\}$
   c) $\left\{1+x, x+x^2\right\}$
   d) $\left\{2-x+x^2, 3+x+x^2, 5+2x^2\right\}$

4. Sea el subespacio $\mathbb{V}$ de $\mathbb{M}_2$ que consta de todas las matrices de la forma $\begin{bmatrix} a & b \\ c & a-b \end{bmatrix}$. Demuestre que $\{E_{11}+E_{22}, E_{12}-E_{22}, E_{21}\}$ es una base de $\mathbb{V}$.

5. Pruebe que si $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ es una base de un espacio vectorial $\mathbb{V}$ y $c \neq 0$, entonces $\{c\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ es también una base de $\mathbb{V}$.

6. En cada caso, halle la dimensión del subespacio $\mathbb{W}$ de $\mathbb{V}$:

   a) $\mathbb{W} = \{(a-b, b, 2a+b) : a, b \in \mathbb{R}\}$; $\mathbb{V} = \mathbb{R}^3$.
   b) $\mathbb{W} = \{(a, b, 0) : a, b \in \mathbb{R}\}$; $\mathbb{V} = \mathbb{R}^3$.
   c) $\mathbb{W} = \{(2a, -b, a+b) : a, b \in \mathbb{R}\}$; $\mathbb{V} = \mathbb{R}^3$.
   d) $\mathbb{W}$ es el conjunto de todos los vectores de $\mathbb{R}^4$ con primera y última componentes iguales a cero; $\mathbb{V} = \mathbb{R}^4$.
   e) $\mathbb{W} = \text{gen}\{(1,2,-1,0),(0,3,1,0),(1,1,1,1),(2,6,1,1)\}$; $\mathbb{V} = \mathbb{R}^4$.
   f) $\mathbb{W} = \left\{\begin{bmatrix} a & b \\ 2b & a-b \end{bmatrix} : a, b \in \mathbb{R}\right\}$; $\mathbb{V} = \mathbb{M}_2$.
   g) $\mathbb{W} = \text{gen}\left\{\begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix}\right\}$; $\mathbb{V} = \mathbb{M}_2$.

$\boxed{h)}$ $\mathbb{W} = \left\{a + bx - (a+b)x^2 : a, b \in \mathbb{R}\right\}$; $\mathbb{V} = \mathbb{P}_2$.

i $\mathbb{W} = \text{gen}\left\{1 + x, 1 - x^2, 2 + x^4, x + x^4\right\}$; $\mathbb{V} = \mathbb{P}$.

$\boxed{7.}$ Indique el valor de a de modo que el subespacio gen $\left\{1 + x, -1 + 2x + ax^2, 2 + x - x^2\right\}$ de $\mathbb{P}_2$ sea de dimensión 2.

8. Halle la dimensión del subespacio $\mathbb{W} = \{a \cos x + b \operatorname{sen} x : a, b \in \mathbb{R}\}$ de $\mathbb{F}(\mathbb{R})$.

$\boxed{9.}$ Encuentre una base para el subespacio gen $\left\{e^x, xe^x, x^2e^x\right\}$ de $\mathbb{F}(\mathbb{R})$ y obtenga su dimensión.

$\boxed{10.}$ Halle la dimensión del subespacio gen $\left\{\cos^2 x, \operatorname{sen}^2 x, 1\right\}$ de $\mathbb{F}(\mathbb{R})$.

11. En cada caso el conjunto S genera al espacio $\mathbb{V}$. Reduzca S hasta obtener una base de $\mathbb{V}$:

a) $S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \right\}$, $\mathbb{V} = \mathbb{R}^3$.

$\boxed{b)}$ $S = \{1 + x, -5x, 3 + 4x\}$, $\mathbb{V} = \mathbb{P}_1$.

12. En cada caso S es un conjunto linealmente independiente del espacio $\mathbb{V}$. Amplíe S hasta obtener una base de $\mathbb{V}$:

a) $S = \{(1, -1, 3)\}$; $\mathbb{V} = \mathbb{R}^3$.

b) $S = \left\{ \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 3 & 0 \end{bmatrix} \right\}$; $\mathbb{V} = \mathbb{M}_2$.

$\boxed{13.}$ Responda verdadero o falso. Justifique adecuadamente su respuesta:

a) Todo conjunto unitario $S = \{a\}$ con $a \neq 0$, es una base del espacio vectorial $\mathbb{R}$.

b) $\mathbb{R}^{10}$ tiene un subespacio de dimensión 9.

c) Un subespacio no cero de $\mathbb{R}^{10}$ puede tener dos bases con diferente número de elementos.

d) Un subespacio no cero de $\mathbb{P}_5$ puede tener una base con 7 elementos.

e) Un subespacio no cero de $\mathbb{P}_5$ puede tener una base con 5 elementos.

f) Si $S = \{\mathbf{u}, \mathbf{v}\}$ es un conjunto linealmente independiente del espacio vectorial $\mathbb{V}$ de dimensión 2, entonces $\{\mathbf{u} - \mathbf{v}, \mathbf{u} + \mathbf{v}\}$ es una base de $\mathbb{V}$.

g) Sea $\mathbb{V}$ un espacio vectorial de dimensión 3. Si $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es conjunto generador de $\mathbb{V}$, entonces $\{2\mathbf{u}, \mathbf{u} - \mathbf{v}, \mathbf{u} + \mathbf{v} + \mathbf{w}\}$ es una base de $\mathbb{V}$.

h) Si $\mathbb{V}$ es un espacio vectorial de dimensión n y k es un número entero tal que $0 \leq k \leq n$, entonces hay un subespacio $\mathbb{W}$ de $\mathbb{V}$ tal que $\dim(\mathbb{W}) = k$.

i) Para cualesquiera $\mathbf{u}$ y $\mathbf{v}$ vectores del espacio $\mathbb{V}$, $\dim(\text{gen}\{\mathbf{u}, \mathbf{v}\}) = \dim(\text{gen}\{\mathbf{u} - \mathbf{v}, \mathbf{u} + \mathbf{v}\})$.

j) Un espacio vectorial $\mathbb{V}$ tal que $\dim(\mathbb{V}) = 1$ no tiene subespacios propios.

k) Si $\mathbb{W}$ es un subespacio no trivial de un espacio vectorial $\mathbb{V}$ de dimensión $n \geq 2$, entonces toda base de $\mathbb{W}$ puede ampliarse a una base de $\mathbb{V}$.

l) Si $\mathbb{W}$ es un subespacio no trivial de un espacio vectorial $\mathbb{V}$ de dimensión $n \geq 2$, entonces toda base de $\mathbb{V}$ contiene una base de $\mathbb{W}$.

4.6. Coordenadas y cambio de base

En esta sección se estudia la relación entre la noción de base y la de sistema de coordenadas.

En la geometría del plano se asocia un par ordenado (a, b) de números reales a un punto P en el plano aplicando dos ejes perpendiculares x e y llamados ejes de coordenadas, a es la coordenada de P con respecto al eje x y b es la coordenada de P con respecto al eje y (véase Figura 4.8).

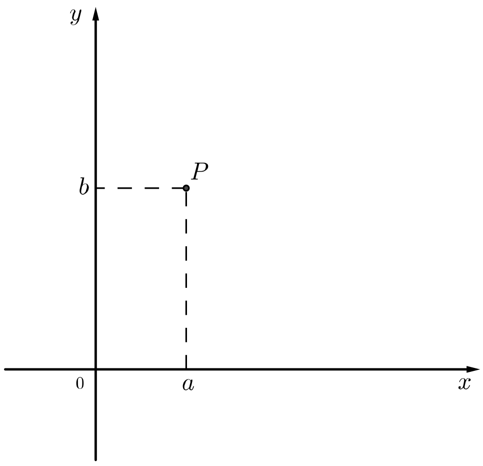


Figura 4.8: Coordenadas (a, b) del punto P en el plano

También se pueden introducir coordenadas utilizando vectores. Por ejemplo, véase Figura 4.9, los versores fundamentales $\mathbf{i}$ y $\mathbf{j}$ forman una base para $\mathbb{R}^2$ y puede verse que al bajar rectas perpendiculares desde P hacia los ejes x e y determinados por los vectores $\mathbf{i}$ y $\mathbf{j}$ se obtienen los vectores $a\mathbf{i}$ y $b\mathbf{j}$, y que

$$\overrightarrow{OP} = a\mathbf{i} + b\mathbf{j}.$$

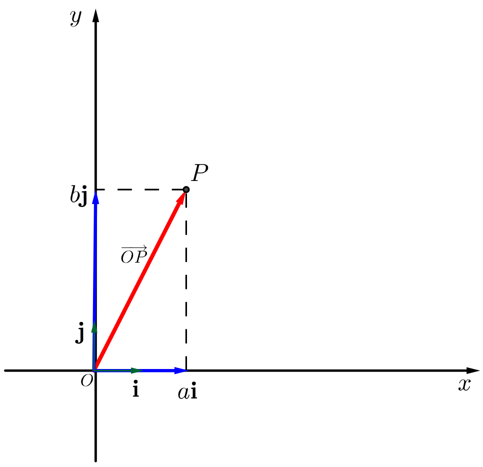


Figura 4.9: Vector $\overline{OP}$ en el plano

Las coordenadas a y b de P se pueden concebir como los números (coeficientes) necesarios para expresar al vector $\mathbf{v} = \overrightarrow{OP}$ como combinación lineal de los vectores de la base $\{\mathbf{i}, \mathbf{j}\}$.

Esta noción de coordenadas se puede extender a espacios más generales. Para esto es necesario un resultado preliminar:

Teorema 4.12. Si $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ es una base para un espacio vectorial $\mathbb{V}$ entonces todo vector $\mathbf{v}$ en $\mathbb{V}$ se puede expresar como combinación lineal de los vectores de S de una única manera. Es decir, para cada $\mathbf{v}$ de $\mathbb{V}$ existen únicos escalares c_1, c_2,...,c_n tales que

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n.$$

Demostración. Dado que S genera a $\mathbb{V}$, todo vector de $\mathbb{V}$ se puede expresar como una combinación lineal de los vectores de S. Se verá que la independencia lineal de S asegura que esta manera es única.

Para ello, supóngase que un vector $\mathbf{v}$ se puede escribir como

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n,$$

y también como

$$\mathbf{v} = k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + ... + k_n\mathbf{u}_n.$$

Por lo tanto:

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n = k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + ... + k_n\mathbf{u}_n,$$

y operando se obtiene:

$$(c_1 - k_1)\,\mathbf{u}_1 + (c_2 - k_2)\,\mathbf{u}_2 + ... + (c_n - k_n)\,\mathbf{u}_n = \mathbf{0},$$

que muestra al vector cero como una combinación lineal de los vectores de S y dado que S es linealmente independiente, necesariamente:

$$c_1 - k_1 = 0, c_2 - k_2 = 0, ..., c_n - k_n = 0,$$

es decir

$$c_1 = k_1,\ c_2 = k_2,\ ...,\ c_n = k_n.$$

Luego, no hay dos maneras distintas de escribir a un vector de $\mathbb{V}$ como combinación lineal de los elementos de una base. Así, quedó probado el teorema. $\square$

Bases ordenadas y vector de coordenadas

Se conoce que si $\mathbb{V}$ es un espacio vectorial de dimensión finita, n, y los vectores $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n$ conforman un conjunto S que es una base de $\mathbb{V}$, no importa el orden en que se los considere. Así, por ejemplo:

$$S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\} = \{\mathbf{u}_2, \mathbf{u}_1, ..., \mathbf{u}_n\} = \{\mathbf{u}_n, \mathbf{u}_{n-1}, ..., \mathbf{u}_1\}.$$

En esta sección interesará el orden en que se toman los vectores de una base y se hablará de **bases ordenadas**.

Una base ordenada B es una base $B = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ juntamente con el orden dado. Cambiar el orden en que están enlistados los elementos de una base ordenada produce una base ordenada diferente. En este sentido $B_1 = \{\mathbf{u}_2, \mathbf{u}_1, ..., \mathbf{u}_n\}$ es una base ordenada distinta a la base ordenada B.

Definición (Vector de coordenadas). Sea $\mathbb{V}$ un espacio vectorial finito dimensional de dimensión n y $B = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ un base ordenada de $\mathbb{V}$. Dado que B es una base de $\mathbb{V}$, para cada vector $\mathbf{v}$ de $\mathbb{V}$ existen únicos escalares $c_1, c_2, ..., c_n$ tales que

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n.$$

Los coeficientes $c_1, c_2, ..., c_n$ de la combinación lineal se denominan *coordenadas del vector* $\mathbf{v}$ *respecto a la base ordenada* B. El coeficiente c_i, correspondiente al i-ésimo vector $\mathbf{u}_i$ de la base B, es *la coordenada* i*-ésima del vector* $\mathbf{v}$ (respecto a la base ordenada B).

El *vector de coordenadas* de $\mathbf{v}$ *respecto a la base ordenada* B se denota por $[\mathbf{v}]_B$ y es el vector columna de $\mathbb{R}^n$ definido por

$$[\mathbf{v}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Observe que un cambio en el orden de los vectores del conjunto B conduce a un cambio en el orden de las componentes del vector de coordenadas de $\mathbf{v}$ y por lo tanto a vectores de coordenadas distintos. De ahí la necesidad de imponer algún orden a los vectores que forman una base B y considerar bases ordenadas de modo que no se produzcan ambigüedades al definir la i-ésima coordenada de $\mathbf{v}$.

Toda base a lo largo de esta sección se considera una base ordenada aún cuando no se lo explicite.

Vale aclarar que el vector de coordenadas se tomó como vector columna aunque bien podría haber sido expresado en la forma

$$(\mathbf{v})_B = (c_1, c_2, ..., c_n).$$

La notación como columna es solo por conveniencia ya que es útil cuando se procede a describir qué le sucede a las coordenadas de un vector $\mathbf{v}$ cuando se cambia de una base ordenada a otra.

Correspondencia entre $\mathbb{V}$ y $\mathbb{R}^n$

Debe observarse que fijada una base ordenada $B = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ de un espacio vectorial $\mathbb{V}$ de dimensión n, existe una correspondencia biunívoca entre el conjunto $\mathbb{V}$ y el conjunto $\mathbb{R}^n$.

En efecto: la asignación

$$\mathbf{v} \to [\mathbf{v}]_B,$$

hace corresponder a cada vector $\mathbf{v}$ de $\mathbb{V}$ un único vector de $\mathbb{R}^n$, el vector $[\mathbf{v}]_B$.

Recíprocamente, cada vector $\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$ de $\mathbb{R}^n$ es el vector de coordenadas, con respecto a la base B, de un único vector $\mathbf{v}$ de $\mathbb{V}$, a saber, el vector $\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n$.

Ejemplo 4.66. Sean las siguientes bases (ordenadas) de $\mathbb{R}^2$, la estándar $E = \{\mathbf{e}_1, \mathbf{e}_2\}$ y la base $B = \{\mathbf{u}_1, \mathbf{u}_2\}$ donde $\mathbf{u}_1 = (1, 2)$ y $\mathbf{u}_2 = (3, 1)$, y el vector $\mathbf{v} = (11, 7)$.

Puede verificarse rápidamente que

$$\mathbf{v} = (11, 7) = 2(1, 2) + 3(3, 1) = 2\mathbf{u}_1 + 3\mathbf{u}_2,$$

y así, 2 y 3 son la primera y segunda coordenada de $\mathbf{v}$ respecto de la base B, respectivamente. Luego,

$$[\mathbf{v}]_B = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Por otra parte,

$$\mathbf{v} = (11, 7) = 11(1, 0) + 7(0, 1) = 11\mathbf{e}_1 + 7\mathbf{e}_2,$$

luego

$$[\mathbf{v}]_E = \begin{bmatrix} 11 \\ 7 \end{bmatrix}.$$

En la Figura 4.10 se muestra al vector $\mathbf{v}$ considerado como combinación lineal de $\mathbf{e}_1$ y $\mathbf{e}_2$ y como combinación lineal de $\mathbf{u}_1$ y $\mathbf{u}_2$.◀

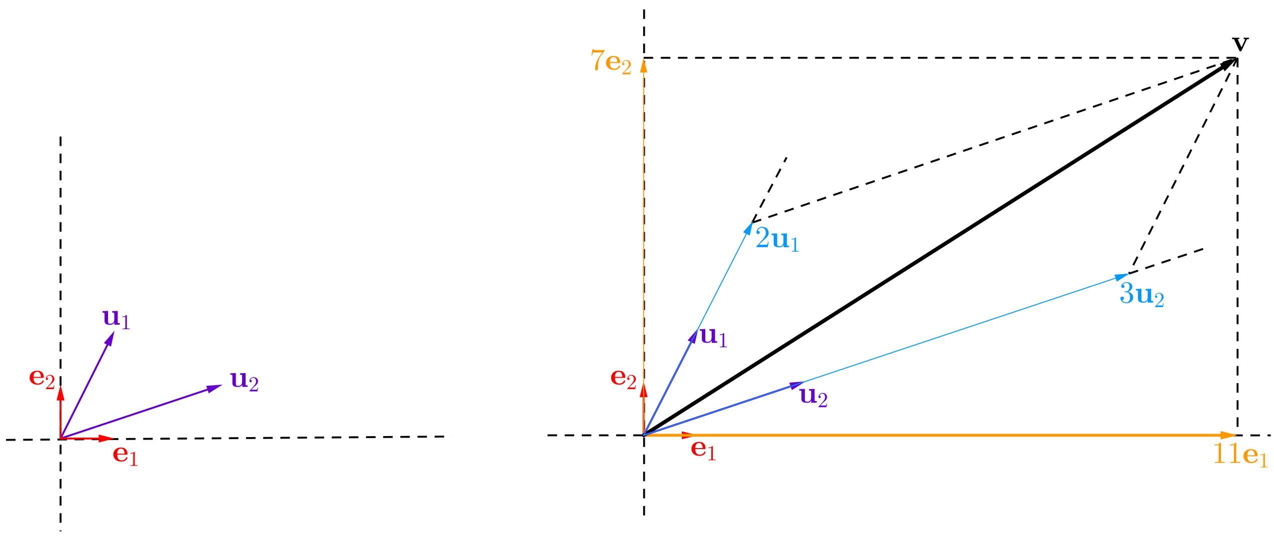


Figura 4.10: Vector $\mathbf{v}$ como combinación lineal de los vectores de E y B

Ejemplo 4.67. Considere los vectores $\mathbf{u}_1$, $\mathbf{u}_2$ y $\mathbf{v}$ del ejemplo anterior. La base ordenada $S = \{\mathbf{u}_2, \mathbf{u}_1\}$ de $\mathbb{R}^2$ difiere de la base $B = \{\mathbf{u}_1, \mathbf{u}_2\}$ por el orden de sus elementos. Entonces dado que

$$\mathbf{v} = (11, 7) = 2\mathbf{u}_1 + 3\mathbf{u}_2 \ \left(= 3\mathbf{u}_2 + 2\mathbf{u}_1\right),$$

resulta

$$[\mathbf{v}]_S = \begin{bmatrix} 3 \\ 2 \end{bmatrix},$$

que es distinto al vector de coordenadas $[\mathbf{v}]_B$. ◀

Ejemplo 4.68. (Obtención de v a partir de $[\mathbf{v}]_B$)

Considere la base $B = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ de $\mathbb{R}^3$ donde $\mathbf{u}_1 = (1, 0, 2)$, $\mathbf{u}_2 = (2, 1, -3)$ y $\mathbf{u}_3 = (3, 1, 0)$. Si el vector de coordenadas de un vector $\mathbf{v}$ de $\mathbb{R}^3$ en la base B es

$$[\mathbf{v}]_B = \begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix},$$

entonces el vector $\mathbf{v}$ es

$$\mathbf{v} = (-1)\mathbf{u}_1 + 2\mathbf{u}_2 + 4\mathbf{u}_3 = (-1, 0, -2) + (4, 2, -6) + (12, 4, 0) = (15, 6, -8). \blacktriangleleft$$

Ejemplo 4.69. (Vector de coordenadas del vector 0)

Sea $B = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ una base de un espacio vectorial $\mathbb{V}$. El vector de coordenadas del cero de $\mathbb{V}$ es el vector $\mathbf{0}$ de $\mathbb{R}^n$. Esto es:

$$[\mathbf{0}]_B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

ya que el vector $\mathbf{0}$ de $\mathbb{V}$ se escribe como combinación lineal de los vectores de B de la siguiente única manera:

$$\mathbf{0} = 0\mathbf{u}_1 + 0\mathbf{u}_2 + ... + 0\mathbf{u}_n. \blacktriangleleft$$

Ejemplo 4.70. Sea $B = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ una base de un espacio vectorial $\mathbb{V}$. Los vectores de coordenadas de los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_n$ con respecto a la base B que los contiene, son los vectores de la base estándar de $\mathbb{R}^n$, $\mathbf{e}_1$, $\mathbf{e}_2$,...,$\mathbf{e}_n$. En efecto, para todo $j = 1, 2, ..., n$:

$$\mathbf{u}_j = 0\mathbf{u}_1 + 0\mathbf{u}_2 + ... + 0\mathbf{u}_{j-1} + 1\mathbf{u}_j + 0\mathbf{u}_{j+1} + ... + 0\mathbf{u}_n,$$

luego

$$[\mathbf{u}_j]_B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow j\text{-ésima coordenada.} \\ \\ \\ \\ \end{matrix}$$

Así, $[\mathbf{u}_j]_B = \mathbf{e}_j$ para todo $j = 1, 2, ..., n$.$\blacktriangleleft$

Ejemplo 4.71. (Coordenadas de un vector x de $\mathbb{R}^n$ con respecto a la base estándar)

Las coordenadas de un vector $\mathbf{x} = (x_1, x_2, ..., x_n)$ de $\mathbb{R}^n$ con respecto a la base estándar coinciden con las componentes de dicho vector.

En efecto, sea $E = \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ la base estándar ordenada de $\mathbb{R}^n$. Dado que:

$$\mathbf{x} = (x_1, x_2, ..., x_n) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n,$$

entonces,

$$[\mathbf{x}]_E = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Puede observarse que el vector de coordenadas $[\mathbf{x}]_E$ es el mismo vector $\mathbf{x}$, escrito como vector columna.◄

Encontrar las coordenadas de un vector particular usualmente involucra resolver un sistema de ecuaciones lineales a menos que la base tenga algo especial como la base canónica o estándar, o las bases ortogonales que se verán en secciones posteriores.

Ejemplo 4.72. Sean las siguientes bases (ordenadas) de $\mathbb{P}_2$: la base canónica $E = \left\{1, x, x^2\right\}$ y la base $B = \{p_1, p_2, p_3\}$ donde $p_1 = 1 + x + x^2$, $p_2 = x + 2x^2$ y $p_3 = 2 + 3x^2$.

Se quiere hallar los vectores de coordenadas $[p]_E$ y $[p]_B$ del polinomio $p = \frac{3}{2} + \frac{7}{2}x - 2x^2$.

Es muy fácil hallar las coordenadas de p en la base canónica E ya que $p = \frac{3}{2}(1) + \frac{7}{2}(x) + (-2)\left(x^2\right)$. Luego:

$$[p]_E = \begin{bmatrix} 3/2 \\ 7/2 \\ -2 \end{bmatrix}.$$

Para hallar $[p]_B$ deben encontrarse los números (únicos) c_1, c_2 y c_3 tales que

$$c_1p_1 + c_2p_2 + c_3p_3 = p,$$

esto es

$$c_1\left(1 + x + x^2\right) + c_2\left(x + 2x^2\right) + c_2\left(2 + 3x^2\right) = \frac{3}{2} + \frac{7}{2}x - 2x^2,$$

operando

$$(c_1 + 2c_3) + (c_1 + c_2)\,x + (c_1 + 2c_2 + 3c_3)\,x^2 = \frac{3}{2} + \frac{7}{2}x - 2x^2,$$

y por igualdad de polinomios

$$\left\{\begin{array}{cccccc} c_1 & & +2c_3 & = & 3/2 \\ c_1 & +c_2 & & = & 7/2 \\ c_1 & +2c_2 & +3c_3 & = & -2 \end{array}\right.$$

La solución de este sistema es: $c_1 = \frac{9}{2}$, $c_2 = -1$ y $c_3 = -\frac{3}{2}$, y por lo tanto

$$[p]_B = \begin{bmatrix} 9/2 \\ -1 \\ -3/2 \end{bmatrix}. ◄$$

Ahora se trata el problema central de esta sección, el problema del cambio de base.

Matriz de transición o de cambio de base

Supóngase que $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ y $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ son dos bases ordenadas de un espacio vectorial $\mathbb{V}$ de dimensión n. Se examinará la relación que existe entre los vectores de coordenadas $[\mathbf{v}]_{B_1}$ y $[\mathbf{v}]_{B_2}$ de cualquier vector $\mathbf{v}$ de $\mathbb{V}$.

Sea $\mathbf{v}$ un vector de $\mathbb{V}$ cuyo vector de coordenadas en la base B_1 es

$$[\mathbf{v}]_{B_1} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix},$$

entonces

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n. \tag{4.8}$$

Si se quiere encontrar el vector de coordenadas de $\mathbf{v}$ con respecto a la base B_2, $[\mathbf{v}]_{B_2}$, deben hallarse los coeficientes que permiten expresar al vector $\mathbf{v}$ como combinación lineal de los vectores de la base B_2. Para ello se expresará cada vector $\mathbf{u}_i$ de la base B_1 como combinación lineal de los vectores de la base B_2. Sean entonces:

$$\begin{aligned} \mathbf{u}_1 &= a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + ... + a_{n1}\mathbf{v}_n \\ \mathbf{u}_2 &= a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + ... + a_{n2}\mathbf{v}_n \\ &\vdots \\ \mathbf{u}_n &= a_{1n}\mathbf{v}_1 + a_{2n}\mathbf{v}_2 + ... + a_{nn}\mathbf{v}_n \end{aligned} \tag{4.9}$$

de modo que

$$[\mathbf{u}_1]_{B_2} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad [\mathbf{u}_2]_{B_2} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, ..., \quad [\mathbf{u}_n]_{B_2} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix}.$$

Reemplazando las expresiones (4.9) en (4.8) y operando (verifique) se obtiene el vector $\mathbf{v}$ expresado como combinación lineal de los vectores de la base B_2:

$$\mathbf{v} = (c_1a_{11} + c_2a_{12} + ... + c_na_{1n})\,\mathbf{v}_1 + (c_1a_{21} + c_2a_{22} + ... + c_na_{2n})\,\mathbf{v}_2 + ... + (c_1a_{n1} + c_2a_{n2} + ... + c_na_{nn})\,\mathbf{v}_n.$$

Los coeficientes que acompañan a los vectores $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_n$ de la expresión anterior son las coordenadas del vector $\mathbf{v}$ con respecto a la base B_2. Así, y por definición de producto de matrices:

$$[\mathbf{v}]_{B_2} = \begin{bmatrix} c_1a_{11} + c_2a_{12} + ... + c_na_{1n} \\ c_1a_{21} + c_2a_{22} + ... + c_na_{2n} \\ \vdots \\ c_1a_{n1} + c_2a_{n2} + ... + c_na_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & ... & \vdots \\ a_{n1} & a_{n2} & ... & a_{nn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Dado que $\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = [\mathbf{v}]_{B_1}$, y llamando P a la matriz que lo premultiplica, entonces:

$$[\mathbf{v}]_{B_2} = P\,[\mathbf{v}]_{B_1},$$

donde

$$P = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

La igualdad anterior muestra que premultiplicando el vector de coordenadas de un vector **v** en la base B_1 por la matriz P, se obtiene el vector de coordenadas del mismo vector **v** en la base B_2.

Debe observarse que la matriz P tiene como columnas a los vectores de coordenadas de los vectores de la base B_1 con respecto a la base B_2 es decir

$$P = \left[\ [\mathbf{u}_1]_{B_2} \quad [\mathbf{u}_2]_{B_2} \quad \dots \quad [\mathbf{u}_n]_{B_2}\ \right].$$

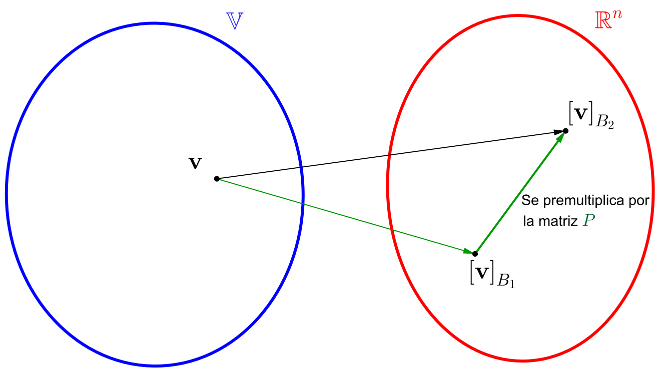
Se ha probado el siguiente teorema:

Teorema 4.13. Sea $\mathbb{V}$ un espacio vectorial de dimensión n y sean $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ y $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ dos bases ordenadas de $\mathbb{V}$. Entonces existe una matriz , P, de $n \times n$ tal que:

$$\forall\ \mathbf{v} \in \mathbb{V}, \quad [\mathbf{v}]_{B_2} = P\,[\mathbf{v}]_{B_1}.$$

En términos de sus columnas, $P = \left[\ [\mathbf{u}_1]_{B_2} \quad [\mathbf{u}_2]_{B_2} \quad \dots \quad [\mathbf{u}_n]_{B_2}\ \right]$.

La matriz P del teorema anterior es la única matriz tal que para cada **v** en $\mathbb{V}$, $[\mathbf{v}]_{B_2} = P\,[\mathbf{v}]_{B_1}$ (puede verse una demostración al final de esta sección). Esta matriz es llamada **matriz de transición o de cambio de base de la base B_1 a la base B_2.**



Ejemplo 4.73. (Cálculo de una matriz de transición) En el espacio vectorial $\mathbb{R}^2$, sean las bases ordenadas $B_1 = \{\mathbf{u}_1, \mathbf{u}_2\}$ y $B_2 = \{\mathbf{v}_1, \mathbf{v}_2\}$ donde $\mathbf{u}_1 = (2, 0)$, $\mathbf{u}_2 = (1, 2)$, $\mathbf{v}_1 = (6, 3)$ y $\mathbf{v}_2 = (4, -1)$.

La matriz P de transición de la base B_1 a la base B_2 tiene por columnas los vectores de coordenadas de los vectores de B_1 con respecto a la base B_2, es decir

$$P = \left[\ [\mathbf{u}_1]_{B_2} \quad [\mathbf{u}_2]_{B_2}\ \right].$$

Para calcular $[\mathbf{u}_1]_{B_2}$, se deben determinar los coeficientes (únicos) a_1 y a_2 tales que

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 = \mathbf{u}_1,$$

es decir

$$a_1(6,3) + a_2(4,-1) = (2,0),$$

o bien

$$\begin{cases} 6a_1 + 4a_2 = 2 \\ 3a_1 - a_2 = 0 \end{cases}.$$

Se trata entonces de encontrar la (única) solución del sistema lineal de matriz aumentada:

$$\begin{bmatrix} 6 & 4 & : & 2 \\ 3 & -1 & : & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & : & \mathbf{u}_1 \end{bmatrix},$$

entendiendo que $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{u}_1$ están expresados como vectores columna.

Similarmente, para hallar $[\mathbf{u}_2]_{B_2}$, se deben determinar los (únicos) coeficientes b_1 y b_2 tales que

$$b_1\mathbf{v}_1 + b_2\mathbf{v}_2 = \mathbf{u}_2,$$

es decir

$$b_1(6,3) + b_2(4,-1) = (1,2),$$

o bien

$$\begin{cases} 6b_1 + 4b_2 = 1 \\ 3b_1 - b_2 = 2 \end{cases}$$

Se tiene entonces que hallar la (única) solución del sistema lineal de matriz aumentada:

$$\begin{bmatrix} 6 & 4 & : & 1 \\ 3 & -1 & : & 2 \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & : & \mathbf{u}_2 \end{bmatrix}.$$

Como la matriz de coeficientes de los dos sistemas es la misma, éstos se pueden resolver en forma simultánea llevando la matriz aumentada $\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & : & \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix}$ a su forma escalonada reducida. Aplicando Gauss-Jordan:

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & : & \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} 6 & 4 & : & 2 & 1 \\ 3 & -1 & : & 0 & 2 \end{bmatrix} \to \begin{bmatrix} 6 & 4 & : & 2 & 1 \\ 0 & -3 & : & -1 & \frac{3}{2} \end{bmatrix} \to \begin{bmatrix} 6 & 4 & : & 2 & 1 \\ 0 & 1 & : & \frac{1}{3} & -\frac{1}{2} \end{bmatrix} \to$$

$$\to \begin{bmatrix} 6 & 0 & : & \frac{2}{3} & 3 \\ 0 & 1 & : & \frac{1}{3} & -\frac{1}{2} \end{bmatrix} \to \begin{bmatrix} 1 & 0 & : & \frac{1}{9} & \frac{1}{2} \\ 0 & 1 & : & \frac{1}{3} & -\frac{1}{2} \end{bmatrix},$$

por lo cual, la solución del primer sistema es $a_1 = \frac{1}{9}$, $a_2 = \frac{1}{3}$ es decir

$$[\mathbf{u}_1]_{B_2} = \begin{bmatrix} \frac{1}{9} \\ \frac{1}{3} \end{bmatrix},$$

y la solución del segundo sistema es $b_1 = \frac{1}{2}$, $b_2 = -\frac{1}{2}$ es decir

$$[\mathbf{u}_2]_{B_2} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}.$$

Así, la matriz de transición de la base B_1 a la base B_2 es

$$P = \begin{bmatrix} \frac{1}{9} & \frac{1}{2} \\ \frac{1}{3} & -\frac{1}{2} \end{bmatrix}. \blacktriangleleft$$

Obsérvese del ejemplo anterior que la matriz P de transición de la base B_1 a la base B_2 se obtuvo al aplicar eliminación de Gauss-Jordan a la matriz

$$\left[\begin{array}{rrcrr} 6 & 4 & : & 2 & 1 \\ 3 & -1 & : & 0 & 2 \end{array}\right] = \left[\begin{array}{ccccc} \mathbf{v}_1 & \mathbf{v}_2 & : & \mathbf{u}_1 & \mathbf{u}_2 \end{array}\right]$$

donde $\mathbf{v}_1$ y $\mathbf{v}_2$ son los vectores de la base final, B_2, y $\mathbf{u}_1$ y $\mathbf{u}_2$ son los vectores de la base inicial, B_1.

La forma escalonada reducida de esta matriz es la matriz de 2×4 que tiene en el lado izquierdo a la matriz identidad, I_2, y en el lado derecho a la matriz P de transición de B_1 a B_2:

$$\left[\begin{array}{cccrr} 1 & 0 & : & \frac{1}{9} & \frac{1}{2} \\ 0 & 1 & : & \frac{1}{3} & -\frac{1}{2} \end{array}\right] = [I : P].$$

Debe notarse que la forma escalonada reducida de la matriz $\left[\begin{array}{cc} \mathbf{v}_1 & \mathbf{v}_2 \end{array}\right]$ es la matriz identidad $I = \left[\begin{array}{cc} \mathbf{e}_1 & \mathbf{e}_2 \end{array}\right]$. Esto no es casual ya que siendo $\{\mathbf{v}_1, \mathbf{v}_2\}$ una base de $\mathbb{R}^2$ hay una única manera de expresar un vector de $\mathbb{R}^2$ como combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$, luego cualquier sistema lineal que tiene a $\left[\begin{array}{cc} \mathbf{v}_1 & \mathbf{v}_2 \end{array}\right]$ como matriz de coeficientes tiene solución única, y por lo tanto la forma escalonada reducida de esta matriz es la matriz identidad I.

El procedimiento visto puede generalizarse como manera práctica para calcular una matriz de cambio de base cuando el espacio vectorial $\mathbb{V}$ es $\mathbb{R}^n$.

Cálculo de una matriz de cambio de base en el espacio $\mathbb{R}^n$

Si $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ y $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ son dos bases ordenadas de $\mathbb{R}^n$, entonces para hallar la matriz P de transición de la base B_1 a la base B_2, el procedimiento es el siguiente:

- Forme la matriz de $n \times 2n$ donde las primeras n columnas son los vectores (ordenados) de la base B_2 y las siguientes n columnas son los vectores (ordenados) de la base B_1:

$$\left[\begin{array}{ccccccccc} \mathbf{v}_1 & \mathbf{v}_2 & ... & \mathbf{v}_n & : & \mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_n \end{array}\right].$$

- Aplique el proceso de eliminación de Gauss-Jordan a dicha matriz. La matriz escalonada reducida que se obtiene tiene en el lado izquierdo a la matriz identidad, I_n, y en el lado derecho a la matriz P de transición buscada:

$$\left[\begin{array}{ccccccccc} \mathbf{e}_1 & \mathbf{e}_2 & ... & \mathbf{e}_n & : & [\mathbf{u}_1]_{B_2} & [\mathbf{u}_2]_{B_2} & ... & [\mathbf{u}_n]_{B_2} \end{array}\right].$$

Ejemplo 4.74. Sean las bases B_1 y B_2 de $\mathbb{R}^2$ del ejemplo anterior, y el vector $\mathbf{v} = (7, 6)$.

Se comprobará la relación $[\mathbf{v}]_{B_2} = P\,[\mathbf{v}]_{B_1}$. Para ello, se obtendrá $[\mathbf{v}]_{B_2}$ en forma directa y se verificará que el mismo resultado se obtiene pre-multiplicando la matriz P por el vector de coordenadas $[\mathbf{v}]_{B_1}$.

Para calcular $[\mathbf{v}]_{B_2}$ en forma directa debe resolverse el sistema lineal que permite encontrar las coordenadas del vector $\mathbf{v}$ respecto a la base B_2. Deben hallarse los coeficientes c_1 y c_2 tales que

$$c_1\,(6, 3) + c_2\,(4, -1) = (7, 6),$$

o bien

$$\begin{cases} 6c_1 + 4c_2 = 7 \\ 3c_1 - c_2 = 6 \end{cases}$$

Resolviendo el sistema se obtiene $c_1 = \frac{31}{18}$ y $c_2 = -\frac{5}{6}$. Así,

$$[\mathbf{v}]_{B_2} = \begin{bmatrix} \frac{31}{18} \\ -\frac{5}{6} \end{bmatrix}.$$

Se utilizará ahora la matriz de transición P. Puede comprobarse fácilmente que:

$$\mathbf{v} = (7,6) = 2\,(2,0) + 3\,(1,2)\,, \text{ por lo que } [\mathbf{v}]_{B_1} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Teniendo calculada la matriz P de transición de la base B_1 a la base B_2 solo un producto de matrices es necesario para obtener $[\mathbf{v}]_{B_2}$:

$$P\,[\mathbf{v}]_{B_1} = \begin{bmatrix} \frac{1}{9} & \frac{1}{2} \\ \frac{1}{3} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{31}{18} \\ -\frac{5}{6} \end{bmatrix}.$$

Así se ha obtenido el vector $[\mathbf{v}]_{B_2}$ de dos maneras distintas.◀

En el Ejemplo 4.73 se halló la matriz P de transición de la base B_1 a la base B_2. Se puede también preguntar por la matriz de transición de la base B_2 a la base B_1, es decir se considera a B_2 como la base inicial (o vieja) y a B_1 como la base final (o nueva). Por definición, esta matriz de transición tendrá como columnas los vectores de coordenadas de los vectores de la base inicial, B_2, con respecto a la base nueva, B_1. Teniendo en cuenta el procedimiento para hallar esta matriz se deberá aplicar eliminación de Gauss-Jordan a la matriz:

$$\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & : & \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & : & 6 & 4 \\ 0 & 2 & : & 3 & -1 \end{bmatrix},$$

obtiéndose (verificar) su forma escalonada reducida:

$$\begin{bmatrix} 1 & 0 & : & \frac{9}{4} & \frac{9}{4} \\ 0 & 1 & : & \frac{3}{2} & -\frac{1}{2} \end{bmatrix} = [I : Q]\,,$$

siendo $Q = \begin{bmatrix} \frac{9}{4} & \frac{9}{4} \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$ la matriz de transición de la base B_2 a la base B_1 .

Esta matriz es la única que satisface la siguiente propiedad:

$$[\mathbf{v}]_{B_1} = Q\,[\mathbf{v}]_{B_2}\,, \quad \forall \mathbf{v} \in \mathbb{V}.$$

Puede observarse que $PQ = \begin{bmatrix} \frac{1}{9} & \frac{1}{2} \\ \frac{1}{3} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{9}{4} & \frac{9}{4} \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$, y por lo tanto P es invertible y $Q = P^{-1}$. En general, se tiene el siguiente teorema:

Teorema 4.14. Sea $\mathbb{V}$ un espacio vectorial de dimensión n, B_1 y B_2 dos bases ordenadas de $\mathbb{V}$ y P la matriz de transición de B_1 a B_2. Entonces P es invertible y P^{-1} es la matriz de transición de B_2 a B_1.

Demostración. Sea Q la matriz de transición de la base B_2 a la base B_1. Se probará que $PQ = I$ y en consecuencia se conluirá que P es invertible y $Q = P^{-1}$.

Supóngase que $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ y que la matriz producto de P por Q, dada en términos de sus columnas, es $PQ = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & ... & \mathbf{c}_n \end{bmatrix}$. Dado que P es la matriz de transición de B_1 a B_2:

$$[\mathbf{v}]_{B_2} = P\,[\mathbf{v}]_{B_1}, \quad \forall \mathbf{v} \in \mathbb{V},$$

y como Q es la matriz de transición de B_2 a B_1 entonces

$$[\mathbf{v}]_{B_1} = Q\,[\mathbf{v}]_{B_2}, \quad \forall \mathbf{v} \in \mathbb{V}.$$

Luego,

$$[\mathbf{v}]_{B_2} = P\,[\mathbf{v}]_{B_1} = P\left(Q\,[\mathbf{v}]_{B_2}\right) = (PQ)\,[\mathbf{v}]_{B_2}, \quad \forall \mathbf{v} \in \mathbb{V},$$

y por lo tanto

$$(PQ)\,[\mathbf{v}]_{B_2} = [\mathbf{v}]_{B_2}, \quad \forall \mathbf{v} \in \mathbb{V}.$$

En particular para los vectores $\mathbf{v}_j$ de la base B_2 se tiene $(PQ)\,[\mathbf{v}_j]_{B_2} = [\mathbf{v}_j]_{B_2}, \quad \forall j = 1, 2, ..., n.$

Dado que $[\mathbf{v}_j]_{B_2} = \mathbf{e}_j$, $j = 1, 2, ..., n$ (véase Ejemplo 4.70) resulta

$$(PQ)\,\mathbf{e}_j = \mathbf{e}_j, \quad \forall j = 1, 2, ..., n.$$

y como

$$(PQ)\,\mathbf{e}_j = \mathbf{c}_j, \quad \forall j = 1, 2, ..., n,$$

entonces $\mathbf{c}_j = \mathbf{e}_j$, $\forall j = 1, 2, ..., n$. Por lo tanto, las columnas de PQ son $\mathbf{e}_1$, $\mathbf{e}_2$,..., $\mathbf{e}_n$ es decir

$$PQ = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & ... & \mathbf{e}_n \end{bmatrix} = I.$$

Así, P es invertible y $Q = P^{-1}$. $\square$

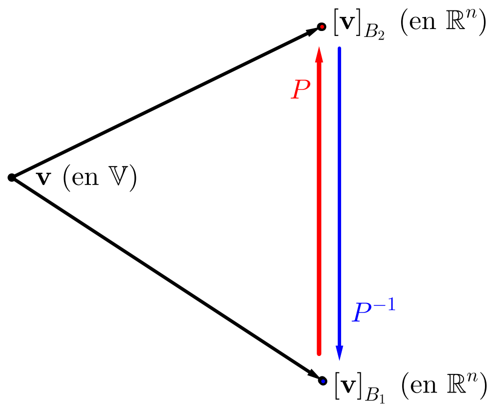


Figura 4.11: Relación entre los vectores de coordenadas en dos bases distintas

El concepto de matriz de transición está definido para cualquier espacio vectorial de dimensión finita. En los ejemplos anteriores se calcularon matrices de transición en el espacio vectorial $\mathbb{R}^n$. En el siguiente ejemplo se obtendrá una matriz de transición en el espacio vectorial $\mathbb{P}_2$:

Ejemplo 4.75. En el espacio vectorial $\mathbb{P}_2$, considere las bases ordenadas $B_1 = \{p_1, p_2, p_3\}$ donde $p_1 = 1,\ \ p_2 = 1 + x\ \ \text{y}\ \ p_3 = 1 + x + x^2$ y $B_2 = \{q_1, q_2, q_3\}$ donde $q_1 = 3 + x - x^2,\ \ q_2 = x,\ q_3 = 1 - x^2$,.

La matriz P de transición de la base B_1 a la base B_2 es

$$P = \left[\begin{array}{ccc} [p_1]_{B_2} & [p_2]_{B_2} & [p_3]_{B_2} \end{array}\right].$$

Para calcular $[p_1]_{B_2}$ se deben determinar los coeficientes a_1, a_2 y a_3 tales que

$$a_1q_1 + a_2q_2 + a_3q_3 = p_1,$$

reemplazando q_1, q_2, q_3, p_1 y operando se tiene

$$(3a_1 + a_3) + (a_1 + a_2)\,x + (-a_1 - a_3)\,x^2 = 1,$$

y por igualdad de polinomios se obtiene:

$$\left\{\begin{array}{rrrcl} 3a_1 & & +a_3 & = & 1 \\ a_1 & +a_2 & & = & 0 \\ -a_1 & & -a_3 & = & 0 \end{array}\right.$$

Procediendo de forma similar, para calcular $[p_2]_{B_2}$ deben hallarse los coeficientes b_1, b_2 y b_3 tales que

$$\left\{\begin{array}{rrrcl} 3b_1 & & +b_3 & = & 1 \\ b_1 & +b_2 & & = & 1 \\ -b_1 & & -b_3 & = & 0 \end{array}\right.$$

y para calcular $[p_3]_{B_2}$ deben hallarse los coeficientes c_1, c_2 y c_3 tales que

$$\left\{\begin{array}{rrrcl} 3c_1 & & c_3 & = & 1 \\ c_1 & +c_2 & & = & 1 \\ -c_1 & & -c_3 & = & 1 \end{array}\right.$$

Hay que resolver entonces tres sistemas lineales que tienen la misma matriz de coeficientes y por lo tanto conviene hacerlo en forma simultánea. Para ello, se aplica Gauss-Jordan a la matriz aumentada

$$\left[\begin{array}{rrrcrrr} 3 & 0 & 1 & : & 1 & 1 & 1 \\ 1 & 1 & 0 & : & 0 & 1 & 1 \\ -1 & 0 & -1 & : & 0 & 0 & 1 \end{array}\right]$$

obteniéndose su forma escalonada reducida:

$$\left[\begin{array}{rrrcrrr} 1 & 0 & 0 & : & \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 1 & 0 & : & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & : & -\frac{1}{2} & -\frac{1}{2} & -2 \end{array}\right].$$

La solución del primer sistema es $a_1 = \frac{1}{2}$, $a_2 = -\frac{1}{2}$ y $a_3 = -\frac{1}{2}$ y por lo tanto

$$[p_1]_{B_2} = \left[\begin{array}{r} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{array}\right].$$

La solución del segundo sistema es $b_1 = \frac{1}{2}$, $b_2 = \frac{1}{2}$ y $b_3 = -\frac{1}{2}$ y así

$$[p_2]_{B_2} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}.$$

Y la solución del último sistema es $c_1 = 1$, $c_2 = 0$ y $c_3 = 2$ por lo que

$$[p_3]_{B_2} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}.$$

Luego, la matriz de transición P de la base B_1 a la base B_2 es

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 1 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & -2 \end{bmatrix}. \blacktriangleleft$$

Observación 4.2. En el ejemplo, la matriz P de transición de la base B_1 a la base B_2 se obtuvo a partir de la matriz

$$\begin{bmatrix} 3 & 0 & 1 & : & 1 & 1 & 1 \\ 1 & 1 & 0 & : & 0 & 1 & 1 \\ -1 & 0 & -1 & : & 0 & 0 & 1 \end{bmatrix}.$$

Puede observarse que la matriz de 3×3 de la izquierda tiene como columnas los vectores de coordenadas de los polinomios de la base nueva, $B_2 = \{q_1, q_2, q_3\}$, con respecto a la base estándar $E = \left\{1, x, x^2\right\}$ de $\mathbb{P}_2$ y la matriz de 3×3 de la derecha tiene como columnas los vectores de coordenadas de los polinomios de la base vieja, $B_1 = \{p_1, p_2, p_3\}$, también con respecto a la base estándar E. Esto es:

$$\begin{bmatrix} 3 & 0 & 1 & : & 1 & 1 & 1 \\ 1 & 1 & 0 & : & 0 & 1 & 1 \\ -1 & 0 & -1 & : & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} [q_1]_E & [q_2]_E & [q_3]_E & : & [p_1]_E & [p_2]_E & [p_3]_E \end{bmatrix}.$$

La forma escalonada reducida es

$$\begin{bmatrix} 1 & 0 & 0 & : & \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 1 & 0 & : & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & : & -\frac{1}{2} & -\frac{1}{2} & -2 \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 : & [p_1]_{B_2} & [p_2]_{B_2} & [p_3]_{B_2} \end{bmatrix}$$

y la matriz de transición P de la base B_1 a la base B_2 es la matriz de 3×3 de la derecha.

Este procedimiento puede generalizarse como método para obtener matrices de transición, por ejemplo, en los espacios vectoriales $\mathbb{M}_{mn}$ y $\mathbb{P}_n$ teniendo cuenta sus bases estándar (véase Ejercicio 4.6).

El siguiente teorema nos permitirá aplicar resultados de los vectores de $\mathbb{R}^n$ a los vectores de cualquier espacio vectorial de dimensión n.

Teorema 4.15. Sea B una base de un espacio vectorial $\mathbb{V}$ de dimensión n y sean $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_k$ vectores de $\mathbb{V}$, entonces:

1. El vector de coordenadas de una combinación lineal es la misma combinación lineal de los vectores de coordenadas:

$$[c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_k\mathbf{v}_k]_B = c_1 [\mathbf{v}_1]_B + c_2 [\mathbf{v}_2]_B + ... + c_k [\mathbf{v}_k]_B$$

   para todo c_1, c_2,..., c_k en $\mathbb{R}$.

2. $\mathbf{v}$ es un vector del espacio generado por $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ si y solo si $[\mathbf{v}]_B$ está en el espacio generado por $\{[\mathbf{v}_1]_B, [\mathbf{v}_2]_B, ..., [\mathbf{v}_k]_B\}$.

3. El conjunto $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ es linealmente independiente (en $\mathbb{V}$) si y solo si $\{[\mathbf{v}_1]_B, [\mathbf{v}_2]_B, ..., [\mathbf{v}_k]_B\}$ es linealmente independiente (en $\mathbb{R}^n$).

Demostración. Al final de la sección se encuentra una prueba de la parte 1. Las demás se han dejado como ejercicio. □

El resultado siguiente es una consecuencia inmediata del teorema anterior:

Corolario 4.3. Sean $\mathbb{V}$ un espacio vectorial de dimensión n y B una base de $\mathbb{V}$. Entonces el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es una base de $\mathbb{V}$ si y solo si $\{[\mathbf{v}_1]_B, [\mathbf{v}_2]_B, ..., [\mathbf{v}_n]_B\}$ es una base de $\mathbb{R}^n$.

Ejemplo 4.76. Para demostrar que el conjunto $S = \{A_1, A_2, A_3, A_4\}$ donde

$$A_1 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},\ A_2 = \begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix},\ A_3 = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} \text{ y } A_4 = \begin{bmatrix} 3 & 0 \\ -1 & -2 \end{bmatrix},$$

es una base de $\mathbb{M}_2$, de acuerdo al corolario anterior, basta con demostrar que los vectores de coordenadas de las matrices de S respecto a una base cualquiera de $\mathbb{M}_2$ forman una base de $\mathbb{R}^4$.

Si en particular se considera la base estándar $E = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ de $\mathbb{M}_2$, entonces el problema consiste en determinar si el conjunto

$$S' = \{[A_1]_E, [A_2]_E, [A_3]_E, [A_4]_E\},$$

es una base de $\mathbb{R}^4$. Por considerarse la base estándar, los vectores de coordenadas con respecto a dicha base se obtienen directamente:

$$[A_1]_E = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix},\ [A_2]_E = \begin{bmatrix} 0 \\ 2 \\ 2 \\ 1 \end{bmatrix},\ [A_3]_E = \begin{bmatrix} 2 \\ 1 \\ 4 \\ 2 \end{bmatrix} \text{ y } [A_4]_E = \begin{bmatrix} 3 \\ 0 \\ -1 \\ -2 \end{bmatrix}.$$

Fácilmente se comprueba que S' es una base de $\mathbb{R}^4$ (verifíquelo) y por lo tanto S es una base de $\mathbb{M}_2$.◄

Unicidad de la matriz P de transición

Demostración. Sean $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ y $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ dos bases ordenadas de un espacio vectorial $\mathbb{V}$ y la matriz de transición de B_1 a B_2, $P = \left[\ [\mathbf{u}_1]_{B_2} \quad [\mathbf{u}_2]_{B_2} \quad ... \quad [\mathbf{u}_n]_{B_2}\ \right]$.

Se probará que P es la única matriz que verifica:

$$[\mathbf{v}]_{B_2} = P\,[\mathbf{v}]_{B_1}\,, \quad \forall \mathbf{v} \in \mathbb{V}.$$

Para ello, supóngase que A es una matriz de $n \times n$ tal que

$$[\mathbf{v}]_{B_2} = A\,[\mathbf{v}]_{B_1}\,,$$

para todo $\mathbf{v} \in \mathbb{V}$. En particular esta última igualdad debe verificarse para los vectores $\mathbf{u}_j$ de la base B_1. Es decir,

$$A\,[\mathbf{u}_j]_{B_1} = [\mathbf{u}_j]_{B_2}\,, \quad \forall j = 1, 2, ..., n.$$

De acuerdo al Ejemplo 4.70 , $[\mathbf{u}_j]_{B_1} = \mathbf{e}_j$, luego

$$A\mathbf{e}_j = [\mathbf{u}_j]_{B_2}\,, \quad \forall j = 1, 2, ..., n.$$

Por otra parte, si $\mathbf{a}_1$, $\mathbf{a}_2$,..., $\mathbf{a}_n$ son las sucesivas columnas de la matriz A, entonces

$$A\mathbf{e}_j = \mathbf{a}_j, \quad \forall j = 1, 2, ..., n.$$

Así,

$$\mathbf{a}_j = [\mathbf{u}_j]_{B_2}\,, \quad \forall j = 1, 2, ..., n.$$

Luego, $A = \left[\ [\mathbf{u}_1]_{B_2} \quad [\mathbf{u}_2]_{B_2} \quad ... \quad [\mathbf{u}_n]_{B_2}\ \right]$ y por lo tanto la matriz A es la matriz P de transición de la base B_1 a la base B_2. □

Teorema 4.15 1. Sea B una base de un espacio vectorial $\mathbb{V}$ finito dimensional y sean $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_k$ vectores de $\mathbb{V}$ Entonces: $[c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_k\mathbf{v}_k]_B = c_1\,[\mathbf{v}_1]_B + c_2\,[\mathbf{v}_2]_B + ... + c_k\,[\mathbf{v}_k]_B$ **para todo c_1, c_2,..., c_k en $\mathbb{R}$**

Demostración. Sea el vector

$$\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_k\mathbf{v}_k \tag{4.10}$$

Supongamos que los vectores de coordenadas de los vectores $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_k$ con respecto a la base B son los siguientes vectores de $\mathbb{R}^n$:

$$[\mathbf{v}_1]_B = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad [\mathbf{v}_2]_B = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, ..., \quad [\mathbf{v}_k]_B = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{bmatrix},$$

por lo cual

$$\begin{array}{c} \mathbf{v}_1 = a_{11}\mathbf{u}_1 + a_{21}\mathbf{u}_2 + ... + a_{n1}\mathbf{u}_n \\ \mathbf{v}_2 = a_{12}\mathbf{u}_1 + a_{22}\mathbf{u}_2 + ... + a_{n2}\mathbf{u}_n \\ \vdots \\ \mathbf{v}_k = a_{1k}\mathbf{u}_1 + a_{2k}\mathbf{u}_2 + ... + a_{nk}\mathbf{u}_n \end{array} \tag{4.11}$$

Reemplazando las expresiones (4.11) en (4.10) y operando (verificar), se obtiene el vector $\mathbf{w}$ expresado como combinación lineal de los vectores de la base B:

$$\mathbf{w} = (c_1a_{11} + c_2a_{12} + ... + c_ka_{1k})\,\mathbf{u}_1 + (c_1a_{21} + c_2a_{22} + ... + c_ka_{2k})\,\mathbf{u}_2 + ... + (c_1a_{n1} + c_2a_{n2} + ... + c_ka_{nk})\,\mathbf{u}_n.$$

Los coeficientes que acompañan a los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_n$ de la expresión anterior son las coordenadas del vector $\mathbf{w}$ con respecto a la base B. Así:

$$\begin{aligned}
[c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_k\mathbf{v}_k]_B &= \begin{bmatrix} c_1a_{11} + c_2a_{12} + ... + c_ka_{1k} \\ c_1a_{21} + c_2a_{22} + ... + c_ka_{2k} \\ \vdots \\ c_1a_{n1} + c_2a_{n2} + ... + c_ka_{nk} \end{bmatrix} \\
&= c_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + c_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + ... + c_k \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{bmatrix} \\
&= c_1\,[\mathbf{v}_1]_B + c_2\,[\mathbf{v}_2]_B + ... + c_k\,[\mathbf{v}_k]_B\,.
\end{aligned}$$

□

Ejercicios 4.6

1. En el espacio $\mathbb{R}^2$, sea $\mathbf{u} = (0, 4)$. Calcule el vector de coordenadas $[\mathbf{u}]_B$ siendo $B = \{(-1, 3), (2, -2)\}$.

2. En cada caso, calcule el vector de coordendas, $[p]_B$, con respecto a la base B en el espacio $\mathbb{P}_n$ indicado:

   a) $B = \left\{1 + 2x, -x^2, x + x^2\right\}$ base de $\mathbb{P}_2$, $p = -1 + 3x + 2x^2$

   b) $B = \{1 + 2x, 2 - 3x\}$ base de $\mathbb{P}_1$, $p = -1 - 4x$

   c) $B = \left\{1, 1 + x, 1 + x + x^2, 1 + x + x^2 + x^3\right\}$ base de $\mathbb{P}_3$, $p = -2 + x - x^2 + 4x^3$

3. Obtenga, a partir de la base B del espacio $\mathbb{R}^3$ y el vector de coordendas $[\mathbf{x}]_B$, el vector $\mathbf{x}$.

   a) $B = \{\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3\}$, $[\mathbf{x}]_B = \begin{bmatrix} 2 \\ 1 \\ -5 \end{bmatrix}$

   b) $B = \{\mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3\}$, $[\mathbf{x}]_B = \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}$

   c) $B = \{(-1, 2, 0), (0, 1, 1), (3, -1, -2)\}$, $[\mathbf{x}]_B = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$

4. Halle el polinomio p en cada caso, siendo B una base del espacio $\mathbb{P}_n$ y $[p]_B$ su vector de coordenadas.

a) $B = \{1 + 2x, -2 + x\}$ una base de $\mathbb{P}_1$, $[p]_B = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$

b) $B = \{x, x^2, 1\}$ una base de $\mathbb{P}_2$, $[p]_B = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$

$\boxed{c)}$ $B = \{1 + x + x^3, x - x^2, 2 - 3x + x^2, x + x^2 - x^3\}$ una base de $\mathbb{P}_3$, $[p]_B = \begin{bmatrix} 0 \\ 2 \\ 3 \\ -4 \end{bmatrix}$

$\boxed{5.}$ Sea $B = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix} \right\}$ una base del espacio $\mathbb{M}_{22}$.

a) Obtenga la matriz A si $(A)_B = (4, 2, 6, -3)$.

b) Halle el vector de coordenadas de la matriz $\begin{bmatrix} -1 & 3 \\ 1 & 2 \end{bmatrix}$ con respecto a la base B.

6. Sea el subespacio finito dimensional $\mathbb{W} = \{ae^x + be^{-x} : a, b \in \mathbb{R}\}$ de $\mathbb{F}(\mathbb{R})$.

a) Obtenga una base para $\mathbb{W}$ y denomínela S.

b) Calcule el vector de coordenadas $[f]_S$ siendo $f(x) = \cosh x$, $x \in \mathbb{R}$.

c) Halle $g(x)$ si $[g]_S = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$

7. Demuestre el recíproco del Teorema 4.12: *Sea $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ un subconjunto de un espacio vectorial $\mathbb{V}$. Si todo vector $\mathbf{v}$ de $\mathbb{V}$ puede escribirse de manera única como combinación lineal de los vectores de S entonces S es una base de $\mathbb{V}$.*

$\boxed{8.}$ Halle la matriz de transición de la base B_1 a la base B_2 del espacio $\mathbb{V}$ que se indica en cada caso:

a) $\mathbb{V} = \mathbb{R}^2$; bases $B_1 = \{\mathbf{u}_1, \mathbf{u}_2\}$ y $B_2 = \{\mathbf{v}_1, \mathbf{v}_2\}$ donde $\mathbf{u}_1 = (-2, 3)$, $\mathbf{u}_2 = (1, -1)$, $\mathbf{v}_1 = (0, -2)$, $\mathbf{v}_2 = (3, 2)$.

b) $\mathbb{V} = \mathbb{R}^3$; bases $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ y $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ donde $\mathbf{u}_1 = \mathbf{e}_2$, $\mathbf{u}_2 = \mathbf{e}_3$, $\mathbf{u}_3 = \mathbf{e}_1$, $\mathbf{v}_1 = \mathbf{e}_1$, $\mathbf{v}_2 = \mathbf{e}_2$, $\mathbf{v}_3 = \mathbf{e}_3$.

c) $\mathbb{V} = \mathbb{P}_1$; bases $B_1 = \{p_1, p_2\}$ y $B_2 = \{q_1, q_2\}$ donde $p_1 = 1$, $p_2 = 2 + x$, $q_1 = 2 - x$, $q_2 = 3 + 4x$.

d) $\mathbb{V} = \mathbb{M}_{22}$; bases $B_1 = \{A_1, A_2, A_3, A_4\}$ y $B_2 = \{C_1, C_2, C_3, C_4\}$ donde

$$\begin{aligned} A_1 &= E_{11} & C_1 &= E_{11} \\ A_2 &= E_{11} + E_{12} & C_2 &= E_{12} \\ A_3 &= E_{11} + E_{12} + E_{21} & C_3 &= E_{21} \\ A_4 &= E_{11} + E_{12} + E_{21} + E_{22} & C_4 &= E_{22} \end{aligned}$$

$\boxed{9.}$ Considere las bases de $\mathbb{R}^3$, $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ y $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, donde $\mathbf{u}_1 = (1, 1, 1)$, $\mathbf{u}_2 = (-1, 3, 0)$, $\mathbf{u}_3 = (2, 0, 3)$, $\mathbf{v}_1 = (0, 0, 2)$, $\mathbf{v}_2 = (-1, 0, 0)$, $\mathbf{v}_3 = (2, -1, 5)$.

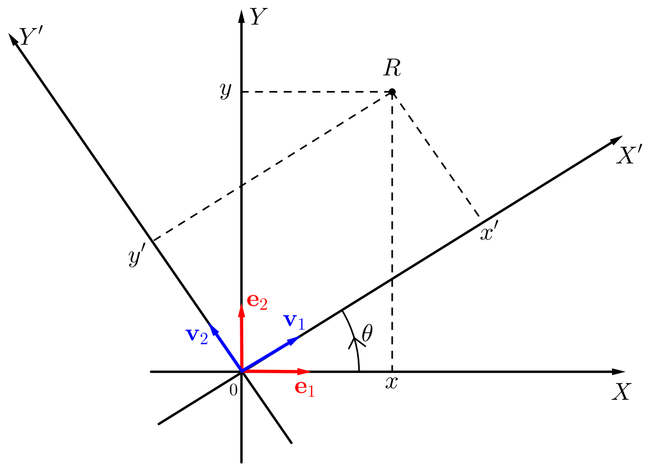
a) Encuentre la matriz de transición, P, de la base B_1 a la base B_2.

b) Halle la matriz de transición de la base B_2 a la base B_1 de dos maneras distintas:

i) utilizando la definición de matriz de transición.

ii) calculando la inversa de la matriz P.

10. **(Rotación de los ejes coordenados)** En muchos problemas se da un sistema de coordenadas rectangulares XY y se obtiene un nuevo sistema de coordenadas $X'Y'$ al hacer girar el sistema XY en sentido contrario a las agujas del reloj alrededor del origen, un ángulo θ. Entonces cada punto R en el plano tiene dos conjuntos de coordenadas, las coordenadas (x, y) con respecto al sistema XY y las coordenadas (x', y') relativas al sistema $X'Y'$. Al introducir los vectores unitarios (longitud 1) $\mathbf{e}_1$ y $\mathbf{e}_2$ a lo largo de los ejes X y Y positivos, y los vectores unitarios $\mathbf{v}_1$, $\mathbf{v}_2$ a lo largo de los ejes X' y Y' positivos, esta rotación se puede considerar como un cambio de base, de la base $E = \{\mathbf{e}_1, \mathbf{e}_2\}$ a la base $B = \{\mathbf{v}_1, \mathbf{v}_2\}$.

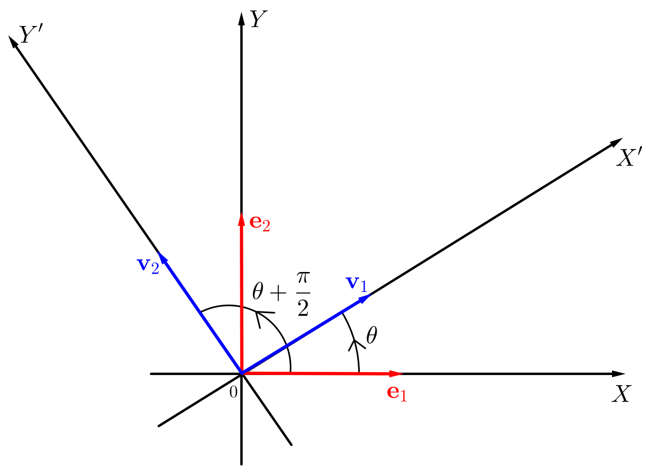


Por tanto las nuevas coordenadas de R cumplen la relación

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = P \begin{bmatrix} x \\ y \end{bmatrix},$$

donde P es la matriz de cambio de base de E a B.

De la figura siguiente puede observarse que:



$$\begin{aligned}\mathbf{v}_1 &= \cos\theta\ \mathbf{e}_1 + \operatorname{sen}\theta\ \mathbf{e}_2,\\ \mathbf{v}_2 &= \cos\left(\theta+\frac{\pi}{2}\right)\mathbf{e}_1 + \operatorname{sen}\left(\theta+\frac{\pi}{2}\right)\mathbf{e}_2 = -\operatorname{sen}\theta\ \mathbf{e}_1 + \cos\theta\ \mathbf{e}_2.\end{aligned}$$

Luego,

$$[\mathbf{v}_1]_E = \begin{bmatrix}\cos\theta\\ \operatorname{sen}\theta\end{bmatrix} \quad \text{y} \quad [\mathbf{v}_2]_E = \begin{bmatrix}-\operatorname{sen}\theta\\ \cos\theta\end{bmatrix}$$

Así, la matriz Q de transición de B a E es

$$Q = \begin{bmatrix}[\mathbf{v}_1]_E & [\mathbf{v}_2]_E\end{bmatrix} = \begin{bmatrix}\cos\theta & -\operatorname{sen}\theta\\ \operatorname{sen}\theta & \cos\theta\end{bmatrix}$$

y la matriz P de transición de E a B es

$$P = Q^{-1} = \begin{bmatrix}\cos\theta & \operatorname{sen}\theta\\ -\operatorname{sen}\theta & \cos\theta\end{bmatrix} = \begin{bmatrix}[\mathbf{e}_1]_B & [\mathbf{e}_2]_B\end{bmatrix}.$$

Por lo tanto, se verifican las siguientes relaciones entre los vectores de coordenadas:

$$\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}\cos\theta & -\operatorname{sen}\theta\\ \operatorname{sen}\theta & \cos\theta\end{bmatrix}\begin{bmatrix}x'\\ y'\end{bmatrix}$$

y

$$\begin{bmatrix}x'\\ y'\end{bmatrix} = \begin{bmatrix}\cos\theta & \operatorname{sen}\theta\\ -\operatorname{sen}\theta & \cos\theta\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}.$$

Considere ahora una rotación de ángulo $\theta = \frac{\pi}{4}$.

a) Si las coordenadas del punto R en el sistema XY son $x = 2$, $y = 4$, obtenga las coordenadas de dicho punto en el sistema final $X'Y'$ rotado.

b) Si las coordenadas del punto S en el sistema $X'Y'$ rotado son $x' = 0$, $y' = \frac{3\sqrt{2}}{2}$, halle las coordenadas x, y de dicho punto en el sistema inicial XY.

$\boxed{11.}$ En el espacio $\mathbb{P}_3$, considere la base estándar E y la base $B = \left\{1, 2x, -2+4x^2, -12x+8x^3\right\}$ (polinomios de Hermite).

a) Obtenga el vector de coordenadas del polinomio $q = -1 + 2x + x^2 - x^3$ con respecto a la base B de manera directa.

b) Halle la matriz de transición, P, de la base E a la base B.

c) Encuentre el vector de coordenadas del polinomio q con respecto a la base B usando la matriz obtenida en el ítem anterior.

d) Verifique que $P^{-1}[q]_B = [q]_E$.

$\boxed{12.}$ **(Caso particular de la parte 1 del Teorema 4.15)** Considere el espacio vectorial $\mathbb{P}_2$ y la base $B = \left\{1, 1+x, 1+x+x^2\right\}$. Para los polinomios $p_1 = 2x + x^2$ y $p_2 = 5 - 3x$, verifique:

a) $[p_1 + p_2]_B = [p_1]_B + [p_2]_B$

b) $[cp_1]_B = c\,[p_1]_B$, para todo $c \in \mathbb{R}$

13. Pruebe los incisos 2 y 3 del Teorema 4.15.

$\boxed{14.}$ Usando el Teorema 4.15 o su corolario pruebe que:

a) Las matrices $\begin{bmatrix} -1 & 3 \\ 1 & 2 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ -1 & 5 \end{bmatrix}$ y $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ forman un conjunto linealmente independiente de $\mathbb{M}_{22}$.

b) gen $\left\{1 + x + x^2, -x - 4x^2, 1 + x + 2x^2, -1 + x + x^2\right\} = \mathbb{P}_2$

c) El conjunto $B = \left\{1 + x - x^2 + x^3, 2 - 3x + 4x^2, 1 - x^2 + x^3, 2 + 3x + 5x^3\right\}$ es una base del espacio $\mathbb{P}_3$.

15. **(Scilab)** En el espacio vectorial $\mathbb{M}_2$ considere las siguientes bases:

$$\begin{aligned} B_1 &= \left\{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right\}, \\ B_2 &= \left\{\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -2 & 0 \\ 4 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -12 \\ 0 & 8 \end{bmatrix}\right\}. \end{aligned}$$

Teniendo en cuenta la Observación 4.2, para hallar la matriz de transición de la base B_1 a la base B_2 se lleva la matriz

$$\begin{bmatrix} 1 & 0 & -2 & 0 & : & 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & -12 & : & 1 & 1 & 1 & 0 \\ 0 & 0 & 4 & 0 & : & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 8 & : & 1 & 0 & 0 & 0 \end{bmatrix}$$

a su forma escalonada reducida. Las primera cuatro columnas de esta matriz son los vectores de coordenadas con respecto a la base estándar de $\mathbb{M}_2$, de las matrices de la base B_2 y las cuatro últimas columnas son los vectores de coordenadas también con respecto a la base estándar de $\mathbb{M}_2$, de las matrices de la base B_1.

Aplicando Gauss-Jordan se obtiene:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & : & 3/2 & 3/2 & 1 & 1 \\ 0 & 1 & 0 & 0 & : & 1/2 & -1/4 & 0 & -1/2 \\ 0 & 0 & 1 & 0 & : & 1/4 & 1/4 & 0 & 0 \\ 0 & 0 & 0 & 1 & : & 1/8 & 0 & 0 & 0 \end{bmatrix}.$$

La matriz 4×4 de la derecha es la matriz de transición de B_1 a B_2

a) Usando el comando “rref”, verifique el resultado anterior.

b) Halle la matriz de transición de la base B_2 a la base B_1.

16. **(Scilab)** En el espacio vectorial $\mathbb{P}_4$, considere las bases ordenadas:

$$B_1 = \left\{1 + x^2 + x^4, 1 + x - x^2 + x^3 + x^4, 2 + 2x - 2x^3 - 2x^4, 3 + 4x + 3x^3 + 4x^4, -1 + x + x^2 + 5x^3\right\},$$

$$B_2 = \left\{1 + x - x^4, -1 + 2x + 2x^2 + 2x^4, 2x - x^2 + 2x^3 + 2x^4, 1 + x + x^2 + x^3, 1 + x^2 + 4x^3\right\}.$$

Obtenga las dos matrices de transición (de la base B_1 a la base B_2 y de la base B_2 a la base B_1).

4.7. Rango y nulidad de una matriz

En esta sección se estudiarán tres espacios vectoriales importantes asociados a una matriz: el *espacio de renglones*, el *espacio de columnas* y el *espacio nulo*.

Definición (Espacios de renglones y columnas). Sea A una matriz de $m \times n$,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Los vectores

$$\mathbf{r}_1 = (a_{11}, a_{12}, ..., a_{1n}), \quad \mathbf{r}_2 = (a_{21}, a_{22}, ..., a_{2n}), \ \dots\ , \ \mathbf{r}_m = (a_{m1}, a_{m2}, ..., a_{mn}),$$

formados a partir de los renglones de A se conocen como *vectores renglón* de A, mientras que los vectores

$$\mathbf{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, ..., \quad \mathbf{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix},$$

son los vectores columna de A.

El *espacio de renglones* de A es el subespacio de $\mathbb{R}^n$ generado por los vectores renglón de A. Se simboliza $\text{ren}(A)$.

El *espacio de columnas* de A es el subespacio de $\mathbb{R}^m$ generado por los vectores columna de A. Se simboliza $\text{col}(A)$.

Ejemplo 4.77. Para la matriz $A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 0 & 3 \end{bmatrix}$, los espacios de columnas y de renglones de A son:

$$\text{col}(A) = \text{gen}\left\{\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}\right\},$$

y

$$\text{ren}(A) = \text{gen}\left\{(1, 2, 1), (3, 0, 3)\right\},$$

respectivamente.◄

El objetivo de esta sección es proveer un método práctico para encontrar bases para el espacio generado por un conjunto de vectores de $\mathbb{R}^n$. Para ello, se tratará primero el problema de la obtención de bases para los espacios de renglones y de columnas de una matriz.

Una base para el espacio de los renglones de una matriz

El lema siguiente es clave para la determinación de una base para el espacio de renglones de una matriz.

Lema 4.2. Las operaciones elementales de renglones no cambian el espacio de renglones de una matriz: si A y B son matrices equivalentes por renglones entonces $\text{ren}(A) = \text{ren}(B)$.

Demostración. Para probar que $\text{ren}(A) = \text{ren}(B)$, se mostrará que todo vector en $\text{ren}(A)$ es un vector de $\text{ren}(B)$ y recíprocamente.

Dado que la matriz B se obtiene de la matriz A a través de una sucesión de operaciones elementales de renglón (suma de vectores y multiplicación por escalares no cero), todo renglón de B es una combinación lineal de los renglones de A. Como todo vector en $\text{ren}(B)$ es una combinación lineal de renglones de B y éstos a su vez son combinaciones lineales de renglones de A, entonces todo vector en $\text{ren}(B)$ es una combinación lineal de renglones de A. Esto prueba que todo vector en $\text{ren}(B)$ es un vector en $\text{ren}(A)$.

Recíprocamente, invirtiendo las operaciones elementales que llevaron la matriz A a la matriz B se obtiene A a partir de B, y por lo tanto todo vector de $\text{ren}(A)$ es una combinación lineal de los renglones de B, es decir, todo vector en $\text{ren}(A)$ es un vector en $\text{ren}(B)$.

□

Teorema 4.16. Los renglones no cero de una matriz en forma escalonada por renglones forman una base para su espacio de renglones.

Una demostración de este teorema se encuentra al final de esta sección.

Teorema 4.17. Los renglones no cero de cualquier forma escalonada por renglones de una matriz A forman una base para el espacio $\text{ren}(A)$.

Demostración. Sea B una forma escalonada de A. Por el teorema anterior, los renglones no cero de B forman una base para su espacio de renglones. Dado que $\text{ren}(A) = \text{ren}(B)$ (Lema 4.2) entonces los renglones no cero de B forman una base para el espacio $\text{ren}(A)$. □

Es inmediato el siguiente resultado:

Corolario 4.4. La dimensión del espacio de renglones de una matriz A es el número de renglones no cero de una forma escalonada por renglones de A.

Estos resultados pueden ser aplicados para facilitar la búsqueda de una base para el espacio generado por un conjunto de vectores de $\mathbb{R}^n$.

Ejemplo 4.78. Supóngase que se quiere encontrar una base para el espacio generado por los siguientes vectores de $\mathbb{R}^5$: $\mathbf{v}_1 = (1, -2, 0, 0, 3)$, $\mathbf{v}_2 = (2, -5, -3, -2, 6)$, $\mathbf{v}_3 = (0, 5, 15, 10, 0)$ y $\mathbf{v}_4 = (2, 6, 18, 8, 6)$.

El espacio generado por el conjunto $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$, es el espacio de renglones de la matriz

$$A = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}.$$

Al llevar esta matriz a una forma escalonada se obtiene (verifíquelo):

$$B = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Los vectores renglón diferentes de cero en esta matriz son: $\mathbf{u}_1 = (1, -2, 0, 0, 3)$, $\mathbf{u}_2 = (0, 1, 3, 2, 0)$ y $\mathbf{u}_3 = (0, 0, 1, 1, 0)$. Estos vectores forman una base para ren(A) = gen(S). Obsérvese que los vectores $\mathbf{u}_2$ y $\mathbf{u}_3$ no son vectores de S.◀

El ejemplo anterior ofrece un método para encontrar una base del espacio generado por un conjunto finito de vectores de $\mathbb{R}^n$:

Cálculo de una base para gen(S)

Sea $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m\}$ un subconjunto de $\mathbb{R}^n$. Entonces se puede determinar una base para gen(S) como sigue:

1. Forme la matriz A de $m \times n$ cuyos renglones son $\mathbf{u}_1$, $\mathbf{u}_2$,...,$\mathbf{u}_m$.
2. Reduzca A hasta una forma escalonada B.
3. Una base para gen(S) es el conjunto de los vectores renglón no cero de B.

Observación 4.3. Este método produce una base de gen(S) formada por vectores que, en general, no son del conjunto S. La ventaja es que este método produce bases “simples” ya que se obtienen vectores con varias componentes cero.

Una base para el espacio de las columnas de una matriz

Es evidente que, excepto por un cambio de notación vertical a horizontal, el espacio de columnas de una matriz A es el mismo que el espacio de renglones de A^t. Por lo tanto, se puede encontrar una base para col(A) hallando una base para ren(A^t) y luego “volviendo” a la notación vertical.

Ejemplo 4.79. Se quiere hallar una base del espacio de columnas de

$$A = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 3 & 2 & 5 & 1 \\ 0 & 4 & 4 & -4 \end{bmatrix}.$$

Al transponerla se obtiene

$$A^t = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 2 & 4 \\ 1 & 5 & 4 \\ 1 & 1 & -4 \end{bmatrix},$$

y llevando A^t a una forma escalonada se llega a (verifíquelo):

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Por lo tanto, los vectores $(1, 3, 0)$ y $(0, 1, 2)$ forman una base para $\text{ren}(A^t)$ o, lo que es equivalente, los vectores

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix},$$

forman una base para $\text{col}(A)$. Nótese que el vector $\mathbf{u}_2$ de la base obtenida no es una columna de A es decir se obtuvo una base de $\text{col}(A)$ que no está compuesta totalmente por columnas de dicha matriz.◀

En el ejemplo anterior se usó el hecho de que $\text{col}(A) = \text{ren}(A^t)$ y se aplicó la técnica del Ejemplo 4.78 a la matriz A^t para hallar una base de $\text{col}(A)$.

A continuación se darán una serie de resultados que brindarán una forma alternativa de hallar una base para el espacio de columnas de una matriz A que esté formada exclusivamente por columnas de dicha matriz. Para ello se verá primero el siguiente ejemplo:

Ejemplo 4.80. Se aplicará una sucesión de operaciones elementales a una matriz A hasta obtener una forma escalonada (eliminación gaussiana). Se mostrará que las columnas de todas las matrices del proceso, satisfacen las mismas relaciones de dependencia lineal.

$$A = \begin{bmatrix} 1 & -2 & 2 & 3 & 0 \\ -1 & 2 & -1 & -2 & 1 \\ 3 & -6 & 5 & 8 & -1 \end{bmatrix} \to B = \begin{bmatrix} 1 & -2 & 2 & 3 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix} \to C = \begin{bmatrix} \boxed{1} & -2 & 2 & 3 & 0 \\ 0 & 0 & \boxed{1} & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

En la matriz escalonada C se vé rápidamente, por la forma en que están dispuestos los ceros, que las columnas pivote $\mathbf{c}_1$ y $\mathbf{c}_3$ son linealmente independientes. Además se puede observar que las otras columnas son combinaciones lineales de éstas dos: $\mathbf{c}_2 = -2\mathbf{c}_1$, $\mathbf{c}_4 = \mathbf{c}_1 + \mathbf{c}_3$, $\mathbf{c}_5 = -2\mathbf{c}_1 + \mathbf{c}_3$. Por lo tanto, reduciendo el conjunto generador, $\text{col}(C) = \text{gen}\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{c}_5\} = \text{gen}\{\mathbf{c}_1, \mathbf{c}_3\}$. Así el conjunto de las columnas pivote de C, $\{\mathbf{c}_1, \mathbf{c}_3\}$ es linealmente independiente y generador de $\text{col}(C)$, y por lo tanto es una base de $\text{col}(C)$.

Puede comprobarse también para la matriz B, que las correspondientes columnas verifican las mismas relaciones: $\mathbf{b}_1$ y $\mathbf{b}_3$ son linealmente independientes y las otras columnas verfican también: $\mathbf{b}_2 = -2\mathbf{b}_1$, $\mathbf{b}_4 = \mathbf{b}_1 + \mathbf{b}_3$, $\mathbf{b}_5 = -2\mathbf{b}_1 + \mathbf{b}_3$. Así, $\{\mathbf{b}_1, \mathbf{b}_3\}$ es una base de $\text{col}(B)$.

Es más difícil ver las relaciones de dependencia lineal entre las columnas de la matriz A pero puede verificarse que las columnas pivote, $\mathbf{a}_1$ y $\mathbf{a}_3$, son linealmente independientes y las otras columnas verifican también las relaciones: $\mathbf{a}_2 = -2\mathbf{a}_1$, $\mathbf{a}_4 = \mathbf{a}_1 + \mathbf{a}_3$, $\mathbf{a}_5 = -2\mathbf{a}_1 + \mathbf{a}_3$. Por lo tanto, el conjunto formado por las columnas pivote de A, $\{\mathbf{a}_1, \mathbf{a}_3\}$, es una base de $\text{col}(A)$. ◀

Observación 4.4. Debe advertirse que el espacio generado por las columnas de una matriz varía, en general, por operaciones elementales de renglón. Nótese del ejemplo anterior, que las columnas pivote de la matriz C forman una base para $\text{col}(C)$ pero no forman una base para el espacio $\text{col}(A)$ ($\text{col}(A) \neq \text{col}(C)$).

Lo mostrado en el ejemplo no es casualidad. Los siguientes teoremas conducen al resultado principal: *las columnas pivote de una matriz forman una base para su espacio de columnas.*

Lema 4.3. Si las matrices A y B son equivalentes por renglones, las columnas de A y B verifican las mismas relaciones de dependencia lineal. Esto es, si $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & ... & \mathbf{a}_n \end{bmatrix}$ y $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & ... & \mathbf{b}_n \end{bmatrix}$,

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + ... + c_n\mathbf{a}_n = \mathbf{0} \;\Leftrightarrow\; c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + ... + c_n\mathbf{b}_n = \mathbf{0}.$$

Demostración. Dado que una columna cero no se modifica por una operación elemental de renglón, la misma sucesión de operaciones elementales que transforman la matriz A en la matriz B, transforman la matriz aumentada $[A : \mathbf{0}]$ en la matriz $[B : \mathbf{0}]$. Luego, estas dos matrices aumentadas son equivalentes por renglones y en consecuencia los sistemas lineales homogéneos $A\mathbf{x} = \mathbf{0}$ y $B\mathbf{x} = \mathbf{0}$ tienen las mismas soluciones. Entonces

$$A\mathbf{c} = \mathbf{0} \;\Leftrightarrow\; B\mathbf{c} = \mathbf{0}.$$

Si las componentes del vector $\mathbf{c}$ son c_1, c_2,..., c_n, por definición de producto de una matriz por un vector, la equivalencia anterior resulta

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + ... + c_n\mathbf{a}_n = \mathbf{0} \;\Leftrightarrow\; c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + ... + c_n\mathbf{b}_n = \mathbf{0}.$$

□

Este lema indica que si A y B son matrices equivalentes por renglones, un conjunto de columnas de A es linealmente dependiente (o independiente) si y solo si las correspondientes columnas de B son linealmente dependientes (o independientes).

Por el momento nos restringiremos a las matrices que están en forma escalonada. Se verá que es fácil encontrar una base para el espacio de columnas de tales matrices.

Teorema 4.18. Las columnas pivote de una matriz que está en forma escalonada por renglones forman una base para su espacio de columnas.

Una demostración de este teorema se puede ver al final de esta sección. Pasemos ahora al caso general:

Teorema 4.19. Las columnas pivote de una matriz A forman una base para el espacio $\mathrm{col}(A)$.

Demostración. Sea $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & ... & \mathbf{a}_n \end{bmatrix}$ una matriz de $m \times n$, y $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & ... & \mathbf{b}_n \end{bmatrix}$ una forma escalonada de A.

Sean $\mathbf{b}_{l_1}$, $\mathbf{b}_{l_2}$,..., $\mathbf{b}_{l_k}$ las columnas pivote de B. Por el teorema anterior, éstas forman una base para $\mathrm{col}(B)$ y por lo tanto son linealmente independientes y las restantes columnas de B son combinaciones lineales de ellas.

De acuerdo al Lema 4.3 las columnas de A y B guardan las mismas relaciones de dependencia lineal. Luego, las correspondientes columnas (pivotes) de A, $\mathbf{a}_{l_1}$, $\mathbf{a}_{l_2}$,..., $\mathbf{a}_{l_k}$, forman un conjunto linealmente independiente y las restantes columnas son combinaciones lineales de ellas. Por lo tanto,

$$\mathrm{col}(A) = \mathrm{gen}\left\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\right\} = \mathrm{gen}\left\{\mathbf{a}_{l_1}, \mathbf{a}_{l_2}, ..., \mathbf{a}_{l_k}\right\}.$$

Así, el conjunto de las columnas pivote de A, $\{\mathbf{a}_{l_1}, \mathbf{a}_{l_2}, ..., \mathbf{a}_{l_k}\}$, es linealmente independiente y generador de $\mathrm{col}(A)$, y por lo tanto es una base de $\mathrm{col}(A)$. $\square$

Corolario 4.5. La dimensión del espacio de columnas de una matriz A es el número de columnas pivote de A.

Ejemplo 4.81. Como aplicación del teorema anterior, suponga que se pide encontrar un subconjunto de $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ donde $\mathbf{v}_1 = (1, -2, 0, 0, 3)$, $\mathbf{v}_2 = (2, -5, -3, -2, 6)$, $\mathbf{v}_3 = (0, 5, 15, 10, 0)$ y $\mathbf{v}_4 = (2, 6, 18, 8, 6)$, que sea una base para $\mathrm{gen}(S)$. Para ello se forma la matriz A cuyas columnas son dichos vectores:

$$A = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 & 2 \\ -2 & -5 & 5 & 6 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 3 & 6 & 0 & 6 \end{bmatrix},$$

entonces $\mathrm{gen}(S) = \mathrm{col}(A)$. De acuardo al Teorema 4.19 el conjunto de las columnas pivote de A forman una base de $\mathrm{col}(A)$. Estas columnas se identifican llevando A a una forma escalonada para obtener las posiciones pivote. Una forma escalonada de A es la matriz B (verifíquelo):

$$B = \begin{bmatrix} \boxed{1} & 2 & 0 & 2 \\ 0 & \boxed{-1} & 5 & 10 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Las posiciones pivote muestran que las columnas pivote son la 1era, la 2da y la 4ta. Por lo tanto, el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ es una base para $\mathrm{col}(A)$. Luego, ésta es una base de $\mathrm{gen}(S)$ formada por vectores de S.◀

El ejemplo anterior ofrece un método para hallar una base del espacio generado por un conjunto finito, S, de vectores de $\mathbb{R}^n$ que esté compuesta exclusivamente por vectores de S.

Cálculo de una base para gen(S) formada por vectores de S

Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m\}$ un subconjunto de $\mathbb{R}^n$. Se puede determinar un subconjunto de S que es una base para $\mathrm{gen}(S)$ como sigue:

1. Forme la matriz A de $n \times m$ cuyas columnas son los vectores $\mathbf{v}_1$, $\mathbf{v}_2$,...,$\mathbf{v}_m$.
2. Reduzca A hasta una forma escalonada B e identifique las columnas pivote de B, y en consecuencia las correspondientes columnas pivote de A.
3. Una base para $\mathrm{gen}(S)$ es el conjunto de los vectores columna de A identificadas en el item anterior.

Se han estado desarrollando métodos que permiten obtener bases para espacios generados por un subconjunto S de $\mathbb{R}^n$. Estos métodos también pueden emplearse con el mismo fin, en espacios vectoriales que no son $\mathbb{R}^n$ teniendo en cuenta que todos los cálculos con vectores de un espacio vectorial $\mathbb{V}$ de dimensión finita pueden hacerse de forma equivalente con sus vectores de coordenadas con respecto a una base dada.

Teniendo en cuenta los resultados del Teorema 4.15, si B es una base de un espacio vectorial $\mathbb{V}$ de dimensión n y S' es el conjunto de los vectores de coordenadas de los vectores de S con respecto a la base B, entonces H es una base de gen(S) si y solo si el correspondiente conjunto de vectores de coordenadas, H', es una base de gen(S').

En el Ejemplo 4.64 se implementó un procedimiento para hallar una base del espacio generado por un conjunto, S, de polinomios de $\mathbb{P}_3$ reduciendo el conjunto S hasta encontrar un subconjunto de él linealmente independiente. Ésto resultó un trabajo tedioso pero teniendo en cuenta lo comentado en el párrafo anterior un problema similar se puede resolver utilizando vectores de coordenadas y trabajando en $\mathbb{R}^n$ lo que facilita los cálculos y el uso de sistemas computacionales. Al respecto veamos el siguiente ejemplo:

Ejemplo 4.82. Al igual que en el Ejemplo 4.64, sea el subconjunto $S = \{p_1, p_2, p_3, p_4, p_5\}$ del espacio vectorial $\mathbb{P}_3$ donde $p_1 = 1 + x + 2x^3$, $p_2 = x + x^2 - x^3$, $p_3 = 2 + 3x + x^2 + 3x^3$, $p_4 = -1 + x^2 - 3x^3$ y $p_5 = 1 + x + x^2 + x^3$. Se desea encontrar una base para el espacio gen(S) que esté contenida en S.

Para ello, basta con hallar una base para el subespacio de $\mathbb{R}^4$ generado por el conjunto S' de los vectores de coordenadas de S respecto a una base cualquiera de $\mathbb{P}_3$. Se elije como base de $\mathbb{P}_3$ a la base estándar $E = \left\{1, x, x^2, x^3\right\}$ ya que no hay que resolver ningún sistema lineal para hallar vectores de coordenadas con respecto a dicha base.

Entonces el problema se reduce a determinar una base para el espacio generado por

$$S' = \{[p_1]_E, [p_2]_E, [p_3]_E, [p_4]_E, [p_5]_E\},$$

que sea un subconjunto de S'. Los vectores de coordenadas se obtienen en forma directa:

$$[p_1]_E = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \quad [p_2]_E = \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \quad [p_3]_E = \begin{bmatrix} 2 \\ 3 \\ 1 \\ 3 \end{bmatrix}, \quad [p_4]_E = \begin{bmatrix} -1 \\ 0 \\ 1 \\ -3 \end{bmatrix}, \quad [p_5]_E = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Se forma la matriz A cuyas columnas son dichos vectores de coordenadas:

$$A = \begin{bmatrix} 1 & 0 & 2 & -1 & 1 \\ 1 & 1 & 3 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 2 & -1 & 3 & -3 & 1 \end{bmatrix}.$$

El conjunto de las columnas pivote de A es una base para gen(S'). Una forma escalonada de A es la matriz

$$B = \begin{bmatrix} \boxed{1} & 0 & 2 & -1 & 1 \\ 0 & \boxed{1} & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Las posiciones pivote indican que las columnas pivote de A son: $[p_1]_E$, $[p_2]_E$ y $[p_5]_E$. Por lo tanto el conjunto $\{[p_1]_E, [p_2]_E, [p_5]_E\}$ es una base para gen(S') y el correspondiente conjunto en $\mathbb{P}_3$, $\{p_1, p_2, p_5\}$ es una base para gen(S) (y está contenida en S).◀

Teorema 4.20. Para cualquier matriz A, $\dim(\text{col}(A)) = \dim(\text{ren}(A))$.

Demostración. La dimensión de col(A) es el número de columnas pivote de A que son aquellas que contienen una posición pivote. Por lo tanto:

$$\dim(\text{col}(A)) = \text{número de pivotes de } A.$$

La dimensión de ren(A) es el número de renglones no cero de una forma escalonada de A. Dado que el número de renglones no cero es igual al número de pivotes:

$$\dim(\text{ren}(A)) = \text{número de pivotes de } A.$$

Por lo tanto,

$$\dim(\text{col}(A)) = \dim(\text{ren}(A)) \ \ (= \text{número de pivotes de } A)\,.$$

□

Definición (Rango de una matriz). El *rango* de una matriz A de $m \times n$ es la dimensión común de los espacios columna y renglón de A, se representa por rango(A) o $\rho(A)$.

Nótese que, de acuerdo al Teorema 4.20, para calcular el rango de una matriz A se debe llevar a ésta a una forma escalonada por renglones, B. El rango de A es el número de renglones no cero de B o, lo que es lo mismo, el número de pivotes de B. Luego, si A es de $m \times n$, $\rho(A) \leq \text{mín}\{m, n\}$. Por ejemplo, una matriz A de 5×7 no puede tener rango 6 pues $\rho(A) \leq 5$.

Puede observarse además que $\rho(A) = 0 \Leftrightarrow A = \mathbf{O}$.

Definición (Espacio nulo y nulidad de una matriz). El *espacio nulo* de una matriz A de $m \times n$, que se simboliza $N(A)$, es el conjunto de todas las soluciones de $A\mathbf{x} = \mathbf{0}$. Es decir,

$$N(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}\,.$$

$N(A)$ es un subespacio de $\mathbb{R}^n$, su dimensión es llamada *nulidad* de A y se denota nulidad(A) o $\nu(A)$.

Nota: En la Sección 4.2 se probó que $N(A)$, el conjunto de todas las soluciones de $A\mathbf{x} = \mathbf{0}$, es un subespacio de $\mathbb{R}^n$.

Ejemplo 4.83. Para hallar el espacio nulo de la matriz

$$A = \begin{bmatrix} 1 & -2 & 3 & 4 \\ 2 & 0 & -1 & 1 \\ 0 & 1 & 1 & -2 \\ 2 & 1 & 0 & -1 \end{bmatrix},$$

se debe resolver el sistema lineal homogéneo

$$\left\{ \begin{array}{rrrrcl} x_1 & -2x_2 & +3x_3 & +4x_4 & = & 0 \\ 2x_1 & & -x_3 & +x_4 & = & 0 \\ & x_2 & +x_3 & -2x_4 & = & 0 \\ 2x_1 & +x_2 & & -x_4 & = & 0 \end{array} \right.$$

La forma escalonada reducida de la matriz ampliada del sistema es

$$\left[\begin{array}{rrrrcr} 1 & 0 & 0 & 5/11 & : & 0 \\ 0 & 1 & 0 & -21/11 & : & 0 \\ 0 & 0 & 1 & -1/11 & : & 0 \\ 0 & 0 & 0 & 0 & : & 0 \end{array}\right],$$

y se observa que x_4 es la única variable libre. La solución general del sistema viene dada por

$$\left\{\begin{array}{rcl} x_1 & = & -\frac{5}{11}t \\ x_2 & = & \frac{21}{11}t \\ x_3 & = & \frac{1}{11}t \\ x_4 & = & t \end{array}\right. , \quad t \in \mathbb{R}.$$

Luego, el espacio nulo de la matriz A es

$$N(A) = \left\{ \begin{bmatrix} -5/11 \\ 21/11 \\ 1/11 \\ 1 \end{bmatrix} t : t \in \mathbb{R} \right\} = \text{gen} \left\{ \begin{bmatrix} -5 \\ 21 \\ 1 \\ 11 \end{bmatrix} \right\}.$$

El conjunto unitario $\left\{ \begin{bmatrix} -5 \\ 21 \\ 1 \\ 11 \end{bmatrix} \right\}$ es linealmente independiente, y por ser generador de $N(A)$, es una base de dicho espacio. Así, $\nu(A) = \dim(N(A)) = 1.$◀

Ejemplo 4.84. Sea la matriz $A = \begin{bmatrix} 2 & 3 \\ 0 & -4 \end{bmatrix}$. Como $\det(A) = -8 \neq 0$, el sistema lineal $A\mathbf{x} = \mathbf{0}$ tiene única solución, la trivial. Luego,

$$N(A) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}.$$

No se define base para $N(A)$ ya que es el espacio cero. La dimensión de $N(A)$ es $\nu(A) = 0.$◀

Puede observarse en los ejemplos anteriores que la nulidad de la matriz A es igual al número de variables libres del sistema lineal $A\mathbf{x} = \mathbf{0}$.

> **Teorema 4.21.** La nulidad de una matriz A es el número de variables libres del sistema lineal homogéneo que tiene a A como matriz de coeficientes.

Una demostración de este teorema se encuentra al final de esta sección.

El teorema siguiente relaciona el rango y la nulidad de una matriz. Es considerado uno de los teoremas más importantes del Álgebra Lineal

> **Teorema 4.22** (del rango)**.** Para toda matriz A de $m \times n$, la suma del rango y la nulidad es igual al número de columnas de A, es decir
>
> $$\rho(A) + \nu(A) = n.$$

Demostración. El rango de una matriz A es el número de columnas pivote de A. Por otra parte la nulidad de A es el número de variables libres de $A\mathbf{x} = \mathbf{0}$. Como se tienen tantas variables libres como columnas no pivote, la nulidad de A es igual a la cantidad de columnas no pivote. Dado que

$$\text{número de columnas pivote} + \text{número de columnas no pivote} = \text{número de columnas},$$

entonces

$$\rho(A) + \nu(A) = n.$$

□

Ejemplo 4.85. En el Ejemplo 4.83, $\nu(A) = 1$. Por otro lado, el número de pivotes de A es 3, por lo que $\rho(A) = 3$. Luego, $\rho(A) + \nu(A) = 3 + 1 = 4$ que es el número de columnas de A, y así se verifica el teorema del rango.

En el Ejemplo 4.84, el número de pivotes de A es 2 por lo que $\rho(A) = 2$. Por otra parte, $\nu(A) = 0$. Luego, $\rho(A) + \nu(A) = 2 + 0 = 2$ que es el número de columnas de A. ◀

Ejemplo 4.86. Si el sistema lineal $A\mathbf{x} = \mathbf{0}$ tiene 25 incógnitas y $\rho(A) = 15$, puede deducirse que A tiene 25 columnas, que $\nu(A) = 10$ y que el número m de renglones de A es $m \geq 15$.◀

Las nociones desarrolladas en esta sección están relacionadas con los sistemas lineales. Dado que el sistema lineal $A\mathbf{x} = \mathbf{b}$ es consistente si y solo si el vector $\mathbf{b}$ es combinación lineal de las columnas de A, se tiene el siguiente teorema:

Teorema 4.23. Un sistema lineal $A\mathbf{x} = \mathbf{b}$ es consistente si y solo si $\mathbf{b}$ está en $\text{col}(A)$.

Y en relación a los sistemas lineales y al rango de una matriz, se tiene el siguiente resultado:

Teorema 4.24. El sistema lineal $A\mathbf{x} = \mathbf{b}$ es consistente si y solo si $\rho(A) = \rho([A : \mathbf{b}])$.

Ejemplo 4.87. Sea $A\mathbf{x} = \mathbf{b}$ un sistema lineal de 20 ecuaciones con 25 incógnitas. Suponga que el espacio nulo de A está generado por cinco vectores linealmente independientes. Se desea saber si el sistema es consistente para cualquier vector $\mathbf{b}$ de $\mathbb{R}^{20}$.

La matriz A es de 20×25. El espacio nulo de A tiene una base que consta de 5 vectores, luego $\nu(A) = 5$. Por el teorema del rango, $\rho(A) = 25 - 5 = 20$ es decir $\dim(\text{col}(A)) = 20$. Esto indica que $\text{col}(A) = \mathbb{R}^{20}$ y por lo tanto todo vector $\mathbf{b}$ de $\mathbb{R}^{20}$ está en $\text{col}(A)$ es decir es una combinación de las columnas de A. Esto prueba que $A\mathbf{x} = \mathbf{b}$ es consistente para todo $\mathbf{b}$ de $\mathbb{R}^{20}$.◀

El siguiente teorema completa el Teorema 3.10 de equivalencias.

Teorema 4.25. Sea A una matriz de $n \times n$. Entonces las siguientes proposiciones son equivalentes:

1. $|A| \neq 0$.
2. A es invertible.
3. $A\mathbf{x} = \mathbf{b}$ es consistente con única solución para todo $\mathbf{b} \in \mathbb{R}^n$.
4. $A\mathbf{x} = \mathbf{0}$ tiene única solución la trivial.
5. A tiene n posiciones pivote.
6. A es equivalente (por renglones) a la matriz identidad I.
7. $\nu(A) = 0$
8. $\rho(A) = n$
9. Los vectores renglón de A son linealmente independientes.
10. Los vectores columna de A son linealmente independientes.
11. $\text{col}(A) = \mathbb{R}^n$.
12. $\text{ren}(A) = \mathbb{R}^n$.

Demostración. Las equivalencias de (1) a (6) han sido probadas en el Teorema 3.10. Se probará aquí las equivalencias restantes. Para ello se seguirá el siguiente orden en las implicancias (6) $\Rightarrow$ (7) $\Rightarrow$ (8) $\Rightarrow$ (9) $\Rightarrow$ (10) $\Rightarrow$ (11) $\Rightarrow$ (12), probando luego que (12) $\Rightarrow$ (6).

(6) $\Rightarrow$ (7) Si A es equivalente por renglones a la matriz identidad, I_n, entonces las n columnas de A son columnas pivote por lo cual $A\mathbf{x} = \mathbf{0}$ no tiene variables libres y así $\nu(A) = 0$.

(7) $\Rightarrow$ (8) Es una consecuencia inmediata del teorema del rango.

(8) $\Rightarrow$ (9) Si $\rho(A) = n$ entonces el espacio de renglones de A es de dimensión n. Ya que los n renglones de A generan el espacio $\text{ren}(A)$ se concluye (Teorema 4.9) que los vectores renglón de A forman una base para $\text{ren}(A)$ y por lo tanto son linealmente independientes.

(9) $\Rightarrow$ (10) Si los n renglones de A forman un conjunto linealmente independiente, como este conjunto es generador de $\text{ren}(A)$, entonces forman una base para $\text{ren}(A)$ y por lo tanto $\dim(\text{ren}(A)) = n$. Por el Teorema 4.20, también $\dim(\text{col}(A)) = n$ y dado que el conjunto de las n columnas de A es generador de $\text{col}(A)$ entonces (Teorema 4.9) las n columnas de A forman una base para $\text{col}(A)$ y por lo tanto son linealmente independientes.

(10) $\Rightarrow$ (11) Las n columnas de A generan $\text{col}(A)$ y dado que son linealmente independientes, entonces forman una base para $\text{col}(A)$. Así $\dim(\text{col}(A)) = n$ y como $\text{col}(A)$ es un subespacio de $\mathbb{R}^n$, entonces (Teorema 4.12) $\text{col}(A) = \mathbb{R}^n$.

(11) $\Rightarrow$ (12) Dado que $\dim(\text{ren}(A)) = \dim(\text{col}(A))$, si $\text{col}(A) = \mathbb{R}^n$ entonces $\dim(\text{ren}(A)) = n$. Como $\text{ren}(A)$ es un subespacio de $\mathbb{R}^n$, necesariamente (Teorema 4.12) $\text{ren}(A) = \mathbb{R}^n$.

(12) $\Rightarrow$ (6) Si ren$(A) = \mathbb{R}^n$ entonces $\dim(\text{ren}(A)) = n$ y por lo tanto A tiene n posiciones pivote. Luego, la forma escalonada reducida de A es la matriz identidad, I_n.

$\square$

Teorema 4.16: Los renglones no cero de una matriz en forma escalonada por renglones forman una base para su espacio de renglones.

Demostración. Sea $B = [b_{ij}]$ una matriz de $m \times n$ en forma escalonada por renglones. Si k es el número de renglones no cero de B, dado que B está en forma escalonada, éstos son los primeros k renglones, mientras que los últimos $m - k$ renglones son renglones cero.

Se probará que el conjunto $S = \{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_k\}$ es una base de ren(B).

Los vectores $\mathbf{r}_1$, $\mathbf{r}_2$,..., $\mathbf{r}_k$ son todos no cero y como B está en forma escalonada, el elemento delantero de cada renglón no cero está más a la derecha que el elemento delantero del renglón anterior. Esto es, si b_{il_i} es el elemento delantero del renglón i-ésimo debe ser $l_1 < l_2 < ... < l_k$. Recordando que el elemento delantero de un renglón no cero es la primera componente no cero de dicho vector, entonces para cada $\mathbf{r}_i$ es $b_{il_i} \neq 0$ mientras que $b_{ij} = 0$ para todo $j < l_i$. Así, por ejemplo, los renglones $\mathbf{r}_1$ y $\mathbf{r}_2$ son de la forma:

$$\begin{aligned} \mathbf{r}_1 &= (0, 0, ..., 0, b_{1l_1}, \times, ..., \times, ..., \times)\,, \\ \mathbf{r}_2 &= (0, 0, ..., 0, 0, ..., 0, b_{2l_2}, \times, ..., \times)\,. \end{aligned}$$

Para probar que S es linealmente independiente debe demostrarse que si c_1, c_2,..., c_k son escalares tales que

$$c_1\mathbf{r}_1 + c_2\mathbf{r}_2 + ... + c_k\mathbf{r}_k = \mathbf{0},$$

necesariamente $c_1 = c_2 = ... = c_k = 0$.

La componente en la posición l_1 del vector $c_1\mathbf{r}_1 + c_2\mathbf{r}_2 + ... + c_k\mathbf{r}_k$ es $c_1 b_{1l_1}$ y las anteriores componentes son ceros. Si este vector es el vector $\mathbf{0}$ necesariamente $c_1 b_{1l_1} = 0$ y dado que $b_{1l_1} \neq 0$ debe ser $c_1 = 0$. Entonces

$$c_1\mathbf{r}_1 + c_2\mathbf{r}_2 + ... + c_k\mathbf{r}_k = \mathbf{0} \Leftrightarrow c_1 = 0 \;\text{ y }\; c_2\mathbf{r}_2 + ... + c_k\mathbf{r}_k = \mathbf{0}.$$

De manera similar, la componente en la posición l_2 del vector $c_2\mathbf{r}_2 + ... + c_k\mathbf{r}_k$ es $c_2 b_{2l_2}$ y las anteriores componentes son ceros. Si este vector es el vector $\mathbf{0}$ necesariamente $c_2 b_{2l_2} = 0$ y dado que $b_{2l_2} \neq 0$ debe ser $c_2 = 0$. Así,

$$c_1\mathbf{r}_1 + c_2\mathbf{r}_2 + ... + c_k\mathbf{r}_k = \mathbf{0} \Leftrightarrow c_1 = 0,\; c_2 = 0 \;\text{ y }\; c_3\mathbf{r}_3 + ... + c_k\mathbf{r}_k = \mathbf{0}.$$

De esta manera se llega en $k - 1$ pasos a

$$c_1\mathbf{r}_1 + c_2\mathbf{r}_2 + ... + c_k\mathbf{r}_k = \mathbf{0} \Leftrightarrow c_1 = 0,\; c_2 = 0,\; ...,\; c_{k-1} = 0 \text{ y }\; c_k\mathbf{r}_k = \mathbf{0},$$

y como $\mathbf{r}_k \neq \mathbf{0}$ necesariamente $c_k = 0$. Luego,

$$c_1\mathbf{r}_1 + c_2\mathbf{r}_2 + ... + c_k\mathbf{r}_k = \mathbf{0} \Leftrightarrow c_1 = c_2 = ... = c_k = 0.$$

Es claro que el conjunto S es generador del espacio ren(B) ya que los últimos $m-k$ renglones de B son vectores cero y por lo tanto pueden eliminarse generando el mismo espacio. Esto es, ren$(B) =$ gen$\{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_k\}$.

Se ha probado entonces que los renglones no cero de una matriz en forma escalonada son linealmente independientes y generan ren(B) y por lo tanto forman una base para el espacio de renglones de B. $\square$

Teorema 4.18: Las columnas pivote de una matriz que está en forma escalonada por renglones forman una base para su espacio de columnas.

Demostración. Sea $B = [b_{ij}]$ una matriz de $m \times n$ que está en forma escalonada por renglones. Si k es el número de renglones no cero de B, dado que B está en forma escalonada, éstos son los primeros k renglones de B, mientras que los últimos $m-k$ renglones son renglones cero.

Sean b_{1l_1}, b_{2l_2},...., b_{kl_k} los elementos delanteros de los k renglones no cero de B. Entonces las columnas pivote de B son $\mathbf{b}_{l_1}$, $\mathbf{b}_{l_2}$,..., $\mathbf{b}_{l_k}$. Se probará que estas columnas forman una base para col(B).

Para ello, consderemos la matriz U de $m \times k$ formada por las columnas pivote de B:

$$U = \begin{bmatrix} \mathbf{b}_{l_1} & \mathbf{b}_{l_2} & ... & \mathbf{b}_{l_k} \end{bmatrix} = \begin{bmatrix} b_{1l_1} & \times & \times & ... & \times \\ 0 & b_{2l_2} & \times & ... & \times \\ 0 & 0 & b_{3l_3} & ... & \times \\ \vdots & \vdots & \vdots & ... & \vdots \\ 0 & 0 & 0 & ... & b_{kl_k} \\ 0 & 0 & 0 & ... & 0 \\ \vdots & \vdots & \vdots & ... & \vdots \\ 0 & 0 & 0 & ... & 0 \end{bmatrix}, \quad \text{donde } b_{il_i} \neq 0.$$

Dado que todos los vectores columna de B tienen sus últimas $m-k$ componentes iguales a cero, todo vector $\mathbf{b}$ de col(B) tendrá sus últimas $m-k$ componentes iguales a cero. Entonces para cualquier $\mathbf{b}$ en col(B), el sistema lineal de matriz aumentada

$$[U : \mathbf{b}] = \begin{bmatrix} \boxed{b_{1l_1}} & \times & \times & ... & \times & : & b_1 \\ 0 & \boxed{b_{2l_2}} & \times & ... & \times & : & b_2 \\ 0 & 0 & \boxed{b_{3l_3}} & ... & \times & : & b_3 \\ \vdots & \vdots & \vdots & ... & \vdots & : & \vdots \\ 0 & 0 & 0 & ... & \boxed{b_{kl_k}} & : & b_k \\ 0 & 0 & 0 & ... & 0 & : & 0 \\ \vdots & \vdots & \vdots & ... & \vdots & : & \vdots \\ 0 & 0 & 0 & ... & 0 & : & 0 \end{bmatrix},$$

es consistente y con única solución ya que la última columna de la matriz aumentada no es columna pivote y todas las columnas de U son pivote. Esto quiere decir que todo vector $\mathbf{b}$ en col(B) puede expresarse de manera única como combinación lineal de las columnas pivote de la matriz B. Esto prueba que las columnas pivote de B forman una base de col(B). $\square$

Teorema 4.21: La nulidad de una matriz A es el número de variables libres del sistema lineal homogéneo que tiene a A como matriz de coeficientes.

Demostración. Si B es cualquier forma escalonada de la matriz A, se conoce que los sistemas homogéneos $A\mathbf{x} = \mathbf{0}$ y $B\mathbf{x} = \mathbf{0}$ tienen las mismas soluciones, en particular si B es la forma escalonada reducida de A.

Si $\rho(A) = r$ y A es de $m \times n$, entonces $r \leq m$ y $r \leq n$. Si U es la forma escalonada reducida de A entonces los primeros r renglones de U son distintos de cero.

Si $r = n$ todas las columnas de U (y por lo tanto de A) son columnas pivote y entonces el sistema $A\mathbf{x} = \mathbf{0}$ tiene única solución, la trivial. Luego, $N(A)$ está formado únicamente por el vector cero y por lo tanto su dimensión es cero, que es también el número de variables libres de $A\mathbf{x} = \mathbf{0}$.

Si $r < n$ no se pierde generalidad si se supone que las columnas pivote son las r primeras columnas. Más aún como los renglones cero de U (si los tiene) no contribuyen a la solución de $U\mathbf{x} = \mathbf{0}$ estos renglones pueden descartarse. De modo que las soluciones de $A\mathbf{x} = \mathbf{0}$ se obtienen del sistema $U'\mathbf{x} = \mathbf{0}$ donde U' es la matriz de $r \times n$ formadas por los r renglones no cero de U. Así las primers r columnas de U' (pivotes) forman la matriz identidad de orden r y las últimas $n - r$ son las columnas no pivote es decir

$$U' = \begin{bmatrix} 1 & 0 & ... & 0 & c_{1\ r+1} & ... & c_{1n} \\ 0 & 1 & ... & 0 & c_{2\ r+1} & ... & c_{2n} \\ \vdots & \vdots & ... & \vdots & \vdots & ... & \vdots \\ 0 & 0 & ... & 1 & c_{r\ r+1} & ... & c_{rn} \end{bmatrix}.$$

Siendo el vector de las incógnitas $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{bmatrix}$, la forma estándar de $U'\mathbf{x} = \mathbf{0}$ es

$$\begin{cases} x_1 + c_{1\ r+1}x_{r+1} + c_{1\ r+1}x_{r+1} + ... + c_{1n}x_n = 0 \\ x_2 + c_{2\ r+1}x_{r+1} + c_{2\ r+1}x_{r+1} + ... + c_{2n}x_n = 0 \\ \vdots \\ x_r + c_{r\ r+1}x_{r+1} + c_{r\ r+1}x_{r+1} + ... + c_{rn}x_n = 0 \end{cases}$$

y despejando las r variables delanteras en función de las $n - r$ libres se tiene

$$\begin{cases} x_1 = -c_{1\ r+1}x_{r+1} - c_{1\ r+1}x_{r+1} - ... - c_{1n}x_n \\ x_2 = -c_{2\ r+1}x_{r+1} - c_{2\ r+1}x_{r+1} - ... - c_{2n}x_n \\ \vdots \\ x_r = -c_{r\ r+1}x_{r+1} - c_{r\ r+1}x_{r+1} - ... - c_{rn}x_n \end{cases}$$

Asignando valores arbitrarios s_1, s_2,...,s_{r-n} a las variables libres x_{r+1}, x_{r+2},...,x_n, respectivamente, la solución general para $A\mathbf{x} = \mathbf{0}$ está dada por:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} -c_{1\ r+1}s_1 - c_{1\ r+1}s_2 - ... - c_{1n}s_{n-r} \\ -c_{2\ r+1}s_1 - c_{2\ r+1}s_2 - ... - c_{2n}s_{n-r} \\ \vdots \\ -c_{r\ r+1}s_1 - c_{r\ r+1}s_2 - ... - c_{rn}s_{n-r} \\ s_1 \\ s_2 \\ \vdots \\ s_{n-r} \end{bmatrix}$$

$$= s_1 \underbrace{\begin{bmatrix} -c_{1\ r+1} \\ -c_{2\ r+1} \\ \vdots \\ -c_{r\ r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\mathbf{u}_1} + s_2 \underbrace{\begin{bmatrix} -c_{1\ r+1} \\ -c_{2\ r+1} \\ \vdots \\ -c_{r\ r+1} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}}_{\mathbf{u}_2} + ... + s_{n-r} \underbrace{\begin{bmatrix} -c_{1n} \\ -c_{2n} \\ \vdots \\ -c_{rn} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}}_{\mathbf{u}_{n-r}},$$

siendo s_1, s_2,..., $s_{n-r} \in \mathbb{R}$.

Luego, $\mathbf{x}$ es una solución de $A\mathbf{x} = \mathbf{0}$ si y solo si $\mathbf{x}$ es una combinación lineal de los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_{n-r}$ es decir $N(A) = \text{gen}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{n-r}\}$. Fácilmente puede probarse que estos vectores son linealmente independientes. En efecto $s_1\mathbf{u}_1 + s_2\mathbf{u}_2 + ... + s_{n-r}\mathbf{u}_{n-r} = \mathbf{0}$ si y solo si

$$\begin{bmatrix} -c_{1\ r+1}s_1 - c_{1\ r+1}s_2 - ... - c_{1n}s_{n-r} \\ -c_{2\ r+1}s_1 - c_{2\ r+1}s_2 - ... - c_{2n}s_{n-r} \\ \vdots \\ -c_{r\ r+1}s_1 - c_{r\ r+1}s_2 - ... - c_{rn}s_{n-r} \\ s_1 \\ s_2 \\ \vdots \\ s_{n-r} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

y por igualdad de vectores necesariamente $s_1 = s_2 = ... = s_{n-r} = 0$.

Así el conjunto $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{n-r}\}$ es una base de $N(A)$. Por lo tanto, $\dim(N(A)) = n - r$ que es el número de variables libres de $A\mathbf{x} = \mathbf{0}$. □

Ejercicios 4.7

1. Para cada matriz determine el rango y la nulidad. Halle además una base para su espacio de renglones, una base para su espacio de columnas y, si es posible, una base para su espacio nulo.

$$a)\begin{bmatrix} 3 & -2 & 8 & 2 & 1 & 4 \\ 0 & 0 & -8 & 5 & 0 & -10 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \qquad b)\begin{bmatrix} 2 & -3 & 5 \\ 0 & 0 & 8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad c)\begin{bmatrix} 3 & -1 & 2 & 1 \\ 0 & -4 & 3 & 5 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 10 \end{bmatrix}$$

$\boxed{2.}$ Para cada matriz encuentre una base para su espacio de renglones, una base para su espacio de columnas y, si es posible, una base para su espacio nulo. Determine además el rango y la nulidad de cada matriz.

$$a)\begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & 8 \\ -1 & -5 & 0 \end{bmatrix} \qquad b)\begin{bmatrix} 0 & 2 & -3 & 1 & 2 \\ 0 & -2 & 3 & 3 & 1 \\ 0 & 4 & -6 & 6 & 7 \end{bmatrix} \qquad c)\begin{bmatrix} 3 & 2 & -1 \\ 6 & 3 & 5 \\ 3 & 1 & 6 \\ 0 & 1 & -7 \end{bmatrix}$$

$\boxed{3.}$ Dada la matriz $A = \begin{bmatrix} -1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 4 \\ 0 & 1 & 1 & 2 \end{bmatrix}$.

a) Describa el espacio de renglones y el de columnas de A.

b) Halle una base para el espacio de renglones de A.

c) Teniendo en cuenta que $\text{col}(A) = \text{ren}(A^t)$, halle una base para el espacio $\text{col}(A)$.

d) Halle una base para el espacio de columnas de A que esté formada exclusivamente por columnas de A.

e) Describa el espacio nulo de A y halle una base, si es posible, del mismo.

f) Verifique el teorema del rango.

g) Indique cuáles de los vectores

$$\mathbf{a} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} \quad \text{y} \quad \mathbf{c} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$

están en el espacio de columnas de la matriz A. Si lo están expréselos como una combinación lineal de columnas de A.

4. Obtenga una base para $\text{gen}(S)$ siendo:

a) $S = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix} \right\}$

b) $S = \{(1, -1, 0), (1, 2, 1), (0, 1, 1)\}$

$\boxed{c)}$ $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 4 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ 1 \\ -4 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 5 \\ -1 \end{bmatrix} \right\}$

5. Utilice vectores de coordenadas para determinar una base del subespacio $\text{gen}(S)$ de $\mathbb{V}$, siendo:

a) $S = \{1 - 3x, -3 + 9x, 4x, 1 - x\}$, $\mathbb{V} = \mathbb{P}_1$.

$\boxed{b)}$ $S = \{-1 + 2x^2, x - x^2, 2 - 4x^2, -1 + x + x^2\}$, $\mathbb{V} = \mathbb{P}_2$.

c) $S = \{1 - 2x - 3x^2 + x^3, x + 2x^2, 3 - 2x - x^2 + 2x^3\}$, $\mathbb{V} = \mathbb{P}_3$.

6. Utilice vectores de coordenadas para hallar una base de gen(S) siendo:

a) $S = \left\{ \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \right\}$

b) $S = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 3 & -2 \end{bmatrix} \right\}$

$\boxed{c)}$ $S = \left\{ \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -2 & 2 \\ 2 & 4 & 6 \end{bmatrix}, \begin{bmatrix} 0 & -2 & 2 \\ 2 & 6 & 7 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix} \right\}$

7. En cada caso, obtenga un subconjunto de S que sea una base del subespacio gen(S) de $\mathbb{V}$.

$\boxed{a)}$ $S = \left\{ \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$, $\mathbb{V} = \mathbb{R}^3$

b) $S = \{1 + 3x - x^2 + x^3, 1 + x + x^3, 2 + x^2 + x^3, -1 + x - x^2, x + x^3\}$, $\mathbb{V} = \mathbb{P}_3$

c) $S = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \right\}$, $\mathbb{V} = \mathbb{M}_2$

8. **(Ampliación de un subconjunto linealmente independiente de $\mathbb{R}^n$ a una base de $\mathbb{R}^n$)** En el Teorema 4.10 se mostró que un subconjunto S linealmente independiente de un espacio vectorial $\mathbb{V}$ de dimensión n, puede ampliarse a una base de $\mathbb{V}$ agregando uno a uno vectores que no son combinaciones lineales de los anteriores hasta obtener un conjunto de n vectores linealmente independientes. Esto puede resultar tedioso ya que en cada paso hay que verificar que el vector que se agrega no pertenezca al espacio generado por los anteriores.

   Un procedimiento más eficiente consiste en reducir un conjunto generador de $\mathbb{V}$ que contenga al conjunto S. Si $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ es un conjunto linealmente independiente en $\mathbb{R}^n$, $k < n$, agregándole a S la base estándar, el conjunto $S' = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k, \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ obtenido es trivialmente generador de $\mathbb{R}^n$. Para obtener una base de $\mathbb{R}^n$ que contenga a S, se forma la matriz A cuyas columnas son los vectores de S' en el orden que se muestran:

$$A = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_k & \mathbf{e}_1 & \mathbf{e}_2 & ... & \mathbf{e}_n \end{bmatrix}.$$

   Luego, $\mathbb{R}^n = \text{gen}(S') = \text{col}(A)$.

   Dado que S es linealmente independiente, entonces el sistema lineal homogéneo cuya matriz de coeficientes es $A_1 = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_k \end{bmatrix}$ tiene solo la solución trivial y por lo tanto no hay variables libres. Luego, las k columnas de A_1 son columnas pivote y por lo tanto la forma escalonada reducida de A_1 es $B_1 = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & ... & \mathbf{e}_k \end{bmatrix}$. Entonces la forma escalonada reducida de la matriz A es

$$B = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & ... & \mathbf{e}_k & \mathbf{v}_1 & \mathbf{v}_2 & ... & \mathbf{v}_n \end{bmatrix}.$$

   Esto muestra que las k primeras columnas de A son columnas pivote y las restantes $n - k$ columnas pivote están entre las n últimas columnas. Así el conjunto de las columnas pivote de A contiene al conjunto S y es una base de $\mathbb{R}^n$.

Por ejemplo, el conjunto $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \right\}$ de $\mathbb{R}^4$ es linealmente independiente. Para obtener una base de $\mathbb{R}^4$ que contenga a S, formamos la matriz

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 \\ 3 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Claramente $\mathrm{col}(A) = \mathbb{R}^4$ por lo que el conjunto de las columnas pivote de A formarán una base para $\mathbb{R}^4$. Una forma escalonada de A es la matriz

$$B = \begin{bmatrix} \boxed{1} & 1 & 1 & 0 & 0 & 0 \\ 0 & \boxed{1} & 2 & -1 & 0 & 0 \\ 0 & 0 & \boxed{5} & -4 & 1 & 0 \\ 0 & 0 & 0 & \boxed{-3} & 2 & 5 \end{bmatrix}$$

que me indica que las primera, segunda, tercera y cuarta columna de A son las columnas pivote de A. Luego, el conjunto

$$S' = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

es una base para $\mathbb{R}^4$ que contiene al conjunto S.

a) Amplíe el conjunto linealmente independiente S hasta obtener una base del espacio vectorial $\mathbb{V}$ indicado en cada caso.

$\boxed{\text{i)}}$ $S = \left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix} \right\}$, $\mathbb{V} = \mathbb{R}^4$

ii) $S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ 2 \\ 1 \end{bmatrix} \right\}$, $\mathbb{V} = \mathbb{R}^5$

b) Utilice vectores de coordenadas para ampliar el conjunto linealmente independiente S hasta obtener una base del espacio $\mathbb{V}$.

$\boxed{\text{i)}}$ $S = \left\{1 + x + x^3, -2 + x + x^2, x^2 + x^3\right\}$, $\mathbb{V} = \mathbb{P}_3$.

ii) $S = \left\{ \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -2 \\ 2 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} \right\}$, $\mathbb{V} = \mathbb{M}_2$.

9. Sea W el plano en $\mathbb{R}^3$ que contiene a la recta $x = y$, $z = 0$ y también al eje z.

a) Encuentre una base de W.

b) Extienda la base hallada a una base de $\mathbb{R}^3$.

c) Sea B la base estándar de $\mathbb{R}^3$. ¿Puede reducirse B a una base de W?, explique.

10. Si el sistema lineal homogéneo $A\mathbf{x} = \mathbf{0}$ tiene 300 incógnitas y su espacio solución está generado por 65 vectores linealmente independientes. ¿Cuál es el rango de A?

11. Demuestre los teoremas 4.23 y 4.24.

$\boxed{12.}$ Demuestre que $\rho(A) = \rho(A^t)$ para toda matriz A de $m \times n$.

$\boxed{13.}$ Sea $A\mathbf{x} = \mathbf{b}$ un sistema lineal de 400 ecuaciones y 480 incógnitas. Si el conjunto solución del sistema lineal homogéneo $A\mathbf{x} = \mathbf{0}$ está generado por 80 vectores linealmente independiente, ¿es consistente el sistema $A\mathbf{x} = \mathbf{b}$, para todo vector $\mathbf{b} \in \mathbb{R}^{400}$?

Capítulo 5

ESPACIOS VECTORIALES CON PRODUCTO INTERNO

5.1. El espacio euclidiano $\mathbb{R}^n$

Se usan pares ordenados de números reales para localizar puntos en el plano y ternas de números reales para localizar puntos en el espacio de tres dimensiones. Aún cuando la concepción geométrica no se extiende más allá del espacio tridimensional, es posible extender muchas ideas conocidas más allá de tal espacio trabajando con propiedades analíticas o numéricas de puntos y vectores en lugar de las propiedades geométricas.

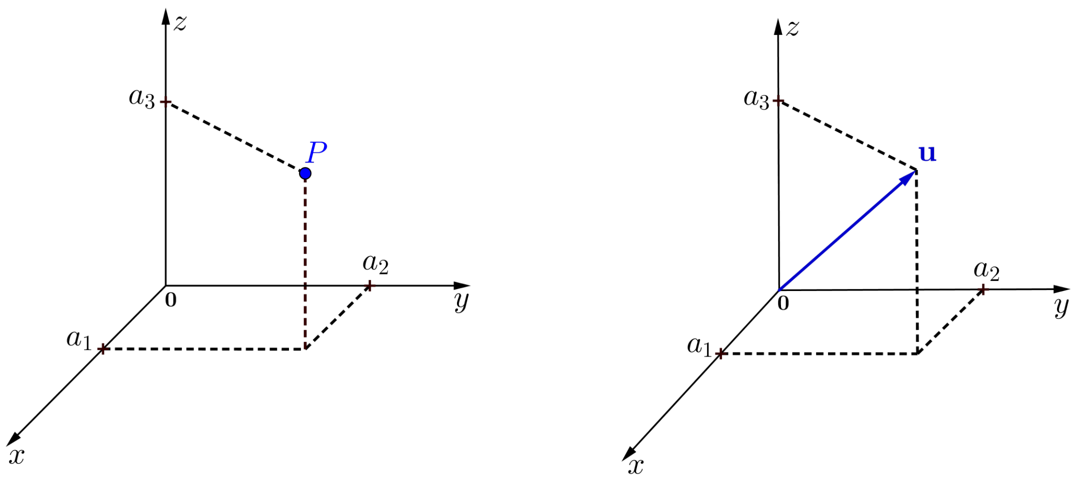


Figura 5.1: Punto $P\,(a_1, a_2, a_3)$ y vector $\mathbf{u} = (a_1, a_2, a_3)$ en $\mathbb{R}^3$

Recordemos que una n-upla ordenada de números reales es una sucesión $(a_1, a_2, ..., a_n)$ de n números reales. El conjunto de todas las n-uplas ordenadas de números reales se conocen como el espacio n dimensional y se denota por $\mathbb{R}^n$.

Si $n = 2$ ó $n = 3$, las n-uplas son pares ordenados o ternas ordenadas, respectivamente. Si $n = 1$ cada n-upla es un número real y por lo tanto $\mathbb{R}^1$ se concibe como el conjunto $\mathbb{R}$ de los números reales y se escribe $\mathbb{R}$ en lugar de $\mathbb{R}^1$.

En el estudio del espacio de dos dimensiones y de tres dimensiones, los símbolos (a_1, a_2) y (a_1, a_2, a_3) tienen dos interpretaciones geométricas. Pueden interpretarse como puntos en el plano y en el espacio, respectivamente, en cuyo caso (a_1, a_2) y (a_1, a_2, a_3) son las coordenadas de dichos puntos. También pueden considerarse como vectores (o segmentos de recta dirigidos) en cuyo caso (a_1, a_2) y (a_1, a_2, a_3) son las componentes de dichos vectores (Figura 5.1).

Por lo tanto una n-upla de números reales $(a_1, a_2, ..., a_n)$ se puede concebir como un "vector generalizado" o como un "punto generalizado". Se tiene la libertad de describir a una n-upla de números reales como un punto en $\mathbb{R}^n$ donde sus coordenadas son a_i con $i = 1, 2, ..., n$ o como un vector donde las a_i con $i = 1, 2, ..., n$ son sus componentes.

Recordemos que el *producto punto* o *producto escalar* entre dos vectores $\mathbf{u} = (u_1, u_2, ..., u_n)$ y $\mathbf{v} = (v_1, v_2, ..., v_n)$ de $\mathbb{R}^n$, es el número real dado por

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + ... + u_n v_n,$$

que es la generalización del producto escalar en $\mathbb{R}^2$ y $\mathbb{R}^3$. Las propiedades fundamentales del producto punto son:

Teorema 5.1. Si $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ son vectores de $\mathbb{R}^n$ y c es un escalar, entonces se cumplen las siguientes propiedades:

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
3. $c\,(\mathbf{u} \cdot \mathbf{v}) = (c\mathbf{u}) \cdot \mathbf{v}$
4. $\mathbf{u} \cdot \mathbf{u} \geq 0$ y $\mathbf{u} \cdot \mathbf{u} = 0$ si solo si $\mathbf{u} = \mathbf{0}$

Demostración. Se probarán aquí las propiedades 1. y 4. dejando las restantes como ejercicio.

Sean $\mathbf{u} = (u_1, u_2, ..., u_n)$ y $\mathbf{v} = (v_1, v_2, ..., v_n)$ entonces

1. $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + ... + u_n v_n = v_1 u_1 + v_2 u_2 + ... + v_n u_n = \mathbf{v} \cdot \mathbf{u}$.

4. Dado que $u_i^2 \geq 0$ para todo $i = 1, 2, ..., n$ entonces $\mathbf{u} \cdot \mathbf{u} = u_1^2 + u_2^2 + ... + u_n^2 \geq 0$.

   Además esta última expresión es cero si y solo si $u_1 = u_2 = ... = u_n = 0$ lo que equivale a decir que $\mathbf{u} = \mathbf{0}$.

$\square$

Definición (Espacio euclidiano $\mathbb{R}^n$). El espacio vectorial $\mathbb{R}^n$ con el producto punto es el *espacio euclidiano* $\mathbb{R}^n$.

Los vectores de $\mathbb{R}^2$ o $\mathbb{R}^3$ pueden caracterizarse como segmentos de recta dirigidos con cierta longitud y dirección. Se usará $\mathbb{R}^2$ (o $\mathbb{R}^3$) como modelo para extender a $\mathbb{R}^n$ los conceptos de longitud de un vector, distancia y ángulo entre vectores.

Longitud o norma euclidiana

Recordemos la definición de longitud de un vector en $\mathbb{R}^2$. Sea $\mathbf{u} = (u_1, u_2)$ un vector en el plano, la longitud de $\mathbf{u}$ denotada por $\|\mathbf{u}\|$ se define como $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2}$. Esta definición corresponde al concepto de longitud de la geometría euclidiana. El vector $\mathbf{u}$ es considerado como la hipotenusa de un triángulo rectángulo cuyos catetos tienen longitudes $|u_1|$ y $|u_2|$ como se ve en la Figura 5.2.

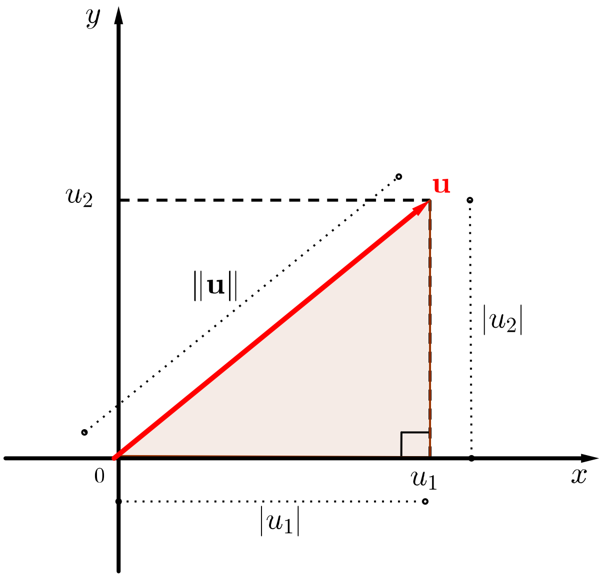


Figura 5.2: Longitud o norma del vector $\mathbf{u}$

Similarmente para vectores en $\mathbb{R}^3$, aplicando el Teorema de Pitágoras dos veces, si $\mathbf{u} = (u_1, u_2, u_3)$ entonces $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2}$.

Nótese que la longitud o norma de un vector $\mathbf{u}$ de $\mathbb{R}^2$ ó $\mathbb{R}^3$ puede expresarse en términos del producto punto del siguiente modo:

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}},$$

o en forma equivalente

$$\mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2 .$$

Tomando $\mathbb{R}^2$ y $\mathbb{R}^3$ como modelos, se define la longitud euclidiana o norma euclidiana de un vector $\mathbf{u}$ de $\mathbb{R}^n$ de la siguiente manera:

Definición (Norma euclidiana). La *norma euclidiana* o *longitud euclidiana* de un vector $\mathbf{u} = (u_1, u_2, ..., u_n)$ de $\mathbb{R}^n$ está dada por

$$\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + ... + u_n^2} = \sqrt{\mathbf{u} \cdot \mathbf{u}}$$

Ejemplo 5.1. Las normas euclidianas de los vectores $\mathbf{u} = (2, -3, 0, 4, -1)$ y $\mathbf{v} = \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)$ son:

$$\|\mathbf{u}\| = \sqrt{2^2 + (-2)^2 + 0^2 + 4^2 + (-1)^2} = \sqrt{30},$$

y

$$\|\mathbf{v}\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2} = 1,$$

respectivamente. ◀

Los vectores que tienen norma euclidiana igual a 1, como el vector $\mathbf{v}$ del ejemplo anterior, se denominan **vectores unitarios**.

Distancia euclidiana

En la Geometría euclidiana la distancia d entre dos puntos en el plano (u_1, u_2) y (v_1, v_2) es

$$d = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2}.$$

En término de vectores esta distancia es la norma o longitud del vector $\mathbf{u} - \mathbf{v}$ siendo $\mathbf{u} = (u_1, u_2)$ y $\mathbf{v} = (v_1, v_2)$ como se observa en la Figura 5.3.

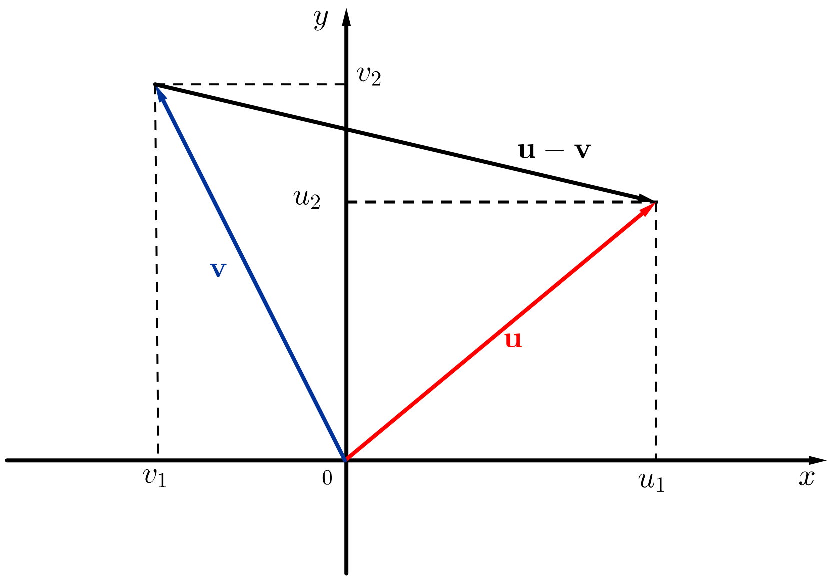


Figura 5.3: Distancia entre $\mathbf{u}$ y $\mathbf{v}$

En efecto: dado que $\mathbf{u} - \mathbf{v} = (u_1 - v_1, u_2 - v_2)$ entonces

$$\sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2} = \|\mathbf{u} - \mathbf{v}\|.$$

Por lo tanto, la distancia entre $\mathbf{u}$ y $\mathbf{v}$, $d(\mathbf{u}, \mathbf{v})$, es

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2}.$$

Similarmente la distancia entre dos vectores $\mathbf{u} = (u_1, u_2, u_3)$ y $\mathbf{v} = (v_1, v_2, v_3)$ de $\mathbb{R}^3$,

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}.$$

Se generaliza este concepto para vectores de $\mathbb{R}^n$:

Definición (Distancia euclidiana). La *distancia euclidiana* entre los vectores de $\mathbb{R}^n$, $\mathbf{u} = (u_1, u_2, ..., u_n)$ y $\mathbf{v} = (v_1, v_2, ..., v_n)$, está dada por

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + ... + (u_n - v_n)^2}.$$

Ejemplo 5.2. La distancia entre los vectores $\mathbf{u} = (1, 3, -2, 7)$ y $\mathbf{v} = (0, 7, 2, 2)$ es

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(1-0)^2 + (3-7)^2 + (-2-2)^2 + (7-2)^2} = \sqrt{58}. \blacktriangleleft$$

Nota: Es fácil comprobar que la norma satisface las siguientes propiedades, para cualesquiera vectores $\mathbf{u}$ y $\mathbf{v}$ de $\mathbb{R}^n$ y cualquier número real c:

1. $\|\mathbf{u}\| \geq 0$ y $\|\mathbf{u}\| = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$.
2. $\|c\mathbf{u}\| = |c| \, \|\mathbf{u}\|$.

A partir de éstas, dado que $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$, se deducen las siguientes propiedades de la distancia en el espacio euclidiano $\mathbb{R}^n$:

1. $d(\mathbf{u}, \mathbf{v}) \geq 0$ y $d(\mathbf{u}, \mathbf{v}) = 0 \Leftrightarrow \mathbf{u} = \mathbf{v}$.
2. $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$.

Actividad 5.1. Demuestre las propiedades de la norma y la distancia enunciadas, a partir de las propiedades del producto punto en $\mathbb{R}^n$.

Ángulo entre dos vectores

Recordemos que en $\mathbb{R}^2$ y en $\mathbb{R}^3$, el ángulo entre dos vectores $\mathbf{u}$ y $\mathbf{v}$ diferentes de cero, es el único ángulo θ tal que

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \, \|\mathbf{v}\|} \quad \text{y} \quad 0 \leq \theta \leq \pi.$$

Para extender la noción de ángulo entre vectores de $\mathbb{R}^n$, es decir usar la misma fórmula para el $\cos\theta$ que en $\mathbb{R}^2$ y $\mathbb{R}^3$, debe mostrarse que está bien definida, vale decir que el cociente $\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}$ es un número entre -1 y 1.

El siguiente teorema nos permite probar la desigualdad $-1 \leq \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} \leq 1$, necesaria para extender la noción de ángulo entre vectores de $\mathbb{R}^n$.

Teorema 5.2 (Desigualdad de Cauchy-Schwartz)**.** Si $\mathbf{u}$ y $\mathbf{v}$ son vectores del espacio euclidiano $\mathbb{R}^n$ entonces

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|.$$

donde $|\mathbf{u} \cdot \mathbf{v}|$ es el valor absoluto de $\mathbf{u} \cdot \mathbf{v}$. Además, la igualdad es válida si y solo si uno de los vectores es un múltiplo escalar del otro.

Demostración. Si $\mathbf{u} = \mathbf{0}$ entonces $\mathbf{u} \cdot \mathbf{v} = \mathbf{0} \cdot \mathbf{v} = 0$ y $\|\mathbf{u}\| = 0$ por lo tanto ambos miembros son ceros y se cumple la igualdad. Así el teorema se verifica si $\mathbf{u} = \mathbf{0}$.

Supongamos ahora que $\mathbf{u} \neq \mathbf{0}$ y consideremos el vector $x\mathbf{u}+\mathbf{v}$ donde x es un número real cualquiera. Por la propiedad 4. del Teorema 5.1

$$(x\mathbf{u} + \mathbf{v}) \cdot (x\mathbf{u} + \mathbf{v}) \geq 0. \tag{5.1}$$

Por otra parte:

$$(x\mathbf{u} + \mathbf{v}) \cdot (x\mathbf{u} + \mathbf{v}) = x^2 (\mathbf{u} \cdot \mathbf{u}) + 2x (\mathbf{u} \cdot \mathbf{v}) + (\mathbf{v} \cdot \mathbf{v}) = (\mathbf{u} \cdot \mathbf{u}) x^2 + 2 (\mathbf{u} \cdot \mathbf{v}) x + (\mathbf{v} \cdot \mathbf{v}). \tag{5.2}$$

De (5.1) y (5.2) se obtiene la siguiente desigualdad para todo $x \in \mathbb{R}$:

$$(\mathbf{u} \cdot \mathbf{u}) x^2 + 2 (\mathbf{u} \cdot \mathbf{v}) x + (\mathbf{v} \cdot \mathbf{v}) \geq 0.$$

$\mathbf{u} \cdot \mathbf{u} \neq 0$ ya que $\mathbf{u} \neq \mathbf{0}$, y por lo tanto el lado izquierdo de esta última desigualdad es un polinomio de segundo grado en x: $p(x) = ax^2 + bx + c$ con $a = \mathbf{u} \cdot \mathbf{u}$, $b = 2 (\mathbf{u} \cdot \mathbf{v})$ y $c = \mathbf{v} \cdot \mathbf{v}$. Como $a = \mathbf{u} \cdot \mathbf{u} > 0$, la gráfica de $p(x)$ es una parábola con las ramas hacia arriba y dado que $p(x) \geq 0$ para todo $x \in \mathbb{R}$, entonces $p(x)$ no tiene raíces reales, o tiene una raíz real doble. Recordando que un polinomio cuadrático tiene una raíz real doble o no tiene raíces reales si y solo si $\Delta = b^2 - 4ac \leq 0$. Entonces:

$$\text{si } \mathbf{u} \neq \mathbf{0}, \qquad 4 (\mathbf{u} \cdot \mathbf{v})^2 - 4 (\mathbf{u} \cdot \mathbf{u}) (\mathbf{v} \cdot \mathbf{v}) \leq 0,$$

o

$$(\mathbf{u} \cdot \mathbf{v})^2 \leq (\mathbf{u} \cdot \mathbf{u}) (\mathbf{v} \cdot \mathbf{v}),$$

que en términos de norma es

$$(\mathbf{u} \cdot \mathbf{v})^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2.$$

Tomando raíz cuadrada a ambos miembros que son no negativos queda probado:

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|.$$

Se deja como ejercicio probar que vale la igualdad en la desigualdad de Cauchy-Schwartz si y solo si uno de los vectores es un múltiplo escalar del otro. □

Actividad 5.2. Verifique la desigualdad de Cauchy-Schwartz para $\mathbf{u} = (1, -2, 3, 0)$ y $\mathbf{v} = (0, 4, -2, 5)$.

Utilizaremos la desigualdad de Cauchy-Schwartz para definir el ángulo entre dos vectores no cero del espacio euclidiano $\mathbb{R}^n$. Si $\mathbf{u}$ y $\mathbf{v}$ son vectores no cero de $\mathbb{R}^n$, de acuerdo a la desigualdad de Cauchy-Schwartz

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|,$$

y por lo tanto

$$-\|\mathbf{u}\| \|\mathbf{v}\| \leq \mathbf{u} \cdot \mathbf{v} \leq \|\mathbf{u}\| \|\mathbf{v}\|.$$

Ya que $\|\mathbf{u}\| \|\mathbf{v}\| > 0$, dado que $\mathbf{u} \neq \mathbf{0}$ y $\mathbf{v} \neq 0$, dividiendo entre $\|\mathbf{u}\| \|\mathbf{v}\|$ se tiene

$$-1 \leq \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1.$$

La desigualdad de Cauchy-Schwartz nos permite entonces dar por bien definida la noción de ángulo entre dos vectores no cero de $\mathbb{R}^n$ como una extensión de la ya conocida noción de ángulo entre vectores de $\mathbb{R}^2$ y $\mathbb{R}^3$:

Definición (Ángulo entre vectores). Sean $\mathbf{u}$ y $\mathbf{v}$ vectores no cero del espacio euclidiano $\mathbb{R}^n$. El *ángulo* θ entre $\mathbf{u}$ y $\mathbf{v}$ es el único ángulo tal que

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} \quad \text{y} \quad 0 \le \theta \le \pi.$$

Actividad 5.3. Encuentre el ángulo entre los vectores $\mathbf{u}$ y $\mathbf{v}$ de la actividad anterior.

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores de $\mathbb{R}^n$ diferentes de cero tales que $\mathbf{u}\cdot\mathbf{v} = 0$, de acuerdo a la definición de ángulo entre vectores, $\cos\theta = 0$ y por lo tanto $\theta = \frac{\pi}{2}$. En este caso decimos que los vectores son **perpendiculares** u **ortogonales**. En base a esta terminología daremos la siguiente definición:

Definición (Vectores ortogonales). Dos vectores $\mathbf{u}$ y $\mathbf{v}$ del espacio euclidiano $\mathbb{R}^n$ son ortogonales si $\mathbf{u}\cdot\mathbf{v} = 0$.

Aunque el ángulo entre el vector cero y otro vector no está definido, se conviene en extender la definición de vectores ortogonales para incluir al vector cero. Así, el vector cero es ortogonal a cualquier otro vector.

Ejemplo 5.3. Los vectores $\mathbf{u} = (3, 2, -1, 4)$ y $\mathbf{v} = (1, -1, 1, 0)$ son ortogonales ya que $\mathbf{u}\cdot\mathbf{v} = 0$. ◀

La desigualdad de Cauchy-Schwartz se utiliza también para demostrar una conocida desigualdad, la *desigualdad triangular*. Este nombre se deriva de la interpretación del teorema en $\mathbb{R}^2$ ilustrado en la Figura 5.4 para los vectores $\mathbf{u}$ y $\mathbf{v}$. Se considera que $\|\mathbf{u}\|$ y $\|\mathbf{v}\|$ son las longitudes de dos lados de un triángulo y por lo tanto la longitud del tercer lado es $\|\mathbf{u}+\mathbf{v}\|$.

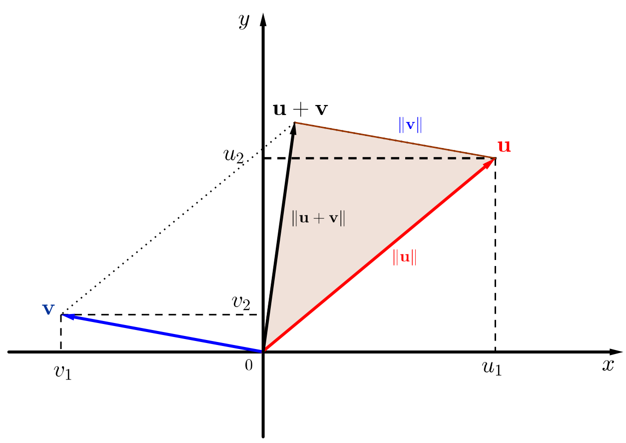


Figura 5.4: Interpretación geométrica de la desigualdad triangular

La longitud de cualquier lado de un triángulo es menor que la suma de las longitudes de los otros lados. En el siguiente teorema se muestra que esta desigualdad se puede generalizar para cualquier par de vectores en el espacio euclidiano $\mathbb{R}^n$:

Teorema 5.3 (Desigualdad triangular)**.** Si $\mathbf{u}$ y $\mathbf{v}$ son vectores del espacio euclidiano $\mathbb{R}^n$ entonces

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Demostración. Aplicando propiedades del producto punto:

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) \\
&= \mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{u}+\mathbf{u}\cdot\mathbf{v}+\mathbf{v}\cdot\mathbf{v} \\
&= \|\mathbf{u}\|^2 + 2\,(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2 \\
&\leq \|\mathbf{u}\|^2 + 2\,|\mathbf{u}\cdot\mathbf{v}| + \|\mathbf{v}\|^2 \qquad \text{(ya que } x \leq |x|\text{, para todo } x\in\mathbb{R}) \\
&\leq \|\mathbf{u}\|^2 + 2\,\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2 \qquad \text{(desigualdad de Cauchy-Schwartz } |\mathbf{u}\cdot\mathbf{v}| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|) \\
&= (\|\mathbf{u}\|+\|\mathbf{v}\|)^2.
\end{aligned}$$

Entonces para todo $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$

$$\|\mathbf{u}+\mathbf{v}\|^2 \leq (\|\mathbf{u}\|+\|\mathbf{v}\|)^2,$$

y tomando raíz cuadrada en ambos miembros que son no negativos:

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

□

Nota: En la desigualdad triangular vale el igual si y solo $\mathbf{u} = c\mathbf{v}$ con $c \geq 0$ (o $\mathbf{v} = k\mathbf{u}$ con $k \geq 0$). (Ejercicio 5).

Concluiremos esta sección con la generalización de un conocido resultado, el llamado teorema de Pitágoras.

Teorema 5.4 (de Pitágoras)**.** Los vectores $\mathbf{u}$ y $\mathbf{v}$ del espacio euclidiano $\mathbb{R}^n$ son ortogonales si y solo si

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

Demostración. Ejercicio 6. □

El nombre de teorema de Pitágoras se deriva de la interpretación del teorema para vectores de $\mathbb{R}^2$. Se considera que $\|\mathbf{u}\|$ y $\|\mathbf{v}\|$ son las longitudes de los catetos de un triángulo rectángulo y la longitud del otro lado, la hipotenusa, es $\|\mathbf{u}+\mathbf{v}\|$. Véase Figura 5.5.

El Teorema de Pitágoras asegura que *el cuadrado de la longitud de la hipotenusa es igual a la suma de los cuadrados de las longitudes de los catetos.*

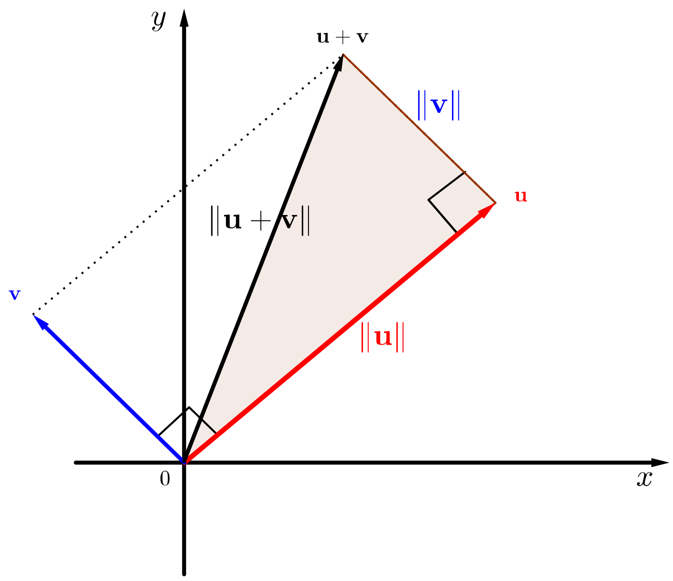


Figura 5.5: Interpretación geométrica del teorema de Pitágoras

Ejercicios 5.1

1. Sean los vectores

$$\mathbf{a} = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ -5 \\ 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 2 \\ -2 \\ 3 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{d} = \begin{bmatrix} 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}.$$

Calcule: $\|\mathbf{a}\|$, $\|\mathbf{c}\|$, $d(\mathbf{c}, \mathbf{d})$, $\mathbf{a} \cdot \mathbf{b}$, $\mathbf{c} \cdot \mathbf{d}$ y $\left(\dfrac{\mathbf{c} \cdot \mathbf{d}}{\mathbf{d} \cdot \mathbf{d}}\right) \mathbf{d}$.

2. Demuestre las restantes propiedades enunciadas en el Teorema 5.1.

3. Encuentre el ángulo que forman los vectores **a** y **b**, y, **c** y **d** del Ejercicio 1.

4. Determine qué pares de vectores son ortogonales:

   a) $\mathbf{u} = (-1, 2, 3, 0)$, $\mathbf{v} = (0, -3, 2, 4)$
   b) $\mathbf{u} = (1, 1, 1)$, $\mathbf{v} = (-1, -1, 0)$
   c) $\mathbf{u} = (0, 0, 1, -1, 5)$, $\mathbf{v} = (1, 2, -1, -1, 0)$

5. Demuestre que se verifica el igual en la desigualdad triangular si y solo si $\mathbf{u} = c\mathbf{v}$, $c \geq 0$.

6. Demuestre el Teorema 5.4 (teorema de Pitágoras).

7. Demuestre la *ley del paralelogramo*:

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2,$$

para todo **u** y **v** del espacio euclidiano $\mathbb{R}^n$.

8. Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores del espacio euclidiano $\mathbb{R}^n$. Determine la veracidad o falsedad de las siguientes proposiciones, justifique.

a) Si la distancia de $\mathbf{u}$ a $\mathbf{v}$ es igual a la distancia de $\mathbf{u}$ a $(-\mathbf{v})$, entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

b) Si $\mathbf{u} \neq \mathbf{0}$ entonces $\dfrac{\mathbf{u}}{\|\mathbf{u}\|}$ es un vector unitario.

c) Si $\mathbf{u}$ y $\mathbf{v}$ forman un ángulo $\theta = \frac{\pi}{4}$, $\mathbf{u}$ y $\mathbf{u}-\mathbf{v}$ son ortogonales y $\|\mathbf{u}\| = 2$, entonces $\|\mathbf{v}\| = 2\sqrt{2}$.

d) Si θ es el ángulo que forman los vectores $\mathbf{u}$ y $\mathbf{v}$, entonces θ es también el ángulo que forman los vectores $c\mathbf{u}$ y $d\mathbf{v}$ (c y d constantes no nulas).

5.2. Producto interno

En la sección anterior se estudió el producto punto o producto interior euclidiano sobre el espacio vectorial $\mathbb{R}^n$. En esta sección se introduce la noción de un producto interior sobre un espacio vectorial arbitrario. Como consecuencia se podrán definir con todo significado las nociones de longitud, distancia y ángulo en espacios vectoriales más generales.

En el Teorema 5.1 se mostraron las propiedades más importantes del producto punto. En un espacio vectorial general, un producto interior se define aplicando estas propiedades como axiomas.

Definición (Producto interior o interno). Un *producto interior*, o *producto interno*, sobre un espacio vectorial real $\mathbb{V}$ es una función que asigna a cada par de vectores $\mathbf{u}$ y $\mathbf{v}$ de $\mathbb{V}$, un número real que simbolizaremos $\langle \mathbf{u}, \mathbf{v} \rangle$ de tal manera que satisface los axiomas siguientes.

Para todos los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ en $\mathbb{V}$ y cualquier escalar c,

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ (axioma de simetría)
2. $\langle \mathbf{u}+\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ (axioma de aditividad)
3. $\langle c\mathbf{u}, \mathbf{v} \rangle = c \langle \mathbf{u}, \mathbf{v} \rangle$ (axioma de homogeneidad)
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$, y $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ si y solo si $\mathbf{u} = \mathbf{0}$ (axioma de positividad)

Todo espacio vectorial con un producto interior se llama *espacio con producto interior* o *espacio con producto interno.*

Ejemplo 5.4. El espacio vectorial $\mathbb{R}^n$ con $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}$ satisface todos los axiomas de un producto interior (Teorema 5.1). Por lo tanto el espacio euclidiano $\mathbb{R}^n$ es un espacio con producto interno. ◄

Las siguientes propiedades se deducen a partir de los cuatro axiomas del producto interno:

Teorema 5.5 (Propiedades del producto interno)**.** Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores en un espacio vectorial con producto interno y c cualquier escalar. Entonces:

1. $\langle \mathbf{u}, \mathbf{0} \rangle = \langle \mathbf{0}, \mathbf{u} \rangle = 0$.
2. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$.
3. $\langle \mathbf{u}, c\mathbf{v} \rangle = c \langle \mathbf{u}, \mathbf{v} \rangle$.
4. $\langle \mathbf{u} - \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle$.

Demostración. Probaremos la propiedad 2. Se dejan las restantes como ejercicio (Ejercicio 4).

$$\begin{aligned}
\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle &= \langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle && \text{(por el axioma de simetría)} \\
&= \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle && \text{(por el axioma de aditividad)} \\
&= \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle && \text{(por el axioma de simetría)}
\end{aligned}$$

□

Actividad 5.4. Verifique que en un espacio vectorial $\mathbb{V}$ con producto interno:

$$\langle c_1\mathbf{u}_1 + c_2\mathbf{u}_2 \ , \ d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + d_3\mathbf{v}_3 \rangle = \sum_{i=1}^{2}\sum_{j=1}^{3} c_i d_j \langle \mathbf{u}_i, \mathbf{v}_j \rangle ,$$

para cualesquiera $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ en $\mathbb{V}$ y c_1, c_2, d_1, d_2 y d_3 escalares.

Ejemplo 5.5. En el espacio vectorial $\mathbb{R}^2$, si $\mathbf{u} = (u_1, u_2)$ y $\mathbf{v} = (v_1, v_2)$, la asignación

$$\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2,$$

define un producto interno. En efecto:

1. Como el producto en $\mathbb{R}$ es conmutativo,

   $$\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2 = 3v_1u_1 + 2v_2u_2 = \langle \mathbf{v}, \mathbf{u} \rangle ,$$

   y se verifica entonces la simetría.

2. Si $\mathbf{w} = (w_1, w_2)$, entonces

   $$\begin{aligned}
   \langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle &= 3(u_1 + v_1)w_1 + 2(u_2 + v_2)w_2 \\
   &= 3u_1w_1 + 3v_1w_1 + 2u_2w_2 + 2v_2w_2 \\
   &= (3u_1w_1 + 2u_2w_2) + (3v_1w_1 + 2v_2w_2) \\
   &= \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle
   \end{aligned}$$

   lo cual establece el axioma de aditividad.

3. Si c es cualquier número real, entonces

$$\begin{aligned} \langle c\mathbf{u}, \mathbf{v} \rangle &= 3(cu_1)v_1 + 2(cu_2)v_2 \\ &= c(3u_1v_1 + 2u_2v_2) \\ &= c\langle \mathbf{u}, \mathbf{v} \rangle , \end{aligned}$$

   y se probó entonces el axioma de homogeneidad.

4. Ya que el cuadrado de un número real es no negativo:

$$\langle \mathbf{u}, \mathbf{u} \rangle = 3u_1u_1 + 2u_2u_2 = 3u_1^2 + 2u_2^2 \geq 0.$$

   Además $\langle \mathbf{u}, \mathbf{u} \rangle = 3u_1^2 + 2u_2^2 = 0$ si y solo si $u_1 = u_2 = 0$, es decir, si y solo $\mathbf{u} = (u_1, u_2) = \mathbf{0}$. Se verifica entonces el axioma de positividad.◀

Ejemplo 5.6. El ejemplo anterior puede generalizarse al espacio vectorial $\mathbb{R}^n$ para demostrar que: dados $\mathbf{u} = (u_1, u_2, ..., u_n)$ y $\mathbf{v} = (v_1, v_2, ..., v_n)$ entonces

$$\langle \mathbf{u}, \mathbf{v} \rangle = c_1u_1v_1 + c_2u_2v_2 + ... + c_nu_nv_n$$

define un producto interno en $\mathbb{R}^n$ si y solo si $c_i > 0$ para todo $i = 1, 2, ..., n$. (Ejercicio 1).◀

Ejemplo 5.7. La siguiente función no define un producto interno en $\mathbb{R}^3$. Sean $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ y la asignación

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 - 2u_2v_2 + u_3v_3.$$

Observe que el axioma de positividad no se satisface. En efecto,

$$\langle \mathbf{u}, \mathbf{u} \rangle = u_1^2 - 2u_2^2 + u_3^2 = (u_1^2 + u_3^2) - 2u_2^2,$$

el cual se hace negativo para aquellos vectores $\mathbf{u}$ cuyas componentes son tales que $u_1^2 + u_3^2 < 2u_2^2$. Por ejemplo, tomando $\mathbf{u} = (0, 1, 0)$ entonces $\langle \mathbf{u}, \mathbf{u} \rangle = -2 < 0$.◀

Ejemplo 5.8. En el espacio vectorial $\mathbb{P}_2$, si $p = a_0 + a_1x + a_2x^2$ y $q = b_0 + b_1x + b_2x^2$, la asignación

$$\langle p, q \rangle = a_0b_0 + a_1b_1 + a_2b_2,$$

define un producto interno (Ejercicio 2).◀

Ejemplo 5.9. El ejemplo anterior puede generalizarse al espacio vectorial $\mathbb{P}_n$ para demostrar que: dados $p = a_0 + a_1x + ... + a_nx^n$ y $q = b_0 + b_1x + ... + b_nx^n$, entonces

$$\langle p, q \rangle = a_0b_0 + a_1b_1 + ... + a_nb_n,$$

define un producto interno que denominamos **producto interno estándar o canónico** en $\mathbb{P}_n$.

Obsérvese que dicho producto interno es el producto punto entre los vectores de coordenadas de p y q con respecto a la base estándar de $\mathbb{P}_n$.◀

Ejemplo 5.10. En el espacio vectorial $\mathbb{M}_2$, si $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ y $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$, la asignación

$$\langle A, B\rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22},$$

define un producto interno (Ejercicio 3).◄

Ejemplo 5.11. En el espacio vectorial $\mathbb{M}_{mn}$, si $A = [a_{ij}]$ y $B = [b_{ij}]$ entonces $\langle A, B\rangle$ definida como la suma de los productos de los elementos de A por los homólogos de B, es un producto interno que denominaremos **producto interno estándar o canónico** en $\mathbb{M}_{mn}$.

Nótese que este producto interno es el producto punto entre los vectores de coordenadas de A y B con respecto a la base estándar de $\mathbb{M}_{mn}$.◄

Ejemplo 5.12. Sean f y g funciones del espacio vectorial $C\,[a, b]$ que consta de las funciones reales continuas en el intervalo cerrado $[a, b]$, la asignación

$$\langle f, g\rangle = \int_a^b f(x)g(x)\,dx,$$

define un producto interno en dicho espacio. En efecto utilizando propiedades del Cálculo se verifican los cuatro axiomas de la definición de producto interno.

1. $\langle f, g\rangle = \displaystyle\int_a^b f(x)g(x)\,dx = \int_a^b g(x)f(x)\,dx = \langle g, f\rangle$

2. $\langle f + g, h\rangle = \displaystyle\int_a^b (f(x) + g(x))\,h(x)\,dx = \int_a^b (f(x)h(x) + g(x)h(x))\;dx =$
   $= \displaystyle\int_a^b f(x)h(x)\,dx + \int_a^b g(x)h(x)\,dx = \langle f, h\rangle + \langle g, h\rangle$

3. $\langle cf, g\rangle = \displaystyle\int_a^b cf(x)g(x)\,dx = c\int_a^b f(x)g(x)\,dx = c\,\langle f, g\rangle$

4. Ya que $f^2(x) = f(x)f(x) \geq 0$ para toda $x \in [a, b]$ y la integral definida de una función continua no negativa es no negativa:

$$\langle f, f\rangle = \int_a^b f(x)f(x)\,dx = \int_a^b f^2(x)\,dx \geq 0.$$

Además $\langle f, f\rangle = \displaystyle\int_a^b f^2(x)\,dx = 0$ si y solo si $f^2(x) = 0$ para toda $x \in [a, b]$ y esto es si y solo si $f(x) = 0$ en $[a, b]$ y por lo tanto f es la función cero en $C\,[a, b]$.◄

La desigualdad de Cauchy-Schwartz permite introducir las nociones de norma (o longitud), distancia y ángulo en cualquier espacio con producto interno:

Teorema 5.6 (Desigualdad de Cauchy-Schwartz)**.** Si $\mathbf{u}$ y $\mathbf{v}$ son vectores de un espacio con producto interno $\mathbb{V}$, entonces

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \sqrt{\langle \mathbf{u}, \mathbf{u}\rangle}\sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$$

Demostración. Es análoga a la del Teorema 5.2. □

Nota: La igualdad en la desigualdad de Cauchy-Schwartz se verifica si y solo si $\mathbf{u}$ y $\mathbf{v}$ son vectores linealmente dependientes.

Longitud y ángulo en espacios con producto interno

Definición (Norma y distancia). Si $\mathbb{V}$ es un espacio con producto interno entonces la *norma* o *longitud* de un vector $\mathbf{u}$, se denota por $\|\mathbf{u}\|$, y es

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}.$$

La *distancia* entre los vectores $\mathbf{u}$ y $\mathbf{v}$, se denota por $d(\mathbf{u}, \mathbf{v})$, y está dada por

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}.$$

Observación 5.1. Nótese que en términos de norma, la desigualdad de Cauchy-Schwartz puede expresarse como:

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\| \, ,$$

para cualesquiera vectores $\mathbf{u}$ y $\mathbf{v}$ de un espacio $\mathbb{V}$ con producto interno.

Ejemplo 5.13. Sean $A = \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$ matrices del espacio $\mathbb{M}_2$ con el producto interno estándar. Entonces, la norma de la matriz A es

$$\|A\| = \sqrt{\langle A, A \rangle} = \sqrt{1(1) + (-2)(-2) + 0(0) + 3(3)} = \sqrt{14},$$

y la distancia entre las matrices A y B es

$$d(A, B) = \|A - B\| = \sqrt{\langle A - B, A - B \rangle} = \sqrt{0(0) + (-3)(-3) + 1(1) + 1(1)} = \sqrt{11}. \blacktriangleleft$$

Ejemplo 5.14. Sea el espacio $\mathbb{R}^2$ con el producto interno $\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2$, analizado en el Ejemplo 5.5. Si $\mathbf{u} = (1, 0)$ y $\mathbf{v} = (0, 1)$ entonces

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{3(1^2) + 2(0^2)} = \sqrt{3},$$

$$\|\mathbf{v}\| = \sqrt{3(0^2) + 2(1^2)} = \sqrt{2},$$

y

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \|(1, -1\| = \sqrt{3(1^2) + 2(-1)^2} = \sqrt{5}. \blacktriangleleft$$

Observación 5.2. Debe tenerse presente que la norma y la distancia definidas están inducidas por un producto interior. Si se cambia el producto interior, las normas y distancias también cambian. Por ejemplo, si en $\mathbb{R}^2$ se considera el producto punto, la norma de $\mathbf{u} = (1, 0)$ es igual a la norma de $\mathbf{v} = (0, 1)$, igual a 1, y la distancia entre $\mathbf{u}$ y $\mathbf{v}$ es $\sqrt{2}$ (véase ejemplo anterior).

Aunque las fórmulas con que se definen la norma y la distancia son imitaciones de las dadas en el espacio euclidiano $\mathbb{R}^n$, los resultados que se obtienen en el Ejemplo 5.14 pueden dar a lugar a dudas acerca de lo adecuado de estas definiciones, se requiere mucha imaginación para afirmar que la longitud del vector $(1,0)$ es $\sqrt{3}$ (y que la de $(0,1)$ es $\sqrt{2}$).

El argumento en apoyo de estas definiciones es que las fórmulas con que se definen norma y distancia en un espacio con producto interno satisfacen las mismas propiedades que la longitud y distancia euclidianas en $\mathbb{R}^n$.

Teorema 5.7. Si $\mathbb{V}$ es un espacio con producto interno, entonces la norma $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u}\rangle}$, y la distancia $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$, satisfacen las propiedades básicas de la norma y la distancia euclidiana en $\mathbb{R}^n$:

Propiedades básicas de la norma
Para cualesquiera vectores $\mathbf{u}$ y $\mathbf{v}$, $c \in \mathbb{R}$:

1. $\|\mathbf{u}\| \geq 0$ y $\|\mathbf{u}\| = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$.
2. $\|c\mathbf{u}\| = |c| \, \|\mathbf{u}\|$.
3. $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ (desigualdad triangular).

Propiedades básicas de la distancia
Para cualesquiera vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$:

1. $d(\mathbf{u}, \mathbf{v}) \geq 0$ y $d(\mathbf{u}, \mathbf{v}) = 0 \Leftrightarrow \mathbf{u} = \mathbf{v}$.
2. $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$.
3. $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$.

Demostración. Son análogas a las demostraciones de las mismas propiedades en el espacio euclidiano $\mathbb{R}^n$ de la sección anterior. □

Definición (Ángulo entre dos vectores). Sea $\mathbb{V}$ un espacio con producto interno, sean $\mathbf{u}$ y $\mathbf{v}$ vectores no cero de $\mathbb{V}$. El *ángulo* θ entre $\mathbf{u}$ y $\mathbf{v}$, es el único ángulo tal que

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|}, \quad \text{y} \quad 0 \leq \theta \leq \pi.$$

La desigualdad de Cauchy-Schwartz garantiza que el cociente $\frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}$ toma valores entre -1 y 1, y por lo tanto $\theta = \arccos \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}$ está bien definido.

Ejemplo 5.15. Para los vectores $\mathbf{u} = (1,-1)$ y $\mathbf{v} = (1,1)$ de $\mathbb{R}^2$ con el producto punto, el coseno del ángulo θ que forman los vectores $\mathbf{u}$ y $\mathbf{v}$ es

$$\cos\theta = \frac{(1,-1)\cdot(1,1)}{\|(1,-1)\| \, \|(1,1)\|} = 0,$$

por lo tanto $\theta = \frac{\pi}{2}$.

En cambio, si se considera el producto interno $\langle \mathbf{u}, \mathbf{v}\rangle = 3u_1v_1 + 2u_2v_2$ entonces el coseno del ángulo θ entre ellos es

$$\cos\theta = \frac{\langle (1,-1),(1,1)\rangle}{\|(1,-1)\| \, \|(1,1)\|} = \frac{1}{\sqrt{5}\sqrt{5}} = \frac{1}{5},$$

y por lo tanto $\theta = \arccos\frac{1}{5} \simeq 78,5°$.◀

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores diferentes de cero tales que $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, se deduce de la definición de ángulo entre vectores que $\cos\theta = 0$ y $\theta = \frac{\pi}{2}$. Esto sugiere la siguiente definición:

Definición (Vectores ortogonales). En un espacio $\mathbb{V}$ con producto interno, dos vectores $\mathbf{u}$ y $\mathbf{v}$ son ortogonales si $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Además si $\mathbf{u}$ es ortogonal a cada vector de un conjunto W de $\mathbb{V}$ se dice que $\mathbf{u}$ es ortogonal al conjunto W.

Debe notarse que la ortogonalidad depende del producto interior. Dos vectores pueden ser ortogonales con respecto a un producto interior y no con respecto a otro (véase Ejemplo 5.15).

Ejemplo 5.16. En el espacio $C\,[-1,1]$ de las funciones continuas en el intervalo cerrado $[-1,1]$, sea el producto interior $\langle f, g \rangle = \displaystyle\int_{-1}^{1} f(x)g(x)\,dx$ analizado en el Ejemplo 5.12. Las funciones f y g tales que $f(x) = x$ y $g(x) = x^2$ son ortogonales. En efecto:

$$\langle f, g \rangle = \int_{-1}^{1} x x^2\,dx = \int_{-1}^{1} x^3\,dx = 0.$$

Claramente $f(x) = x$ es ortogonal al conjunto W formado por todas las funciones de la forma $g_{2n}(x) = x^{2n}$ con $n \in \mathbb{N}$.◄

El siguiente teorema es una generalización de un conocido resultado:

Teorema 5.8 (de Pitágoras generalizado)**.** Sea $\mathbb{V}$ un espacio con producto interno y $\mathbf{u}$ y $\mathbf{v}$ vectores de $\mathbb{V}$. Entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales si y solo si $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

Demostración. Es análoga a la del Teorema 5.4. □

Proyecciones ortogonales en espacios con producto interno

Recordemos que geométricamente para vectores $\mathbf{u} = \overrightarrow{OP}$ y $\mathbf{v} = \overrightarrow{OQ}$, $\mathbf{v} \neq \mathbf{0}$, del espacio euclidiano $\mathbb{R}^2$ (similarmente para vectores del espacio euclidiano $\mathbb{R}^3$), el vector **proyección ortogonal de u sobre v**, es el vector $\overrightarrow{OP'}$, que simbolizamos $\text{proy}_{\mathbf{v}}\mathbf{u}$, donde P' es el punto de intersección de la recta sostén de $\mathbf{v}$ y la perpendicular a ella que contiene al punto P (véase Figura 5.6).

En Álgebra y Geometría Analítica, se vió que este vector está dado por

$$\text{proy}_{\mathbf{v}}\mathbf{u} = \left(\mathbf{u} \cdot \frac{\mathbf{v}}{\|\mathbf{v}\|}\right) \frac{\mathbf{v}}{\|\mathbf{v}\|} \qquad \text{o bien} \qquad \text{proy}_{\mathbf{v}}\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}\mathbf{v}.$$

El concepto de proyección ortogonal se extiende de manera natural a un espacio general con producto interno.

Definición (Proyección ortogonal sobre un vector). Sea $\mathbb{V}$ un espacio con producto interno, la proyección ortogonal de un vector $\mathbf{u}$ sobre un vector $\mathbf{v}$ no cero, es el vector

$$\text{proy}_{\mathbf{v}}\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\mathbf{v}.$$

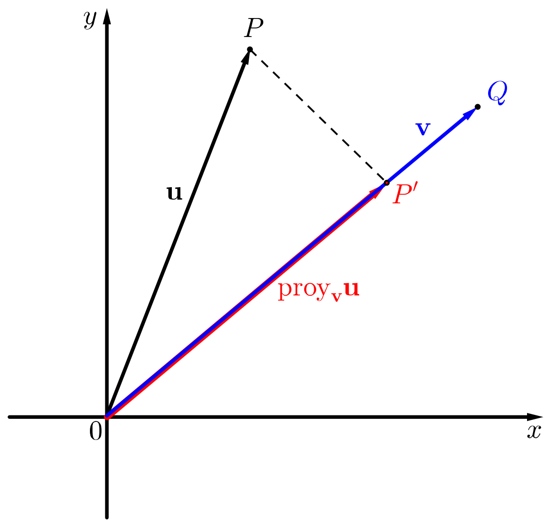
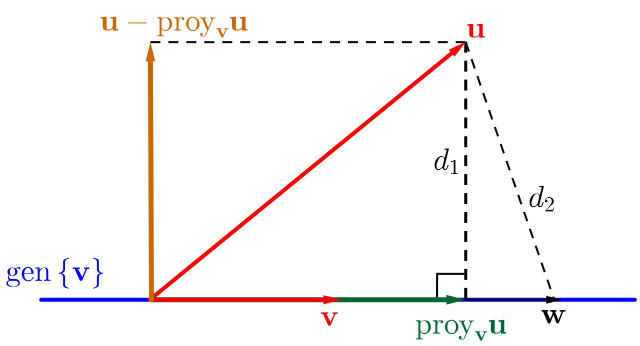


Figura 5.6: Proyección ortogonal de un vector sobre otro en $\mathbb{R}^2$

El vector proyección $\text{proy}_{\mathbf{v}}\mathbf{u}$ es un múltiplo escalar de $\mathbf{v}$, por lo tanto es un vector del subespacio $\text{gen}\,\{\mathbf{v}\}$. Así, mas apropiadamente, a este vector se lo puede considerar como la proyección de $\mathbf{u}$ sobre el subespacio generado por $\mathbf{v}$. La proyección ortogonal de $\mathbf{u}$ sobre $\mathbf{v}$ goza de las siguientes propiedades:

1. Es el único vector $\mathbf{w}$ en $\text{gen}\,\{\mathbf{v}\}$ tal que $\mathbf{u}-\mathbf{w}$ es ortogonal al espacio $\text{gen}\,\{\mathbf{v}\}$.
2. Es el (único) vector de $\text{gen}\,\{\mathbf{v}\}$ "más próximo" al vector $\mathbf{u}$.

Estas ideas se visualizan en la Figura 5.7 y se prueban en el siguiente teorema:



$$d_1 = d\left(\mathbf{u}, \text{proy}_{\mathbf{v}}\mathbf{u}\right)$$

$$d_2 = d\left(\mathbf{u}, \mathbf{w}\right)$$

Figura 5.7: Proyección ortogonal sobre un vector

Teorema 5.9. Sea $\mathbb{V}$ un espacio con producto interno y $\mathbf{u}$ y $\mathbf{v}$ vectores de $\mathbb{V}$ con $\mathbf{v} \neq \mathbf{0}$. Entonces:

1. $\langle \mathbf{u} - \text{proy}_{\mathbf{v}}\mathbf{u}, \mathbf{w} \rangle = 0$, para todo $\mathbf{w} \in \text{gen}\,\{\mathbf{v}\}$.
2. Si $\mathbf{w} \in \text{gen}\,\{\mathbf{v}\}$ y $\mathbf{w} \neq \text{proy}_{\mathbf{v}}\mathbf{u}$, entonces $d\,(\mathbf{u}, \text{proy}_{\mathbf{v}}\mathbf{u}) < d\,(\mathbf{u}, \mathbf{w})$.

Demostración.

1. Sea $\mathbf{w}$ un vector de $\text{gen}\,\{\mathbf{v}\}$. Entonces, $\mathbf{w} = c\mathbf{v}$ para algún $c \in \mathbb{R}$. Luego: $\langle \mathbf{u} - \text{proy}_{\mathbf{v}}\mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{u} - \text{proy}_{\mathbf{v}}\mathbf{u}, c\mathbf{v} \rangle = c\,\langle \mathbf{u}, \mathbf{v} \rangle - c\,\langle \text{proy}_{\mathbf{v}}\mathbf{u}, \mathbf{v} \rangle = c\left(\langle \mathbf{u}, \mathbf{v} \rangle - \left\langle \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\mathbf{v}, \mathbf{v} \right\rangle\right) =$
$= c\left(\langle \mathbf{u}, \mathbf{v} \rangle - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\langle \mathbf{v}, \mathbf{v} \rangle\right) = c\,(\langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{v} \rangle) = 0.$

   Es fácil probar que no existe otro vector en $\text{gen}\,\{\mathbf{v}\}$ que satisface esta propiedad.

2. La justificación de este hecho la proporciona el teorema generalizado de Pitágoras.

   Sea $\mathbf{w} \in \text{gen}\,\{\mathbf{v}\}$ y $\mathbf{w} \neq \text{proy}_{\mathbf{v}}\mathbf{u}$. Probar que $d\,(\mathbf{u}, \text{proy}_{\mathbf{v}}\mathbf{u}) < d\,(\mathbf{u}, \mathbf{w})$, por definición de distancia, significa mostrar que

$$\|\mathbf{u} - \text{proy}_{\mathbf{v}}\mathbf{u}\| < \|\mathbf{u} - \mathbf{w}\|\,.$$

   Para ello, sumando y restando $\text{proy}_{\mathbf{v}}\mathbf{u}$ al vector diferencia $\mathbf{u} - \mathbf{w}$ se tiene:

$$\mathbf{u} - \mathbf{w} = (\mathbf{u} - \text{proy}_{\mathbf{v}}\mathbf{u}) + (\text{proy}_{\mathbf{v}}\mathbf{u} - \mathbf{w})\,.$$

   El vector $\text{proy}_{\mathbf{v}}\mathbf{u} - \mathbf{w}$ es un vector de $\text{gen}\,\{\mathbf{v}\}$ por ser resta de dos vectores de $\text{gen}\,\{\mathbf{v}\}$. Dado que el vector $\mathbf{u} - \text{proy}_{\mathbf{v}}\mathbf{u}$ es ortogonal a todo vector de $\text{gen}\,\{\mathbf{v}\}$, entonces $(\mathbf{u} - \text{proy}_{\mathbf{v}}\mathbf{u})$ y $(\text{proy}_{\mathbf{v}}\mathbf{u} - \mathbf{w})$ son ortogonales. En consecuencia, puede aplicarse el teorema generalizado de Pitágoras para obtener:

$$\|\mathbf{u} - \mathbf{w}\|^2 = \|(\mathbf{u} - \text{proy}_{\mathbf{v}}\mathbf{u}) + (\text{proy}_{\mathbf{v}}\mathbf{u} - \mathbf{w})\|^2 = \|(\mathbf{u} - \text{proy}_{\mathbf{v}}\mathbf{u})\|^2 + \|(\text{proy}_{\mathbf{v}}\mathbf{u} - \mathbf{w})\|^2\,.$$

   Como $\mathbf{w} \neq \text{proy}_{\mathbf{v}}\mathbf{u}$, entonces $\|\text{proy}_{\mathbf{v}}\mathbf{u} - \mathbf{w}\|^2 > 0$ y por lo tanto

$$\|\mathbf{u} - \mathbf{w}\|^2 > \|\mathbf{u} - \text{proy}_{\mathbf{v}}\mathbf{u}\|^2\,,$$

   y tomando raíz cuadrada a ambos miembros

$$\|\mathbf{u} - \mathbf{w}\| > \|\mathbf{u} - \text{proy}_{\mathbf{v}}\mathbf{u}\|\,,$$

   es decir

$$d\,(\mathbf{u}, \mathbf{w}) > d\,(\mathbf{u}, \text{proy}_{\mathbf{v}}\mathbf{u})\,.$$

□

Ejercicios 5.2

1. Demuestre lo enunciado en el Ejemplo 5.6.

2. Demuestre lo enunciado en el Ejemplo 5.8

3. Demuestre lo enunciado en el Ejemplo 5.10.

4. Demuestre las restantes propiedades del Teorema 5.5.

5. Suponga que el espacio $\mathbb{R}^2$ tiene el producto interno definido en el Ejemplo 5.5, y sean los vectores $\mathbf{u} = (1, 1)$ y $\mathbf{v} = (5, -1)$.

   a) Calcule $\|\mathbf{u}\|$, $\|\mathbf{v}\|$, $\langle \mathbf{u}, \mathbf{v} \rangle$, la distancia entre $\mathbf{u}$ y $\mathbf{v}$ y el ángulo entre $\mathbf{u}$ y $\mathbf{v}$.

   b) Verifique la desigualdad de Cauchy-Schwartz para los vectores $\mathbf{u}$ y $\mathbf{v}$.

   c) Describa todos los vectores $\mathbf{w} = (w_1, w_2)$ de $\mathbb{R}^2$ que son ortogonales a $\mathbf{v}$.

   d) Determine la proyección ortogonal de $\mathbf{u}$ sobre $\mathbf{v}$.

6. En el espacio $\mathbb{P}_2$ considere el producto interno estándar, y sean los polinomios p y q dados por $p(x) = 4 + x$ y $q(x) = 3 - 2x + x^2$.

   a) Calcule $\langle p, q \rangle$, $\|p\|$ y $\|q\|$.

   b) Halle el ángulo entre los polinomios p y q.

   c) ¿Cuáles de los polinomios p y/o q es ortogonal al polinomio r? siendo $r(x) = -1 + 4x + 5x^2$.

   d) Determine la proyección ortogonal de q sobre p.

   e) Halle el polinomio de gen $\{p\}$ “más próximo” al polinomio q.

7. Sean x_0, x_1,..., x_n números reales distintos. En el espacio vectorial $\mathbb{P}_n$, considere la función definida por

$$\langle p, q \rangle = p(x_0)q(x_0) + p(x_1)q(x_1) + ... + p(x_n)q(x_n),$$

   donde p y q son polinomios de $\mathbb{P}_n$.

   a) Demuestre que dicha función es un producto interno en $\mathbb{P}_n$.

   b) Considere el espacio vectorial $\mathbb{P}_2$ con el producto interno definido siendo $x_0 = 1$, $x_1 = 2$ y $x_3 = -1$. Sean los polinomios $p = 12x^2$ y $q = -1 + 2x$.

      i) Calcule $\langle p, q \rangle$, $\langle q, q \rangle$, $\|p\|$ y $\|q\|$.

      ii) Halle el ángulo entre p y q.

      iii) Obtenga la proyección ortogonal de p sobre q.

      iv) Encuentre un polinomio r que sea ortogonal a p.

8. En el espacio vectorial $C\,[0, \pi]$, considere el producto interno definido en el Ejemplo 5.12, y sean las funciones f, g y h dadas por $f(x) = x$, $g(x) = e^x$ y $h(x) = \text{sen }x$.

   a) Obtenga las normas de las funciones f, g y h.

   b) Calcule la distancia entre f y h.

   c) Halle la proyección de f sobre g.

   d) ¿Es la función f ortgonal a la función g y/o h?, justifique su respuesta.

9. Sea $\mathbb{V}$ un espacio con producto interno. Demuestre que $\mathbf{u}$ es ortogonal al espacio gen $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ si y solo si es ortogonal a cada uno de los vectores $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k$.

10. Sea $\mathbb{V}$ un espacio con producto interno. Pruebe que el único vector de $\mathbb{V}$ que es ortogonal a cada vector de $\mathbb{V}$ es el vector cero.

11. Sea $\mathbb{V}$ un espacio con producto interno, demuestre que si $\mathbf{u}$ y $\mathbf{v}$ son ortogonales y $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$, entonces $d(\mathbf{u}, \mathbf{v}) = \sqrt{2}$.

12. Sea $\mathbf{u}$ un vector de un espacio con producto interno $\mathbb{V}$ y sea W el conjunto de todos los vectores de $\mathbb{V}$ que son ortogonales a $\mathbf{u}$.

    a) Pruebe que W es un subespacio de $\mathbb{V}$.

    b) Describa geométricamente el subespacio W en el caso en que $\mathbb{V}$ es el espacio euclidiano $\mathbb{R}^2$ o $\mathbb{R}^3$ y $\mathbf{u} \neq \mathbf{0}$.

13. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores de un espacio con producto interno. Demuestre que

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}\|\mathbf{u}+\mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u}-\mathbf{v}\|^2 .$$

5.3. Bases ortogonales y ortonormales. Proceso de Gram-Schmidt

En muchos problemas referentes a espacios vectoriales la selección de una base se hace según la conveniencia. La mejor estrategia es elegir una base que simplifique la solución del problema en cuestión.

En los espacios con productos interiores el mejor procedimiento, a menudo, es elegir una base en la que todos los vectores sean ortogonales entre sí. En esta sección se verá cómo contruir tales bases.

Definición (Conjuntos ortogonales y ortonormales). Un conjunto de vectores en un espacio $\mathbb{V}$ con producto interno es un *conjunto ortogonal* si todos los pares de vectores distintos son ortogonales. Un conjunto ortogonal en el que cada vector tiene norma uno se conoce como *conjunto ortonormal*.

Así el conjunto $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ es *ortogonal* si $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ para $i \neq j$; y es *ortonormal* si además $\langle \mathbf{u}_i, \mathbf{u}_i \rangle = 1$ para todo $i = 1, 2, ..., k$.

Ejemplo 5.17. En el espacio euclidiano $\mathbb{R}^3$, el conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ donde $\mathbf{u}_1 = (0, 1, 0)$, $\mathbf{u}_2 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$, $\mathbf{u}_3 = (\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$ es un conjunto ortonormal ya que:

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \mathbf{u}_1 \cdot \mathbf{u}_2 = 0,$$

$$\langle \mathbf{u}_1, \mathbf{u}_3 \rangle = \mathbf{u}_1 \cdot \mathbf{u}_3 = 0,$$

$$\langle \mathbf{u}_2, \mathbf{u}_3 \rangle = \mathbf{u}_2 \cdot \mathbf{u}_3 = 0,$$

y además $\|\mathbf{u}_1\| = \|\mathbf{u}_2\| = \|\mathbf{u}_3\| = 1$.◀

Ejemplo 5.18. En el espacio $\mathbb{M}_2$ con el producto interno estándar, el conjunto $S = \{A, B, C\}$ donde

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -2 & -2 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix},$$

no es un conjunto ortogonal. En efecto: $\langle A, B \rangle = 0$, $\langle B, C \rangle = 0$ pero $\langle A, C \rangle = 2 \neq 0$.◀

Nota: Debe observarse que todo conjunto unitario, $S = \{\mathbf{v}\}$, en un espacio con producto interno es un conjunto ortogonal dado que no existen dos vectores distintos en S cuyo producto interno no sea cero.

Ejemplo 5.19. En un espacio con producto interior, si $\mathbf{v} \neq \mathbf{0}$, entonces $\frac{1}{\|\mathbf{v}\|}\mathbf{v}$ es un vector de norma uno dado que

$$\left\| \frac{1}{\|\mathbf{v}\|}\mathbf{v} \right\| = \left| \frac{1}{\|\mathbf{v}\|} \right| \|\mathbf{v}\| = 1. \blacktriangleleft$$

El proceso de multiplicar un vector $\mathbf{v}$ no cero por el recíproco de su longitud para obtener un vector de norma uno, se denomina **normalización del vector v**. Es claro que cualquier conjunto ortogonal de vectores no ceros se transforma en un conjunto ortonormal si se normaliza cada vector del conjunto.

Los vectores ortogonales tienen muchas propiedades, la principal se dá en el siguiente teorema:

Teorema 5.10. Si $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ es un conjunto ortogonal de vectores no cero en un espacio $\mathbb{V}$ con producto interno, entonces S es linealmente independiente.

En particular, todo conjunto ortonormal es linealmente independiente.

Demostración. Suponga que c_1, c_2,..., c_k son escalares tales que

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k = \mathbf{0}. \tag{5.3}$$

Para demostrar que S es linealmente independiente, debe probarse que $c_1 = c_2 = ... = c_k = 0$.

Para cada $\mathbf{u}_i$ de S, de acuerdo a (5.3):

$$\langle c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k, \mathbf{u}_i\rangle = \langle \mathbf{0}, \mathbf{u}_i\rangle = 0.$$

Desarrollando el producto interno del primer miembro se obtiene

$$c_1 \langle \mathbf{u}_1, \mathbf{u}_i\rangle + c_2 \langle \mathbf{u}_2, \mathbf{u}_i\rangle + ... + c_k \langle \mathbf{u}_k, \mathbf{u}_i\rangle = 0.$$

Dado que S es un conjunto ortogonal $\langle \mathbf{u}_j, \mathbf{u}_i\rangle = 0$ si $j \neq i$, por lo tanto la suma del primer miembro se reduce al único término $c_i \langle \mathbf{u}_i, \mathbf{u}_i\rangle$ y la última igualdad se reduce a

$$c_i \langle \mathbf{u}_i, \mathbf{u}_i\rangle = 0.$$

Dado que los vectores de S son no cero, por el axioma de positividad es $\langle \mathbf{u}_i, \mathbf{u}_i\rangle > 0$, y por lo tanto $c_i = 0$. Como esto se verifica para todo $\mathbf{u}_i$ de S, entonces $c_1 = c_2 = ... = c_k = 0$ y luego S es linealmente independiente.

Se completa la prueba del teorema observando que todo conjunto ortonormal es un conjunto ortogonal de vectores no cero y por lo tanto es linealmente independiente. □

Corolario 5.1. Si $\mathbb{V}$ es un espacio con producto interno de dimensión n entonces todo conjunto ortogonal de n vectores no cero es una base de $\mathbb{V}$.

Actividad 5.5. Justifique la afirmación del corolario anterior.

Definición (Bases ortogonales y ortonormales). Una *base ortogonal* en un espacio $\mathbb{V}$ con producto interno es un conjunto ortogonal que constituye una base del mismo. Si además dichos vectores tienen norma uno, dicha base es una *base ortonormal*.

Ejemplo 5.20. El conjunto S del Ejemplo 5.17 es una base ortonormal del espacio euclidiano $\mathbb{R}^3$ ya que $\dim\left(\mathbb{R}^3\right) = 3$ y S es un conjunto ortonormal que consta de 3 vectores.◀

El interés en encontrar bases ortonormales para espacios con producto interno está en parte motivado por el teorema que sigue, el cual muestra que puede ser sencillo expresar un vector en términos de una base ortogonal.

Teorema 5.11. Si $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ es una base ortogonal para un espacio $\mathbb{V}$ con producto interno y $\mathbf{v}$ es cualquier vector de $\mathbb{V}$ entonces

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 + \frac{\langle \mathbf{v}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2 + ... + \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\mathbf{u}_n.$$

Si la base S es ortonormal entonces la expresión anterior se reduce a

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{v}, \mathbf{u}_2 \rangle \mathbf{u}_2 + ... + \langle \mathbf{v}, \mathbf{u}_n \rangle \mathbf{u}_n.$$

Demostración. Sea $\mathbf{v}$ un vector de $\mathbb{V}$. Dado que S es una base de $\mathbb{V}$, existen únicos escalares c_1, c_2,..., c_n tales que

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n. \tag{5.4}$$

Se demostrará que $c_i = \dfrac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle}$ para cada $i = 1, 2, ..., n$.

De acuerdo a (5.4) y por propiedades del producto interno, para cada vector $\mathbf{u}_i$ en S:

$$\langle \mathbf{v}, \mathbf{u}_i \rangle = \langle c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n, \mathbf{u}_i \rangle = c_1 \langle \mathbf{u}_1, \mathbf{u}_i \rangle + c_2 \langle \mathbf{u}_2, \mathbf{u}_i \rangle + ... + c_n \langle \mathbf{u}_n, \mathbf{u}_i \rangle .$$

Dado que S es un conjunto ortogonal, $\langle \mathbf{u}_j, \mathbf{u}_i \rangle = 0$ si $j \neq i$. Por lo tanto la suma anterior se reduce al único término $c_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle$. Así

$$\langle \mathbf{v}, \mathbf{u}_i \rangle = c_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle .$$

Por ser S una base, $\mathbf{u}_i \neq \mathbf{0}$, luego $\langle \mathbf{u}_i, \mathbf{u}_i \rangle \neq 0$ por lo cual

$$c_i = \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle}, \qquad \text{para cada } i = 1, 2, ..., n.$$

Si S es una base ortonormal entonces $\langle \mathbf{u}_i, \mathbf{u}_i \rangle = 1$ y en este caso

$$c_i = \langle \mathbf{v}, \mathbf{u}_i \rangle , \qquad \text{para todo } i = 1, 2, ..., n.$$

□

Ejemplo 5.21. Los vectores $\mathbf{u}_1 = (0,1,0)$, $\mathbf{u}_2 = \left(-\frac{4}{5}, 0, \frac{3}{5}\right)$ y $\mathbf{u}_3 = \left(\frac{3}{5}, 0, \frac{4}{5}\right)$ forman una base ortonormal para el espacio euclidiano $\mathbb{R}^3$ (verifíquelo).

Consideremos el vector $\mathbf{v} = (1,1,1)$. Por el teorema anterior, el vector $\mathbf{v}$ expresado en términos de la base ortonormal es

$$\mathbf{v} = (\mathbf{v}\cdot\mathbf{u}_1)\,\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\,\mathbf{u}_2 + (\mathbf{v}\cdot\mathbf{u}_3)\,\mathbf{u}_3.$$

Se verifica que $\mathbf{v}\cdot\mathbf{u}_1 = 1$, $\mathbf{v}\cdot\mathbf{u}_2 = -\frac{1}{5}$ y $\mathbf{v}\cdot\mathbf{u}_3 = \frac{7}{5}$. Así,

$$\mathbf{v} = (1,1,1) = 1\,(0,1,0) - \frac{1}{5}\left(-\frac{4}{5}, 0, \frac{3}{5}\right) + \frac{7}{5}\left(\frac{3}{5}, 0, \frac{4}{5}\right) = 1\mathbf{u}_1 - \frac{1}{5}\mathbf{u}_2 + \frac{7}{5}\mathbf{u}_3. \blacktriangleleft$$

Es evidente la utilidad del teorema si se tiene presente que por lo general es necesario resolver un sistema de ecuaciones lineales para expresar un vector en términos de una base que no es ortogonal.

Ahora se considera el problema de construir bases ortonormales. Para ello se deben considerar los siguientes resultados preliminares:

Teorema 5.12. Sea $\mathbb{V}$ un espacio con producto interno y $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ un conjunto ortogonal de vectores no cero en $\mathbb{V}$. Si $\mathbb{W} = \text{gen}(S)$ entonces todo vector $\mathbf{v}$ en $\mathbb{V}$ se puede expresar en la forma

$$\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2,$$

donde $\mathbf{w}_1$ está en $\mathbb{W}$ y $\mathbf{w}_2 = \mathbf{v} - \mathbf{w}_1$ es ortogonal a $\mathbb{W}$ siendo

$$\mathbf{w}_1 = \frac{\langle \mathbf{v}, \mathbf{u}_1\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 + \frac{\langle \mathbf{v}, \mathbf{u}_2\rangle}{\langle \mathbf{u}_2, \mathbf{u}_2\rangle}\mathbf{u}_2 + ... + \frac{\langle \mathbf{v}, \mathbf{u}_k\rangle}{\langle \mathbf{u}_k, \mathbf{u}_k\rangle}\mathbf{u}_k. \tag{5.5}$$

En particular $\mathbf{w}_1$ y $\mathbf{w}_2$ son ortogonales.

Demostración. Claramente $\mathbf{w}_1$ está en $\mathbb{W}$ ya que es una combinación lineal de vectores de S. Debe probarse que $\mathbf{w}_2 = \mathbf{v} - \mathbf{w}_1$ es ortogonal a $\mathbb{W}$. Para ello, basta probar que $\mathbf{w}_2$ es ortogonal a cada uno de los vectores de S (véase Ejercicio 9 de la sección anterior).

Sea $\mathbf{u}_i$ un vector de S entonces

$$\begin{aligned}
\langle \mathbf{v} - \mathbf{w}_1, \mathbf{u}_i\rangle &= \langle \mathbf{v}, \mathbf{u}_i\rangle - \langle \mathbf{w}_1, \mathbf{u}_i\rangle \\
&= \langle \mathbf{v}, \mathbf{u}_i\rangle - \left\langle \frac{\langle \mathbf{v}, \mathbf{u}_1\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 + \frac{\langle \mathbf{v}, \mathbf{u}_2\rangle}{\langle \mathbf{u}_2, \mathbf{u}_2\rangle}\mathbf{u}_2 + ... + \frac{\langle \mathbf{v}, \mathbf{u}_k\rangle}{\langle \mathbf{u}_k, \mathbf{u}_k\rangle}\mathbf{u}_k, \mathbf{u}_i \right\rangle \\
&= \langle \mathbf{v}, \mathbf{u}_i\rangle - \left(\frac{\langle \mathbf{v}, \mathbf{u}_1\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\langle \mathbf{u}_1, \mathbf{u}_i\rangle + \frac{\langle \mathbf{v}, \mathbf{u}_2\rangle}{\langle \mathbf{u}_2, \mathbf{u}_2\rangle}\langle \mathbf{u}_2, \mathbf{u}_i\rangle + ... + \frac{\langle \mathbf{v}, \mathbf{u}_k\rangle}{\langle \mathbf{u}_k, \mathbf{u}_k\rangle}\langle \mathbf{u}_k, \mathbf{u}_i\rangle \right).
\end{aligned}$$

Como S es un conjunto ortogonal, $\langle \mathbf{u}_j, \mathbf{u}_i\rangle = 0$ si $j \neq i$, y la expresión entre paréntesis se reduce a $\frac{\langle \mathbf{v}, \mathbf{u}_i\rangle}{\langle \mathbf{u}_i, \mathbf{u}_i\rangle}\langle \mathbf{u}_i, \mathbf{u}_i\rangle = \langle \mathbf{v}, \mathbf{u}_i\rangle$. Por lo tanto:

$$\langle \mathbf{v} - \mathbf{w}_1, \mathbf{u}_i\rangle = \langle \mathbf{v}, \mathbf{u}_i\rangle - \langle \mathbf{v}, \mathbf{u}_i\rangle = 0.$$

Queda probado que si $\mathbf{w}_1$ es la combinación lineal dada en la ecuación (5.5), entonces $\mathbf{w}_2 = \mathbf{v} - \mathbf{w}_1$ es ortogonal a todo vector del espacio $\mathbb{W}$. En particular $\mathbf{w}_1$ y $\mathbf{w}_2$ son ortogonales. $\square$

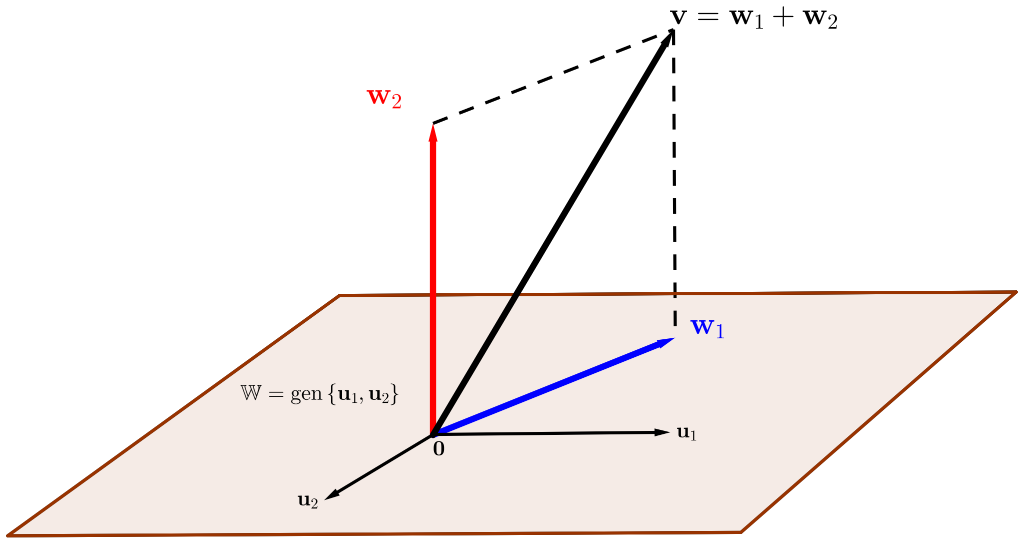


Figura 5.8: Interpretación geométrica en el espacio euclidiano $\mathbb{R}^3$ del Teorema 5.12

Con la motivación de la Figura 5.8, al vector $\mathbf{w}_1$ se le dá el nombre de *proyección ortogonal de* $\mathbf{v}$ *sobre* $\mathbb{W}$ y se lo denota $\text{proy}_{\mathbb{W}}\mathbf{v}$. El vector $\mathbf{w}_2$ se conoce como *componente de* $\mathbf{v}$ *ortogonal a* $\mathbb{W}$. Se tiene entonces la siguiente definición:

Definición. Sean $\mathbf{v}$ un vector de un espacio $\mathbb{V}$ con producto interno y $\mathbb{W}$ un subespacio de $\mathbb{V}$ con una base ortogonal $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$. La *proyección ortogonal de* $\mathbf{v}$ *sobre* $\mathbb{W}$ es el vector

$$\text{proy}_{\mathbb{W}}\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 + \frac{\langle \mathbf{v}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2 + ... + \frac{\langle \mathbf{v}, \mathbf{u}_k \rangle}{\langle \mathbf{u}_k, \mathbf{u}_k \rangle}\mathbf{u}_k.$$

El vector diferencia $\mathbf{v} - \text{proy}_{\mathbb{W}}\mathbf{v}$ se llama *componente de* $\mathbf{v}$ *ortogonal a* $\mathbb{W}$.

Ejemplo 5.22. Sea el espacio euclidiano $\mathbb{R}^3$ y $\mathbb{W}$ el subespacio (el plano) generado por los vectores ortonormales $\mathbf{u}_1 = (0, 1, 0)$ y $\mathbf{u}_2 = \left(\frac{4}{5}, 0, \frac{3}{5}\right)$. La proyección ortogonal del vector $\mathbf{v} = (1, 1, 1)$ sobre $\mathbb{W}$ es

$$\text{proy}_{\mathbb{W}}\mathbf{v} = (\mathbf{v} \cdot \mathbf{u}_1)\, \mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\, \mathbf{u}_2 = 1\,(0, 1, 0) + \frac{7}{5}\left(\frac{4}{5}, 0, \frac{3}{5}\right) = \left(\frac{28}{25}, 1, \frac{21}{25}\right).$$

La componente de $\mathbf{v}$ ortogonal a $\mathbb{W}$ es

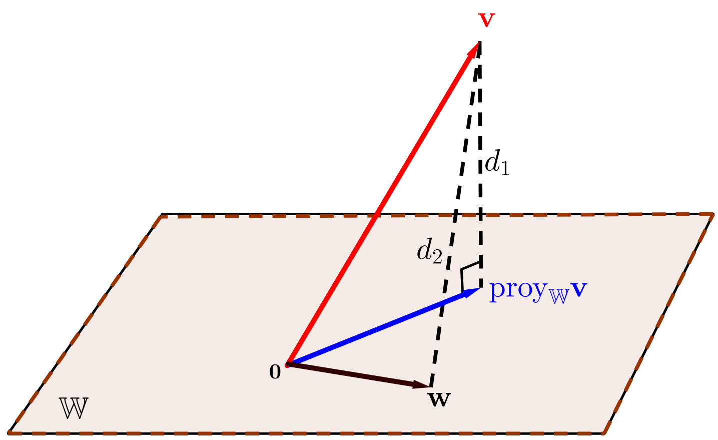
$$\mathbf{v} - \text{proy}_{\mathbb{W}}\mathbf{v} = (1, 1, 1) - \left(\frac{28}{25}, 1, \frac{21}{25}\right) = \left(-\frac{3}{25}, 0, \frac{4}{25}\right).$$

Obsérvese que $\mathbf{v} - \text{proy}_{\mathbb{W}}\mathbf{v}$ es ortogonal tanto a $\mathbf{u}_1$ como a $\mathbf{u}_2$, de manera que este vector es ortogonal a cada vector en el espacio $\mathbb{W}$, como debe ser.◀

Actividad 5.6. Pruebe que el conjunto $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ donde $\mathbf{v}_1 = (0, 2, 0)$ y $\mathbf{v}_2 = (4, 0, 3)$ es una base ortogonal del subespacio $\mathbb{W}$ del ejemplo anterior. Y muestre que si calcula el vector proyección de $\mathbf{v} = (1, 1, 1)$ sobre $\mathbb{W}$ usando la base S se obtiene el mismo vector que el del ejemplo anterior, que fue calculado utilizando la base $\{\mathbf{u}_1, \mathbf{u}_2\}$.

Este no es casual, el cálculo de un vector proyección sobre un subespacio es independiente de la base ortogonal considerada.

En la sección anterior se notó que el vector proyección de $\mathbf{u}$ sobre $\mathbf{v}$ es el vector del espacio gen $\{\mathbf{v}\}$ más próximo al vector $\mathbf{u}$ (véase Figura 5.7). Similarmente, el vector $\text{proy}_{\mathbb{W}}\mathbf{v}$ tiene la misma interesante propiedad. Es el vector en $\mathbb{W}$ más cercano a $\mathbf{v}$, es decir, es la mejor aproximación de los vectores de $\mathbb{W}$ al vector $\mathbf{v}$ (véase Figura 5.9).



$$d_1 = d(\mathbf{v}, \text{proy}_{\mathbb{W}}\mathbf{v}) = \|\mathbf{v} - \text{proy}_{\mathbb{W}}\mathbf{v}\|$$
$$d_2 = d(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\|$$

Figura 5.9: La proyección ortogonal es el vector de $\mathbb{W}$ más cercano a $\mathbf{v}$

Teorema 5.13. (de la mejor aproximación) Si $\mathbb{W}$ es un subespacio de dimensión finita de un espacio $\mathbb{V}$ con producto interno, y si $\mathbf{v}$ es un vector en $\mathbb{V}$, entonces $\text{proy}_{\mathbb{W}}\mathbf{v}$ es el vector de $\mathbb{W}$ más cercano a $\mathbf{v}$, esto es:

$$\|\mathbf{v} - \text{proy}_{\mathbb{W}}\mathbf{v}\| < \|\mathbf{v} - \mathbf{w}\|,$$

para todo vector $\mathbf{w}$ en $\mathbb{W}$ distinto de $\text{proy}_{\mathbb{W}}\mathbf{v}$.

Así, $\|\mathbf{v} - \text{proy}_{\mathbb{W}}\mathbf{v}\|$ representa la distancia del vector $\mathbf{v}$ al subespacio $\mathbb{W}$.

Demostración. Para cualquier vector $\mathbf{w}$ en $\mathbb{W}$ se puede escribir

$$\mathbf{v} - \mathbf{w} = (\mathbf{v} - \text{proy}_{\mathbb{W}}\mathbf{v}) + (\text{proy}_{\mathbb{W}}\mathbf{v} - \mathbf{w}).$$

El vector $\text{proy}_{\mathbb{W}}\mathbf{v} - \mathbf{w}$ pertenece a $\mathbb{W}$ ya que es diferencia de dos vectores del espacio $\mathbb{W}$. Por otra parte, de acuerdo al Teorema 5.12, el vector $\mathbf{v} - \text{proy}_{\mathbb{W}}\mathbf{v}$ es ortogonal a todo vector de $\mathbb{W}$ de modo que los dos vectores del segundo miembro son ortogonales. Por el Teorema generalizado de Pitágoras:

$$\|\mathbf{v}-\mathbf{w}\|^2 = \|\mathbf{v}-\mathrm{proy}_{\mathbb{W}}\mathbf{v}\|^2 + \|\mathrm{proy}_{\mathbb{W}}\mathbf{v}-\mathbf{w}\|^2,$$

o bien,

$$\|\mathbf{v}-\mathrm{proy}_{\mathbb{W}}\mathbf{v}\|^2 = \|\mathbf{v}-\mathbf{w}\|^2 - \|\mathrm{proy}_{\mathbb{W}}\mathbf{v}-\mathbf{w}\|^2.$$

Como $\|\mathrm{proy}_{\mathbb{W}}\mathbf{v}-\mathbf{w}\|^2 > 0$ para todo $\mathbf{w} \neq \mathrm{proy}_{\mathbb{W}}\mathbf{v}$, entonces

$$\|\mathbf{v}-\mathrm{proy}_{\mathbb{W}}\mathbf{v}\|^2 < \|\mathbf{v}-\mathbf{w}\|^2,$$

y tomando raíz cuadrada a ambos miembros

$$\|\mathbf{v}-\mathrm{proy}_{\mathbb{W}}\mathbf{v}\| < \|\mathbf{v}-\mathbf{w}\|.$$

□

El teorema de la mejor aproximación implica que la proyección ortogonal sobre un subespacio es única, no depende de la base ortonormal que se usó para calcularla (Ejercicio 13).

Ahora se tienen los elementos necesarios para probar el resultado principal de esta sección: *Todo espacio con producto interno de dimensión finita tiene una base ortogonal.* Este resultado se deduce como consecuencia de un teorema cuya demostración enseña a construir bases ortogonales para cualquier espacio con producto interno de dimensión finita. Esta contrucción es el *método de Gram-Schmidt*, en memoria de J. P. Gram (1850-1916) y E. Smith (1845-1921), que permite "ortogonalizar" cualquier base S de un subespacio $\mathbb{W}$ finito dimensional de un espacio vectorial $\mathbb{V}$ con producto interno.

Teorema 5.14. (Proceso de Gram-Schmidt) Sea $\mathbb{W}$ un subespacio de dimensión finita de un espacio vectorial $\mathbb{V}$ con producto interno. Si $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ es una base de $\mathbb{W}$ entonces $S' = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ es una base ortogonal de $\mathbb{W}$ siendo

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 \\
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{u}_1\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 \\
\mathbf{u}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{u}_1\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 - \frac{\langle \mathbf{v}_3, \mathbf{u}_2\rangle}{\langle \mathbf{u}_2, \mathbf{u}_2\rangle}\mathbf{u}_2 \\
&\vdots \\
\mathbf{u}_k &= \mathbf{v}_k - \frac{\langle \mathbf{v}_k, \mathbf{u}_1\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 - \frac{\langle \mathbf{v}_k, \mathbf{u}_2\rangle}{\langle \mathbf{u}_2, \mathbf{u}_2\rangle}\mathbf{u}_2 - ... - \frac{\langle \mathbf{v}_k, \mathbf{u}_{k-1}\rangle}{\langle \mathbf{u}_{k-1}, \mathbf{u}_{k-1}\rangle}\mathbf{u}_{k-1}.
\end{aligned}$$

Una base ortonormal S'' de $\mathbb{W}$ se obtiene normalizando la base S':

$$S'' = \left\{ \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, ..., \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|} \right\}.$$

Demostración. La prueba es constructiva, se construirá gradualmente la base S' deseada siguiendo los siguientes pasos:

Paso 1. Se elige cualquier vector de S, podemos comenzar con $\mathbf{v}_1$ y hacemos

$$\mathbf{u}_1 = \mathbf{v}_1.$$

Claramente $\mathbf{u}_1 \neq \mathbf{0}$ ya que $\mathbf{v}_1$ es parte de una base.

Paso 2. Seguidamente buscamos un vector $\mathbf{u}_2$ en $\mathbb{W}$ que sea ortogonal a $\mathbf{u}_1$. Para ello, hacemos que $\mathbf{u}_2$ sea la componente de $\mathbf{v}_2$ ortogonal al subespacio $\mathbb{W}_1 = \text{gen}\,\{\mathbf{u}_1\}$. Así:

$$\mathbf{u}_2 = \mathbf{v}_2 - \text{proy}_{\mathbb{W}_1}\mathbf{v}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1.$$

Claramente $\mathbf{u}_2$ es ortogonal a $\mathbf{u}_1$ y pertenece al subespacio $\mathbb{W}$ ya que es combinación lineal de los vectores $\mathbf{v}_2$ y $\mathbf{u}_1 = \mathbf{v}_1$ del espacio $\mathbb{W}$.

Debe notarse que $\mathbf{u}_2 \neq \mathbf{0}$ ya que de otro modo se tendría $\mathbf{v}_2 = \dfrac{\langle \mathbf{v}_2, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1$ y sería $\mathbf{v}_2$ un múltiplo escalar de $\mathbf{v}_1$ contradiciendo la independencia lineal de los vectores de S.

Paso 3. Para construir un vector $\mathbf{u}_3$ que sea ortogonal tanto a $\mathbf{u}_1$ como a $\mathbf{u}_2$, se toma $\mathbf{u}_3$ como la componente de $\mathbf{v}_3$ ortogonal al subespacio $\mathbb{W}_2 = \text{gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$. Esto es:

$$\mathbf{u}_3 = \mathbf{v}_3 - \text{proy}_{\mathbb{W}_2}\mathbf{v}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 - \frac{\langle \mathbf{v}_3, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2.$$

El vector $\mathbf{u}_3$ pertenece a $\mathbb{W}$ ya que es combinación lineal de vectores de $\mathbb{W}$. Además $\mathbf{u}_3 \neq \mathbf{0}$ ya que en caso contrario se tendría $\mathbf{v}_3 = \dfrac{\langle \mathbf{v}_3, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 + \dfrac{\langle \mathbf{v}_3, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2$ y sería $\mathbf{v}_3$ una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ (a través de $\mathbf{u}_1$ y $\mathbf{u}_2$), contradiciendo la independencia lineal de los vectores de S.

Paso 4. De manera similar se obtiene el vector $\mathbf{u}_4$ de $\mathbb{W}$ que es ortogonal a $\mathbf{u}_1$, $\mathbf{u}_2$ y $\mathbf{u}_3$ calculando la componente de $\mathbf{v}_4$ ortogonal al subespacio $\mathbb{W}_3 = \text{gen}\,\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$:

$$\mathbf{u}_4 = \mathbf{v}_4 - \text{proy}_{\mathbb{W}_3}\mathbf{v}_4 = \mathbf{v}_4 - \frac{\langle \mathbf{v}_4, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 - \frac{\langle \mathbf{v}_4, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2 - \frac{\langle \mathbf{v}_4, \mathbf{u}_3 \rangle}{\langle \mathbf{u}_3, \mathbf{u}_3 \rangle}\mathbf{u}_3.$$

Otra vez la independencia lineal de S garantiza $\mathbf{u}_4 \neq \mathbf{0}$.

Al continuar de esta manera se obtiene un conjunto ortogonal, $S' = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$, de k vectores no cero en $\mathbb{W}$. Dado que $\dim(\mathbb{W}) = k$ y todo conjunto ortogonal de vectores no cero es linealmente independiente, entonces S' es una base ortogonal de $\mathbb{W}$.

Una base ortonormal S'' de $\mathbb{W}$ se obtiene normalizando los vectores de S'. □

Nota 1: Puede comprobarse que en cada paso de este proceso los vectores $\mathbf{u}_1$, $\mathbf{u}_2$,..., $\mathbf{u}_l$ forman una base ortogonal para el subespacio generado por los vectores $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_l$ de la base original S.

Nota 2: Si la base S ya es ortogonal, el proceso de Gram-Schmidt "devuelve" la misma base S; esto es $\mathbf{u}_1 = \mathbf{v}_1$, $\mathbf{u}_2 = \mathbf{v}_2$,..., $\mathbf{u}_k = \mathbf{v}_k$.

En la Figura 5.10 se muestra una visualización geométrica del proceso de Gram-Schmidt en el espacio euclidiano $\mathbb{R}^3$. En `https://es.wikipedia.org/wiki/Archivo:Gram-Schmidt_orthonormalization_process.gif` puede verse una construcción dinámica del mismo en dicho espacio.

Ejemplo 5.23. Sea el subespacio $\mathbb{W} = \text{gen}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ donde $\mathbf{v}_1 = (1, 1, 1, 1)$, $\mathbf{v}_2 = (0, 1, 1, 1)$ y $\mathbf{v}_3 = (0, 0, 1, 1)$ del espacio euclidiano $\mathbb{R}^4$.

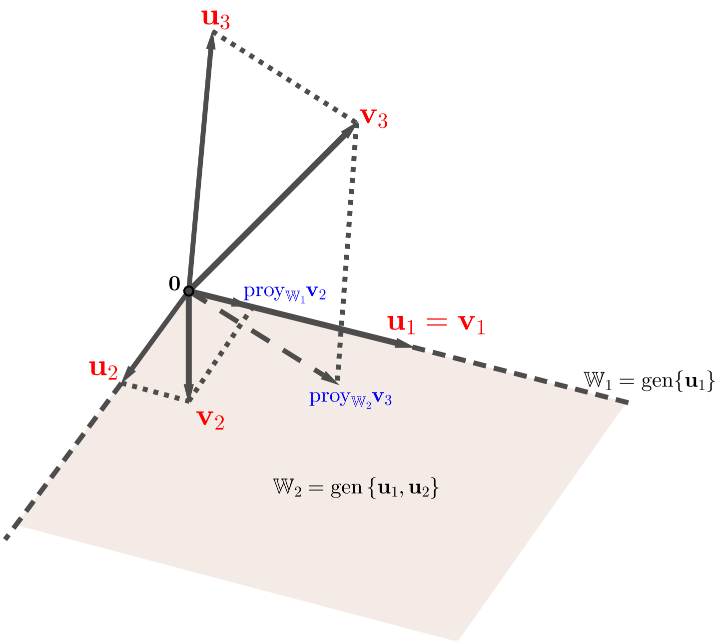


Figura 5.10: Proceso de Gram-Schmidt en el espacio euclidiano $\mathbb{R}^3$

El conjunto $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente (y generador de $\mathbb{W}$) por lo que constituye una base para $\mathbb{W}$. Claramente S no es una base ortogonal de $\mathbb{W}$ por lo que para hallar una base ortonormal del mismo se aplica el proceso de Gram-Schmidt a la base S.

Primero, se obtendrá, a partir de la base (no ortogonal) S, la base ortogonal $S' = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ de la siguiente forma:

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 = (1,1,1,1), \\
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 = (0,1,1,1) - \frac{3}{4}(1,1,1,1) = \left(-\frac{3}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right), \\
\mathbf{u}_3 &= \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 = (0,0,1,1) - \frac{2}{4}(1,1,1,1) - \frac{1/2}{3/4}\left(-\frac{3}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) = \\
&= (0,0,1,1) - \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) - \left(-\frac{1}{2}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right) = \left(0, -\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right).
\end{aligned}$$

Dado que $\|\mathbf{u}_1\| = 2$, $\|\mathbf{u}_2\| = \frac{\sqrt{3}}{2}$ y $\|\mathbf{u}_3\| = \frac{\sqrt{6}}{3}$, el conjunto $S'' = \left\{\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}\right\}$ donde

$$\begin{aligned}
\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} &= \frac{1}{2}(1,1,1,1) = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right), \\
\frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} &= \frac{2}{\sqrt{3}}\left(-\frac{3}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) = \left(-\frac{3}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}\right), \\
\frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} &= \frac{3}{\sqrt{6}}\left(0, -\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right) = \left(0, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right),
\end{aligned}$$

es una base ortonormal para el subespacio $\mathbb{W}$.◀

Ejemplo 5.24. Sea el subespacio $\mathbb{W} = \text{gen}\{f_1, f_2, f_3\}$ siendo $f_1(x) = 1$, $f_2(x) = x$ y $f_3(x) = x^2$, del espacio $C[-1, 1]$ de las funciones continuas en el intervalo cerrado $[-1, 1]$, con el producto interno dado por $\langle f, g\rangle = \displaystyle\int_{-1}^{1} f(x)g(x)\,dx$, donde f y g son funciones de $C[-1, 1]$.

El conjunto $S = \{f_1, f_2.f_3\}$ es linealmente independiente y por lo tanto constituye una base para $\mathbb{W}$ pero no es un conjunto ortogonal. En efecto:

$$\langle f_1, f_3\rangle = \int_{-1}^{1} 1(x^2)\,dx = \int_{-1}^{1} x^2\,dx = \frac{2}{3} \neq 0.$$

Se obtendrá, aplicando el proceso de Gram-Schmidt, una base ortogonal para $\mathbb{W}$ a partir de la base S:

$$\begin{aligned}
g_1 &= f_1,\\
g_2 &= f_2 - \frac{\langle f_2, g_1\rangle}{\langle g_1, g_1\rangle} g_1 = f_2 - \frac{\displaystyle\int_{-1}^{1} x\,dx}{\displaystyle\int_{-1}^{1} dx} g_1 = f_2 - \left(\frac{0}{2}\right) g_1 = f_2,\\
g_3 &= f_3 - \frac{\langle f_3, g_1\rangle}{\langle g_1, g_1\rangle} g_1 - \frac{\langle f_3, g_2\rangle}{\langle g_2, g_2\rangle} g_2 = f_3 - \frac{\displaystyle\int_{-1}^{1} x^2\,dx}{\displaystyle\int_{-1}^{1} dx} g_1 - \frac{\displaystyle\int_{-1}^{1} x^3\,dx}{\displaystyle\int_{-1}^{1} x^2\,dx} g_2 =\\
&= f_3 - \left(\frac{2/3}{2}\right) g_1 - \left(\frac{0}{2/3}\right) g_2 = f_3 - \frac{1}{3} g_1.
\end{aligned}$$

Por lo tanto el conjunto $S' = \{g_1, g_2, g_3\}$ donde $g_1(x) = 1$, $g_2(x) = x$ y $g_3(x) = x^2 - \frac{1}{3}$, es una base ortogonal para $\mathbb{W}$.◀

Ejemplo 5.25. En el espacio euclidiano $\mathbb{R}^4$, sea $\mathbb{W} = \text{gen}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5\}$ donde $\mathbf{w}_1 = (1, 1, 0, 2)$, $\mathbf{w}_2 = (0, 1, 1, -1)$, $\mathbf{w}_3 = (2, 3, 1, 3)$, $\mathbf{w}_4 = (-1, 0, 1, -3)$ y $\mathbf{w}_5 = (1, 1, 1, 1)$. Se desea hallar el vector proyección de $\mathbf{v} = (0, -1, 0, 1)$ sobre el subespacio $\mathbb{W}$, la componente de $\mathbf{v}$ ortogonal a $\mathbb{W}$ y la distancia del vector $\mathbf{v}$ al subespacio $\mathbb{W}$.

Se debe obtener una base ortogonal para el subespacio $\mathbb{W}$. Claramente $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5\}$ no es una base de $\mathbb{W}$ (¿por qué?).

Para hallar una base de $\mathbb{W}$, se forma la matriz A cuyos renglones son los vectores $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ y $\mathbf{w}_5$:

$$A = \begin{bmatrix} 1 & 1 & 0 & 2\\ 0 & 1 & 1 & -1\\ 2 & 3 & 1 & 3\\ -1 & 0 & 1 & -3\\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Los renglones no cero de cualquier forma escalonada de A forman una base para $\text{ren}(A) = \mathbb{W}$. Una forma escalonada de A es

$$B = \begin{bmatrix} 1 & 1 & 0 & 2\\ 0 & 1 & 1 & -1\\ 0 & 0 & 1 & -1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Así, el conjunto $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ donde $\mathbf{v}_1 = (1,1,0,2)$, $\mathbf{v}_2 = (0,1,1,-1)$ y $\mathbf{v}_3 = (0,0,1,-1)$ es una base para $\mathbb{W}$. S no es una base ortogonal, aplicando el proceso de Gram-Schmidt a S se obtiene una base ortogonal para $\mathbb{W}$:

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 = (1,1,0,2)\,,\\
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 = (0,1,1,-1) - \left(\frac{-1}{6}\right)(1,1,0,2) = \left(\frac{1}{6},\frac{7}{6},1,-\frac{2}{3}\right),\\
\mathbf{u}_3 &= \mathbf{v}_3 - \frac{\mathbf{v}_3\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{v}_3\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = (0,0,1,-1) - \left(\frac{-2}{6}\right)(1,1,0,2) - \frac{5/3}{17/6}\left(\frac{1}{6},\frac{7}{6},1,-\frac{2}{3}\right) =\\
&= \left(\frac{4}{17},-\frac{6}{17},\frac{7}{17},\frac{1}{17}\right),
\end{aligned}$$

resultando $S' = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ una base ortogonal para $\mathbb{W}$.

La proyección del vector $\mathbf{v}$ sobre el subespacio $\mathbb{W}$ es el vector

$$\begin{aligned}
\mathrm{proy}_{\mathbb{W}}\mathbf{v} &= \frac{\mathbf{v}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{v}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{v}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\mathbf{u}_3 =\\
&= \frac{1}{6}(1,1,0,2) - \frac{11}{17}\left(\frac{1}{6},\frac{7}{6},1,-\frac{2}{3}\right) + \frac{7}{6}\left(\frac{4}{17},-\frac{6}{17},\frac{7}{17},\frac{1}{17}\right) =\\
&= \left(\frac{1}{3},-1,-\frac{1}{6},\frac{5}{6}\right).
\end{aligned}$$

La componente de $\mathbf{v}$ ortogonal a $\mathbb{W}$ es el vector

$$\begin{aligned}
\mathbf{w} &= \mathbf{v} - \mathrm{proy}_{\mathbb{W}}\mathbf{v} =\\
&= (0,-1,0,1) - \left(\frac{1}{3},-1,-\frac{1}{6},\frac{5}{6}\right) =\\
&= \left(-\frac{1}{3},0,\frac{1}{6},\frac{1}{6}\right).
\end{aligned}$$

La distancia del vector $\mathbf{v}$ al subespacio $\mathbb{W}$ es

$$\|\mathbf{v} - \mathrm{proy}_{\mathbb{W}}\mathbf{v}\| = \sqrt{\frac{6}{36}} = \frac{\sqrt{6}}{6}. \blacktriangleleft$$

Complemento ortogonal y teorema de la descomposición ortogonal

Se han visto vectores que son ortogonales a todos los vectores de un subespacio. Por ejemplo, para cada vector $\mathbf{v}$ del espacio $\mathbb{V}$, el vector $\mathbf{v} - \mathrm{proy}_{\mathbb{W}}\mathbf{v}$ es ortogonal al espacio $\mathbb{W}$. El conjunto de todos los vectores ortogonales a un espacio $\mathbb{W}$ es el llamado **complemento ortogonal** de $\mathbb{W}$ y se simboliza $\mathbb{W}^\perp$. Más precisamente:

Definición. (Complemento ortogonal) Sea $\mathbb{W}$ un subespacio de un espacio $\mathbb{V}$ con producto interno. El conjunto de todos los vectores de $\mathbb{V}$ que son ortogonales a $\mathbb{W}$, se designa con $\mathbb{W}^\perp$ y se llama *complemento ortogonal de* $\mathbb{W}$,

$$\mathbb{W}^\perp = \{\mathbf{v}\in\mathbb{V} : \langle\mathbf{v},\mathbf{w}\rangle = 0,\ \text{para todo } \mathbf{w}\in\mathbb{W}\}\,.$$

Se deja como ejercicio la demostración del siguiente teorema:

Teorema 5.15. Sea $\mathbb{W}$ un subespacio de un espacio vectorial $\mathbb{V}$ con producto interno, entonces:

1. $\mathbb{W}^{\perp}$ es un subespacio de $\mathbb{V}$.
2. $\mathbb{W} \cap \mathbb{W}^{\perp} = \{\mathbf{0}\}$

Ejemplo 5.26. Complemento ortogonal de los subespacios triviales.

Si $\mathbb{W} = \{\mathbf{0}\}$ entonces $\mathbb{W}^{\perp} = \mathbb{V}$ ya que $\langle \mathbf{0}, \mathbf{v} \rangle = 0$ para todo $\mathbf{v} \in \mathbb{V}$.

Si $\mathbb{W} = \mathbb{V}$ entonces $\mathbb{W}^{\perp} = \{\mathbf{0}\}$. ◄

Ejemplo 5.27. Sea $\mathbb{W}$ el subespacio del espacio euclidiano $\mathbb{R}^3$ generado por los vectores linealmente independientes $\mathbf{u}_1 = (1, -2, 3)$ y $\mathbf{u}_2 = (1, 2, 5)$. Geométricamente $\mathbb{W}$ es un plano que contiene al origen de coordenadas.

Un vector $\mathbf{v}$ de $\mathbb{R}^3$ es ortogonal a $\mathbb{W}$ si y solo si es ortogonal a cada uno de los vectores de un conjunto generador de $\mathbb{W}$ (véase Ejercicio 9 de la sección anterior).

Así, $\mathbf{v} = (v_1, v_2, v_3)$ es ortogonal a $\mathbb{W}$ si y solo si $\mathbf{u}_1 \cdot \mathbf{v} = 0$ y $\mathbf{u}_2 \cdot \mathbf{v} = 0$. Esto es:

$$\begin{cases} v_1 - 2v_2 + 3v_3 = 0 \\ v_1 + 2v_2 + 5v_3 = 0 \end{cases}$$

Por lo tanto $\mathbb{W}^{\perp}$ es el conjunto de soluciones del sistema homogéneo de matriz de coeficientes

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 1 & 2 & 5 \end{bmatrix}.$$

La solución general del sistema es

$$\begin{cases} v_1 = -4t \\ v_2 = -\frac{1}{2}t \quad , \quad t \in \mathbb{R} \\ v_3 = t \end{cases}$$

por lo cual el complemento ortogonal de $\mathbb{W}$ es

$$\mathbb{W}^{\perp} = \left\{(-4t, -\frac{1}{2}t, t) : t \in \mathbb{R}\right\} = \text{gen}\left\{\left(-4, -\frac{1}{2}, 1\right)\right\},$$

que geométricamente es una recta que contiene al origen de coordenadas.

Observe que $\dim(\mathbb{W}) = 2$, $\dim(\mathbb{W}^{\perp}) = 1$ y $\dim(\mathbb{W}) + \dim(\mathbb{W}^{\perp}) = 3 = \dim(\mathbb{R}^3)$.◄

El teorema de Gram-Schmidt asegura la existencia de bases ortogonales para un subespacio no cero finito dimensional en un espacio con producto interno. Es posible ahora reenunciar el Teorema 5.12 y "hacerlo más fuerte" ya que se puede aplicar a cualquier subespacio finito dimensional de un espacio con producto interno.

Teorema 5.16. (Descomposición ortogonal) Si $\mathbb{W}$ es un subespacio de dimensión finita de un espacio vectorial $\mathbb{V}$ con producto interno entonces todo vector $\mathbf{v}$ de $\mathbb{V}$ se puede expresar de manera única como

$$\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2,$$

donde $\mathbf{w}_1 \in \mathbb{W}$ y $\mathbf{w}_2 \in \mathbb{W}^{\perp}$.

Demostración. Si $\mathbb{W} = \{\mathbf{0}\}$ entonces $\mathbb{W}^{\perp} = \mathbb{V}$. En este caso, para todo $\mathbf{v} \in \mathbb{V}$:

$$\mathbf{v} = \mathbf{0} + \mathbf{v},$$

donde $\mathbf{0} \in \mathbb{W}$ y $\mathbf{v} \in \mathbb{W}^{\perp}$, y ésta es la única manera de descomponer a $\mathbf{v}$ como suma de un vector de $\mathbb{W}$ y otro de $\mathbb{W}^{\perp}$.

Supóngase ahora que $\mathbb{W} \neq \{\mathbf{0}\}$ y $\mathbb{W}$ es finito dimensional. La demostración se realiza en dos etapas.

La primera consiste en probar la existencia de dos vectores, $\mathbf{w}_1$ y $\mathbf{w}_2$, con las propiedades enunciadas. Por el proceso de Gram-Schmidt se sabe que existe una base ortogonal para $\mathbb{W}$, entonces los vectores $\mathbf{w}_1 = \text{proy}_{\mathbb{W}}\mathbf{v}$ y $\mathbf{w}_2 = \mathbf{v} - \text{proy}_{\mathbb{W}}\mathbf{v}$ tienen las propiedades requeridas puesto que $\mathbf{w}_1 \in \mathbb{W}$ y, por el Teorema 5.12, $\mathbf{w}_2$ es ortogonal a todo vector de $\mathbb{W}$. Así:

$$\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2.$$

Para hacer ver que éstos son los únicos vectores con estas propiedades supóngase que también $\mathbf{v}$ se puede descomponer:

$$\mathbf{v} = \mathbf{w}_1' + \mathbf{w}_2',$$

donde $\mathbf{w}_1' \in \mathbb{W}$ y $\mathbf{w}_2' \in \mathbb{W}^{\perp}$.

Así,

$$\mathbf{w}_1 + \mathbf{w}_2 = \mathbf{w}_1' + \mathbf{w}_2'\ ,$$

por lo tanto

$$\mathbf{w}_1 - \mathbf{w}_1' = \mathbf{w}_2' - \mathbf{w}_2\ .$$

El vector $\mathbf{w}_1 - \mathbf{w}_1' \in \mathbb{W}$ ya que es resta de dos vectores de $\mathbb{W}$ y éste es un subespacio. El vector $\mathbf{w}_2' - \mathbf{w}_2 \in \mathbb{W}^{\perp}$ ya que es resta de dos vectores de $\mathbb{W}^{\perp}$ y éste es un subespacio. Luego:

$$\mathbf{u} = \mathbf{w}_1 - \mathbf{w}_1' = \mathbf{w}_2' - \mathbf{w}_2\ ,$$

es un vector que está tanto en $\mathbb{W}$ como en $\mathbb{W}^{\perp}$ por lo cual, del Teorema 5.15 parte 2, $\mathbf{u} = \mathbf{0}$ esto es $\mathbf{w}_1 = \mathbf{w}_1'$ y $\mathbf{w}_2 = \mathbf{w}_2'$.

□

Ejercicios 5.3

1. Halle el vector de coordenadas del vector $\mathbf{v} = (8, 2, -3)$ del espacio euclidiano $\mathbb{R}^3$ con respecto a la base ortogonal $B = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ donde $\mathbf{u}_1 = (4, -1, 1)$, $\mathbf{u}_2 = (0, 1, 1)$ y $\mathbf{u}_3 = \left(-\frac{1}{2}, -1, 1\right)$. (No hay que resolver ningún sistema).

2. Compruebe que el conjunto formado por los vectores $\mathbf{u}_1$ y $\mathbf{u}_2$ es un conjunto ortogonal del espacio euclidiano $\mathbb{R}^3$. Luego encuentre la proyección de $\mathbf{v}$ sobre gen $\{\mathbf{u}_1, \mathbf{u}_2\}$, siendo:

   $\boxed{a)}$ $\mathbf{v} = (-1, 4, 3)$; $\mathbf{u}_1 = (1, 1, 0)$ y $\mathbf{u}_2 = (-1, 1, 0)$.

   b) $\mathbf{v} = (6, 4, 1)$; $\mathbf{u}_1 = (-4, -1, 1)$ y $\mathbf{u}_2 = (0, 1, 1)$.

3. Sea $\mathbb{W}$ el subespacio del espacio euclidiano $\mathbb{R}^4$ generado por los vectores $\mathbf{u}_1$, $\mathbf{u}_2$ y $\mathbf{u}_3$. Escriba el vector $\mathbf{v}$ como la suma de un vector de $\mathbb{W}$ y un vector ortogonal a $\mathbb{W}$. Calcule además la distancia del vector $\mathbf{v}$ al subespacio $\mathbb{W}$:

   $\boxed{a)}$ $\mathbf{u}_1 = (1, 1, 0, 1)$, $\mathbf{u}_2 = (-1, 3, 1, -2)$, $\mathbf{u}_3 = (-1, 0, 1, 1)$ y $\mathbf{v} = (4, 3, 3, -1)$.

   b) $\mathbf{u}_1 = (1, 1, 0, -1)$, $\mathbf{u}_2 = (1, 0, 1, 1)$, $\mathbf{u}_3 = (0, -1, 1, -1)$ y $\mathbf{v} = (3, 4, 5, 6)$.

4. Sea el subespacio $\mathbb{W} = \text{gen}\{\mathbf{u}_1, \mathbf{u}_2\}$ del espacio euclidiano $\mathbb{R}^4$. Determine el vector del subespacio $\mathbb{W}$ más cercano a $\mathbf{v}$ y calcule dicha distancia:

   $\boxed{a)}$ $\mathbf{u}_1 = (3, 1, -1, 1)$, $\mathbf{u}_2 = (1, -1, 1, -1)$ y $\mathbf{v} = (3, 1, 5, 1)$.

   b) $\mathbf{u}_1 = (1, -2, -1, 2)$, $\mathbf{u}_2 = (-4, 1, 0, 3)$ y $\mathbf{v} = (3, -1, 1, 13)$.

5. Halle, en cada caso, una base ortonormal para el subespacio $\mathbb{W}$ del espacio euclidiano $\mathbb{R}^n$ correspondiente:

   a) $\mathbb{W} = \text{gen}\{\mathbf{v}_1, \mathbf{v}_2\}$; $\mathbf{v}_1 = (0, 4, 2)$, $\mathbf{v}_2 = (8, 5, -6)$.

   $\boxed{b)}$ $\mathbb{W} = \text{gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$; $\mathbf{v}_1 = (3, 1, -1, 3)$, $\mathbf{v}_2 = (-5, 1, 5, -7)$, $\mathbf{v}_3 = (1, 1, -2, 8)$.

   $\boxed{c)}$ $\mathbb{W} = \{(a, a+b, -b) \in \mathbb{R}^3 : a, b \in \mathbb{R}\}$

   d) $\mathbb{W} = N(A)$ donde $A = \begin{bmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 2 & 6 & 16 \\ 1 & 3 & 2 \end{bmatrix}$.

   e) $\mathbb{W} = \text{ren}(A)$ donde A es la matriz del ítem *d*).

   f) $\mathbb{W} = \text{col}(A)$ donde A es la matriz del ítem *d*).

   $\boxed{g)}$ $\mathbb{W}$ es el espacio solución de la ecuación $x_1 - 2x_2 + 3x_3 - x_4 = 0$.

6. Halle una base ortogonal para el subespacio $\mathbb{W}$ del espacio $\mathbb{V} = \mathbb{P}_n$ con el producto interno estándar.

   $\boxed{a)}$ $\mathbb{W} = \text{gen}\{p_1, p_2, p_3\}$; $\mathbb{V} = \mathbb{P}_3$; $p_1 = 1 + x - x^2$, $p_2 = 1 + x + x^2 - x^3$, $p_3 = 1 + x^3$.

   b) $\mathbb{W} = \{a + (a-b)x + bx^2 \in \mathbb{P}_2 : a, b \in \mathbb{R}\}$; $\mathbb{V} = \mathbb{P}_2$.

7. Halle una base ortogonal para el subespacio $\mathbb{W}$ del espacio $\mathbb{M}_{mn}$ correspondiente con el producto interno estándar.

   a) $\mathbb{W} = \text{gen}\{A_1, A_2, A_3\}$; $A_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, $A_3 = \begin{bmatrix} 0 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$

   $\boxed{b)}$ $\mathbb{W} = \left\{ \begin{bmatrix} a & b & -b \\ 0 & a+b & 0 \\ 0 & 0 & -a \end{bmatrix} \in \mathbb{M}_3 : a, b \in \mathbb{R} \right\}$.

8. Halle una base ortogonal para el subespacio $\mathbb{W}$ del espacio $C[0,1]$ con el producto interno dado por $\langle f, g\rangle = \displaystyle\int_0^1 f(x)g(x)\,dx$, donde f y g son funciones continuas en el intervalo cerrado $[0,1]$.

   $\boxed{a)}$ $\mathbb{W} = \text{gen}\{f_1, f_2\}$; $f_1(x) = x$, $f_2(x) = e^x$.

   b) $\mathbb{W} = \text{gen}\{f_1, f_2, f_3\}$; $f_1(x) = x$, $f_2(x) = 1 + 2x$, $f_3(x) = x^5$.

   c) $\mathbb{W} = \{a \operatorname{sen}\, x + b\cos x : a, b \in \mathbb{R}\}$.

9. Calcule la proyección del vector $\mathbf{v}$ sobre el subespacio $\mathbb{W}$ indicado en cada caso y la componente de $\mathbf{v}$ ortogonal a $\mathbb{W}$.

   a) $\mathbf{v} = (1,1,1)$ y el subespacio $\mathbb{W}$ del Ejercicio 5.a.

   $\boxed{b)}$ $\mathbf{v} = (1,-1,1,1)$ y el subespacio $\mathbb{W} = \left\{(a, -b, a+b+c, c) \in \mathbb{R}^4 : a,b,c \in \mathbb{R}\right\}$ del espacio euclidiano $\mathbb{R}^4$.

10. Descomponga el polinomio p como suma de un polinomio en $\mathbb{W}$ y otro en $\mathbb{W}^{\perp}$.

   $\boxed{a)}$ $p = 1 + x + x^2 + x^3$ y el subespacio $\mathbb{W}$ del Ejercicio 6.a.

   a) $p = x - x^2$ y el subespacio $\mathbb{W} = \left\{a - bx + ax^2 \in \mathbb{P}_2 : a,b \in \mathbb{R}\right\}$ del espacio $\mathbb{P}_2$ con el producto interno estándar.

11. Halle la matriz del subespacio $\mathbb{W}$ más cercana a la matriz A:

   a) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $\mathbb{W}$ el subespacio del Ejercicio 7.a.

   $\boxed{b)}$ $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$; $\mathbb{W}$ el subespacio del Ejercicio 7.b.

$\boxed{12.}$ Obtenga la proyección de $f(x) = 1 + x$ sobre el subespacio $\mathbb{W}$ del Ejercicio 8.a.

13. Sea $\mathbb{V}$ un espacio vectorial con producto interno y $\mathbb{W}$ un subespacio de dimensión finita. Demuestre que la proyección ortogonal de un vector $\mathbf{v}$ sobre $\mathbb{W}$ es única.
   Sugerencia: Utilice el teorema de la mejor aproximación para probar que no hay dos vectores proyección de $\mathbf{v}$ sobre $\mathbb{W}$ distintos.

$\boxed{14.}$ Demuestre el Teorema 5.15.

$\boxed{15.}$ Sea A una matriz de $m \times n$. Para los subespacios $\text{ren}(A)$ y $\text{col}(A)$ de los espacios euclidianos $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente, se tienen las siguientes relaciones con respecto a sus complementos ortogonales:

$$(\text{ren}(A))^{\perp} = N(A),$$

y dado que $\text{col}(A) = \text{ren}(A^t)$, resulta

$$(\text{col}(A))^{\perp} = N(A^t).$$

   a) Demuestre que $(\text{ren}(A))^{\perp} = N(A)$ (un vector de $\mathbb{R}^n$ es ortogonal a cada uno de los renglones de A si y solo si es un vector de $N(A)$).

b) Utilice el resultado del ítem anterior para probar que para todo subespacio $\mathbb{W}$ en el espacio euclidiano $\mathbb{R}^n$, $\dim(\mathbb{W}) + \dim(\mathbb{W}^{\perp}) = n$.

c) Halle el complemento ortogonal y una base del mismo para los siguientes subespacios $\mathbb{W}$ del espacio euclidiano $\mathbb{V}$.

i) $\mathbb{W} = \text{gen}\{(1,1,1),(2,1,3),(3,2,4)\}$; $\mathbb{V} = \mathbb{R}^3$.

ii) $\mathbb{W} = \text{gen}\{(1,0,0,1),(0,1,1,1),(1,1,1,2),(1,2,2,3)\}$; $\mathbb{V} = \mathbb{R}^4$.

iii) $\mathbb{W} = \text{gen}\{(1,0,0,1),(0,1,1,1),(1,1,1,2),(1,2,2,2)\}$; $\mathbb{V} = \mathbb{R}^4$.

$\boxed{16.}$ Sea $\mathbb{W}$ un subespacio de dimensión finita de un espacio vectorial $\mathbb{V}$ con producto interno y $\mathbf{v}$ un vector de $\mathbb{V}$. Demuestre:

a) $\text{proy}_{\mathbb{W}}\mathbf{v} = \mathbf{v} \;\Leftrightarrow\; \mathbf{v} \in \mathbb{W}$.

b) $\text{proy}_{\mathbb{W}}\mathbf{v} = \mathbf{0} \;\Leftrightarrow\; \mathbf{v} \in \mathbb{W}^{\perp}$.

17. **(Polinomios y series de Fourier)** Con frecuencia las funciones continuas se aproximan mediante combinaciones lineales de funciones seno y coseno. Por ejemplo, una función continua podría representar una onda sonora, una señal eléctrica de algún tipo, el movimiento de un sistema mecánico vibratorio, etc. Esta idea data de Euler, sin embargo, floreció con el trabajo de Fourier (1768-1830).

Sea B el conjunto de las siguientes funciones trigonométricas definidas en el intervalo $[-\pi, \pi]$.

$$B = \{1, \cos x, \cos(2x), ..., \cos(nx), \text{sen }\; x, \text{sen}(2x), ..., \text{sen}(nx)\}.$$

Un *polinomio trigonométrico de orden* n es una combinación lineal de los elementos de B:

$$a_0 + a_1 \cos x + a_2 \cos(2x) + ... + a_n \cos(nx) + b_1 \,\text{sen}\, x + b_2 \,\text{sen}(2x) + ... + b_n \,\text{sen}(nx), \qquad (5.6)$$

si a_n y b_n no son cero simultáneamente.

Cualquier función f en $C[-\pi, \pi]$ puede aproximarse mediante un polinomio trigonométrico. Por aproximar queremos decir que f y algún polinomio trigonométrico "están cerca" con la norma inducida por el producto interno integral definido en el Ejemplo 5.12 de la sección anterior.

Sea el subespacio $\text{gen}(B)$ de $C[-\pi, \pi]$ formado por todos los polinomios trigonométricos cuyo orden es a lo sumo n.

a) Demuestre que B es una base ortogonal de $\text{gen}(B)$. Para ello, pruebe que:

- $\langle 1, \cos(nx)\rangle = 0; \quad n = 1, 2, 3, ...$
- $\langle 1, \text{sen}(nx)\rangle = 0; \quad n = 1, 2, 3, ...$
- $\langle \cos(mx), \cos(nx)\rangle = 0; \quad \text{si } m \neq n$
- $\langle \cos(mx), \text{sen}(nx)\rangle = 0; \quad m, n = 1, 2, 3, ...$
- $\langle \text{sen}(mx), \text{sen}(nx)\rangle = 0; \quad \text{si } m \neq n$

b) Pruebe que $\|1\| = \sqrt{2\pi}$, $\|\cos(kx)\| = \sqrt{\pi}$ y $\|\text{sen}(kx)\| = \sqrt{\pi}$, $k \geq 1$.

Sea f una función de $C[-\pi,\pi]$. Una aproximación de la función f mediante un polinomio trigonométrico p_n es

$$\begin{aligned} f(x) \simeq p_n(x) = \text{proy}_{\text{gen}(B)} f \quad &= \frac{\langle f,1\rangle}{\langle 1,1\rangle}1 + \frac{\langle f,\cos x\rangle}{\langle \cos x,\cos x\rangle}\cos x + \frac{\langle f,\cos 2x\rangle}{\langle \cos 2x,\cos 2x\rangle}\cos 2x + ... \\ &+ \frac{\langle f,\cos nx\rangle}{\langle \cos nx,\cos nx\rangle}\cos nx + ... + \frac{\langle f,\operatorname{sen} x\rangle}{\langle \operatorname{sen} x,\operatorname{sen} x\rangle}\operatorname{sen} x + \\ &+ \frac{\langle f,\operatorname{sen} 2x\rangle}{\langle \operatorname{sen} 2x,\operatorname{sen} 2x\rangle}\operatorname{sen} 2x + ... + \frac{\langle f,\operatorname{sen} nx\rangle}{\langle \operatorname{sen} nx,\operatorname{sen} nx\rangle}\operatorname{sen} nx. \end{aligned}$$

c) Demuestre que los coeficientes de la aproximación anterior son:

$$\begin{aligned} a_0 &= \frac{\langle f,1\rangle}{\langle 1,1\rangle} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,dx \\ a_k &= \frac{\langle f,\cos kx\rangle}{\langle \cos kx,\cos kx\rangle} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\,dx \qquad k=1,2,...,n \\ b_k &= \frac{\langle f,\operatorname{sen} kx\rangle}{\langle \operatorname{sen} kx,\operatorname{sen} kx\rangle} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\operatorname{sen} kx\,dx \qquad k=1,2,...,n \end{aligned}$$

Las fórmulas dadas en el ítem anterior se conocen como *fórmulas de Euler* que utilizó Fourier para resolver la ecuación del calor.

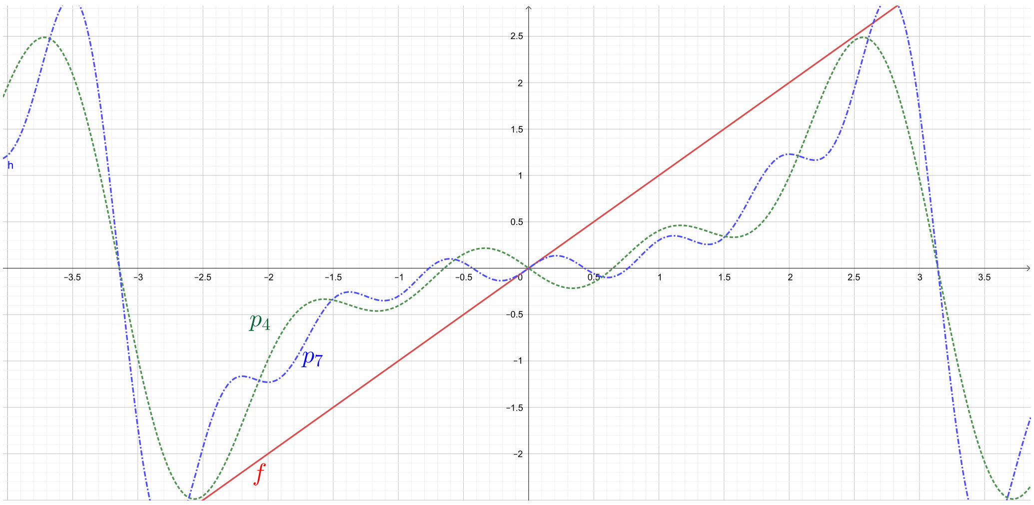


Figura 5.11: Polinomios de Fourier p_4 y p_7 de $f(x)=x$, $-\pi \le x \le \pi$

El polinomio trigonométrico p_n que aproxima a la función f definido por (5.6) y cuyos coeficientes se obtienen de las fórmulas de Euler, se llama *polinomio o aproximación de Fourier* de orden n de f en $[-\pi,\pi]$.

d) Pruebe que el polinomio de Fourier de orden n de la función dada por $f(x)=x$, $x\in[-\pi,\pi]$ está dado por:

$$p_n(x) = \operatorname{sen} x - \operatorname{sen} 2x + \frac{2}{3}\operatorname{sen} 3x + ... + \frac{2(-1)^{n+1}}{n}\operatorname{sen} nx.$$

En la Figura 5.11 se representan la función f y los polinomios de Fourier de orden 4, p_4 y de orden 7, p_7 de f.

A medida que n crece, los polinomios p_n se acercan cada vez más a la función f. Tomando el límite cuando $n \to +\infty$ se obtiene la serie

$$f(x) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos(nx) + b_n \operatorname{sen}(nx)\right),$$

llamada *serie de Fourier* de la función f en el intervalo $[-\pi, \pi]$.

Capítulo 6

TRANSFORMACIONES LINEALES

En este capítulo se estudiará un tipo particular de funciones denominadas *transformaciones lineales* que se encuentran con frecuencia en el Álgebra Lineal y en otras áreas de la Matemática. Éstas poseen una gran variedad de aplicaciones.

6.1. Transformaciones lineales

Recordemos que dados dos conjuntos no vacíos, A y B, una función (mapeo, transformación o aplicación) T de A en B, representada por $T : A \to B$, es una ley o regla que asigna a cada elemento x de A un único elemento y de B. El elemento y asignado a x es llamado *imagen* de x por la transformación T y se denota $y = T(x)$. El conjunto A recibe el nombre de *dominio* y el conjunto B de *codominio*.

El subconjunto de B formado por todas las imágenes de los elementos de A se llama *recorrido de* T o *conjunto imagen de* A y se denota por $R(T)$. Estos es:

$$R(T) = \{y \in B : y = T(x), \text{ para algún } x \in A\} .$$

Una clase especial de funciones son las transformaciones lineales, las cuales son transformaciones entre espacios vectoriales que "conservan" la suma y la multiplicación por escalares. Más precisamente:

Definición (Transformación lineal). Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales. Una función T de $\mathbb{V}$ en $\mathbb{W}$, $T : \mathbb{V} \to \mathbb{W}$, es una *transformación lineal* si

a) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$, para todo $\mathbf{u}, \mathbf{v}$ en $\mathbb{V}$.

b) $T(k\mathbf{u}) = kT(\mathbf{u})$, para todo $\mathbf{u}$ en $\mathbb{V}$ y k en $\mathbb{R}$.

Observe que, en a) de la definición anterior, el signo "+" en $\mathbf{u} + \mathbf{v}$ se refiere a la operación suma en $\mathbb{V}$, mientras que el signo "+" en $T(\mathbf{u}) + T(\mathbf{v})$ es la operación suma en $\mathbb{W}$.

De manera análoga, en b), el producto por escalar $k\mathbf{u}$ se refiere a la multiplicación por un escalar en $\mathbb{V}$ mientras que el producto por escalar $kT(\mathbf{u})$ se refiere a la multiplicación por un escalar en $\mathbb{W}$.

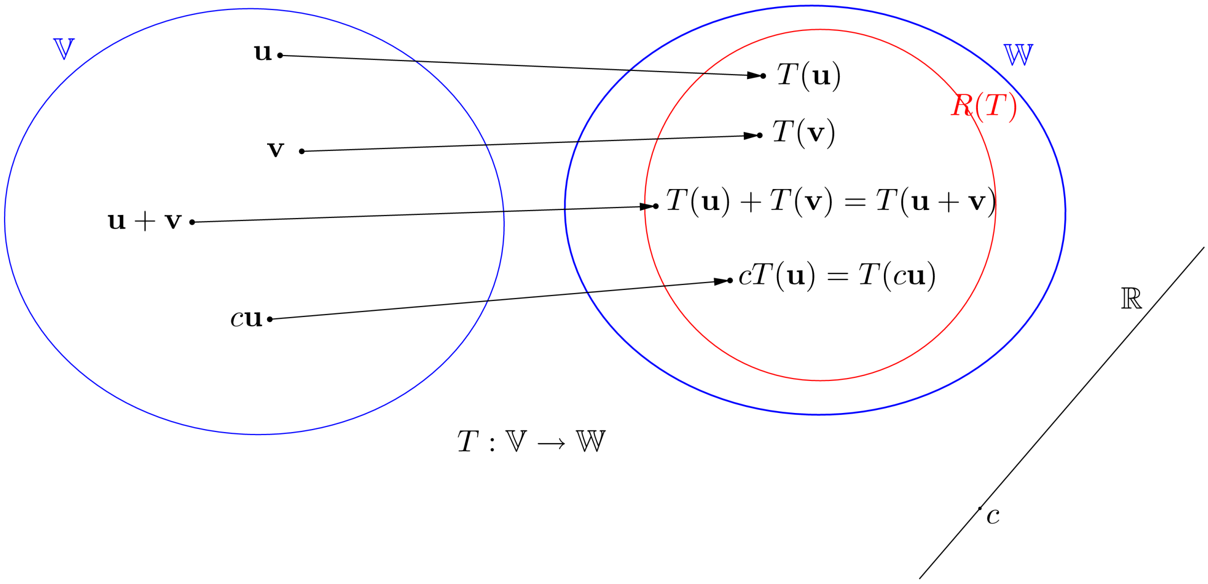


Figura 6.1: Transformación lineal T de $\mathbb{V}$ en $\mathbb{W}$

Ejemplo 6.1. Sea la transformación $T:\mathbb{P}_2\to\mathbb{R}^2$ definida por

$$T\left(a+bx+cx^2\right)=(a+b,c)\,.$$

T es la función que transforma cada polinomio $p=a+bx+cx^2$ del espacio vectorial $\mathbb{P}_2$ en el único vector $(a+b,c)$ del espacio $\mathbb{R}^2$. Por ejemplo, la imagen del polinomio $p=-1+2x+3x^2$ es el vector $T(p)=T(-1+2x+3x^2)=(1,3)$ de $\mathbb{R}^2$. Es fácil ver que la imagen del polinomio cero de $\mathbb{P}_2$ es el vector cero de $\mathbb{R}^2$. Puede comprobarse que el recorrido de T es $\mathbb{R}^2$ ($R(T)=\mathbb{R}^2$).

Para determinar si T es una transformación lineal, deben verificarse a) y b) de la definición. Para ello, sean $p=a_1+b_1x+c_1x^2$ y $q=a_2+b_2x+c_2x^2$ polinomios de $\mathbb{P}_2$ y k un escalar real. Dado que:

$$\begin{aligned}
T(p+q) &= T\left((a_1+b_1x+c_1x^2)+(a_2+b_2x+c_2x^2)\right)\\
&= T\left((a_1+a_2)+(b_1+b_2)x+(c_1{+}_2)x^2\right)\\
&= ((a_1+a_2)+(b_1+b_2),c_1+c_2)\\
&= ((a_1+b_1)+(a_2+b_2),c_1+c_2)\\
&= (a_1+b_1,c_1)+(a_2+b_2,c_2)\\
&= T\left(a_1+b_1x+c_1x^2\right)+T\left(a_2+b_2x+c_2x^2\right)\\
&= T(p)+T(q),
\end{aligned}$$

entonces T cumple con a). De manera similar como:

$$\begin{aligned}
T(kp) &= T\left(k(a_1 + b_1x + c_1x^2)\right) \\
&= T\left((ka_1) + (kb_1)x + (kc_1)x^2\right) \\
&= (ka_1 + kb_1, kc_1) \\
&= (k(a_1 + b_1), k(c_1)) \\
&= k\,(a_1 + b_1, c_1) \\
&= kT\left(a_1 + b_1x + c_1x^2\right) \\
&= kT(p),
\end{aligned}$$

se verifica la parte b) de la definición.

Así, la transformación T es una transformación lineal.◄

Ejemplo 6.2 (**Una transformación que no es lineal**). Sea la transformación

$$T : \mathbb{R} \to \mathbb{R}^2 \text{ tal que } T(x) = (x, x^2).$$

Para probar que no es una transformación lineal, se deberá verificar que no se cumple a) o b) de la definición. Para ello, basta con encontrar un par de elementos x, y en $\mathbb{R}$ tales que $T(x + y) \neq T(x) + T(y)$, o un escalar real k tal que $T(kx) \neq kT(x)$. Por ejemplo si $x = -1$, $y = 1$ se tiene

$$T(x + y) = T(-1 + 1) = T(0) = (0, 0),$$

mientras que

$$T(x) + T(y) = T(-1) + T(1) = (-1, 1) + (1, 1) = (0, 2),$$

por lo que no se satisface a) de la definición.

Ejemplo 6.3. Sea la transformación $T : \mathbb{M}_2 \to \mathbb{M}_2$ tal que

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a+b & 0 \\ 0 & c+d \end{bmatrix}.$$

Es claro que T es una función con dominio y codominio $\mathbb{M}_2$ que asigna a cada matriz $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ de $\mathbb{M}_2$, la única matriz $\begin{bmatrix} a+b & 0 \\ 0 & c+d \end{bmatrix}$ de $\mathbb{M}_2$.

Para determinar si T es lineal, deben verificarse a) y b) de la definición de transformación lineal. Sean las matrices $A = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}$ y $B = \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}$ de $\mathbb{M}_2$.

$$\begin{aligned}
T(A + B) &= T\left(\begin{bmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{bmatrix}\right) \\
&= \begin{bmatrix} (a_1 + a_2) + (b_1 + b_2) & 0 \\ 0 & (c_1 + c_2) + (d_1 + d_2) \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} (a_1+b_1)+(a_2+b_2) & 0 \\ 0 & (c_1+d_1)+(c_2+d_2) \end{bmatrix} \\
&= \begin{bmatrix} a_1+b_1 & 0 \\ 0 & c_1+d_1 \end{bmatrix} + \begin{bmatrix} a_2+b_2 & 0 \\ 0 & c_2+d_2 \end{bmatrix} \\
&= T\left(\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}\right) \\
&= T(A)+T(B),
\end{aligned}$$

con lo que se satisface a) de la definición de transformación lineal. Además, para cualquier $k \in \mathbb{R}$,

$$\begin{aligned}
T(kA) &= T\left(\begin{bmatrix} ka_1 & kb_1 \\ kc_1 & kd_1 \end{bmatrix}\right) \\
&= \left(\begin{bmatrix} ka_1+kb_1 & 0 \\ 0 & kc_1+kd_1 \end{bmatrix}\right) \\
&= \begin{bmatrix} k(a_1+b_1) & 0 \\ 0 & k(c_1+d_1) \end{bmatrix} \\
&= k\begin{bmatrix} a_1+b_1 & 0 \\ 0 & c_1+d_1 \end{bmatrix} \\
&= kT\left(\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}\right) \\
&= kT(A),
\end{aligned}$$

entonces se satisface b) de la definición. Así, T es una transformación lineal.

Es claro que el recorrido de T es el conjunto de todas las matrices diagonales de $\mathbb{M}_2$ ya que la imagen de cualquier matriz de $\mathbb{M}_2$ es una matriz diagonal y toda matriz diagonal $D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$ es imagen de alguna matriz de $\mathbb{M}_2$, en particular D es imagen de sí misma, $D = T(D)$.◀

Ejemplo 6.4 (**La transformación lineal cero**). Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales. La *transformación constante cero*, $O : \mathbb{V} \to \mathbb{W}$, definida por

$$O(\mathbf{v}) = \mathbf{0}, \quad \text{para todo } \mathbf{v} \in \mathbb{V},$$

es una transformación lineal. En efecto, si $\mathbf{u}$ y $\mathbf{v}$ son vectores de $\mathbb{V}$ y k es un escalar, entonces $O(\mathbf{u}+\mathbf{v}) = \mathbf{0}$, $O(\mathbf{u}) = \mathbf{0}$, $O(\mathbf{v}) = \mathbf{0}$ y $O(k\mathbf{u}) = \mathbf{0}$. Por lo tanto:

$$O(\mathbf{u}+\mathbf{v}) = \mathbf{0} = \mathbf{0}+\mathbf{0} = O(\mathbf{u})+O(\mathbf{v}),$$

$$O(k\mathbf{u}) = \mathbf{0} = k\mathbf{0} = kO(\mathbf{u}).$$

Nótese que el recorrido de la transformación cero es $R(O) = \{\mathbf{0}\}$. ◀

Actividad 6.1. Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales. Pruebe que una *transformación constante no cero*, esto es, $T : \mathbb{V} \to \mathbb{W}$ tal que $T(\mathbf{x}) = \mathbf{w}$ donde $\mathbf{w}$ es un vector fijo no cero, no es una transformación lineal.

Ejemplo 6.5 (La transformación lineal identidad). Sea $\mathbb{V}$ un espacio vectorial. La *transformación identidad* definida por

$$I(\mathbf{v}) = \mathbf{v}, \quad \text{para todo } \mathbf{v} \in \mathbb{V},$$

es una transformación lineal del espacio $\mathbb{V}$ en sí mismo.

En efecto: sean $\mathbf{u}$ y $\mathbf{v}$ vectores de $\mathbb{V}$ y k un escalar, entonces $I(\mathbf{u}+\mathbf{v}) = \mathbf{u}+\mathbf{v}$, $I(\mathbf{u}) = \mathbf{u}$, $I(\mathbf{v}) = \mathbf{v}$ y $I(k\mathbf{u}) = k\mathbf{u}$. Por lo tanto:

$$I(\mathbf{u}+\mathbf{v}) = \mathbf{u}+\mathbf{v} = I(\mathbf{u}) + I(\mathbf{v}),$$

$$I(k\mathbf{u}) = k\mathbf{u} = kI(\mathbf{u}).$$

Nótese que el recorrido de la transformación identidad es $R(I) = \mathbb{V}$. ◀

Ejemplo 6.6. Considere los espacios vectoriales $\mathbb{M}_{mn}$ y $\mathbb{M}_{nm}$ y la transformación $T : \mathbb{M}_{mn} \to \mathbb{M}_{nm}$ tal que

$$T(A) = A^t.$$

Esta transformación (transposición de matrices) es lineal. En efecto, por propiedades de la transposición de matrices:

$$T(A+B) = (A+B)^t = A^t + B^t = T(A) + T(B),$$

$$T(kA) = (kA)^t = kA^t = kT(A),$$

para todo par de matrices A y B de $m \times n$ y para cualquier escalar k. ◀

Nota: Si como en los ejemplos 6.3, 6.5 y 6.6, T es una transformación lineal de un espacio vectorial en sí mismo; esto es $T : \mathbb{V} \to \mathbb{V}$, se dice que T es un **operador lineal** sobre $\mathbb{V}$.

Las condiciones a) y b) que debe satisfacer una transformación para que sea lineal, pueden escribirse de manera simultánea, por lo tanto la definición de transformación lineal puede expresarse como sigue:

Teorema 6.1. Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales y una transformación $T : \mathbb{V} \to \mathbb{W}$. Entonces

$$T \text{ es una transformación lineal si y solo si } T(c\mathbf{u} + k\mathbf{v}) = cT(\mathbf{u}) + kT(\mathbf{v}),$$

para todo $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{V}$ y escalares reales c y k.

Demostración. Por un lado, si T es una transformación lineal, por la parte a) de la definición, se tiene que si $\mathbf{u}$ y $\mathbf{v}$ son vectores de $\mathbb{V}$ y c y k escalares en $\mathbb{R}$:

$$T(c\mathbf{u} + k\mathbf{v}) = T(c\mathbf{u}) + T(k\mathbf{v}), \tag{6.1}$$

y por b)

$$T(c\mathbf{u}) = cT(\mathbf{u}) \;\text{ y }\; T(k\mathbf{v}) = kT(\mathbf{v}). \tag{6.2}$$

Reemplazando las expresiones de (6.2) en la expresión dada en (6.1) se obiene

$$T(c\mathbf{u} + k\mathbf{v}) = cT(\mathbf{u}) + kT(\mathbf{v}).$$

Recíprocamente, si $T(c\mathbf{u} + k\mathbf{v}) = cT(\mathbf{u}) + kT(\mathbf{v})$, tomando $c = k = 1$ se prueba la parte a) de la definición de transformación lineal y tomando $c = 0$ se verifica b). Por lo tanto, T es una transformación lineal. □

Ejemplo 6.7 (**Proyección de un vector sobre una recta que pasa por el origen**). Sea $\mathbf{u}$ un vector fijo distinto de cero del espacio euclidiano $\mathbb{R}^3$. La transformación $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$T(\mathbf{v}) = \text{proy}_{\mathbf{u}}\mathbf{v},$$

es una transformación lineal. En efecto:

$$\text{proy}_{\mathbf{u}}\mathbf{v} = \frac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u},$$

$$T(c\mathbf{v}_1 + k\mathbf{v}_2) = \frac{(c\mathbf{v}_1 + k\mathbf{v}_2)\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{c\mathbf{v}_1\cdot\mathbf{u} + k\mathbf{v}_2\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = c\frac{\mathbf{v}_1\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} + k\frac{\mathbf{v}_2\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = cT(\mathbf{v}_1) + kT(\mathbf{v}_2),$$

para todo $\mathbf{v}_1$ y $\mathbf{v}_2$ en $\mathbb{R}^3$ y c, k escalares.◄

Ejemplo 6.8 (**Transformaciones matriciales: multiplicación por una matriz**). Sea, por ejemplo, la matriz $A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. Para un vector cualquiera $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ de $\mathbb{R}^3$, al efectuar el producto $A\mathbf{x}$ se obtiene el único vector de $\mathbb{R}^2$:

$$A\mathbf{x} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y \\ x+y+z \end{bmatrix}.$$

Se puede decir entonces que la matriz A define la transformación $T : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $T(\mathbf{x}) = A\mathbf{x}$.

En general si A es una matriz de $m \times n$ y se utiliza la notación de vector columna para vectores en $\mathbb{R}^n$ y $\mathbb{R}^m$, entonces se puede definir una función $T : \mathbb{R}^n \to \mathbb{R}^m$ por medio de:

$$T(\mathbf{x}) = A\mathbf{x}$$

ya que siendo $\mathbf{x}$ un vector de $\mathbb{R}^n$ (matriz de $n \times 1$) y A una matriz de $m \times n$, entonces $A\mathbf{x}$ es una matriz de $m \times 1$ (vector de $\mathbb{R}^m$). Diremos que esta transformación es **la multiplicación por** A. Este tipo de transformaciones se denominan **transformaciones matriciales** y la matriz A se llama **matriz estándar de** T.

Las transformaciones matriciales son transformaciones lineales. En efecto, sea T una transformación matricial de matriz estándar A de $m \times n$, y sean $\mathbf{x}_1$, $\mathbf{x}_2$ vectores de $\mathbb{R}^n$ y k_1, k_2 escalares reales, entonces por propiedades del producto de matrices:

$$T(k_1\mathbf{x}_1 + k_2\mathbf{x}_2) = A(k_1\mathbf{x}_1 + k_2\mathbf{x}_2) = k_1(A\mathbf{x}_1) + k_2(A\mathbf{x}_2) = k_1T(\mathbf{x}_1) + k_2T(\mathbf{x}_2).$$ ◄

Son inmediatas las siguientes propiedades de las transformaciones lineales:

Teorema 6.2. Sea $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Entonces:

1. $T(\mathbf{0}) = \mathbf{0}$ (donde el $\mathbf{0}$ del primer miembro es el cero en $\mathbb{V}$ y el $\mathbf{0}$ del segundo miembro es el cero en $\mathbb{W}$).
2. $T(-\mathbf{v}) = -T(\mathbf{v})$, para todo $\mathbf{v} \in \mathbb{V}$
3. $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$, para todo $\mathbf{u}, \mathbf{v} \in \mathbb{V}$

Demostración. Se probará la propiedad 1. Dado que $0\mathbf{v} = \mathbf{0}$ para cualquier vector $\mathbf{v}$ de $\mathbb{V}$ y de la parte b) de la definición de transformación lineal se tiene:

$$T(\mathbf{0}) = T(0\mathbf{v}) = 0T(\mathbf{v}) = \mathbf{0}.$$

□

Actividad 6.2. Demuestre las propiedades 2. y 3. del Teorema 6.2.

El teorema anterior dá condiciones necesarias para que una transformación $T : \mathbb{V} \to \mathbb{W}$ sea una transformación lineal. Si alguna de éstas no se cumple, se puede afirmar que T no es una transformación lineal. En particular es fácil verificar si $T(\mathbf{0})$ es, o no, el cero de $\mathbb{W}$.

Ejemplo 6.9 (**Una transformación que no es lineal**)**.** Sea la transformación $T : \mathbb{R}^2 \to \mathbb{R}^3$ tal que $T(x, y) = (x, y, x + 1)$. Dado que

$$T(0, 0) = (0, 0, 1) \neq (0, 0, 0),$$

no se verifica la propiedad 1 del teorema anterior y por lo tanto T no es una transformación lineal.◀

Determinación de una transformación lineal conocidos sus valores sobre una base del espacio dominio

El Teorema 6.1 puede generalizarse para un número finito de vectores. Esto es, si T es una transformación lineal de $\mathbb{V}$ en $\mathbb{W}$,

$$T(k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + ... + k_n\mathbf{v}_n) = k_1T(\mathbf{u}_1) + k_2T(\mathbf{v}_2) + ... + k_nT(\mathbf{v}_n), \tag{6.3}$$

para todo $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_n$ de $\mathbb{V}$ y k_1, k_2,..., k_n escalares reales.

Una función de un conjunto $\mathbb{V}$ en un conjunto $\mathbb{W}$ puede especificarse mediante una fórmula que asigne a cada elemento de $\mathbb{V}$ un único elemento de $\mathbb{W}$. También se puede especificar una función indicando qué elemento de $\mathbb{W}$ se le asigna a cada elemento de $\mathbb{V}$. Parece imposible describir de esta última forma una transformación lineal $T : \mathbb{V} \to \mathbb{W}$ donde $\mathbb{V} \neq \{\mathbf{0}\}$ pues todo espacio vectorial distinto del espacio cero tiene infinitos elementos. Sin embargo, la propiedad (6.3) permite determinar la imagen $T(\mathbf{v})$ para cada $\mathbf{v}$ del espacio $\mathbb{V}$ conociendo solo las imágenes de los vectores de una base de $\mathbb{V}$.

Al respecto vale el siguiente teorema cuya demostración es directa.

Teorema 6.3. Sea $\mathbb{V}$ un espacio vectorial de dimensión n y T una transformación lineal del espacio $\mathbb{V}$ en un espacio $\mathbb{W}$. Si $B = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es una base de $\mathbb{V}$ y $\mathbf{w}_1$, $\mathbf{w}_2$,...,$\mathbf{w}_n$ son las imágenes de $\mathbf{v}_1$, $\mathbf{v}_2$,...,$\mathbf{v}_n$, respectivamente, entonces para cada $\mathbf{v}$ en $\mathbb{V}$,

$$T(\mathbf{v}) = k_1\mathbf{w}_1 + k_2\mathbf{w}_2 + ... + k_n\mathbf{w}_n,$$

donde k_1, k_2, ..., k_n son los (únicos) escalares tales que $\mathbf{v} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + ... + k_n\mathbf{v}_n$.

Demostración. Sea $\mathbf{v}$ un vector del espacio $\mathbb{V}$. Como B es una base de $\mathbb{V}$, existen únicos escalares k_1, k_2,..., k_n tal que que

$$\mathbf{v} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + ... + k_n\mathbf{v}_n.$$

Dado que T es una transformación lineal y $T(\mathbf{v}_i) = \mathbf{w}_i$ para todo $i = 1, 2, ..., n$, aplicando T a la expresión anterior se obtiene:

$$T(\mathbf{v}) = k_1T(\mathbf{v}_1) + k_2T(\mathbf{v}_2) + ... + k_nT(\mathbf{v}_n) = k_1\mathbf{w}_1 + k_2\mathbf{w}_2 + ... + k_n\mathbf{w}_n$$

□

Ejemplo 6.10. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal tal que

$$T(1, 0) = (2, 4) \quad \text{y} \quad T(1, 1) = (1, 3).$$

Para hallar la ley de T, es decir, $T(x, y)$ para todo (x, y) de $\mathbb{R}^2$, se escribe al vector genérico (x, y) como combinación lineal de los vectores $(1, 0)$ y $(1, 1)$ que conforman una base de $\mathbb{R}^2$. Esto es, se buscan los (únicos) valores de k_1 y k_2 tales que

$$(x, y) = k_1(1, 0) + k_2(1, 1),$$

o bien

$$(x, y) = (k_1 + k_2, k_2)$$

que resultan de resolver el sistema lineal

$$\begin{cases} k_1 + k_2 = x \\ \qquad k_2 = y \end{cases}$$

La solución es

$$\begin{cases} k_1 & = & x - y \\ k_2 & = & y \end{cases}$$

Así,

$$(x, y) = (x - y)(1, 0) + y(1, 1),$$

y dado que T es una transformación lineal se tiene

$$\begin{array}{rcl} T(x, y) & = & T\left((x - y)(1, 0) + y(1, 1)\right) \\ & = & (x - y)T(1, 0) + yT(1, 1) \\ & = & (x - y)(2, 4) + y(1, 3) \\ & = & (2x - y, 4x - y). \blacktriangleleft \end{array}$$

Actividad 6.3. Sea $f : \mathbb{R} \to \mathbb{R}$ una transformación lineal del espacio vectorial $\mathbb{R}$ en sí mismo. Suponga que $f(1) = a$. Halle $f(x)$ para todo $x \in \mathbb{R}$.

Esto muestra que toda transformación lineal f de $\mathbb{R}$ en $\mathbb{R}$ es de la forma $f(x) = ax$ para algún $a \in \mathbb{R}$.

Si no se dan los valores de T sobre una base no se puede determinar la imagen de cada elemento de $\mathbb{V}$. Si en el ejemplo anterior se dieran solo las imágenes del $(1,1)$ y $(-3,-3)$ no se podría calcular la imagen, por ejemplo, del vector $(2,4)$ ya que no es combinación lineal de ellos.

En general, vale el siguiente teorema cuya demostración queda como ejercicio.

Teorema 6.4. Si $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ es un conjunto generador de un espacio vectorial $\mathbb{V}$ y $T : \mathbb{V} \to \mathbb{W}$ es una transformación lineal, entonces $T(S) = \{T(\mathbf{v}_i) : \ \mathbf{v}_i \in S\}$ es un generador del recorrido $R(T)$.

Una consecuencia inmediata es la siguiente:

Corolario 6.1. Si $T : \mathbb{V} \to \mathbb{W}$ es una transformación lineal y $\mathbb{V}$ es finito dimensional, entonces el recorrido $R(T)$ es de dimensión finita y $\dim(R(T)) \leq \dim(\mathbb{V})$.

Ejercicios 6.1

1. Sea la transformación T que asigna a cada vector (x, y) de $\mathbb{R}^2$, el vector $(x - y, y, 0)$ de $\mathbb{R}^3$.

   a) Identifique el dominio y codominio de T.
   b) Halle las imágenes de los vectores $\mathbf{u} = (-1, 3)$, $\mathbf{v} = (2, -1)$ y $\mathbf{0}$.
   c) Determine todos los vectores del dominio que se aplican al vector cero de $\mathbb{R}^3$.
   d) ¿Existe algún vector $\mathbf{v}$ en $\mathbb{R}^2$ tal que $T(\mathbf{v}) = (-1, 1, 1)$?
   e) Describa el recorrido de T.
   f) Pruebe que T es una transformación lineal.

2. Determine si la transformación T es lineal.

   a) $T : \mathbb{R}^2 \to \mathbb{R}^2, \quad T(x, y) = (3x + y, x - 2y)$.
   b) $T : \mathbb{R}^3 \to \mathbb{R}, \quad T(x, y, z) = xyz$.
   c) $T : \mathbb{P}_1 \to \mathbb{R}^2, \quad T(a + bx) = (a + b, 0)$.
   d) $T : \mathbb{R}^3 \to \mathbb{P}_2, \quad T(a, b, c) = a + (b + c)x + cx^2$.
   e) $T : \mathbb{P}_2 \to \mathbb{M}_2, \quad T(a + bx + cx^2) = \begin{bmatrix} a & -b \\ 0 & ac \end{bmatrix}$.
   f) $T : \mathbb{P}_2 \to \mathbb{P}_1, \quad T(a + bx + cx^2) = b + 2cx$.
   g) $T : C(\mathbb{R}) \to \mathbb{R}, \quad T(f) = \lim\limits_{x \to a} f(x)$, donde a es un número real fijo.
   h) $T : \mathbb{M}_2 \to \mathbb{R}, \quad T(A) = \det(A)$.
   i) $T : \mathbb{M}_n \to \mathbb{M}_n, \quad T(A) = P^{-1}AP$, donde P es una matriz fija invertible de orden n.
   j) $T : \mathbb{R}^n \to \mathbb{R}, \quad T(\mathbf{v}) = \mathbf{v} \cdot \mathbf{u}$, donde $\mathbf{u}$ es un vector fijo de $\mathbb{R}^n$.
   k) $T : \mathbb{V} \to \mathbb{R}, \quad T(\mathbf{v}) = \langle \mathbf{u}, \mathbf{v} \rangle$, donde $\langle , \rangle$ es un producto interno en el espacio $\mathbb{V}$ y $\mathbf{u}$ es un vector fijo de $\mathbb{V}$.

l) $T : \mathbb{R}^3 \to \mathbb{R}^3$, $\quad T(\mathbf{v}) = \mathbf{v} \times \mathbf{u}$, donde $\times$ indica el producto vectorial entre el vector $\mathbf{v}$ y un vector fijo $\mathbf{u}$ de $\mathbb{R}^3$.

$\boxed{3.}$ Sea $\mathbb{V}$ un espacio vectorial de dimensión n y $B = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ una base de $\mathbb{V}$. Considere la transformación que asigna a cada vector $\mathbf{v}$ de $\mathbb{V}$, su vector de coordenadas respecto a la base B. Esto es:

$$T : \mathbb{V} \to \mathbb{R}^n \quad \text{tal que} \quad T(\mathbf{v}) = [\mathbf{v}]_B \,,$$

a) Demuestre que T es una transformación lineal (véase demostración del Teorema 4.15, 1.)

b) Si $\mathbb{V} = \mathbb{M}_2$ $(n = 4)$ y $B = \{E_{11}, E_{11} + E_{12}, E_{11} + E_{12} + E_{21}, E_{11} + E_{12} + E_{21} + E_{22}\}$, obtenga la imagen de la matriz $\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$.

4. En cada caso sea $T : \mathbb{R}^n \to \mathbb{R}^m$ la transformación matricial de matriz estándar A. Halle n y m y la imagen de un vector genérico $\mathbf{x}$ de $\mathbb{R}^n$.

$\boxed{a)}$ $A = \begin{bmatrix} 2 & 3 \\ -1 & 5 \\ 0 & 4 \end{bmatrix}$

b) $A = \begin{bmatrix} 1 & -2 & 5 & 3 \end{bmatrix}$

5. Sea T la transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$ tal que $T(-1, 1) = (5, 1)$ y $T(1, 1) = (0, 2)$.

a) Determine la ley de T, es decir, $T(x, y)$ para todo (x, y) de $\mathbb{R}^2$.

b) Halle el recorrido de T.

$\boxed{6.}$ Sea la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^2$ tal que

$$T(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) = (-1, 3), \quad T(-\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) = (1, 1), \quad T(\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3) = (0, 1),$$

halle $T(\mathbf{x})$ para todo $\mathbf{x} \in \mathbb{R}^3$.

$\boxed{7.}$ Sea $T : \mathbb{P}_1 \to \mathbb{P}_2$ la transformación lineal tal que $T(-1+3x) = 1+x-x^2$ y $T(2x) = x^2$. Obtenga la ley de T; esto es, $T(a + bx)$ para todo polinomio $p = a + bx$ de $\mathbb{P}_1$.

$\boxed{8.}$ Sea $T : \mathbb{P} \to \mathbb{P}$ la transformación lineal tal que satisface

$$T(x^n) = \frac{1}{n+1} x^{n+1},$$

para todo $n \geq 0$. Calcule: $T(1 + x + x^2)$, $T(-1 + x^3)$ y $T(2x)$.

9. Demuestre el Teorema 6.4 y su corolario (Corolario 6.1).

$\boxed{10.}$ Indique si las siguientes proposiciones son verdaderas o falsas. Jusitifique sus respuestas.

a) Existe una transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(-1, 1) = (0, 1)$ y $T(2, -2) = (1, 4)$.

b) Si la transformación $T : \mathbb{V} \to \mathbb{W}$ es tal que $T(\mathbf{0}) = \mathbf{0}$, entonces T es lineal.

c) Sea $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal y $\{\mathbf{v}_1, \mathbf{v}_2\}$ un subconjunto linealmente dependiente de $\mathbb{V}$. Entonces $\{T(\mathbf{v}_1), T(\mathbf{v}_2)\}$ es un subconjunto linealmente dependiente de $\mathbb{W}$.

11. Las siguientes son algunas de las transformaciones más utilizadas en el Cálculo. Estas transformaciones están definidas en diferentes espacios vectoriales:

- El espacio vectorial $\dot{C}(\mathbb{R})$ *de las funciones de* $\mathbb{R}$ *en* $\mathbb{R}$ *derivables con las operaciones habituales de suma y multiplicación por escalar definidas en* $\mathbb{F}(\mathbb{R})$.
- *El espacio vectorial* $\mathbb{F}(\mathbb{R}^3)$ *de los campos escalares de* $\mathbb{R}^3$ *en* $\mathbb{R}$ *con las operaciones habituales de suma y multiplicación por escalar.*
- *El espacio vectorial* $\dot{\mathbb{F}}(\mathbb{R}^3)$ *de los campos escalares de* $\mathbb{R}^3$ *en* $\mathbb{R}$ *tal que existen las derivadas parciales de primer orden con las operaciones habituales de suma y multiplicación por escalar en* $\mathbb{F}(\mathbb{R}^3)$.
- *El espacio vectorial* $\mathbb{F}^3(\mathbb{R}^3)$ *de los campos vectoriales de* $\mathbb{R}^3$ *en* $\mathbb{R}^3$ *con las operaciones habituales de suma y multiplicación por escalar.*
- *El espacio vectorial* $\dot{\mathbb{F}}^3(\mathbb{R}^3)$ *de los campos vectoriales de* $\mathbb{R}^3$ *en* $\mathbb{R}^3$ *cuyas funciones componentes tienen derivadas parciales de primer orden con las operaciones habituales de suma y multiplicación por escalar en* $\mathbb{F}^3(\mathbb{R}^3)$.
- *El espacio vectorial* $A = \left\{ f \in C[0,+\infty) : \int_0^{+\infty} e^{-xt} f(t)\, dt < \infty \right\}$ *con las operaciones habituales de suma y multiplicación por escalar definidas en* $\mathbb{F}(\mathbb{R})$.

Demuestre que son transformaciones lineales:

$\boxed{a)}$ (Integral) $I : C[a,b] \to \mathbb{R}$ tal que $I(f) = \displaystyle\int_a^b f(x)\, dx$.

$\boxed{b)}$ (Derivada) $D : \dot{C}(\mathbb{R}) \to \mathbb{F}(\mathbb{R})$ tal que $D(f) = f'$.

c) (Gradiente) $\nabla : \dot{\mathbb{F}}(\mathbb{R}^3) \to \mathbb{F}^3(\mathbb{R}^3)$ tal que $\nabla f = \left(\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}, \dfrac{\partial f}{\partial z} \right)$.

d) (Rotacional) $\mathrm{rot} : \dot{\mathbb{F}}^3(\mathbb{R}^3) \to \mathbb{F}^3(\mathbb{R}^3)$ tal que $\mathrm{rot}(\mathbf{F}) = \nabla \times \mathbf{F}$.

e) (Divergencia) $\mathrm{div} : \dot{\mathbb{F}}^3(\mathbb{R}^3) \to \mathbb{F}(\mathbb{R}^3)$ tal que $\mathrm{div}(\mathbf{F}) = \nabla \cdot \mathbf{F}$.

f) (Transformada de Laplace) $\mathfrak{L} : A \to \mathbb{F}(\mathbb{R})$ tal que $\mathfrak{L}\{f(t)\} = \displaystyle\int_0^{+\infty} e^{-xt} f(t)\, dt$.

6.2. Núcleo y recorrido de una transformación lineal

En la sección anterior se definió el *recorrido* $R(T)$ de una transformación $T : \mathbb{V} \to \mathbb{W}$ como

$$R(T) = \{\mathbf{w} \in \mathbb{W} : \mathbf{w} = T(\mathbf{v}),\ \text{para algún } \mathbf{v} \in \mathbb{V}\}.$$

En esta sección definiremos un subconjunto importante de $\mathbb{V}$, denominado *núcleo* o *kernel* de T.

Definición (Núcleo o kernel). Sea $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal. El *núcleo* o *kernel* de T es el conjunto de los vectores $\mathbf{v}$ de $\mathbb{V}$ que tienen por imagen al vector $\mathbf{0}$ (de $\mathbb{W}$). Se lo denota $\ker(T)$.

$$\ker(T) = \{\mathbf{v} \in \mathbb{V} : T(\mathbf{v}) = \mathbf{0}\}$$

Se verá que los conjuntos $\ker(T)$ y $R(T)$ son subespacios de $\mathbb{V}$ y $\mathbb{W}$, respectivamente.

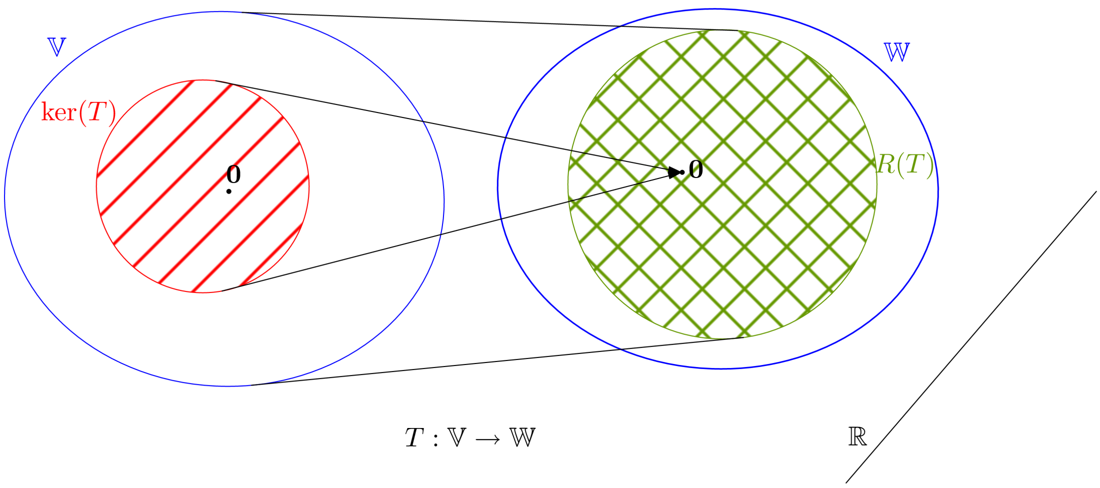


Figura 6.2: $T : \mathbb{V} \to \mathbb{W}$, núcleo $\ker(T)$ y recorrido $R(T)$

Teorema 6.5. Sea $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Entonces:

1. $\ker(T)$ es un subespacio de $\mathbb{V}$.
2. $R(T)$ es un subespacio de $\mathbb{W}$.

Demostración. Dado que $T : \mathbb{V} \to \mathbb{W}$ es una transformación lineal se sabe que $T(\mathbf{0}) = \mathbf{0}$, luego el $\mathbf{0}$ de $\mathbb{V}$ está en $\ker(T)$ y el $\mathbf{0}$ de $\mathbb{W}$ está en $R(T)$. Por lo tanto $R(T)$ y $\ker(T)$ son subconjuntos no vacíos de $\mathbb{V}$ y $\mathbb{W}$, respectivamente.

1. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores de $\ker(T)$ y k un número real. Como $\mathbf{u}$ y $\mathbf{v}$ están en $\ker(T)$ entonces $T(\mathbf{u}) = \mathbf{0}$ y $T(\mathbf{v}) = \mathbf{0}$. Luego:
$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0},$$
y
$$T(k\mathbf{u}) = kT(\mathbf{u}) = k\mathbf{0} = \mathbf{0}.$$
Dado que $T(\mathbf{u} + \mathbf{v}) = \mathbf{0}$ y $T(k\mathbf{u}) = \mathbf{0}$, entonces tanto $\mathbf{u} + \mathbf{v}$ como $k\mathbf{u}$ están en $\ker(T)$. Así, $\ker(T)$ es cerrado bajo la suma y la multiplicación por escalares definidas en $\mathbb{V}$, lo que prueba que $\ker(T)$ es un subespacio de $\mathbb{V}$.

2. Sean $\mathbf{w}_1$ y $\mathbf{w}_2$ vectores de $R(T)$ y k un escalar real. Debe probarse que $\mathbf{w}_1 + \mathbf{w}_2$ y $k\mathbf{w}_1$ están en $R(T)$; esto es, que $\mathbf{w}_1 + \mathbf{w}_2$ y $k\mathbf{w}_1$ son imágenes de vectores de $\mathbb{V}$.

   Supuesto que $\mathbf{w}_1$ y $\mathbf{w}_2$ están en $R(T)$, entonces existen vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{V}$ tales que $T(\mathbf{u}) = \mathbf{w}_1$ y $T(\mathbf{v}) = \mathbf{w}_2$. Luego y dado que T es lineal:
$$\mathbf{w}_1 + \mathbf{w}_2 = T(\mathbf{u}) + T(\mathbf{v}) = T(\mathbf{u} + \mathbf{v}),$$
esto prueba que $\mathbf{w}_1 + \mathbf{w}_2$ está en $R(T)$ ya que es la imagen del vector $\mathbf{u} + \mathbf{v}$ de $\mathbb{V}$.

   Además,
$$k\mathbf{w}_1 = kT(\mathbf{u}) = T(k\mathbf{u}),$$

por lo que el vector $k\mathbf{w}_1$ está en $R(T)$ ya que es la imagen del vector $k\mathbf{u}$ de $\mathbb{V}$.

Así, $R(T)$ es cerrado bajo la suma y la multiplicación por escalares y por lo tanto es un subespacio de $\mathbb{W}$.

$\square$

Definición (Rango y nulidad). Sea $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal.

Si $R(T)$ es finito dimensional, a su dimensión se la llama *rango de* T y se representa por $\text{rango}(T)$ o $\rho(T)$.

Si $\ker(T)$ es finito dimensional, a su dimensión se la llama *nulidad* de T y se representa por $\text{nulidad}(T)$ o $\nu(T)$.

Ejemplo 6.11. Sea la transformación lineal $T : \mathbb{P}_2 \to \mathbb{R}^2$ tal que $T\left(a + bx + cx^2\right) = (a + b, c)$ del Ejemplo 6.1. El núcleo de T es el siguiente subespacio de $\mathbb{P}_2$:

$$
\begin{aligned}
\ker(T) &= \left\{a + bx + cx^2 \in \mathbb{P}_2 : T(a + bx + cx^2) = (0, 0)\right\} \\
&= \left\{a + bx + cx^2 \in \mathbb{P}_2 : (a + b, c) = (0, 0)\right\} \\
&= \left\{a + bx + cx^2 \in \mathbb{P}_2 : a + b = 0,\ c = 0\right\} \\
&= \left\{a + bx + cx^2 \in \mathbb{P}_2 : a = -b,\ c = 0,\ b \in \mathbb{R}\right\} \\
&= \{-b + bx \in \mathbb{P}_2 : b \in \mathbb{R}\} \\
&= \{b(-1 + x) : b \in \mathbb{R}\} \\
&= \text{gen}\,\{-1 + x\}
\end{aligned}
$$

El conjunto unitario $\{-1 + x\}$ es linealmente independiente y generador de $\ker(T)$ por lo tanto es una base de $\ker(T)$. Así, $\nu(T) = \dim(\ker(T)) = 1$.

Es fácil ver que el recorrido de T es todo el espacio $\mathbb{R}^2$. Por ejemplo, el vector $(3, 5)$ es la imagen del polinomio $3 + 5x^2$ ya que $T(3 + 5x^2) = T(3 + 0x + 5x^2) = (3 + 0, 5) = (3, 5)$, también es la imagen del polinomio $1+2x+5x^2$. En general, para cualquier vector (a, b) de $\mathbb{R}^2$ es $(a, b) = T((a-c)+cx+bx^2)$ con $c \in \mathbb{R}$, por lo tanto todo vector de $\mathbb{R}^2$ es un vector de $R(T)$. Así, $R(T) = \mathbb{R}^2$ y dado que $\dim(\mathbb{R}^2) = 2$, entonces $\rho(T) = 2$.◀

Ejemplo 6.12. Para la transformación lineal cero $O : \mathbb{V} \to \mathbb{W}$, dado que $O(\mathbf{v}) = \mathbf{0}$ para todo $\mathbf{v}$ en $\mathbb{V}$ su núcleo es el conjunto $\mathbb{V}$. Su recorrido es el conjunto unitario $\{\mathbf{0}\}$ puesto que $\mathbf{0}$ es la única imagen. Esto es:

$$\ker(O) = \mathbb{V} \quad \text{y} \quad R(O) = \{\mathbf{0}\}\,.$$

Así, $\rho(O) = 0$ y, si $\mathbb{V}$ es de dimensión finita $\nu(O) = \dim(\mathbb{V})$.

Para la transformación lineal identidad $I : \mathbb{V} \to \mathbb{V}$ dado que $I(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ en $\mathbb{V}$ entonces el único vector que se transforma en cero es el mismo cero, y todo vector $\mathbf{v}$ de $\mathbb{V}$ es imagen de sí mismo. Por lo tanto

$$\ker(I) = \{\mathbf{0}\} \quad \text{y} \quad R(I) = \mathbb{V}.$$

Así, $\nu(I) = 0$ y, si $\mathbb{V}$ es finito dimensional $\rho(I) = \dim(\mathbb{V})$.◀

En el caso en que T es una transformación matricial, vale el siguiente resultado.

Teorema 6.6. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación matricial de matriz estándar A. Entonces:

1. $\ker(T) = N(A)$
2. $R(T) = \mathrm{col}(A)$
3. $\rho(T) = \rho(A)$
4. $\nu(T) = \nu(A)$

Demostración. La transformación T es la multiplicación por la matriz A es decir $T(\mathbf{x}) = A\mathbf{x}$ para todo $\mathbf{x}$ de $\mathbb{R}^n$.

1. Recordando que el espacio nulo de una matriz A, $N(A)$, es el conjunto de los $\mathbf{x}$ en $\mathbb{R}^n$ tal que $A\mathbf{x} = \mathbf{0}$ y $T(\mathbf{x}) = A\mathbf{x}$, resulta

$$\mathbf{x} \in \ker(T) \Leftrightarrow T(\mathbf{x}) = \mathbf{0} \Leftrightarrow A\mathbf{x} = \mathbf{0} \Leftrightarrow \mathbf{x} \in N(A).$$

2. El espacio de las columnas de A, $\mathrm{col}(A)$, es el conjunto de todos los vectores de $\mathbb{R}^m$ que son combinaciones lineales de las columnas de A. Así,

$$\mathbf{b} \in R(T) \Leftrightarrow \mathbf{b} = T(\mathbf{x}), \text{ para algún } \mathbf{x} \in \mathbb{R}^n \Leftrightarrow \mathbf{b} = A\mathbf{x}, \text{ para algún } \mathbf{x} \in \mathbb{R}^n \Leftrightarrow \mathbf{b} \in \mathrm{col}(A).$$

3. Como $\ker(T) = N(A)$ entonces $\nu(T) = \dim(\ker(T)) = \dim(N(A)) = \nu(A)$.
4. $\rho(A) = \dim(\mathrm{col}(A))$ y como $R(T) = \mathrm{col}(A)$ entonces $\rho(T) = \dim(R(T)) = \rho(A)$.

□

Se sabe, por el teorema del rango, que para toda matriz A de $m \times n$, la suma del rango y la nulidad es igual al número de columnas de A (Teorema 4.22). Se tiene entonces el siguiente resultado.

Corolario 6.2. Si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación matricial entonces la suma del rango de T y la nulidad de T es igual a n. Esto es:

$$\rho(T) + \nu(T) = n.$$

Demostración. Sea A la matriz estándar de T. Entonces A es una matriz de $m \times n$ y por el teorema del rango (Teorema 4.22):

$$\rho(A) + \nu(A) = n.$$

Por el teorema anterior, $\rho(T) = \rho(A)$ y $\nu(T) = \nu(A)$, luego

$$\rho(T) + \nu(T) = n.$$

□

Ejemplo 6.13. Sea T la transformación matricial de matriz estándar $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$.

Entonces $T : \mathbb{R}^3 \to \mathbb{R}^2$ y

$$T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_1 + x_2 \end{bmatrix}.$$

La forma escalonada reducida de A es la matriz $B = \begin{bmatrix} \boxed{1} & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. El número de pivotes da el rango de A, $\rho(A) = 1$, luego $\rho(T) = 1$. El número de columnas no pivotes de A es $\nu(A) = 3 - 1 = 2$ y así $\nu(T) = 2$.

$\mathrm{col}(A)$ es el espacio generado por las columnas pivote de A, así

$$R(T) = \mathrm{col}(A) = \mathrm{gen}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\} = \left\{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{R}^2 : y_1 = y_2\right\}.$$

El espacio nulo de A, $N(A)$, es el conjunto de las soluciones del sistema $A\mathbf{x} = \mathbf{0}$ que es equivalente a la ecuación $x_1 + x_2 + 0x_3 = 0$. La solución general es

$$\begin{cases} x_1 = -t \\ x_2 = t \\ x_3 = r \end{cases}, \quad t, r \in \mathbb{R}.$$

Así,

$$\ker(T) = N(A) = \left\{\begin{bmatrix} -t \\ t \\ r \end{bmatrix}, \; t, r \in \mathbb{R}\right\} = \mathrm{gen}\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}. \blacktriangleleft$$

El siguiente teorema, cuya demostración se omitirá en este momento, es uno de los más importantes del Álgebra Lineal, y generaliza al Corolario 6.2.

Teorema 6.7 (de la dimensión)**.** Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales tal que $\mathbb{V}$ es finito dimensional. Si $T : \mathbb{V} \to \mathbb{W}$ es una transformación lineal entonces

$$\dim(R(T)) + \dim(\ker(T)) = \dim(\mathbb{V}).$$

Esto es: $\rho(T) + \nu(T) = \dim(\mathbb{V})$.

Debe observarse que una de las hipótesis del teorema es que $\mathbb{V}$ es un espacio de dimensión finita. Esto permite asegurar que ambos, $\ker(T)$ (por ser un subespacio de $\mathbb{V}$) y $R(T)$ (Teorema 6.4 y su Corolario 6.1), son de dimensión finita.

Ejercicios 6.2

1. Obtenga el núcleo (kernel) de las transformaciones lineales $a)$, $c)$, $d)$, $f)$ y $i)$ del Ejercicio 2 de la Sección 6.1.

2. Para cada transformación lineal T, halle los subespacios $\ker(T)$ y $R(T)$, determine bases (si es posible) de los mismos y verifique el teorema de la dimensión.

   a) $T:\mathbb{R}^2 \to \mathbb{R}^3$, $\quad T(x,y)=(x-y,0,-2y)$.

   $\boxed{b)}$ $T:\mathbb{R}^3 \to \mathbb{R}^3$, $\quad T(\mathbf{x})=-3\mathbf{x}$.

   c) $T:\mathbb{P}_2 \to \mathbb{P}_1$, $\quad T(a+bx+cx^2)=(a+b+c)x$.

   d) $T:\mathbb{M}_2 \to \mathbb{M}_2$, $\quad T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)=\begin{bmatrix} d & b+c \\ 0 & -a \end{bmatrix}$.

   $\boxed{e)}$ $T:\mathbb{P}_2 \to \mathbb{R}^2$, $\quad T(a+bx+cx^2)=(a+b,a-c)$.

   $\boxed{f)}$ $T:\mathbb{R}^3 \to \mathbb{M}_{23}$, $\quad T(x,y,z)=\begin{bmatrix} x-y & 0 & 0 \\ z & 0 & -z \end{bmatrix}$.

3. Halle el rango y nulidad de la transformación lineal $T:\mathbb{M}_2 \to \mathbb{R}^2$ tal que

$$T(E_{11})=\mathbf{e}_1, \quad T(E_{12})=\mathbf{0}, \quad T(E_{21})=\mathbf{e}_1+\mathbf{e}_2, \quad T(E_{22})=\mathbf{e}_2.$$

$\boxed{4.}$ Halle el núcleo y el recorrido de la transformación lineal $T:\mathbb{P}_2 \to \mathbb{P}_2$ tal que $T\left(p(x)\right)=p(1-x)$. Calcule el rango y nulidad de T.

$\boxed{5.}$ Dada la transformación $T:\mathbb{M}_n \to \mathbb{M}_n$ tal que $T(X)=AX-XA$, donde A es una matriz cuadrada fija de orden n.

   a) Demuestre que T es una transformación lineal.

   b) Considere $n=2$ y $A=\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Indique el rango y la nulidad de T.

6. Halle el núcleo, el recorrido, el rango y la nulidad de la transformación matricial $T:\mathbb{R}^n \to \mathbb{R}^m$ donde la matriz estándar, A, de T es:

   a) $A=\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

   $\boxed{b)}$ $A=\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$

   c) $A=\begin{bmatrix} 1 & -2 & 1 \\ 2 & 2 & 0 \end{bmatrix}$

   $\boxed{d)}$ $A=\begin{bmatrix} 1 & 1 \\ -2 & 2 \\ 1 & -1 \end{bmatrix}$

   $\boxed{e)}$ $A=\begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & -3 \\ -1 & -1 & 2 \end{bmatrix}$

   f) $A=\begin{bmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \\ -1 & 1 & -2 \end{bmatrix}$

   $\boxed{g)}$ $A=\begin{bmatrix} 1 & 0 & -1 & 1 \\ -1 & 0 & -1 & -1 \end{bmatrix}$

$\boxed{7.}$ Compruebe que $\rho(T)+\nu(T)=n$ donde T son las transformaciones matriciales del ejercicio anterior.

$\boxed{8.}$ En cada ítem utilice la información dada a fin de encontrar la nulidad de la siguiente transformación lineal.

   a) $T:\mathbb{R}^5 \to \mathbb{R}^3$, $\rho(T)=3$.

   b) $T:\mathbb{P}_4 \to \mathbb{P}_3$, $\rho(T)=4$.

c) $T : \mathbb{M}_2 \to \mathbb{M}_{23}$ tal que $R(T) = \text{gen}\left\{ \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \right\}$.

9. Sea $T : \mathbb{M}_2 \to \mathbb{R}^4$ una tranformación lineal tal que $\nu(T) = 0$. Encuentre el recorrido de T.

10. Sea $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal donde $\dim(\mathbb{V}) = \dim(\mathbb{W}) = n$. Pruebe que

$$R(T) = \mathbb{W} \Leftrightarrow \ker(T) = \{\mathbf{0}\}$$

11. Sea $D : \mathbb{P}_2 \to \mathbb{P}_1$ la transformación lineal derivación; esto es, $D(p) = p'$. Obtenga el núcleo de D.

12. Sea $J : \mathbb{P}_1 \to \mathbb{R}$ la transformación lineal integración dada por $J(p) = \displaystyle\int_{-1}^{1} p(x)\,dx$. Obtenga su núcleo.

6.3. Transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$

En la sección anterior se estudió un tipo particular de transformación lineal: las transformaciones matriciales, que son aquellas transformaciones T de $\mathbb{R}^n$ en $\mathbb{R}^m$ tales que $T(\mathbf{x}) = A\mathbf{x}$ para alguna matriz A de $m \times n$.

En esta sección se demostrará que toda transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación matricial, es decir, se probará la existencia de una matriz A de $m \times n$ tal que la transformación T es la multiplicación por A.

Debe tenerse en cuenta que en esta sección todos los vectores deben ser considerados como vectores columna.

Teorema 6.8. Si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal entonces T es una transformación matricial. Esto es, existe una matriz A de $m \times n$ tal que

$$T(\mathbf{x}) = A\mathbf{x}, \quad \text{para todo } x \in \mathbb{R}^n.$$

Esta matriz A es la única matriz que satisface la igualdad enunciada.

Demostración. Sean $B = \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ la base estándar de $\mathbb{R}^n$ y $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ un vector de $\mathbb{R}^n$. Dado que B es la base estándar de $\mathbb{R}^n$, entonces

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n.$$

Aplicando T a la expresión anterior y teniendo en cuenta que es una transformación lineal se obtiene

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_nT(\mathbf{e}_n) \\ &= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + ... + x_nT(\mathbf{e}_n). \end{aligned} \tag{6.4}$$

Los n vectores $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$,..., $T(\mathbf{e}_n)$ son vectores de $\mathbb{R}^m$. Si se considera la matriz A cuyas columnas son dichos vectores, y recordando el producto de una matriz por un vector:

$$A\mathbf{x} = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + ... + x_nT(\mathbf{e}_n). \tag{6.5}$$

Así, comparando (6.4) y (6.5) se prueba que para todo $\mathbf{x}$ de $\mathbb{R}^n$,

$$T(\mathbf{x}) = A\mathbf{x},$$

donde $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$.

Para probar que la matriz A hallada es la única matriz que satisface la igualdad requerida, supóngase que $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}$ es una matriz de $m \times n$ tal que $T(\mathbf{x}) = B\mathbf{x}$, para todo $\mathbf{x}$ de $\mathbb{R}^n$. En particular esta igualdad se verifica para los vectores $\mathbf{e}_1$, $\mathbf{e}_2$,..., $\mathbf{e}_n$ de la base estándar de $\mathbb{R}^n$. Esto es

$$T(\mathbf{e}_j) = B\mathbf{e}_j, \quad j = 1, 2, ..., n$$

y dado que

$$B\mathbf{e}_j = \mathbf{b}_j, \quad j = 1, 2, ..., n$$

entonces

$$\mathbf{b}_j = T(\mathbf{e}_j), \quad j = 1, 2, .., n.$$

Por lo tanto las columnas de B son las mismas que las de A. Así, $B = A$. □

Definición (Matriz estándar). Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal, la matriz $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$ del teorema anterior se llama *matriz estándar de la transformación T*. Esta es la única matriz tal que $T(\mathbf{x}) = A\mathbf{x}$, para todo $\mathbf{x}$ en $\mathbb{R}^n$.

Toda transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ es entonces una transformación matricial, la multiplicación por la matriz estándar $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$. La matriz A proporciona toda la información que se requiera de la transformación. Esto es muy útil porque trabajar con matrices facilita los cálculos y permite el uso de herramientas computacionales. Se puede evaluar la imagen $T(\mathbf{x})$ de cualquier $\mathbf{x}$ de $\mathbb{R}^n$ mediante una simple multiplicación de matrices. Además dado que $\ker(T) = N(A)$ y $R(T) = \mathrm{col}(A)$ (Teorema 6.6) se pueden encontrar bases para estos subespacios por los métodos ya conocidos de obtención de bases para el espacio de columnas y el espacio nulo de una matriz.

De la definición de transformación matricial y del teorema anterior, se tiene el siguiente importante resultado.

Corolario 6.3. Una transformación $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal si y solo si la matriz $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$ satisface $T(\mathbf{x}) = A\mathbf{x}$, para todo $\mathbf{x}$ de $\mathbb{R}^n$.

De acuerdo al corolario, puede comprobarse si una transformación $T : \mathbb{R}^n \to \mathbb{R}^m$ es, o no, lineal del siguiente modo: se obtiene la matriz $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$ y se realiza el producto $A\mathbf{x}$ para un $\mathbf{x}$ genérico de $\mathbb{R}^n$. Si $A\mathbf{x} = T(\mathbf{x})$ entonces T es una transformación lineal y A es su matriz estándar, en caso contrario, T no es lineal.

Ejemplo 6.14. Sea la transformación $T : \mathbb{R}^3 \to \mathbb{R}^4$ tal que $T(x, y, z) = (x + y, 2y - z, z, x - y)$.

De acuerdo al comentario anterior, para comprobar si T es o no una transformación lineal, obtendremos la matriz $A = \left[\begin{array}{ccc} T(\mathbf{e}_1) & T(\mathbf{e}_2) & T(\mathbf{e}_3) \end{array}\right]$. De acuerdo a la ley de T:

$$T(\mathbf{e}_1) = T(1, 0, 0) = (1, 0, 0, 1), \quad T(\mathbf{e}_2) = T(0, 1, 0) = (1, 2, 0, -1), \quad T(\mathbf{e}_3) = T(0, 0, 1) = (0, -1, 1, 0).$$

En este momento debe aclararse que aunque se ha usado la notación horizontal en la definición de T para los vectores de $\mathbb{R}^3$ y $\mathbb{R}^4$, se hace necesario utilizar la notación vertical. Así,

$$A = \left[\begin{array}{ccc} T(\mathbf{e}_1) & T(\mathbf{e}_2) & T(\mathbf{e}_3) \end{array}\right] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}.$$

Haciendo el producto $A\mathbf{x}$:

$$A\mathbf{x} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + y \\ 2y - z \\ z \\ x - y \end{bmatrix},$$

se comprueba que $A\mathbf{x} = T(\mathbf{x})$ y por lo tanto T es una transformación lineal de $\mathbb{R}^3$ en $\mathbb{R}^4$ (transformación matricial) y A es su matriz estándar.◀

Actividad 6.4. Para la transformación matricial del ejemplo anterior, obtenga su recorrido, su núcleo, el rango y la nulidad.

Geometría de las transformaciones lineales de $\mathbb{R}^2$ en $\mathbb{R}^2$

A continuación se presentan algunas transformaciones matriciales geométricas del plano (de $\mathbb{R}^2$ en $\mathbb{R}^2$) que son muy interesantes: reflexiones, compresiones (expansiones) y rotaciones.

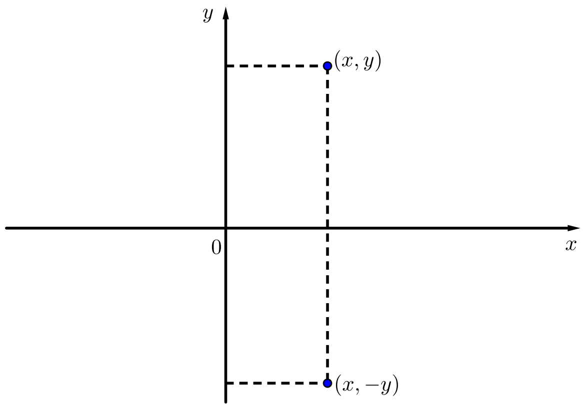
Ejemplo 6.15 (Reflexiones). Las reflexiones en el plano se definen respecto a cualquier recta. Las más interesantes son las reflexiones respecto a los ejes coordenados y a la recta $y = x$. Éstas se definen como sigue:

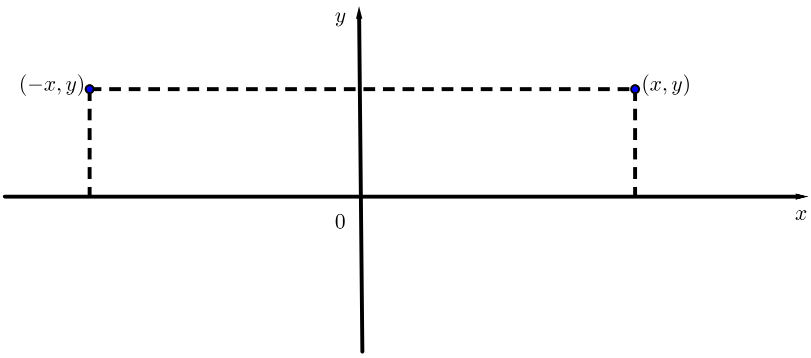
- *Reflexión con respecto al eje x:*

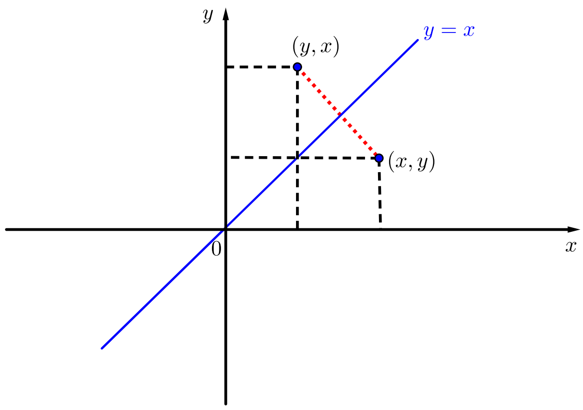
$$T_x \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix} \quad \text{cuya matriz estándar es} \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{(verifíquelo).}$$

- *Reflexión con respecto al eje y:*

$$T_y \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ y \end{bmatrix} \quad \text{cuya matriz estándar es} \quad \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$





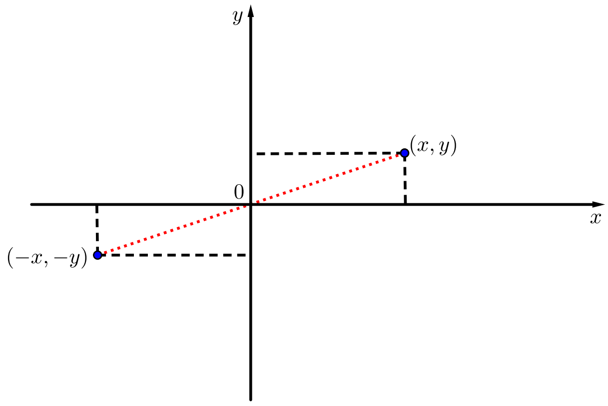


- *Reflexión con respecto a la recta $y = x$*:

$$T_{id}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix} \quad \text{cuya matriz estándar es} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Otra reflexión interesante es la *reflexión respecto al origen o*:

$$T_o\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix} \quad \text{cuya matriz estándar es} \quad \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$



Ejemplo 6.16 (**Compresiones y expansiones**)**.** Las compresiones y expansiones son escalamientos a lo largo de los ejes coordenados.

- *Escalamiento en la dirección de la coordenada* x:

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} cx \\ y \end{bmatrix} \quad \text{cuya matriz estándar es} \quad \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix},$$

  donde $c > 0$. Si $0 < c < 1$ se trata de una compresión en la dirección de la coordenada x, un factor c, mientras que si $c > 1$ se trata de una expansión en la dirección de la coordenada x, un factor c.

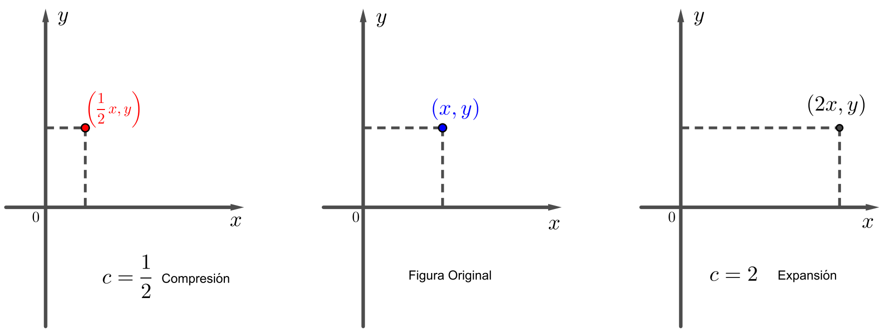


Figura 6.3: Escalamiento en la dirección de la coordenada x

- *Escalamiento en la dirección de la coordenada* y:

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ dy \end{bmatrix} \quad \text{cuya matriz estándar es} \quad \begin{bmatrix} 1 & 0 \\ 0 & d \end{bmatrix},$$

  donde $d > 0$. Si $0 < d < 1$ se trata de una compresión en la dirección de la coordenada y, un factor d, mientras que si $d > 1$ se trata de una expansión en la dirección de la coordenada y, un factor d.

- *Escalamiento en ambas direcciones*:

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} cx \\ dy \end{bmatrix} \quad \text{cuya matriz estándar es} \quad \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix},$$

  donde $c > 0$ y $d > 0$. Esto puede resultar en una expansión o en una compresión en ambas direcciones o bien en una expansión en una dirección y una compresión en la otra.

Ejemplo 6.17 (**Rotación**). Como caso especial de una transformación matricial están las transformaciones llamadas de rotación. Sea θ un ángulo fijo. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación matricial de matriz estándar

$$A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

Si $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ entonces

$$T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\operatorname{sen}\theta \\ x\operatorname{sen}\theta + y\cos\theta \end{bmatrix}.$$

Se verá que $T(\mathbf{x})$ es el vector que se obtiene si se hace rotar al vector $\mathbf{x}$ (respecto al origen) un ángulo θ en sentido contrario a las agujas de reloj.

En efecto, si r denota la longitud del vector $\mathbf{x}$ y ϕ el ángulo que forma con el eje positivo de las abscisas, entonces

$$\begin{cases} x = r\cos\phi \\ y = r\operatorname{sen}\phi \end{cases}$$

Si $\mathbf{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$ es el vector que se obtiene al hacer girar $\mathbf{x}$ un ángulo θ se demostrará que $\mathbf{x}' = T(\mathbf{x})$.

El vector $\mathbf{x}'$ tiene la misma longitud que $\mathbf{x}$ y el ángulo que describe con el sentido positivo del eje de las abscisas es $\phi + \theta$. Luego:

$$\begin{cases} x' = r\cos(\phi + \theta) \\ y' = r\operatorname{sen}(\phi + \theta) \end{cases}$$

Por lo tanto:

$$\begin{aligned}
\mathbf{x}' &= \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} r\cos(\phi+\theta) \\ r\operatorname{sen}(\phi+\theta) \end{bmatrix} = \begin{bmatrix} r(\cos\phi\cos\theta - \operatorname{sen}\phi\operatorname{sen}\theta) \\ r(\cos\phi\operatorname{sen}\theta + \operatorname{sen}\phi\cos\theta) \end{bmatrix} \\
&= \begin{bmatrix} (r\cos\phi)\cos\theta - r(\operatorname{sen}\phi)\operatorname{sen}\theta \\ (r\cos\phi)\operatorname{sen}\theta + (r\operatorname{sen}\phi)\cos\theta \end{bmatrix} \\
&= \begin{bmatrix} x\cos\theta & -y\operatorname{sen}\theta \\ x\operatorname{sen}\theta & y\cos\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\
&= A\mathbf{x} = T(\mathbf{x}).
\end{aligned}$$

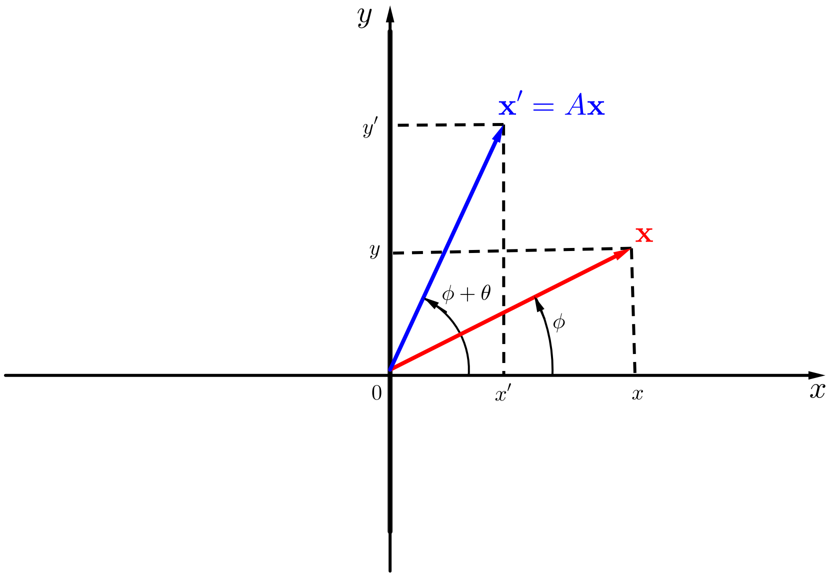


Figura 6.4: Rotación

Ejercicios 6.3

1. Para las siguientes transformaciones de $\mathbb{R}^n$ en $\mathbb{R}^m$, identifique n y m, y para las que sean lineales encuentre su matriz estándar.

   a) $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 \\ x_1 + 3x_2 \end{bmatrix}$

   b) $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y + 3 \\ x \end{bmatrix}$

   c) $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3x_1 + x_2x_3 \\ 2x_2 \end{bmatrix}$

   d) $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3x_1$

   e) $T\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} x + 2y - z + w \\ y + z \\ -x \end{bmatrix}$

   f) $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

2. Halle la matriz estándar para la transformación lineal en el plano $T : \mathbb{R}^2 \to \mathbb{R}^2$ que aplica un punto (x, y) en:

   a) su reflexión con respecto a la recta $y = -x$.
   b) su proyección ortogonal sobre el eje x.
   c) su proyección ortogonal sobre el eje y.
   d) su proyección ortogonal sobre la recta que pasa por el origen y contiene al punto $(1, 1)$.

   Verifique las respuestas gráficamente situando los puntos $(2, 1)$ y $T(2, 1)$.

3. Halle la matriz estándar para la transformación lineal T de $\mathbb{R}^2$ en $\mathbb{R}^2$ que haga girar un punto (x, y) alrededor del origen hasta describir un ángulo θ de:

   $\boxed{a)}$ 45°
   b) 60°
   c) 90°
   d) −30°

   Verifique las respuestas gráficamente situando los puntos $(2, 1)$ y $T(2, 1)$.

4. Encuentre la matriz que escala un factor de:

   $\boxed{a)}$ $\frac{1}{2}$ en la dirección del eje x.
   b) 3 en la dirección del eje y.

$\boxed{5.}$ En cada caso describa el efecto geométrico de la multiplicación por la matriz dada:

   a) $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$
   b) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
   c) $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$
   d) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
   e) $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$
   f) $\begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$

6.4. Matriz de una transformación lineal

Como se ha visto en la sección anterior toda transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación matricial, $T(\mathbf{x}) = A\mathbf{x}$ siendo $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$.

Si $\mathbb{V}$ y $\mathbb{W}$ son espacios vectoriales cualesquiera de dimensión finita y $T : \mathbb{V} \to \mathbb{W}$ es una transformación lineal, también es posible "trabajar" esta transformación como una transformación matricial. De hecho, las transformaciones matriciales son de $\mathbb{R}^n$ en $\mathbb{R}^m$ y $\mathbb{V}$ y $\mathbb{W}$ pueden ser otros espacios finito dimensionales, por ejemplo $\mathbb{V} = \mathbb{P}_2$ y $\mathbb{W} = \mathbb{M}_{23}$, y no se concibe la multiplicación de una matriz por un polinomio. La idea es elegir bases ordenadas B y B' de $\mathbb{V}$ y $\mathbb{W}$, respectivamente y trabajar con los vectores de coordenadas con respecto a estas bases. Si $\dim(\mathbb{V}) = n$ y $\dim(\mathbb{W}) = m$, entonces para cada $\mathbf{v}$ de $\mathbb{V}$, el vector de coordenadas $[\mathbf{v}]_B$ es un vector de $\mathbb{R}^n$ y el vector de coordenadas de su transformado, $[T(\mathbf{v})]_{B'}$ es un vector de $\mathbb{R}^m$.

La asignación $\mathbf{v} \to T(\mathbf{v})$ de $\mathbb{V}$ en $\mathbb{W}$ genera la transformación de $\mathbb{R}^n$ en $\mathbb{R}^m$ tal que $[\mathbf{v}]_B \to [T(\mathbf{v})]_{B'}$. Se verá que existe una matriz A de $m \times n$ tal que $[T(\mathbf{v}]_{B'} = A\,[\mathbf{v}]_B$ para todo $\mathbf{v}$ en $\mathbb{V}$.

De esta manera la transformación lineal $T : \mathbb{V} \to \mathbb{W}$ se podrá trabajar indirectamente a través de una transformación matricial. Al respecto se tiene el siguiente teorema.

Teorema 6.9. Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales no cero de dimensiones finita n y m respectivamente, y $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Sea $B = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ una base de $\mathbb{V}$ y $B' = \{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m\}$ una base de $\mathbb{W}$. Entonces existe una matriz A de $m \times n$ tal que

$$[T(\mathbf{v})]_{B'} = A\,[\mathbf{v}]_B\,,$$

para todo $\mathbf{v} \in \mathbb{V}$. Más aún esta es la única matriz que satisface la igualdad anterior.

Demostración. Sea $\mathbf{v}$ un vector del espacio $\mathbb{V}$. Como $B = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es una base de $\mathbb{V}$, existen únicos escalares c_1, c_2,..., c_n tales que

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n,$$

es decir, el vector de coordenadas de $\mathbf{v}$ en la base B es

$$[\mathbf{v}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Por la linealidad de T, $T(\mathbf{v})$ es el siguiente vector de $\mathbb{W}$:

$$T(\mathbf{v}) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + ... + c_nT(\mathbf{v}_n) \tag{6.6}$$

Para hallar el vector de coordenadas de $T(\mathbf{v})$ con respecto a la base B' debe expresarse a este vector como combinación lineal de los vectores $\mathbf{w}_1$, $\mathbf{w}_2$,..., $\mathbf{w}_m$ de la base B'. Dado que cada $T(\mathbf{v}_j)$ con $j = 1, 2, ..., n$ es un vector de $\mathbb{W}$ y $B' = \{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m\}$ es una base de $\mathbb{W}$, entonces existen únicos escalares a_{1j}, a_{2j},...,a_{mj} tales que

$$T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + ... + a_{mj}\mathbf{w}_m, \tag{6.7}$$

es decir, los vectores de coordenadas de $T(\mathbf{v}_j)$ con respecto a la base B' son

$$[T(\mathbf{v}_j)]_{B'} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix},$$

para todo $j = 1, 2, ..., n$.

Sustituyendo las expresiones dadas en (6.7) en la igualdad (6.6) y operando, se obtiene el vector $T(\mathbf{v})$ expresado como combinación lineal de los vectores de la base B':

$$\begin{aligned} T(\mathbf{v}) &= c_1\left(a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + ... + a_{m1}\mathbf{w}_m\right) + c_2\left(a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + ... + a_{m2}\mathbf{w}_m\right) + ... \\ &\quad + c_n\left(a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + ... + a_{mn}\mathbf{w}_m\right) \\ &= \left(c_1a_{11} + c_2a_{12} + ... + c_na_{1n}\right)\mathbf{w}_1 + \left(c_1a_{21} + c_2a_{22} + ... + c_na_{2n}\right)\mathbf{w}_2 + ... \\ &\quad + \left(c_1a_{m1} + c_2a_{m2} + ... + c_na_{mn}\right)\mathbf{w}_m. \end{aligned}$$

Los coeficientes que acompañan a los vectores $\mathbf{w}_1$, $\mathbf{w}_2$,..., $\mathbf{w}_m$ son las coordenadas de $T(\mathbf{v})$ en la base B'. Así, y por definición del producto de una matriz por un vector, se tiene:

$$\begin{aligned} [T(\mathbf{v})]_{B'} &= \begin{bmatrix} c_1a_{11} + c_2a_{12} + ... + c_na_{1n} \\ c_1a_{21} + c_2a_{22} + ... + c_na_{2n} \\ \vdots \\ c_1a_{m1} + c_2a_{m2} + ... + c_na_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}. \end{aligned}$$

Observe que $[T(\mathbf{v})]_{B'}$ es el producto de una matriz de $m \times n$ por un vector. Las columnas de la matriz son los vectores de coordenadas $[T(\mathbf{v}_1)]_{B'}$, $[T(\mathbf{v}_2)]_{B'}$, ..., $[T(\mathbf{v}_n)]_{B'}$ y el vector es el vector de coordenadas $[\mathbf{v}]_B$. Luego:

$$[T(\mathbf{v})]_{B'} = A\,[\mathbf{v}]_B\,, \text{ para cada } \mathbf{v} \in \mathbb{V},$$

donde $A = \begin{bmatrix} [T(\mathbf{v}_1)]_{B'} & [T(\mathbf{v}_2)]_{B'} & \cdots & [T(\mathbf{v}_n)]_{B'} \end{bmatrix}$.

De esta forma se ha probado la existencia de una tal matriz A. La demostración de la unicidad es análoga a la demostración de la unicidad de la matriz estándar del Teorema 6.8.

□

Definición (Matriz de una transformación lineal). Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales no cero de dimensiones n y m respectivamente y $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Sea $B = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ una base de $\mathbb{V}$ y $B' = \{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m\}$ una base de $\mathbb{W}$. La *matriz de T con respecto a las bases B y B'* es la matriz de $m \times n$ dada por

$$A = \begin{bmatrix} [T(\mathbf{v}_1)]_{B'} & [T(\mathbf{v}_2)]_{B'} & \cdots & [T(\mathbf{v}_n)]_{B'} \end{bmatrix}.$$

Esta es la única matriz tal que $[T(\mathbf{v})]_{B'} = A\,[\mathbf{v}]_B$ para todo $\mathbf{v}$ en $\mathbb{V}$.
En el caso particular en que $\mathbb{W} = \mathbb{V}$ (de modo que T es un operador lineal sobre $\mathbb{V}$) es común usar la misma base ordenada, esto es hacer $B' = B$, y se dice que la matriz A es la *matriz de T con respecto a la base B*.

Nota: Debe notarse que las bases B y B' deben considerarse bases ordenadas ya que se está trabajando con vectores de coordenadas.

Obsérvese que en general existen infinitas matrices para una transformación lineal T dado que existen infinitas bases para los espacios $\mathbb{V}$ y $\mathbb{W}$. Sin embargo, fijadas las bases ordenadas B y B' de $\mathbb{V}$ y $\mathbb{W}$, respectivamente, la matriz A es única.

Ejemplo 6.18. Considere la transformación lineal $T : \mathbb{M}_2 \to \mathbb{P}_2$ tal que

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a+2b) - cx + dx^2,$$

y las bases (ordenadas) $B = \{A_1, A_2, A_3, A_4\}$ de $\mathbb{M}_2$ donde $A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ y $A_4 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ y $B' = \{1,\ x,\ x^2\}$ la base estándar de $\mathbb{P}_2$.

La matriz de T con respecto a B y B' es

$$A = \begin{bmatrix} [T(A_1)]_{B'} & [T(A_2)]_{B'} & [T(A_3)]_{B'} & [T(A_4)]_{B'} \end{bmatrix}.$$

Los vectores de coordenadas de $T(A_1) = 3 - x + x^2$, $T(A_2) = 3 - x$, $T(A_3) = 3$ y $T(A_4) = 1$ con respecto a la base B' se encuentran directamente ya que B' es la base estándar de $\mathbb{P}_2$:

$$[T(A_1)]_{B'} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}, \quad [T(A_2)]_{B'} = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}, \quad [T(A_3)]_{B'} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} \quad \text{y} \quad [T(A_4)]_{B'} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Luego,

$$A = \begin{bmatrix} 3 & 3 & 3 & 1 \\ -1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \blacktriangleleft$$

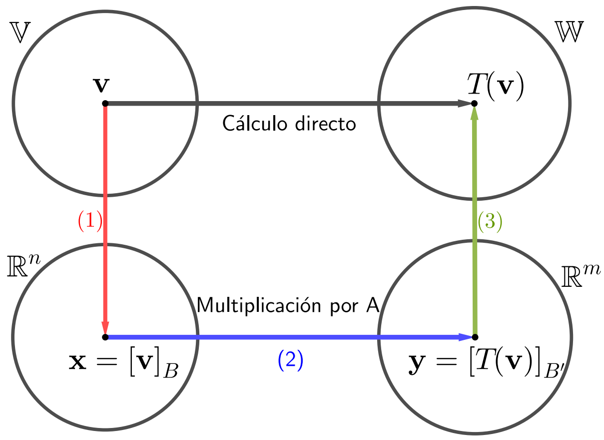


Figura 6.5: Matriz de la transformación

El esquema de la Figura 6.5 muestra que dada la matriz A de la transformación lineal T con respecto a las bases B y B' se puede hallar $T(\mathbf{v})$ en tres pasos por medio del procedimiento indirecto que se indica.

Dado un vector $\mathbf{v}$ de $\mathbb{V}$:

(1) Hallar $[\mathbf{v}]_B$

(2) Premultiplicar $[\mathbf{v}]_B$ por A para obtener $[T(\mathbf{v})]_{B'}$

(3) Obtener $T(\mathbf{v})$ a partir de su vector de coordenadas $[T(\mathbf{v})]_{B'}$

Ejemplo 6.19. Sea $T : \mathbb{P}_1 \to \mathbb{P}_2$ definida por $T(p) = xp$, es decir $T(a + bx) = ax + bx^2$.

Puede comprobarse rápidamente que T es una transformación lineal.

Considere las bases $B = \{p_1, p_2\}$ y $B' = \{p_1, p_2, p_3\}$ de $\mathbb{P}_1$ y $\mathbb{P}_2$, respectivamente, siendo $p_1 = 1$, $p_2 = 1 + x$ y $p_3 = x^2$.

Se desea:

1. Obtener la matriz, A, de T con respecto a las bases B y B'.
2. Si $p = 2 - 5x$, hallar $T(p)$ de dos maneras distintas:
   a) en forma directa (usando la ley de T).
   b) en forma indirecta (usando la matriz de A).

Solución:

1. Por definición, $A = \left[\ [T(p_1)]_{B'} \quad [T(p_2)]_{B'}\ \right]$.

   De acuerdo a la ley de T: $T(p_1) = T(1) = x(1) = x$ y $T(p_2) = T(1 + x) = x(1 + x) = x + x^2$.

   Para hallar $[T(p_1)]_{B'}$ y $[T(p_2)]_{B'}$, se deben encontrar los coeficientes a_1, a_2, a_3 y b_1, b_2, b_3 tales que:

$$\begin{aligned} a_1p_1 + a_2p_2 + a_3p_3 &= T(p_1), \\ b_1p_1 + b_2p_2 + b_3p_3 &= T(p_2). \end{aligned}$$

   Reemplazando p_1, p_2, p_3, $T(p_1)$ y $T(p_2)$, y operando se obtienen:

$$\begin{aligned} (a_1 + a_2) + a_2x + a_3x^2 &= x, \\ (b_1 + b_2) + b_2x + b_3x^2 &= x + x^2, \end{aligned}$$

   y por igualdad de polinomios las ecuaciones anteriores se convierten en los siguientes sistemas lineales:

$$\begin{cases} a_1 \ +a_2 \qquad = 0 \\ \qquad a_2 \qquad = 1 \\ \qquad\qquad a_3 = 0 \end{cases} \quad \text{y} \quad \begin{cases} b_1 \ +b_2 \qquad = 0 \\ \qquad b_2 \qquad = 1 \\ \qquad\qquad b_3 = 1 \end{cases}$$

   Estos sistemas tienen la misma matriz de coeficientes y pueden resolverse en forma simultánea. En este caso las soluciones se encuentran directamente pues los sistemas están escalonados. Éstas son:

$$\begin{cases} a_1 = -1 \\ a_2 = 1 \\ a_3 = 0 \end{cases} \quad \text{y} \quad \begin{cases} b_1 = -1 \\ b_2 = 1 \\ b_3 = 1 \end{cases}$$

   Por lo tanto, $[T(p_1)]_{B'} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ y $[T(p_2)]_{B'} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$. Luego, la matriz A buscada es

$$A = \begin{bmatrix} -1 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

2. Sea $p = 2 - 5x$, para hallar $T(p)$:

   a) De manera directa (usando la ley de T):

$$T(p) = T(2 - 5x) = x(2 - 5x) = 2x - 5x^2.$$

   b) De manera indirecta:

   (1) Se halla $[p]_B$. Para ello, se deben encontrar a_1 y a_2 tales que

$$a_1(1) + a_2(1 + x) = 2 - 5x,$$

   y operando se obtiene $a_1 = 7$ y $a_2 = -5$. Así,

$$[p]_B = [2 - 5x]_B = \begin{bmatrix} 7 \\ -5 \end{bmatrix}.$$

(2) Se utiliza la matriz A para hallar $[T(p)]_{B'}$. Sabiendo que $[T(p)]_{B'} = A\,[\mathbf{v}]_B$, entonces

$$[T(p)]_{B'} = [T(2-5x)]_{B'} = \begin{bmatrix} -1 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 7 \\ -5 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ -5 \end{bmatrix}.$$

(3) $[T(p)]_{B'} = [T(2-5x)]_{B'} = \begin{bmatrix} -2 \\ 2 \\ -5 \end{bmatrix}$ y $B' = \left\{1,\ 1+x,\ x^2\right\}$, luego:

$$T(p) = T(2-5x) = -2(1) + 2(1+x) - 5(x^2) = 2x - 5x^2. \blacktriangleleft$$

En lo que sigue se muestra que la matriz de una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ con respecto a las bases estándar de ambos espacios es precisamente la que fue llamada matriz estándar de T.

Ejemplo 6.20 (Matriz con respecto a las bases estándar). Si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal y B y B' son las bases estándar para $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente, entonces la matriz de T con respecto a estas bases es $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$.

En efecto, la matriz de T con respecto a las bases estándar B y B' tiene como columnas a los vectores de coordenadas $[T(\mathbf{e}_1)]_{B'}$, $[T(\mathbf{e}_2)]_{B'}$,..., $[T(\mathbf{e}_n)]_{B'}$.

Como B' es la base estándar de $\mathbb{R}^m$, entonces $[T(\mathbf{e}_j)]_{B'} = T(\mathbf{e}_j)$, para todo $j = 1, 2, ..., n$.

Así, la matriz de T con respecto a estas bases es

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix},$$

que es la matriz estándar de T definida en la sección anterior. Considerando esta matriz la imagen de un vector $\mathbf{v}$ de $\mathbb{R}^n$ se obtiene directamente:

$$T(\mathbf{v}) = A\mathbf{v}. \blacktriangleleft$$

Ejemplo 6.21 (Obtención de la ley de T conocida una matriz que la representa). Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ la transformación lineal tal que la matriz de T con respecto a la base B y B' de $\mathbb{R}^3$ y $\mathbb{R}^2$, respectivamente es

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix}.$$

Si tanto B como B' fueran las bases estándar, la matriz A sería la matriz estándar de la transformación. En dicho caso, la ley de T se obtendría directamente, $T(\mathbf{x}) = A\mathbf{x}$.

Supongamos, en cambio, $B = \{(1,-1,1), (0,1,2), (0,0-1)\}$ y $B' = \{(1,-1), (0,1)\}$. Para hallar la ley de T usando la matriz A, es decir, teniendo en cuenta que

$$[T(\mathbf{x})]_{B'} = A\,[\mathbf{x}]_B\,,$$

debe calcularse primero el vector de coordenadas $[\mathbf{x}]_B$ de un vector genérico $\mathbf{x}$ de $\mathbb{R}^3$. Luego, efectuar el producto $A\,[\mathbf{x}]_B$ para así obtener $[T(\mathbf{x})]_B$. Por último, a partir de la definición de vector de coordenadas, se determina $T(\mathbf{x})$.

(1) Sea $\mathbf{x} = (x, y, z)$ un vector de $\mathbb{R}^3$. Las coordenadas de $\mathbf{x}$ con respecto a la base B son los valores de c_1, c_2 y c_3 tales que

$$\mathbf{x} = c_1(1, -1, 1) + c_2(0, 1, 2) + c_3(0, 0, -1),$$

es decir

$$(x, y, z) = (c_1, -c_1 + c_2, c_1 + 2c_2 - c_3),$$

que por igualdad de vectores conduce al sistema lineal

$$\left\{ \begin{array}{rcl} c_1 &=& x \\ -c_1 + c_2 &=& y \\ c_1 + 2c_2 - c_3 &=& z \end{array} \right. ,$$

de donde se obtiene

$$\left\{ \begin{array}{rcl} c_1 &=& x \\ c_2 &=& x + y \\ c_3 &=& 3x + 3y - z \end{array} \right.$$

Por lo tanto, $[\mathbf{x}]_B = \begin{bmatrix} x \\ x + y \\ 3x + 2y - z \end{bmatrix}$.

(2) $[T(\mathbf{x})]_{B'} = A\,[\mathbf{x}]_B = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ x + y \\ 3x + 2y - z \end{bmatrix} = \begin{bmatrix} 6x + 4y - z \\ 3x + y - z \end{bmatrix}$.

(3) Por último, se obtiene $T(\mathbf{x})$ a partir de su vector de coordenadas, $[T(\mathbf{x})]_{B'}$.

$$T(\mathbf{x}) = (6x + 4y - z)\,(1, -1) + (3x + y - z)\,(0, 1)\,,$$

resultando

$$T(x, y, z) = (6x + 4y - z, -3x - 3y)\,,$$

que es la ley de T.◀

Se ha visto en el Teorema 6.6 y su Corolario que si T es una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$, y A es la matriz estándar de T:

$$\ker(T) = N(A), \quad R(T) = \text{col}(A), \quad \rho(T) = \rho(A), \quad \nu(T) = \nu(A), \quad \rho(T) + \nu(T) = n = \dim(\mathbb{R}^n).$$

El siguiente teorema generaliza este resultado:

Teorema 6.10. Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales no cero de dimensiones finita n y m respectivamente, y $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Sea B una base de $\mathbb{V}$ y B' una base de $\mathbb{W}$. Si A es la matriz de T con respecto a las bases B y B', entonces:

1. $\mathbf{v} \in \ker(T) \Leftrightarrow [\mathbf{v}]_B \in N(A)$
2. $\mathbf{w} \in R(T) \Leftrightarrow [\mathbf{w}]_{B'} \in \text{col}(A)$
3. $\nu(T) = \nu(A)$
4. $\rho(T) = \rho(A)$

Demostración. Por ser A la matriz de T con respecto a las bases B y B',

$$[T(\mathbf{v})]_{B'} = A\,[\mathbf{v}]_B$$

para cada $\mathbf{v}$ de $\mathbb{V}$.

1. La equivalencia es válida porque:

$$\begin{aligned}
\mathbf{v} \in \ker(T) &\Leftrightarrow T(\mathbf{v}) = \mathbf{0} \quad (\text{en } \mathbb{W}) \\
&\Leftrightarrow [T(\mathbf{v})]_{B'} = [\mathbf{0}]_{B'} \quad (\text{en } \mathbb{R}^m) \\
&\Leftrightarrow A\,[\mathbf{v}]_B = \mathbf{0} \quad (\text{en } \mathbb{R}^m) \\
&\Leftrightarrow [\mathbf{v}]_B \in N(A).
\end{aligned}$$

2. Análogamente esta equivalencia es válida ya que:

$$\begin{aligned}
\mathbf{w} \in R(T) &\Leftrightarrow \text{existe } \mathbf{v} \in \mathbb{V} : \mathbf{w} = T(\mathbf{v}) \\
&\Leftrightarrow \text{existe } \mathbf{v} \in \mathbb{V} : [\mathbf{w}]_{B'} = [T(\mathbf{v})]_{B'} \\
&\Leftrightarrow \text{existe } \mathbf{v} \in \mathbb{V} : [\mathbf{w}]_{B'} = A\,[\mathbf{v}]_B \\
&\Leftrightarrow \text{el sistema lineal } A\mathbf{x} = [\mathbf{w}]_{B'} \text{ es consistente} \\
{\scriptstyle(\text{Teorema 4.23})} &\Leftrightarrow [\mathbf{w}]_{B'} \in \operatorname{col}(A).
\end{aligned}$$

Se sabe que la asignación $\mathbf{v} \to [\mathbf{v}]_B$ establece una correspondencia biunívoca entre los vectores de $\mathbb{V}$ y los vectores de $\mathbb{R}^n$.

3. Dado que $\mathbf{v} \in \ker(T) \Leftrightarrow \mathbf{x} = [\mathbf{v}]_B \in N(A)$, todos los vectores $\mathbf{x} \in \mathbb{R}^n$ que están en el espacio nulo de A, $N(A)$, son las coordenadas, con respecto a la base B, de los vectores $\mathbf{v} \in \mathbb{V}$ que están en el núcleo de T, $\ker(T)$. Esto es, $N(A) = \{[\mathbf{v}]_B : \mathbf{v} \in \ker(T)\}$. De acuerdo al Teorema 4.15, $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ es una base de $\ker(T)$ si y solo si $\{[\mathbf{v}_1]_B, [\mathbf{v}_2]_B, ..., [\mathbf{v}_k]_B\}$ es una base de $N(A)$. Por lo tanto: $\dim(\ker(T)) = \dim(N(A))$ esto es; $\nu(T) = \nu(A)$.

4. Similarmente, dado que $\mathbf{w} \in R(T) \Leftrightarrow \mathbf{y} = [\mathbf{w}]_{B'} \in \operatorname{col}(A)$, todos los vectores $\mathbf{y} \in \mathbb{R}^m$ que están en el espacio generado por las columnas de A, $\operatorname{col}(A)$, son las coordenadas, con respecto a la base B', de los vectores $\mathbf{w} \in \mathbb{W}$ que están en el recorrido de T, $R(T)$. Esto es, $\operatorname{col}(A) = \{[\mathbf{w}]_{B'} : \mathbf{w} \in R(T)\}$. De acuerdo al Teorema 4.15, $\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_k\}$ es una base de $R(T)$ si y solo si $\{[\mathbf{w}_1]_{B'}, [\mathbf{w}_2]_{B'}, ..., [\mathbf{w}_k]_{B'}\}$ es una base de $\operatorname{col}(A)$. Por lo tanto: $\dim(R(T)) = \dim(\operatorname{col}(A))$ esto es; $\rho(T) = \rho(A)$.

$\square$

Como consecuencia del teorema del rango y del teorema anterior se tiene una versión particular del teorema de la dimensión.

Corolario 6.4. Si $\mathbb{V}$ y $\mathbb{W}$ son espacios vectoriales no cero de dimensiones n y m, respectivamente, y $T : \mathbb{V} \to \mathbb{W}$ es una transformación lineal entonces:

$$\rho(T) + \nu(T) = \dim(\mathbb{V}).$$

Demostración. Sea A cualquier matriz de $m \times n$ representante de la transformación T. Por el teorema teorema del rango (Teorema 4.22), $\rho(A) + \nu(A) = n$. Por el teorema anterior $\rho(A) = \rho(T)$ y $\nu(A) = \nu(T)$, y como $n = \dim(\mathbb{V})$ entonces

$$\rho(T) + \nu(T) = \dim(\mathbb{V}).$$

□

Ejemplo 6.22. Sea la transformación lineal $T : \mathbb{P}_2 \to \mathbb{R}^2$ tal que la matriz de T con respecto a las bases $B = \{1,\ 1+x,\ x+x^2\}$ y $B' = \{(1,1), (-1,1)\}$ es

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

De acuerdo al teorema anterior, el núcleo de T es el conjunto de los polinomios p de $\mathbb{P}_2$ cuyos vectores de coordenadas en la base B, $[p]_B$, están en el espacio nulo de A. Puede verificarse que:

$$N(A) = \text{gen}\left\{ \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} \right\}.$$

Así, $\ker(T) = \text{gen}\left\{ p : [p]_B = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} \right\}$. Dado que $[p]_B = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$ entonces

$$p = -\frac{1}{2}(1) - \frac{1}{2}(1+x) + 1(x+x^2) = -1 + \frac{1}{2}x + x^2.$$

Por lo tanto,

$$\ker(T) = \text{gen}\left\{ -1 + \frac{1}{2}x + x^2 \right\}.$$

Así, el conjunto $\left\{-1 + \frac{1}{2}x + x^2\right\}$ es una base de $\ker(T)$, por lo cual $\nu(T) = \dim(\ker(T)) = 1$.

El rango de T puede calcularse directamente por la ecuación

$$\rho(T) + \nu(T) = \dim(\mathbb{P}_2),$$

resultando $\rho(T) = 3 - 1 = 2$. Así, $\dim(R(T)) = 2$ y, como $R(T)$ es un subespacio del codominio ($\mathbb{R}^2$), necesariamente $R(T) = \mathbb{R}^2$.◀

Ejemplo 6.23. Sea la transformación lineal $T : \mathbb{P}_5 \to \mathbb{M}_2$ tal que la matriz de T con respecto a ciertas bases dadas B y B' de $\mathbb{P}_5$ y $\mathbb{M}_2$, respectivamente, es

$$A = \begin{bmatrix} 1 & -1 & 2 & 2 & 3 & 1 \\ 0 & 1 & -1 & 0 & 1 & -1 \\ 2 & -2 & 0 & 0 & 1 & 0 \\ 3 & -3 & 2 & 2 & 4 & 1 \end{bmatrix}.$$

El rango de T y la nulidad de T son respectivamente iguales al rango y la nulidad de cualquier matriz que la represente (no depende de la elección de las bases B y B'). Llevando a la matriz A a una forma escalonada, el rango de T es el número de columnas pivote de A y la nulidad el número de columnas no pivote. La forma escalonada reducida de A (que se obtiene rápidamente usando una herramienta computacional) es

$$B = \begin{bmatrix} \boxed{1} & 0 & 0 & 1 & \frac{11}{4} & -\frac{1}{2} \\ 0 & \boxed{1} & 0 & 1 & \frac{9}{4} & -\frac{1}{2} \\ 0 & 0 & \boxed{1} & 1 & \frac{5}{4} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Luego, $\rho(T) = \rho(A) = 3$ y $\nu(T) = \nu(A) = 3$.

Puede verificarse que $\rho(T) + \nu(T) = 3 + 3 = 6 = \dim(\mathbb{P}_5)$.◀

Demostración del Teorema de la dimensión (Teorema 6.7)

El Corolario 6.4 es el teorema de la dimensión (Teorema 6.7) para el caso en que $\mathbb{V}$ y $\mathbb{W}$ son espacios vectoriales no ceros y $\mathbb{W}$ también es finito dimensional. Unos pocos detalles completan la demostración para el caso general.

El teorema de la dimensión establece que si $T : \mathbb{V} \to \mathbb{W}$ es una transformación lineal y $\mathbb{V}$ es finito dimensional, entonces

$$\rho(T) + \nu(T) = \dim(\mathbb{V}).$$

En efecto:

Si T es la transformación cero entonces $R(T) = \{\mathbf{0}\}$ y $\ker(T) = \mathbb{V}$, así $\rho(T) = 0$ y $\nu(T) = \dim(\mathbb{V})$. Por lo tanto se verifica el teorema ya que $\rho(T) + \nu(T) = \dim(\mathbb{V})$.

Supongamos que T es una transformación lineal distinta de la transformación cero. En este caso, $\mathbb{V} \neq \{\mathbf{0}\}$ y $R(T) \neq \{\mathbf{0}\}$, y dado que $\mathbb{V}$ es finito dimensional, $R(T)$, es un subespacio de dimensión finita de $\mathbb{W}$. No se piede información si se restringe el codominio, $\mathbb{W}$, de T a su recorrido y se considera entonces $T : \mathbb{V} \to R(T)$.

Así, T es una transformación lineal entre dos espacios vectoriales finito dimensionales y no ceros. Por lo tanto, aplicando el Corolario 6.4, es:

$$\rho(T) + \nu(T) = \dim(\mathbb{V}).$$

□

Matriz de un operador lineal. Cambio de base

Estamos interesados ahora en la representación por matrices de las transformaciones lineales de un espacio vectorial en sí mismo, es decir, de los operadores lineales $T : \mathbb{V} \to \mathbb{V}$ donde $\mathbb{V}$ es un espacio finito dimensional no cero. En tal caso es más conveniente fijar la misma base ordenada para el dominio y el codominio, esto es, hacer $B' = B$. La matriz entonces tendrá como columnas a los vectores de coordenadas de los transformados de la base B con respecto a la misma base B y se llamará **matriz de T con respecto a la base B**. Debe tenerse en cuenta que esta matriz depende de la base B de $\mathbb{V}$ fijada, para no olvidar esta dependencia la simbolizaremos $[T]_B$. Resumiendo:

Si $T : \mathbb{V} \to \mathbb{V}$ es un operador lineal y $B = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es una base (ordenada) de $\mathbb{V}$, la *matriz de T con respecto a la base B* es la siguiente matriz de $n \times n$

$$[T]_B = \left[\ [T(\mathbf{v}_1)]_B \quad [T(\mathbf{v}_2)]_B \quad \cdots \quad [T(\mathbf{v}_n)]_B \ \right].$$

Esta es la única matriz tal que

$$[T(\mathbf{v})]_B = [T]_B \, [\mathbf{v}]_B$$

para todo $\mathbf{v}$ en $\mathbb{V}$.

Se quiere ahora investigar el modo en que queda afectada una matriz representante de T cuando se cambia la base. Uno de los problemas del Álgebra Lineal es analizar si es posible hallar una base para $\mathbb{V}$ de modo que la matriz de T con respecto a dicha base sea sencilla, en particular una matriz diagonal ya que las matrices diagonales son las más simples de manejar. Este problema se estudiará en el capítulo siguiente y en esta sección se presentarán los fundamentos para resolver el problema.

El teorema que sigue fundamenta como determinar una nueva representación matricial (potencialmente más simple) de un operador lineal T a partir de una dada.

Teorema 6.11. Sea $T : \mathbb{V} \to \mathbb{V}$ un operador lineal sobre un espacio vectorial $\mathbb{V}$ de dimensión n. Sean B y B' dos bases ordenadas de $\mathbb{V}$. Entonces las matrices $[T]_{B'}$ y $[T]_B$ están relacionadas del siguiente modo

$$[T]_{B'} = P^{-1} \, [T]_B \, P,$$

donde P es la matriz de transición de B' a B.

Demostración. Se conoce que la matriz $[T]_{B'}$ es la única matriz tal que

$$[T(\mathbf{v})]_{B'} = [T]_{B'} \, [\mathbf{v}]_{B'}\,, \quad \text{para todo } \mathbf{v} \in \mathbb{V} \tag{6.8}$$

Se mostrará que la matriz $P^{-1} \, [T]_B \, P$ también verifica (6.8). En efecto:

Si $\mathbf{v}$ es un vector de $\mathbb{V}$, se verifican las siguientes relaciones:

(1) $P \, [\mathbf{v}]_{B'} = [\mathbf{v}]_B$ (P matriz de transición de B' a B)

(2) $[T]_B \, [\mathbf{v}]_B = [T(\mathbf{v})]_B$ ($[T]_B$ matriz de T con respecto a B)

(3) $P^{-1} \, [T(\mathbf{v})]_B = [T(\mathbf{v})]_{B'}$ (P^{-1} matriz de transición de B a B')

Por propiedad asociativa del producto de matrices y combinando (1), (2) y (3):

$$\begin{aligned} \left(P^{-1} \, [T]_B \, P\right) [\mathbf{v}]_{B'} &= \left(P^{-1} \, [T]_B\right) P \, [\mathbf{v}]_{B'} \overset{(1)}{=} \left(P^{-1} \, [T]_B\right) [\mathbf{v}]_B \\ &= P^{-1} \left([T]_B \, [\mathbf{v}]_B\right) \overset{(2)}{=} P^{-1} \, [T(\mathbf{v}]_B \\ &\overset{(3)}{=} [T(\mathbf{v})]_{B'} \end{aligned}$$

Así para cualquier vector $\mathbf{v}$ de $\mathbb{V}$,

$$[T(\mathbf{v})]_{B'} = \left(P^{-1} \, [T]_B \, P\right) [\mathbf{v}]_{B'}$$

por lo tanto, y debido a la unicidad de la matriz de T en la base B', debe ser $P^{-1} \, [T]_B \, P = [T]_{B'}$. $\square$

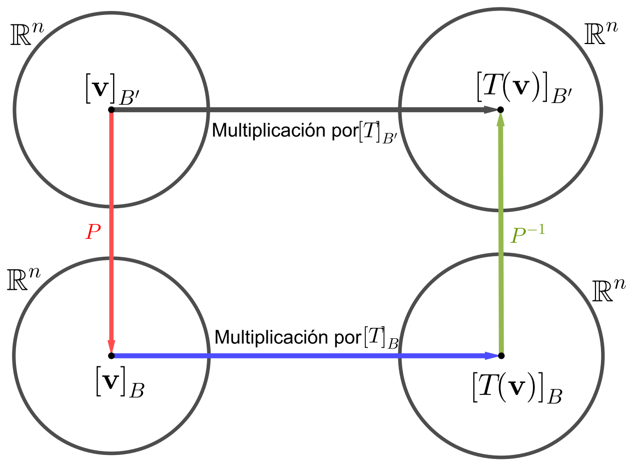


Figura 6.6: Ilustración del Teorema 6.11

Al aplicar el teorema puede que no se recuerde rápidamente si la matriz P es la matriz de transición de B a B' (error) o de B' a B (correcto). Puede ser de ayuda llamar a B base vieja y a B' base nueva, a $[T]_B$ matriz vieja y a $[T]_{B'}$ matriz nueva, y observar el diagrama de la Figura 6.6.

Ejemplo 6.24. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

$$T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4x_1 - 6x_2 \\ 3x_1 - 5x_2 \end{bmatrix}.$$

Fácilmente se puede verificar que T es un operador lineal sobre $\mathbb{R}^2$ con matriz estándar

$$A = \begin{bmatrix} 4 & -6 \\ 3 & -5 \end{bmatrix}.$$

Esto es, $A = [T]_B$ donde B es la base estándar de $\mathbb{R}^2$.

Sea la siguiente base de $\mathbb{R}^2$, $B' = \{\mathbf{u}_1, \mathbf{u}_2\}$ donde $\mathbf{u}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ y $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$. Para hallar la matriz de T con respecto a la base B' se puede proceder de dos maneras: de manera directa, por definición de $[T]_{B'}$, o de manera indirecta utilizando la matriz $A = [T]_B$.

- *De manera directa*: por definición, $[T]_{B'} = \begin{bmatrix} [T(\mathbf{u}_1)]_{B'} & [T(\mathbf{u}_2)]_{B'} \end{bmatrix}$.

  Aplicando la ley de T: $\quad T(\mathbf{u}_1) = T \begin{bmatrix} 1 \\ -3 \end{bmatrix} = \begin{bmatrix} 22 \\ 18 \end{bmatrix}; \quad T(\mathbf{u}_2) = T \begin{bmatrix} -2 \\ 5 \end{bmatrix} = \begin{bmatrix} -38 \\ -31 \end{bmatrix}.$

  Para hallar los vectores de coordenadas $[T(\mathbf{u}_1)]_{B'}$ y $[T(\mathbf{u}_2)]_{B'}$, se deben encontrar los valores de a_1, a_2 y b_1, b_2 tales que:

  $$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 = T(\mathbf{u}_1),$$
  $$b_1\mathbf{u}_1 + b_2\mathbf{u}_2 = T(\mathbf{u}_2).$$

  Reemplazando $\mathbf{u}_1$, $\mathbf{u}_2$, $T(\mathbf{u}_1)$ y $T(\mathbf{u}_2)$, las ecuaciones anteriores se convierten en dos sistemas de ecuaciones lineales con la misma matriz de coeficientes , $\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix}$. La matriz ampliada que permite resolver los dos sistemas en forma simultánea es:

  $$\begin{bmatrix} 1 & -2 & : & 22 & -38 \\ -3 & 5 & : & 18 & -31 \end{bmatrix}.$$

Puede comprobarse que la forma escalonada reducida de esta matriz es:

$$\left[\begin{array}{cccrr} 1 & 0 & : & -146 & 252 \\ 0 & 1 & : & -84 & 145 \end{array}\right].$$

Las soluciones se leen directamente de esta matriz. La tercera y cuarta columna son, respectivamente, $[T(\mathbf{u}_1)]_{B'}$ y $[T(\mathbf{u}_2)]_{B'}$. Por lo tanto:

$$[T]_{B'} = \left[\begin{array}{rr} -146 & 252 \\ -84 & 145 \end{array}\right].$$

- *De manera indirecta*: utilizando la matriz estándar $A = [T]_B$. Se necesita la matriz P de transición de la base B' a la base B.

  Por definición, P tiene como columnas los vectores de coordenadas de los vectores de la base B' con respecto a la base B, $[\mathbf{u}_1]_B$ y $[\mathbf{u}_2]_B$. Por ser B la base estándar, $[\mathbf{u}_1]_B = \mathbf{u}_1$ y $[\mathbf{u}_2]_B = \mathbf{u}_2$. Luego,

$$P = \left[\begin{array}{cc} \mathbf{u}_1 & \mathbf{u}_2 \end{array}\right] = \left[\begin{array}{rr} 1 & -2 \\ -3 & 5 \end{array}\right].$$

  Se verifica fácilmente que la inversa de P es:

$$P^{-1} = \left[\begin{array}{rr} -5 & -2 \\ -3 & -1 \end{array}\right].$$

  Por el Teorema 6.11 es:

$$[T]_{B'} = P^{-1}\,[T]_B\,P = \left[\begin{array}{rr} -5 & -2 \\ -3 & -1 \end{array}\right]\left[\begin{array}{rr} 4 & -6 \\ 3 & -5 \end{array}\right]\left[\begin{array}{rr} 1 & -2 \\ -3 & 5 \end{array}\right] = \left[\begin{array}{rr} -146 & 252 \\ -84 & 145 \end{array}\right],$$

  verificándose que es la misma matriz obtenida por el modo directo.

Consideremos ahora la base de $\mathbb{R}^2$, $B'' = \{\mathbf{v}_1, \mathbf{v}_2\}$ donde $\mathbf{v}_1 = \left[\begin{array}{c} 2 \\ 1 \end{array}\right]$ y $\mathbf{v}_2 = \left[\begin{array}{c} 1 \\ 1 \end{array}\right]$. Se obtendrá la matriz de T con respecto a la base B'' de manera directa. Por definición:

$$[T]_{B''} = \left[\begin{array}{cc} [T(\mathbf{v}_1)]_{B''} & [T(\mathbf{v}_2)]_{B''} \end{array}\right].$$

Aplicando la ley de T: $T(\mathbf{v}_1) = T\left[\begin{array}{c} 2 \\ 1 \end{array}\right] = \left[\begin{array}{c} 2 \\ 1 \end{array}\right] = \mathbf{v}_1$; $T(\mathbf{v}_2) = T\left[\begin{array}{c} 1 \\ 1 \end{array}\right] = \left[\begin{array}{r} -2 \\ -2 \end{array}\right] = -2\mathbf{v}_2$.

Los vectores $T(\mathbf{v}_1)$ y $T(\mathbf{v}_2)$ son múltiplos escalares de $\mathbf{v}_1$ y $\mathbf{v}_2$, respectivamente, por lo tanto no se requiere ningún cálculo para hallar sus coordenadas con respecto a la base B''. En efecto:

$T(\mathbf{v}_1) = \mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2$, así $[T(\mathbf{v}_1)]_{B''} = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$,

$T(\mathbf{v}_2) = -2\mathbf{v}_2 = 0\mathbf{v}_1 + (-2)\mathbf{v}_2$, así $[T(\mathbf{v}_2)]_{B''} = \left[\begin{array}{r} 0 \\ -2 \end{array}\right]$.

Rápidamente se obtuvo:

$$[T]_{B''} = \left[\begin{array}{rr} 1 & 0 \\ 0 & -2 \end{array}\right]. \blacktriangleleft$$

En el ejemplo anterior, se mostraron tres matrices que representan el mismo operador lineal en bases distintas $[T]_B$, $[T]_{B'}$ y $[T]_{B''}$. Puede observarse que la matriz más sencilla de las tres es $[T]_{B''}$ ya que es diagonal. *Las bases que producen matrices representantes diagonales son aquellas formadas por vectores que se transforman en múltiplos de sí mismos*, como ocurre con la base B''. El problema de obtener tales bases, si existen, se abordará en el capítulo siguiente.

También es interesante notar que si se conoce la matriz de un operador lineal en una base y se desea encontrar la matriz representante en otra base es a veces más conveniente efectuar el cambio de coordenadas usando la matriz de transición P, pero puede ser más simple obtener la nueva matriz por aplicación directa de la definición.

Ahora bien, *¿por qué decimos que las matrices diagonales son más "sencillas" de manipular?...* Una matriz diagonal tiene muchas ventajas computacionales. Repasemos algunas:

- es directo el cálculo del producto de dos matrices si una de ellas es diagonal. Por ejemplo:

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 2a & 3b & 5c \\ 2d & 3e & 5f \end{bmatrix},$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} 2a & 2b & 2c \\ 3d & 3e & 3f \end{bmatrix}.$$

- es directo el cálculo de las potencias. Por ejemplo:

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \Rightarrow D^k = \begin{bmatrix} 2^k & 0 \\ 0 & 3^k \end{bmatrix}.$$

- el determinante es el producto de los elementos de la diagonal.

- el rango es el número de elementos no cero en su diagonal mientras que la nulidad es el número de elementos cero de su diagonal.

- es invertible si y solo si en la diagonal no hay ceros, y su inversa es la matriz diagonal que tiene en su diagonal los recíprocos de los elementos correpondientes de la matriz original. Por ejemplo:

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \quad D^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/3 \end{bmatrix}.$$

- es igual a su transpuesta.

Semejanza de matrices

Definición (Matrices semejantes). Sean A y A' dos matrices cuadradas de orden n. A' es *semejante* a A si existe una matriz invertible P tal que

$$A' = P^{-1}AP.$$

Si A' es semejante a A, también A es semejante a A' como se muestra en el teorema que sigue. Por lo tanto se dirá simplemente que A y A' son **semejantes**.

Teorema 6.12. Sean A, B y C matrices cuadradas de orden n. Entonces:

1. A es semejante a A (A es semejante a sí misma).
2. Si A es semejante a B entonces B es semejante a A.
3. Si A es semejante a B y B es semejante a C, entonces A es semejante a C.

Demostración. La primera propiedad se compueba fácilmente ya que $A = IAI$ (I es invertible y $I^{-1} = I$). □

Actividad 6.5. Demuestre las propiedades 2. y 3. del teorema anterior.

Es importante observar lo siguiente. Las matrices que representan a un operador lineal T en dos bases distintas, B y B', son semejantes ya que, como lo establece el Teorema 6.11,

$$[T]_{B'} = P^{-1}\,[T]_B\,P$$

donde P es la matriz de transición de la base B' a la base B.

Recíprocamente, las matrices semejantes pueden considerarse representantes de un mismo operador lineal. En efecto: si A' y A son matrices semejantes de orden n, entonces existe una matriz P invertible tal que

$$A' = P^{-1}AP.$$

Consideremos el operador lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ tal que $T(\mathbf{x}) = A\mathbf{x}$. Entonces A es la matriz estándar de T. Esto es, $A = [T]_B$, donde B es la base estándar de $\mathbb{R}^n$. La matriz $P = \left[\begin{array}{cccc} \mathbf{p}_1 & \mathbf{p}_2 & ... & \mathbf{p}_n \end{array}\right]$ es invertible y por lo tanto sus columnas son linealmente independientes y forman una base de $\mathbb{R}^n$. Sea B' la base de $\mathbb{R}^n$ formada por las columnas de P. La matriz P es entonces la matriz de transición de la base B' a la base estándar B ya que $\left[\mathbf{p}_j\right]_B = \mathbf{p}_j$, $j = 1, 2, ..., n$. Así, de acuerdo al Teorema 6.11,

$$[T]_{B'} = P^{-1}AP.$$

Por lo tanto, $A' = [T]_{B'}$ es decir, es también una matriz representante de T pero en la base B'.

Ejemplo 6.25. Las matrices A, A' y A'' del Ejemplo 6.24 donde $A = [T]_B$, $A' = [T]_{B'}$ y $A'' = [T]_{B''}$, son semejantes ya que representan al mismo operador lineal en bases diferentes.

Por ejemplo, $A' = P^{-1}AP$ donde P es la matriz de transición de la base B' a la base B. De manera similar, se puede encontrar las matrices invertibles que relacionan A'' y A', y A'' con A.◀

Las matrices que son semejantes entre sí conforman un conjunto que comparte ciertas propiedades. Se demostrarán algunas de estas propiedades:

Teorema 6.13.

1. Las matrices semejantes tienen el mismo determinante.
2. Las matrices semejantes tienen el mismo rango.

Demostración. 1. Si A' y A son matrices semejantes entonces existe una matriz P invertible tal que $A' = P^{-1}AP$. Luego,

$$\det(A') = \det\left(P^{-1}AP\right) = \det\left(P^{-1}\right)\det(A)\det(P) = \frac{1}{\det(P)}\det(A)\det(P) = \det(A)$$

2. Si A' es semejante a A entonces A' y A son dos matrices representantes de un mismo operador lineal T. De acuerdo al Teorema 6.10, el rango de T es igual al rango de cualquier matriz representante de T. Esto es, $\rho(T) = \rho(A)$ y $\rho(T) = \rho(A')$ por lo tanto $\rho(A) = \rho(A')$.

□

Actividad 6.6. Verifique el teorema anterior para las matrices A, A' y A'' del Ejemplo 6.24 donde $A = [T]_B$, $A' = [T]_{B'}$ y $A'' = [T]_{B''}$.

Ejercicios 6.4

1. Para cada transformación lineal, obtenga la matriz de T con respecto a las bases que se indican.

   $\boxed{a)}$ $T : \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (-x, 0, -y)$
   $B = \{(1,0),(1,1)\}$, $B' = \{(1,1,1),(0,1,1),(0,0,1)\}$.

   b) $T : \mathbb{R}^2 \to \mathbb{R}^2$ $T(x,y) = (2x, x+y)$
   $B = B' = \{(-1,2),(0,3)\}$.

   c) $T : \mathbb{P}_2 \to \mathbb{R}^2$ $T(a+bx+cx^2) = (a+b, b-c)$
   B la base estándar de $\mathbb{P}_2$, $B' = \{(1,-1),(2,3)\}$.

   $\boxed{d)}$ $T : \mathbb{M}_2 \to \mathbb{P}_1$ $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a+b) + (c+d)x$
   $B = \{E_{11}, E_{11}+E_{12}, E_{21}, E_{21}+E_{22}\}$, B' la base estándar de $\mathbb{P}_1$.

$\boxed{2.}$ Sea $T : \mathbb{P}_2 \to \mathbb{P}_1$ una transformación lineal tal que

$$A = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix}$$

es la matriz de T con respecto a las bases $B = \left\{1,\ 1+x,\ 2-x+x^2\right\}$ y $B' = \{-1+x,\ 2x\}$. Halle la ley de T.

3. Repita el ejercicio anterior siendo B y B' las bases estándar de $\mathbb{P}_2$ y $\mathbb{P}_1$, respectivamente.

$\boxed{4.}$ Obtenga la ley de la transformación lineal $T : \mathbb{M}_2 \to \mathbb{R}^4$ si la matriz de T con respecto a las bases

$$B = \{E_{11}-E_{12},\ E_{12}-E_{21},\ E_{21}-E_{22},\ E_{22}+E_{11}\}, \quad \text{y}$$
$$B' = \{(1,-1,0,0),\ (0,1,-1,0),\ (0,0,1,-1),\ (1,0,0,1)\}$$

es

$$A = \begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

5. Haciendo uso del Teorema 6.10 obtenga el rango y la nulidad de las transformaciones lineales de los ejercicios 2 y 4.

6. Use el Teorema 6.10 para hallar el núcleo y el recorrido de la transformación lineal $T : \mathbb{R}^3 \to \mathbb{P}_2$ tal que

$$A = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

es la matriz de T con respecto a las bases

$$B = \{(-1, 2, 0), (0, 1, 0), (1, 1, 3)\}, \quad \text{y}$$

$$B' = \left\{1 + x, \ -1 + 3x, \ 2x^2\right\}.$$

7. Sea $T : \mathbb{P}_2 \to \mathbb{P}_4$ la transformación lineal definida por $T(p) = x^2 p$.

a) Halle la matriz de T con respecto a las bases $B = \{p_1, \ p_2, \ p_3\}$ y B' donde $p_1 = 1 + x^2$, $p_2 = 1 + 2x + 3x^2$ y $p_3 = 4 + 5x + x^2$, y B' es la base estándar de $\mathbb{P}_4$.

b) Obtenga $T(-3 + 5x - 2x^2)$ de dos maneras distintas: de forma directa (usando la ley de T) y de forma indirecta (usando la matriz de T hallada).

8. Sea $A = \begin{bmatrix} 3 & -2 & 1 & 0 \\ 1 & 6 & 2 & 1 \\ -3 & 0 & 7 & 1 \end{bmatrix}$ la matriz de una transformación lineal de $\mathbb{R}^4$ en $\mathbb{R}^3$ con respecto a las bases $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ y $B' = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ en donde $\mathbf{v}_1 = (0, 1, 1, 1)$, $\mathbf{v}_2 = (2, 1, -1, 1,)$, $\mathbf{v}_3 = (1, 4, -1, 2)$ y $\mathbf{v}_4 = (6, 9, 4, 2)$, $\mathbf{w}_1 = (0, 8, 8)$, $\mathbf{w}_2 = (-7, 8, 1)$ y $\mathbf{w}_3 = (-6, 9, 1)$.

a) Obtenga $[T(\mathbf{v}_1)]_{B'}$, $[T(\mathbf{v}_2)]_{B'}$, $[T(\mathbf{v}_3)]_{B'}$ y $[T(\mathbf{v}_4)]_{B'}$.

b) Encuentre $T(\mathbf{v}_1)$, $T(\mathbf{v}_2)$, $T(\mathbf{v}_3)$ y $T(\mathbf{v}_4)$.

c) Halle la ley de T.

9. Halle dos matrices distintas A' y A'' semejantes a la matriz $A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$. Verifique además que A'' y A' resultan también semejantes.

10. Halle todas las matrices semejantes a la matriz identidad de orden 3.

11. En cada ítem explique por qué no son semejantes las matrices A y B.

a) $A = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 3 \\ 4 & 3 \end{bmatrix}$.

b) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 4 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 5 & 5 & 5 \end{bmatrix}$.

12. Sea $T : \mathbb{V} \to \mathbb{V}$ un operador lineal. Encuentre la matriz $A' = [T]_{B'}$ y verifique que A' es semejante a $A = [T]_B$ donde B es la base estándar de $\mathbb{V}$ siendo:

a) $T : \mathbb{P}_1 \to \mathbb{P}_1$ tal que $T(a + bx) = (2a - b) + (b - a)x$, $B' = \{1 - 2x, \ 3x\}$.

b) $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $T(x, y, z) = (x, x+2y, x+y+3z)$, $B' = \{(1, -1, 0), \ (0, 0, 1), \ (0, 1, -1)\}$.

c) $T:\mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(x,y)=(4x+y,0)$, $B'=\{(1,1),\ (1,-1)\}$

$\boxed{13.}$ Si $A=\begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$ es la matriz del operador lineal T sobre $\mathbb{R}^2$ en la base estándar B y $P=\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ es la matriz de transición de una base B' a la base estándar B, halle la base B' y la matriz A' de T con respecto a la base B'.

14. Si $A=\begin{bmatrix} 2 & 0 & 2 \\ 0 & -3 & 1 \\ 0 & 2 & 0 \end{bmatrix}$ es la matriz del operador lineal T sobre $\mathbb{P}_2$ en la base estándar B y $P=\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$ es la matriz de transición de una base B' a la base estándar B, halle la base B' y la matriz A' de T con respecto a la base B'.

$\boxed{15.}$ Sea el operador lineal T sobre $\mathbb{R}^2$ cuya matriz en la base $B=\{(1,-1),\ (1,2)\}$ es $A=[T]_B=\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$. Sea $B'=\{(2,0),(-4,1)\}$ otra base de $\mathbb{R}^2$.

a) Halle P y Q, las matrices de transición de B' a B y de B a B', respectivamente.

b) Halle la matriz de T en la base B', $[T]_{B'}$.

c) Si $\mathbf{v}=(0,1)$, halle $[T(\mathbf{v})]_B$ y $[T(\mathbf{v})]_{B'}$.

$\boxed{16.}$ **(Matriz del operador identidad)** Sea $\mathbb{V}$ un espacio vectorial de dimensión n. Demuestre que la matriz con respecto a cualquier base del operador operador lineal identidad I sobre $\mathbb{V}$ es la matriz identidad I_n. Esto es, $[I]_B=I_n$ para toda base B de $\mathbb{V}$.

17. **(Matriz del operador cero)** Halle la matriz con respecto a cualquier base del operador lineal cero sobre un espacio vectorial $\mathbb{V}$ de dimensión n.

18. Sean A y B matrices semejantes. Demuestre:

a) A^t y B^t son matrices semejantes.

b) Si A es invertible entonces también B es invertible, y A^{-1} y B^{-1} son matrices semejantes.

c) A^2 y B^2 son matrices semejantes. En general, si k es un entero positivo entonces A^k y B^k son matrices semejantes.

Capítulo 7

AUTOVALORES Y AUTOVECTORES

7.1. Autovalores y autovectores

Como se ha visto en el capítulo anterior, si A es una matriz real de $n \times n$ la función $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(\mathbf{x}) = A\mathbf{x}$ es una transformación lineal (transformación matricial). Una cuestión de importancia en una gran variedad de problemas de aplicación es la determinación de vectores $\mathbf{x} \neq \mathbf{0}$, si los hay, tales que $\mathbf{x}$ y su transformado, $A\mathbf{x}$, sean "paralelos".

Definición (Autovalor y autovector). Sea A una matriz real de $n \times n$, un número real λ es un *autovalor* de A si existe un vector $\mathbf{u}$, distinto del vector cero, en $\mathbb{R}^n$ tal que

$$A\mathbf{u} = \lambda\mathbf{u}.$$

El vector $\mathbf{u}$ es un *autovector* de A correspondiente al autovalor λ.

Así, un vector $\mathbf{u} \neq \mathbf{0}$ en $\mathbb{R}^n$ es un autovector de la matriz A si existe un número real λ tal que $A\mathbf{u} = \lambda\mathbf{u}$. Esto es, $A\mathbf{u}$ es un múltiplo escalar de $\mathbf{u}$. El número λ es un autovalor de A.

A los autovalores también se los llama *valores característicos*, *valores propios* o *eigenvalores* ("eigen" significa "propio" en alemán). Similarmente los autovectores también son llamados *vectores característicos*, *vectores propios* o *eigenvectores*.

Debe observarse que en la definición de autovalor no se tiene en cuenta al vector $\mathbf{u} = \mathbf{0}$. Esto se debe a que el vector cero satisface $A\mathbf{0} = \lambda\mathbf{0}$ para cualquier número λ y entonces todo número real sería un autovalor. El vector cero no es considerado un autovector en cambio, sí es posible un autovalor $\lambda = 0$.

Ejemplo 7.1. Sea $A = \begin{bmatrix} 2 & 0 \\ 6 & -1 \end{bmatrix}$ y los vectores $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ y $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$. Para determinar si éstos vectores son autovectores de A, se efectúan los productos $A\mathbf{u}_1$, $A\mathbf{u}_2$ y $A\mathbf{u}_3$:

$$A\mathbf{u}_1 = \begin{bmatrix} 2 & 0 \\ 6 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2\mathbf{u}_1,$$

luego $\mathbf{u}_1$ es un autovector de A asociado al autovalor $\lambda_1 = 2$.

$$A\mathbf{u}_2 = \begin{bmatrix} 2 & 0 \\ 6 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = (-1) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = (-1)\mathbf{u}_2,$$

por lo tanto $\mathbf{u}_2$ es un autovector de A asociado al autovalor $\lambda_2 = -1$.

En cambio el vector $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ no es un autovector de A dado que

$$A\mathbf{u}_3 = \begin{bmatrix} 2 & 0 \\ 6 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix},$$

no es un múltiplo escalar de $\mathbf{u}_3$.◀

Nota: Debe notarse que un autovalor tiene asociado infinitos autovectores. De hecho, *si* $\mathbf{u}$ *es un* autovector de A *correspondiente al autovalor* λ, *cualquier múltiplo escalar no cero de* $\mathbf{u}$ *es un autovector de* A *asociado al mismo autovalor* λ.

En efecto, sea $\mathbf{v} = k\mathbf{u}$ donde $k \neq 0$. Entonces $\mathbf{v} \neq \mathbf{0}$ y

$$A\mathbf{v} = A\,(k\mathbf{u}) = k\,(A\mathbf{u}) = k\,(\lambda\mathbf{u}) = \lambda\,(k\mathbf{u}) = \lambda\mathbf{v},$$

y así $k\mathbf{u}$ es también un autovector de A correspondiente al autovalor λ.

Geométricamente esto muestra que en $\mathbb{R}^2$ o $\mathbb{R}^3$, si $\mathbf{u}$ es un autovector de A correspondiente a un autovalor λ, todos los vectores no ceros que se encuentran en la recta (en el plano o en el espacio) que contiene al origen de coordenadas y es paralela al vector $\mathbf{u}$, son autovectores de A correspondientes al mismo autovalor λ.

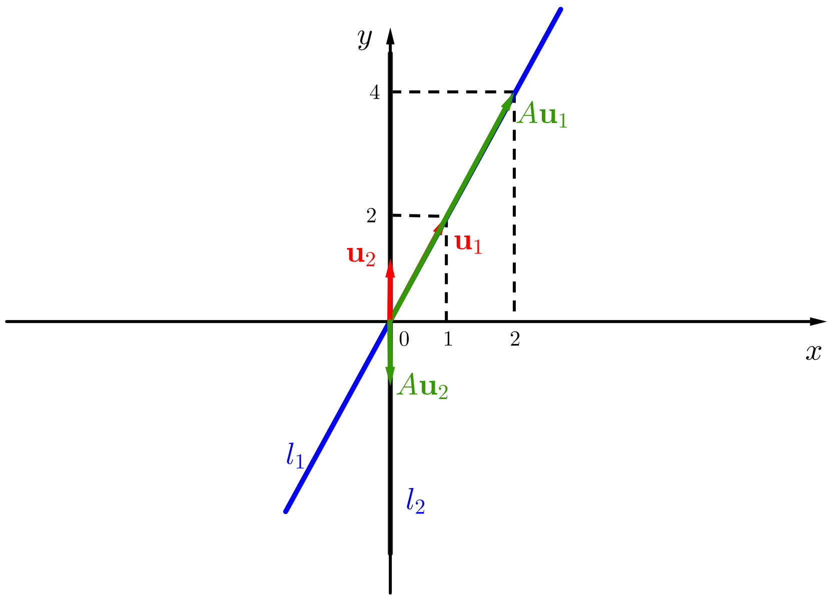


Figura 7.1: Interpretación geométrica del Ejemplo 7.1

Observe que en el ejemplo anterior todos los vectores distintos de cero que están en las rectas l_1 y l_2, paralelas a los vectores $\mathbf{u}_1$ y $\mathbf{u}_2$ respectivamente, son autovectores de A. La transformación matricial $\mathbf{x} \to A\mathbf{x}$ "estira" los vectores no cero de l_1 en un factor de 2 y los vectores de la recta l_2 se reflejan respecto al origen (véase Figura 7.1).

Ejemplo 7.2. Sea la matriz $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. A es la matriz estándar del operador lineal reflexión, en $\mathbb{R}^2$ con respecto a la recta de ecuación $y = x$. Esto es

$$A\mathbf{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}.$$

Una interpretación visual de la transformación es suficiente para comprender que los únicos vectores que permanecen en la misma recta después de la reflexión son los que se encuentran en las rectas de ecuaciones $y = x$ y $y = -x$. Estos vectores (excepto el vector cero) son los únicos que se transforman en múltiplos de sí mismos. Si $\mathbf{u}$ está en la recta $y = x$ entonces $\mathbf{u} = \begin{bmatrix} u_1 \\ u_1 \end{bmatrix}$ y $A\mathbf{u} = \mathbf{u} = 1\mathbf{u}$, así $\mathbf{u}$ es un autovector correspondiente al autovalor $\lambda_1 = 1$.

Si $\mathbf{v}$ está en la recta $y = -x$ entonces $\mathbf{v} = \begin{bmatrix} v_1 \\ -v_1 \end{bmatrix}$ por lo que $A\mathbf{v} = -\mathbf{v} = (-1)\mathbf{v}$, así $\mathbf{v}$ es un autovector correspondiente al autovalor $\lambda_2 = -1$.◀

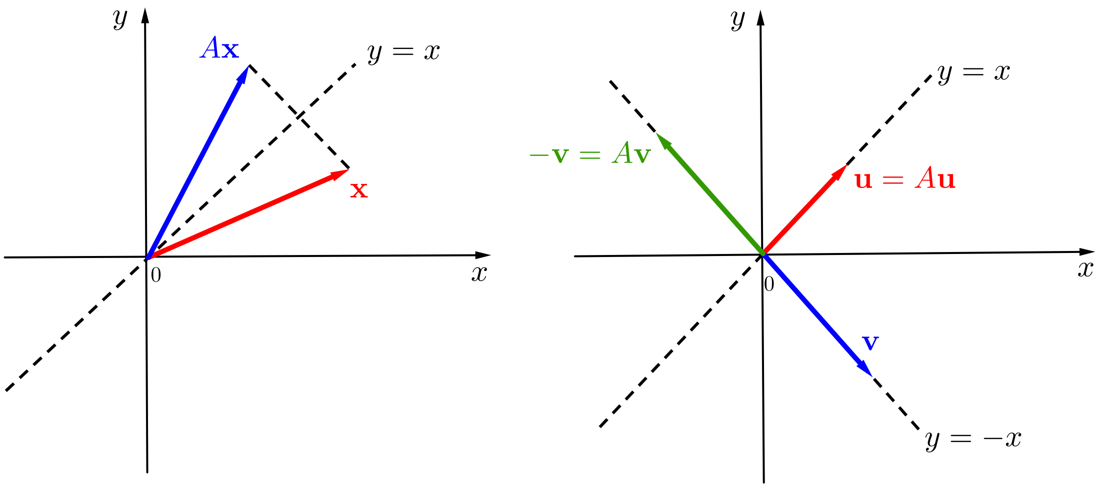


Figura 7.2: Reflexiones

En el ejemplo anterior se pudieron determinar los autovalores y sus correspondientes autovectores basándose en un enfoque geométrico. Si A es una matriz cuadrada de orden $n > 3$ ya no es posible usar la intuición geométrica y, aún para $n = 3$, hay transformaciones complicadas para visualizarlas geométricamente.

Ejemplo 7.3 (**Autovalores y autovectores de la matriz identidad**). Si $A = I$ (matriz identidad de orden n) entonces $A\mathbf{x} = I\mathbf{x} = \mathbf{x} = 1\mathbf{x}$, para todo $x \in \mathbb{R}^n$.

Luego, todo vector no cero de $\mathbb{R}^n$ es un autovector de A correspondiente al único autovalor $\lambda = 1$.◀

En los ejemplos vistos se han obtenido los autovalores y autovectores por medio de una simple inspección, argumentos geométricos o enfoques algebraicos muy sencillos. En el ejemplo siguiente se calcularán los autovalores y autovectores de una matriz de manera sistemática.

Ejemplo 7.4. Se desean determinar los autovalores y autovectores de la matriz $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$.

Por definición, un número real λ es un autovalor de A si existe un vector $\mathbf{u}$ no cero en $\mathbb{R}^2$ tal que $A\mathbf{u} = \lambda\mathbf{u}$. Por lo tanto, λ es un autovalor de A si y solo si la ecuación

$$A\mathbf{x} = \lambda\mathbf{x},$$

es decir:

$$\begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

tiene soluciones no triviales. La ecuación anterior se convierte en el sistema lineal

$$\begin{cases} 6x_1 + 5x_2 = \lambda x_1 \\ x_1 + 2x_2 = \lambda x_2 \end{cases}$$

y una forma canónica del mismo es

$$\begin{cases} (\lambda - 6)\, x_1 - 5x_2 = 0 \\ -x_1 + (\lambda - 2)\, x_2 = 0 \end{cases} \tag{7.1}$$

De acuerdo al Teorema 3.10 este sistema homogéneo tiene soluciones no triviales si y solo si el determinante de su matriz de coeficientes, $\lambda I - A$, es cero. Esto es, si y solo si

$$\begin{vmatrix} \lambda - 6 & -5 \\ -1 & \lambda - 2 \end{vmatrix} = 0$$

es decir

$$(\lambda - 6)\,(\lambda - 2) - 5 = 0,$$

o

$$\lambda^2 - 8\lambda + 7 = 0,$$

cuyas soluciones son $\lambda_1 = 7$ y $\lambda_2 = 1$. Estos valores son autovalores de A puesto que son números reales para los cuales la ecuación $A\mathbf{x} = \lambda\mathbf{x}$ tiene soluciones no triviales.

Para determinar los autovectores correspondientes al autovalor $\lambda_1 = 7$, se resuelve

$$A\mathbf{x} = 7\mathbf{x}$$

que es equivalente a resolver el sistema homogéno

$$\begin{cases} x_1 - 5x_2 = 0 \\ -x_1 + 5x_2 = 0 \end{cases}$$

Observe que se obtiene este último sistema simplemente sustituyendo $\lambda = 7$ en (7.1). La solución general del sistema es

$$\begin{cases} x_1 = 5t \\ x_2 = t \end{cases}, \quad t \in \mathbb{R}.$$

Entonces los (infinitos) autovectores de A asociados al autovalor $\lambda_1 = 7$ están dados por $\begin{bmatrix} 5t \\ t \end{bmatrix}$ para cualquier t real, $t \neq 0$. En particular, $\mathbf{u} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ es un autovector de A asociado al autovalor $\lambda_1 = 7$.

De manera similar se obtienen los autovectores correspondientes a $\lambda_2 = 1$ resolviendo $A\mathbf{x} = \mathbf{x}$ que equivale a resolver el sistema homogéneo

$$\begin{cases} -5x_1 - 5x_2 = 0 \\ \quad -x_1 - x_2 = 0 \end{cases}$$

cuya solución general es

$$\begin{cases} x_1 = -t \\ x_2 = t \end{cases}, \quad t \in \mathbb{R}.$$

Entonces los (infinitos) autovectores de A asociados al autovalor $\lambda_2 = 1$ están dados por $\begin{bmatrix} -t \\ t \end{bmatrix}$ para cualquier t real, $t \neq 0$. En particular, $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ es un autovector de A asociado al autovalor $\lambda_2 = 1$.◀

El siguiente teorema resume lo realizado en el ejemplo anterior para el cálculo de autovalores y autovectores de una matriz.

Teorema 7.1 (Caracterización de autovalores y autovectores)**.** Sea A una matriz cuadrada de orden n y λ un número real, entonces:

1. λ es un autovalor de A si y solo si $\det(\lambda I - A) = 0$.
2. Si λ es un autovalor de A, un vector $\mathbf{u} \in \mathbb{R}^n$ es un autovector de A asociado a λ si y solo si $\mathbf{u}$ es una solución no trivial del sistema homogéneo $(\lambda I - A)\,\mathbf{x} = \mathbf{0}$.

Demostración. Nótese que, por definición, $\lambda \in \mathbb{R}$ es un autovalor de A si y solo si la ecuación $A\mathbf{x} = \lambda\mathbf{x}$ admite soluciones distintas de la trivial. Esta ecuación, como vimos en el ejemplo, y ahora se generaliza se transforma en un sistema lineal homogéneo de matriz de coeficientes $\lambda I - A$. En efecto:

$$\begin{aligned} A\mathbf{x} = \lambda\mathbf{x} \quad &\Leftrightarrow \quad \lambda\mathbf{x} - A\mathbf{x} = \mathbf{0} \\ &\Leftrightarrow \quad (\lambda I\mathbf{x}) - A\mathbf{x} = \mathbf{0} \\ &\Leftrightarrow \quad (\lambda I - A)\,\mathbf{x} = \mathbf{0}. \end{aligned}$$

Luego, un número real λ es un autovalor de A si y solo si el sistema homogéno de matriz de coeficientes $\lambda I - A$ tiene soluciones no triviales. Como esta matriz es cuadrada, el sistema tiene soluciones no triviales si y solo si $\lambda I - A$ es singular o lo que es equivalente $\det(\lambda I - A) = 0$. Esto prueba la parte 1.

Por otra parte, un vector $\mathbf{u}$ de $\mathbb{R}^n$ es un autovector de A asociado a un autovalor λ si y solo si $\mathbf{u}$ es una solución no trivial de la ecuación $A\mathbf{x} = \lambda\mathbf{x}$, o lo que es equivalente, $\mathbf{u}$ es una solución no trivial del sistema homogéneo $(\lambda I - A)\,\mathbf{x} = \mathbf{0}$. Esto prueba la parte 2. del Teorema. □

Polinomio característico y ecuación característica. Espacios característicos

El problema de encontrar los autovalores de una matriz real se resuelve con la ayuda de los determinantes ya que un autovalor de la matriz A es un número real λ tal que $\det(\lambda I - A) = 0$.

Esto sugiere que $\det(\lambda I - A)$ debería estudiarse como una función de λ. En el ejemplo anterior se obtuvo para la matriz cuadrada de orden 2, $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$:

$$\det(\lambda I - A) = \lambda^2 - 8\lambda + 7.$$

En general, si se desarrolla $\det(\lambda I - A)$ para una matriz A de orden n se obtiene un polinomio de grado n en la variable λ. Al respecto vale el siguiente teorema que aceptaremos sin demostración y que es fácilmente verificable si $n = 2$ o $n = 3$.

Teorema 7.2. Si A es una matriz cuadrada de orden n entonces

$$\det(\lambda I - A) = \lambda^n + c_{n-1}\lambda^{n-1} + ... + c_1\lambda + c_0,$$

donde $c_i \in \mathbb{R}$, $i = 1, 2, ..., n$. Es decir, $\det(\lambda I - A)$ es un polinomio de grado n en la variable λ a coeficientes reales. Además su coeficiente principal (coeficiente que acompaña a λ^n) es 1.

En relación al polinomio que se obtiene del teorema anterior, se tienen las siguientes definiciones.

Definición (Polinomio característico y ecuación característica). Sea A una matriz cuadrada de orden n y I la matriz identidad de orden n.

1. Se denota y define el *polinomio característico* de la matriz A como

$$p_A(\lambda) = \det(\lambda I - A).$$

2. A la ecuación

$$\det(\lambda I - A) = 0,$$

esto es $p_A(\lambda) = 0$, se la llama *ecuación característica* de la matriz A.

Con esta terminología los autovalores de una matriz (real) A son las soluciones reales de la ecuación característica, es decir, las raíces reales del polinomio característico $p_A(\lambda)$.

Dado que $p_A(\lambda)$ es un polinomio de grado n a coeficientes reales, éste tiene exactamente n raíces (contando las repeticiones) algunas de las cuales (o todas) pueden ser números complejos. Las raíces reales son los autovalores de A, por lo tanto una matriz (real) A, de orden n tiene a lo sumo n autovalores, no necesariamente todos distintos. Es importante señalar que A puede carecer de autovalores (si todas las raíces del polinomio característico son complejas).

Al número de veces que se repite un autovalor λ_i como raíz del polinomio característico se lo llama **multiplicidad algebraica** de λ_i y se denotará $m_a(\lambda_i)$.

Por ejemplo si A es una matriz cuadrada de orden 6 y las raíces de $p_A(\lambda)$ son: 4, 4, 4, -7, $2 - i$ y $2 + i$, entonces A tiene sólo dos autovalores distintos 4 y -7 con multiplicidades algebraicas $m_a(4) = 3$ y $m_a(-7) = 1$.

El problema de hallar los autovectores asociados a un autovalor se reduce a la resolución de un sistema lineal homogéneo con infinitas soluciones. Sabemos que si λ es un autovalor de la matriz A, los autovectores correspondientes a λ son las soluciones no triviales del sistema $(\lambda I - A)\,\mathbf{x} = \mathbf{0}$ es decir son los vectores no cero del espacio nulo de la matriz $\lambda I - A$.

El espacio nulo de la matriz $\lambda I - A$ se llama **espacio característico** o **autoespacio** de la matriz A asociado al autovalor λ y lo simbolizaremos E_λ. Así

$$E_\lambda = N(\lambda I - A) = \{\mathbf{x} \in \mathbb{R}^n : (\lambda I - A)\,\mathbf{x} = \mathbf{0}\}.$$

Se ha visto (Capítulo 4) que el espacio nulo de una matriz es un subespacio de $\mathbb{R}^n$. Luego, E_λ es el subespacio de $\mathbb{R}^n$ que consta de todos los autovectores de A asociados al autovalor λ, juntamente con el vector cero.

A la dimensión de $E_\lambda = N(\lambda I - A)$ se la denomina **multiplicidad geométrica** del autovalor λ y lo denotaremos $m_g(\lambda)$. Entonces,

$$m_g(\lambda) = \nu(\lambda I - A),$$

es el número máximo de autovectores linealmente independientes que están asociados al autovalor λ. Claramente $m_g(\lambda) \geq 1$ ya que $E_\lambda = N(\lambda I - A) \neq \{\mathbf{0}\}$. Es un hecho básico que la *multiplicidad geométrica es menor o igual que la multiplicidad algebraica*, es decir si λ es un autovalor de A se tiene que

$$m_g(\lambda) \leq m_a(\lambda).$$

Resumiendo:

Para obtener los autovalores, los autovectores y las dimensiones de los autoespacios de una matriz A de orden n, se procede de la siguiente manera:

1. Se obtiene el polinomio característico $p_A(\lambda) = \det(\lambda I - A)$.

2. Se buscan las soluciones reales de la ecuación característica $\det(\lambda I - A) = 0$. Estos números son los autovalores de A.

3. Para cada autovalor distinto λ_i, se resuelve el sistema homogéneo $(\lambda_i I - A)\,\mathbf{x} = \mathbf{0}$. El conjunto solución es el autoespacio E_{λ_i} y las soluciones no triviales son los autovectores de A asociados al autovalor λ_i.

4. Se busca una base para cada autoespacio $E_{\lambda_i} = N(\lambda_i I - A)$. La multiplicidad geométrica $m_g(\lambda_i)$ es la dimensión de E_{λ_i}.

Nota: El problema de encontrar los autovalores de una matriz puede ser muy difícil, en particular si $n \geq 3$ ya que se trata de determinar las raíces de un polinomio de grado n. En cursos de Análisis Numérico, se estudian métodos numéricos eficientes para hallar aproximaciones de las raíces de un polinomio. De hecho muchos programas de computadora lo hacen.

Un resultado útil es el siguiente: si el polinomio característico de la matriz A es

$$p_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + ... + c_1\lambda + c_0,$$

donde c_1, c_2,..., c_{n-1} son números enteros entonces sus raíces racionales, si las tiene, deben estar entre los números enteros divisores de c_0 (teorema de Gauss). Por supuesto, $p_A(\lambda)$ podrá tener raíces irracionales o complejas, sin embargo, por conveniencia muchos de los polinomios característicos en este texto tendrán sólo raíces enteras.

Ejemplo 7.5. Sea $A = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 3 & 1 \end{bmatrix}$. El polinomio característico de A es

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I - A) = \begin{vmatrix} \lambda-2 & -1 & 0 \\ 1 & \lambda & -1 \\ -1 & -3 & \lambda-1 \end{vmatrix} \\ &= (\lambda-2)\lambda(\lambda-1) - 1 - 3(\lambda-2) + (\lambda-1) \\ &= \lambda^3 - 3\lambda^2 + 4. \end{aligned}$$

Dado que $p_A(\lambda)$ es a coeficientes enteros y su coeficiente principal es igual a 1, entonces si el polinomio tiene raíces racionales éstas deben estar entre los enteros divisores del término independiente (Teorema de Gauss). Las posibles raíces racionales son: ± 1, ± 2 y ± 4.

Se observa que $p_A(-1) = 0$ por lo tanto $\lambda + 1$ es un divisor de $p_A(\lambda)$. Así, $p_A(\lambda) = (\lambda+1)q(\lambda)$ donde $q(\lambda)$ es el cociente de la división del polinomio $p_A(\lambda)$ por $\lambda + 1$. Aplicando la regla de Ruffini, $q(\lambda) = \lambda^2 - 4\lambda + 4 = (\lambda-2)^2$. Así,

$$p_A(\lambda) = (\lambda+1)(\lambda-2)^2.$$

Los autovalores de A son las soluciones reales de la ecuación característica:

$$p_A(\lambda) = (\lambda+1)(\lambda-2)^2 = 0,$$

que por simple inspección son 2 (raíz doble) y -1 (raíz simple). Por lo tanto los autovalores distintos de A son:

$$\lambda_1 = 2 \quad \text{y} \quad \lambda_2 = -1,$$

con multiplicidades algebraicas $m_a(2) = 2$ y $m_a(-1) = 1$.

Los autovectores correpondientes a cada autovalor λ_i son las soluciones no triviales del sistema homogéneo $(\lambda_i I - A)\,\mathbf{x} = \mathbf{0}$.

Para $\lambda_1 = 2$: $\quad \lambda_1 I - A = 2I - A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 2 & -1 \\ -1 & -3 & 1 \end{bmatrix}$ y el sistema lineal homogéneo es

$$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 2 & -1 \\ -1 & -3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Aplicando eliminación gaussiana a la matriz ampliada del sistema se obtiene la siguiente forma escalonada

$$\begin{bmatrix} 1 & 2 & -1 & : & 0 \\ 0 & -1 & 0 & : & 0 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}.$$

La solución general del sistema es

$$\begin{cases} x_1 = t \\ x_2 = 0 \\ x_3 = t \end{cases}, \quad t \in \mathbb{R}.$$

Los autovectores de A asociados a $\lambda_1 = 2$ son todos los vectores $\begin{bmatrix} t \\ 0 \\ t \end{bmatrix}$ con $t \neq 0$. El autoespacio correspondiente a $\lambda_1 = 2$ es

$$E_2 = N(2I - A) = \left\{ \begin{bmatrix} t \\ 0 \\ t \end{bmatrix} : t \in \mathbb{R} \right\} = \text{gen} \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

El conjunto $S_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ es linealmente independiente y generador de E_2 por lo que constituye una base para éste. Así, $m_g(2) = \dim(E_2) = 1$.

Para $\lambda_2 = -1$: $\quad \lambda_2 I - A = -I - A = \begin{bmatrix} -3 & -1 & 0 \\ 1 & -1 & -1 \\ -1 & -3 & -2 \end{bmatrix}$ y el sistema lineal homogéneo es

$$\begin{bmatrix} -3 & -1 & 0 \\ 1 & -1 & -1 \\ -1 & -3 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Aplicando eliminación gaussiana a la matriz ampliada del sistema se obtiene la siguiente forma escalonada

$$\begin{bmatrix} 1 & -1 & -1 & : & 0 \\ 0 & -4 & -3 & : & 0 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}.$$

La solución general del sistema es

$$\left\{ \begin{array}{r} x_1 = t \\ x_2 = -3t \\ x_3 = 4t \end{array} \right. , \quad t \in \mathbb{R}.$$

Los autovectores de A asociados a $\lambda_2 = -1$ son todos los vectores $\begin{bmatrix} t \\ -3t \\ 4t \end{bmatrix}$ con $t \neq 0$. El autoespacio correspondiente a $\lambda_2 = -1$ es

$$E_{-1} = N(-I - A) = \left\{ \begin{bmatrix} t \\ -3t \\ 4t \end{bmatrix} : t \in \mathbb{R} \right\} = \text{gen} \left\{ \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix} \right\}.$$

El conjunto $S_2 = \left\{ \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix} \right\}$ es linealmente independiente y generador de E_{-1} por lo que constituye una base para éste. Así, $m_g(-1) = \dim(E_{-1}) = 1$. ◀

Ejemplo 7.6. Sea la matriz $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}$ cuyo polinomio característico es

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I - A) = \begin{vmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -4 & 17 & \lambda - 8 \end{vmatrix} \\ &= \lambda^3 - 8\lambda^2 + 17\lambda - 4. \end{aligned}$$

Los autovalores de A son las soluciones reales de la ecuación característica

$$p_A(\lambda) = \lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0.$$

Para resolver esta ecuación se comienza por buscar soluciones racionales que, si las tiene, deben ser números enteros divisores del término constante -4. Así, las únicas soluciones racionales posibles son: ± 1, ± 2 y ± 4.

Se comprueba que ± 1 y ± 2 no solucionan la ecuación pero $p_A(4) = 0$ por lo tanto 4 es raíz y $\lambda - 4$ es un divisor de $p_A(\lambda)$. Así, $p_A(\lambda) = (\lambda - 4)q(\lambda)$ donde $q(\lambda)$ es el cociente de la división del polinomio $p_A(\lambda)$ por $\lambda - 4$. Aplicando la regla de Ruffini, $q(\lambda) = \lambda^2 - 4\lambda + 1$ lo que indica que la ecuación característica se puede escribir como

$$p_A(\lambda) = (\lambda - 4)\left(\lambda^2 - 4\lambda + 1\right) = 0.$$

Por consiguiente una solución es el número entero 4 y las restantes son las que satisfacen la ecuación cuadrática $\lambda^2 - 4\lambda + 1 = 0$ que se puede resolver por medio de la fórmula correspondiente obteniéndose los raíces irracionales $2 + \sqrt{3}$ y $2 - \sqrt{3}$.

Por lo tanto los autovalores de A son $\lambda_1 = 4$, $\lambda_2 = 2 + \sqrt{3}$ y $\lambda_3 = 2 - \sqrt{3}$, todos de multiplicidad algebraica igual a 1.◀

Actividad 7.1. Para cada autovalor de la matriz A del ejemplo anterior, exhiba dos autovectores distintos y verifique que las multiplicidades geométricas son todas iguales a 1.

Ejemplo 7.7 (Una matriz que no tiene autovalores). Sea $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. El polinomio característico de A es

$$p_A(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda & 1 \\ -1 & \lambda \end{vmatrix} = \lambda^2 + 1.$$

La ecuación característica $\lambda^2 + 1 = 0$ no tiene soluciones reales. Por lo tanto A no tiene autovalores y en consecuencia no tiene autovectores. Puede verse geométricamente la razón. Esta matriz es una matriz de rotación de ángulo $\theta = \frac{\pi}{2}$. Así, todo vector no cero en $\mathbb{R}^2$ es girado, por la matriz A, un ángulo de $90°$ y por lo tanto no se transforma en múltiplo de sí mismo (véase Figura 7.3).

Hay aplicaciones que requieren escalares complejos. En dichos casos las raíces complejas del polinomio característico de una matriz A son también autovalores de A y habrá autovectores asociados con componentes complejas. En este caso, las únicas soluciones de la ecuación $\lambda^2 + 1 = 0$ son los complejos conjugados $\lambda_1 = i$ y $\lambda_2 = -i$. Si se permiten escalares complejos, éstos son autovalores de A. Puede verificarse que

$$\mathbf{v}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

son autovectores correspondientes a $\lambda_1 = i$ y $\lambda_2 = -i$, respectivamente.◀

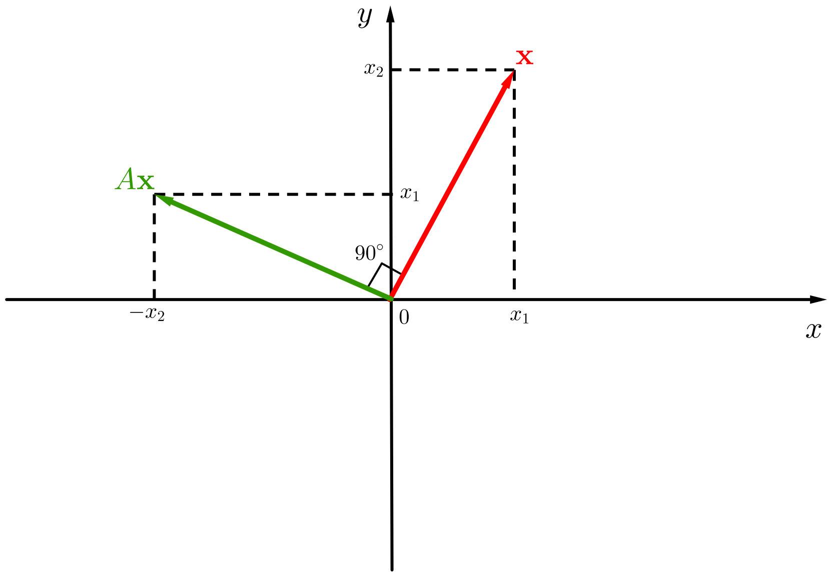


Figura 7.3: Interpretación geométrica del Ejemplo 7.7

Para finalizar con esta sección se verá que para ciertas matrices es sencillo hallar sus autovalores. No se necesita hacer ningún cálculo para hallar los autovalores de una matriz triangular, en particular, para las matrices diagonales.

Teorema 7.3. Los autovalores de una matriz triangular son los elementos de la diagonal.

Demostración. Sea $A = [a_{ij}]$ una matriz triangular de orden n. Se probará que los elementos a_{ii}, $i = 1, 2, ..., n$ son las raíces del polinomio característico $p_A(\lambda) = \det(\lambda I - A)$.

Por ser A triangular, la matriz $\lambda I - A$ es también triangular y por lo tanto su determinante es el producto de los elementos de la diagonal. Los elementos de la diagonal de $\lambda I - A$ son $\lambda - a_{ii}$, $i = 1, 2, ..., n$. Luego,

$$p_A(\lambda) = \det(\lambda I - A) = (\lambda - a_{11})(\lambda - a_{22}) \ldots (\lambda - a_{nn}).$$

Claramente, las raíces de $p_A(\lambda)$ son los números: $\lambda_1 = a_{11}$, $\lambda_2 = a_{22}$,..., $\lambda_n = a_{nn}$ (no todos necesariamente distintos) que, como son números reales, todos son autovalores de A. □

Para las matrices triangulares, en particular para las matrices diagonales, valen las siguientes propiedades que son inmediatas a partir del teorema anterior:

Corolario 7.1. Si A es una matriz triangular entonces:

1. El producto de los autovalores de A es igual al determinante de A.
2. La suma de los autovalores de A es igual a la suma de los elementos de la diagonal de A. La suma de los elementos de la diagonal de una matriz cuadrada recibe el nombre de *traza* de A y se la denota $\mathrm{tr}(A)$.

Demostración. Sea $A = [a_{ij}]$ una matriz triangular de orden n. Dado que los autovalores de A son $\lambda_1 = a_{11}$, $\lambda_2 = a_{22}$,...., $\lambda_n = a_{nn}$, y como el determinante de una matriz triangular es igual al producto de los elementos de la diagonal y la traza es la suma de dichos elementos, entonces

$$\det(A) = a_{11}a_{22}...a_{nn} = \lambda_1\lambda_2...\lambda_n,$$

y

$$\operatorname{tr}(A) = a_{11} + a_{22} + ... + a_{nn} = \lambda_1 + \lambda_2 + ... + \lambda_n.$$

□

Estos resultados son válidos también para una matriz no triangular si se consideran las n raíces de su polinomio característico.

Teorema 7.4. Si A es una matriz cuadrada de orden n y las n raíces de su polinomio característico son λ_1, λ_2,..., λ_n (no necesariamente todas distintas) entonces:

1. $\det(A) = \lambda_1\lambda_2...\lambda_n$.
2. $\operatorname{tr}(A) = \lambda_1 + \lambda_2 + ... + \lambda_n$.

Demostración. Se probará aquí la propiedad 1. El polinomio característico de A es

$$p_A(\lambda) = \det(\lambda I - A) = \lambda^n + c_{n-1}\lambda^{n-1} + ... + c_1\lambda + c_0.$$

Dado que λ_1, λ_2,..., λ_n son las n raíces de $p_A(\lambda)$ se tiene la siguiente factorización del polinomio característico

$$p_A(\lambda) = \det(\lambda I - A) = (\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_n).$$

Tomando $\lambda = 0$ en la expresión factorizada de $p_A(\lambda)$:

$$p_A(0) = \det(-A) = (-1)^n\lambda_1\lambda_2...\lambda_n,$$

y como

$$\det(-A) = (-1)^n \det(A),$$

entonces

$$\det(A) = \lambda_1\lambda_2...\lambda_n.$$

Además puede observarse en la expresión del polinomio característico de A no factorizada, que el término independiente c_0 es $p_A(0)$ por lo tanto,

$$c_0 = (-1)^n \det(A).$$

La prueba de 2. no es tan inmediata (Véase Ejercicio 7.1). □

Observación 7.1. Los resultados establecidos por el teorema anterior son condiciones necesarias que deben satisfacer las raíces del polinomio característico y puede usarse como una comprobación numérica más en el cálculo de las raíces de $p_A(\lambda)$. Es sencillo comprobar si la suma de las raíces de $p_A(\lambda)$ coincide con la suma de los elementos de la diagonal (traza) de A.

Observación 7.2. Debe notarse que si una matriz tiene todos sus elementos reales su determinante y su traza son números reales. Pero el determinante es el producto de las raíces del polinomio característico y la traza es la suma de éstas, entonces aunque algunas de las raíces sean complejas su producto y su suma deben ser números reales. Esto no debe sorprender; es un hecho conocido que las raíces complejas de polinomios a coeficientes reales aparecen como pares conjugados, es decir, si $z = a + bi$ es una raíz también lo es $\overline{z} = a - bi$. Por lo tanto, puesto que $z\overline{z} = a^2 + b^2 \in \mathbb{R}$ y $z + \overline{z} = 2a \in \mathbb{R}$, resultan también $\lambda_1\lambda_2...\lambda_n \in \mathbb{R}$ y $\lambda_1 + \lambda_2 + ... + \lambda_n \in \mathbb{R}$.

Actividad 7.2. Para cada matriz A de los ejemplos 7.5, 7.6 y 7.7, verifique los resultados del teorema anterior.

Con respecto a las matrices invertibles, se tiene el siguiente resultado como consecuencia del teorema anterior.

Corolario 7.2. Una matriz A de orden n es invertible si y solo si 0 no es un autovalor de A.

Demostración. Sabemos que A es invertible si y solo si $\det(A) \neq 0$. Por el teorema anterior, $\det(A) = \lambda_1\lambda_2...\lambda_n$ donde λ_i, $i = 1, 2, ..., n$ es una raíz del polinomio característico. Por lo tanto $\det(A) \neq 0$ si y sólo $\lambda_i \neq 0$, para todo $i = 1, 2, ..., n$ es decir A no tiene autovalores 0. □

Ejercicios 7.1

1. Sean $A = \begin{bmatrix} 3 & -2 \\ -3 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Aplique la definición para determinar si los vectores dados son autovectores de A y, en caso afirmativo, halle sus autovalores correspondientes.

2. Para la matriz A y los vectores $\mathbf{u}$ y $\mathbf{v}$ del ejercicio anterior, ¿es $\mathbf{u} + \mathbf{v}$ un autovector de A?, y ¿$c\mathbf{u}$ con $c \neq 0$?

3. Obtenga, mentalemente, los autovalores de las siguientes matrices:
$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 8 & -1 & 0 \\ 0 & 17 & -1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 5 \\ 5 & 0 \end{bmatrix}.$$

4. Pruebe que $\lambda_1 = -6$, $\lambda_2 = 4$ y $\lambda_3 = 7$ son autovalores de la matriz
$$A = \begin{bmatrix} 7 & 0 & 0 \\ 0 & -2 & 4 \\ 0 & 6 & 0 \end{bmatrix}$$
y obtenga, para cada autovalor, el conjunto de sus autovectores asociados.

5. Sean $A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$. Utilice la definición de autovector para probar que $\mathbf{v}$ es un autovector de A y obtenga su correspondiente autoespacio.

6. Para cada matriz halle: el polinomio característico, los autovalores, bases para los autoespacios e indique las multiplicidades algebraica y geométrica de cada autovalor.

a) $\begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix}$

b) $\begin{bmatrix} 0 & -9 \\ 1 & -6 \end{bmatrix}$

c) $\begin{bmatrix} a & b \\ a & b \end{bmatrix}$

d) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

e) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 4 & 0 \\ 2 & 0 & 0 \end{bmatrix}$

f) $\begin{bmatrix} 0 & -1 & 0 \\ -4 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

g) $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ 3 & 1 & -1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

7. Sea A una matriz de 3×3 tal que $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$ son autovectores de A asociados a los autovalores $\lambda_1 = -1$ y $\lambda_2 = 2$, respectivamente. Halle el vector $A\,(2\mathbf{u} - 3\mathbf{v})$.

8. Sea A una matriz real de $n \times n$. Demuestre las siguientes proposiciones.

   a) A y A^t tienen el mismo polinomio característico y por lo tanto los mismos autovalores.

   b) Si $\mathbf{v}$ un autovector de A asociado al autovalor λ entonces $\mathbf{v}$ es un autovector de A^2 y su autovalor correspondiente es λ^2. (Esto puede generalizarse para A^k con $k \in \mathbb{Z}^+$).

9. Sea A una matriz de $n \times n$. Demuestre:

   a) Usando la definición de autovalor que A es invertible si y solo si 0 no es un autovalor de A.

   b) Si A es invertible y $\mathbf{v}$ es un autovector de A correspondiente al autovalor λ, entonces $\mathbf{v}$ es un autovector de A^{-1} y su correspondiente autovalor es $\frac{1}{\lambda}$.

10. Demuestre que si A es una matriz de $n \times n$ y $\mathbf{v}$ es un autovector de A correspondiente al autovalor λ, entonces $\mathbf{v}$ es un autovector de $A - cI$ con correspondiente autovalor $\lambda - c$, para todo $c \in \mathbb{R}$.

11. Demuestre que si A es una matriz de $n \times n$ tal que $A^2 = -I$, entonces A no tiene autovalores.

12. Demuestre la parte 2. del Teorema 7.4.

7.2. Diagonalización

Definición (**Matriz diagonalizable**). Una matriz A cuadrada de orden n es *diagonalizable* si existe una matriz P invertible y una matriz D diagonal tal que

$$P^{-1}AP = D.$$

Es decir, A es diagonalizable si es semejante a una matriz diagonal.

También suele decirse que las matrices P y D *diagonalizan* a A.

Debe tenerse en cuenta que estamos trabajamos solo con matrices reales. Esto es, la diagonalización debe ser considerada en $\mathbb{R}$, las matrices P y D que diagonalizan a una matriz real A, deben ser matrices reales.

Ejemplo 7.8 (**Las matrices diagonales son diagonalizables**). Esto es así ya que son semejantes a sí mismas: Si D es una matriz diagonal, $D = I^{-1}DI$.◀

Ejemplo 7.9. La matriz $A = \begin{bmatrix} 4 & -6 \\ 3 & -5 \end{bmatrix}$ es diagonalizable.

En efecto, la matriz invertible $P = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ y la matriz diagonal $D = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$ satisfacen:

$$P^{-1}AP = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 & -6 \\ 3 & -5 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} = D.$$

Puede comprobarse que la matriz invertible $P_1 = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ y la matriz diagonal $D_1 = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}$ también diagonzalizan a A.

Se verá luego que las únicas matrices diagonales semejantes a A son estas dos: D y D_1.◀

Los autovalores de las matrices diagonales D y D_1 del ejemplo anterior, son -2 y 1. Puede comprobarse que éstos son también los autovalores de la matriz A. El siguiente teorema permite generalizar este resultado.

Teorema 7.5. Las matrices semejantes tienen el mismo polinomio característico, y por lo tanto los mismos autovalores.

Demostración. Sean A y B matrices semejantes es decir $B = P^{-1}AP$, para alguna matriz invertible P. De acuerdo a propiedades de los determinantes:

$$\begin{aligned} p_B(\lambda) = \det(\lambda I - B) &= \det(\lambda I - P^{-1}AP) = \det(P^{-1}(\lambda I)P - P^{-1}AP) \\ &= \det(P^{-1}(\lambda I - A)P) = \det(P^{-1})\det(\lambda I - A)\det(P) \\ &= \tfrac{1}{\det(P)}\det(\lambda I - A)\det(P) \\ &= \det(\lambda I - A) = p_A(\lambda). \end{aligned}$$

Se ha probado entonces que las matrices semejantes tienen el mismo polinomio característico, por lo tanto tienen los mismos autovalores puesto que éstos son raíces del polinomio característico. □

Si D es una matriz diagonal semejante a una matriz A, ambas matrices tienen el mismo polinomio característico, $p(\lambda)$. Por ser D diagonal, las raíces de $p(\lambda)$ son los elementos de su diagonal y como D debe ser una matriz real, todas estas raíces deben ser números reales.

Luego, si una matriz A de $n \times n$ es diagonalizable, todas las raíces de su polinomio característico deben ser reales; esto es, A debe tener n autovalores (no necesariamente todos distintos) y toda matriz D diagonal semejante a A debe tener en su diagonal a dichos autovalores.

Ejemplo 7.10. Las siguientes matrices no son diagonalizables:

a) $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. El polinomio característico de A, $p_A(\lambda) = \lambda^2 + 1$, no tiene raíces reales (A no tiene autovalores). Esto significa que no existe ninguna matriz (real) diagonal semejante a A.

b) $A = \begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}$, $a \in \mathbb{R}$. El polinomio característico de A es $p_A(\lambda) = (\lambda - a)^2$. Luego, a es una raíz real doble (la multiplicidad algebraica de a es 2) por lo que A tiene dos autovalores (contando multiplicidades).

Si A fuese diagonalizable, la única matriz D semejante a A debe ser $D = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$. Pero puede probarse que no existe una matriz P invertible tal que $D = P^{-1}AP$. Por lo tanto A no es diagonalizable.◀

Actividad 7.3. Pruebe la afirmación de la parte b) del ejemplo anterior. Sugerencia: pruebe que no existe una matriz invertible P tal que $AP = PD$.

En los Ejemplos 7.9 y 7.10 vimos que:

- no todas las matrices cuadradas pueden diagonalizarse.
- las matrices P y D que diagonalizan a A no son únicas.

En lo que sigue se verá cuándo es diagonalizable una matriz A y cómo determinar las matrices P y D que la diagonalizan.

Teorema 7.6. Sean A, P y D matrices cuadradas de orden n donde D es una matriz diagonal. Entonces:

P es invertible y $P^{-1}AP = D$ si y solo si las columnas de P son autovectores linealmente independientes de A y la matriz D tiene en su diagonal los autovalores correspondientes.

Demostración. Sean $P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & ... & \mathbf{v}_n \end{bmatrix}$ y $D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{e}_1 & \lambda_2\mathbf{e}_2 & ... & \lambda_n\mathbf{e}_n \end{bmatrix}$.

Supongamos que $P^{-1}AP = D$ entonces (premultiplicando por P a ambos miembros), $AP = PD$.

Describiendo los productos AP y PD en términos de sus columnas:

$$AP = A\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & ... & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} A\mathbf{v}_1 & A\mathbf{v}_2 & ... & A\mathbf{v}_n \end{bmatrix}, \quad \text{y}$$

$$PD = \begin{bmatrix} P(\lambda_1\mathbf{e}_1) & P(\lambda_2\mathbf{e}_2) & ... & P(\lambda_n\mathbf{e}_n) \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & ... & \lambda_n\mathbf{v}_n \end{bmatrix}.$$

Dado que $AP = PD$ entonces sus correspondientes columnas son iguales, así:

$$A\mathbf{v}_j = \lambda_j\mathbf{v}_j, \quad j = 1, 2, ..., n.$$

Las igualdades anteriores dicen que si $\mathbf{v}_j \neq \mathbf{0}$, entonces $\mathbf{v}_j$ es un autovector de A y λ_j es su correspondiente autovalor. Los vectores $\mathbf{v}_j$ son las columnas de la matriz invertible P, por lo tanto son todos distintos de cero, y más aún, (Teorema 4.25) son linealmente independientes. Luego, las columnas de P, $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_n$ forman un conjunto de n autovectores linealmente independientes de A, y los elementos λ_1, λ_2,..., λ_n de la diagonal de D son sus correspondientes autovalores.

Recíprocamente, supongamos que las columnas de la matriz P, $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_n$, son autovectores de A (linealmente independientes) y los elementos en la diagonal de la matriz D, λ_1, λ_2,..., λ_n son sus correspondientes autovalores. Luego,

$$A\mathbf{v}_j = \lambda_j \mathbf{v}_j, \quad j = 1, 2, ..., n.$$

Como $AP = \begin{bmatrix} A\mathbf{v}_1 & A\mathbf{v}_2 & ... & A\mathbf{v}_n \end{bmatrix}$ y $PD = \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & ... & \lambda_n\mathbf{v}_n \end{bmatrix}$, entonces las ecuaciones anteriores muestran que las columnas de AP y PD son respectivamente iguales, lo que prueba que $AP = PD$. Por hipótesis, las columnas de P son linealmente independientes y por lo tanto P es invertible. Premultiplicando por P^{-1} a ambos miembros de $AP = PD$ se obtiene

$$P^{-1}AP = D,$$

que es lo que se quería probar. □

Actividad 7.4. Compruebe que para las matrices A, P y D del Ejemplo 7.9, las columnas de la matriz P son autovectores linealmente independientes de A y la matriz D tiene en su diagonal los correspondientes autovalores.

Del Teorema 7.6 se desprende el siguiente corolario que da una condición necesaria y suficiente para que una matriz sea diagonalizable.

Corolario 7.3. Una matriz cuadrada A de orden n es diagonalizable si y solo si A tiene n autovectores linealmente independientes.

Demostración. Si A es diagonalizable entonces existe una matriz P invertible y una matriz D diagonal tal que $P^{-1}AP = D$. Luego, por el teorema anterior, las columnas de P son n autovectores linealmente independientes de A.

Recíprocamente, si A tiene n autovectores linealmente independientes, entonces A es diagonalizable ya que por el teorema anterior, una matriz P cuyas columnas son dichos autovectores es invertible y satisface que $P^{-1}AP = D$ siendo D la matriz diagonal de sus correspondientes autovalores. □

Corolario 7.4. Una matriz cuadrada A de orden n es diagonalizable si y solo si existe una base de $\mathbb{R}^n$ formado por autovectores de A.

Demostración. Es una consecuencia inmediata del corolario anterior ya que los autovectores de A son vectores de $\mathbb{R}^n$, y n vectores de $\mathbb{R}^n$ linealmente independientes forman una base de $\mathbb{R}^n$. □

Ejemplo 7.11. En la parte b) del Ejemplo 7.10 se vió que la matriz

$$A = \begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}, \quad a \in \mathbb{R},$$

no es diagonalizable demostrando que no existe una matriz invertible P tal que $P^{-1}AP$ es una matriz diagonal.

Aplicando el Corolario 7.3 se verá ahora que A no es diagonalizable mostrando que no es posible encontrar dos autovectores de A que sean linealmente independientes.

El único autovalor de A es $\lambda = a$ (de multiplicidad algebraica 2) por lo tanto los únicos autovectores de A son las soluciones no triviales del sistema homogéneo $(aI - A)\,\mathbf{x} = \mathbf{0}$, es decir $N(aI - A)$. Resolviendo el sistema,

$$\begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

se obtiene

$$\begin{cases} x_1 = t \\ x_2 = 0 \end{cases}, \quad t \in \mathbb{R}.$$

Los autovectores de A son todos los vectores $\begin{bmatrix} t \\ 0 \end{bmatrix}$ con $t \neq 0$. Por lo tanto,

$$E_a = \left\{ \begin{bmatrix} t \\ 0 \end{bmatrix} : t \in \mathbb{R} \right\} = \text{gen} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}.$$

Así, $m_g(a) = \nu(aI - A) = 1$ lo que prueba que el número máximo de autovectores linealmente independientes asociados a a es 1. Así, no es posible encontrar dos autovectores de A linealmente independientes y A no es diagonalizable.◄

En lo que sigue se darán una serie de criterios útiles de diagonalización. Por ejemplo, una matriz cuadrada de orden n que tiene n autovalores distintos es diagonalizable. Para probar este resultado, se requiere el siguiente lema que establece que autovectores correspondientes a autovalores distintos son linealmente independientes.

Lema 7.1. Sea A una matriz de $n \times n$. Si $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_k$ son autovectores de A asociados a autovalores distintos λ_1, λ_2,..., λ_k, entonces $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_k$ son linealmente independientes.

Demostración. Como los vectores $\mathbf{v}_j$ $(j = 1, 2, ..., k)$ son autovectores, éstos son distintos de cero. En particular, $\mathbf{v}_1 \neq \mathbf{0}$ y por lo tanto $\{\mathbf{v}_1\}$ es linealmente independiente. Consideremos el mayor entero m tal que $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m\}$ es linealmente independiente. Claramente $1 \leq m \leq k$, si se demuestra que $m = k$ se habrá probado el teorema. Así que supondremos $m < k$ y llegaremos a una contradicción.

Si $m < k$, el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m, \mathbf{v}_{m+1}\}$ es linealmente dependiente y dado que $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_m$ son linealmente independientes, entonces $\mathbf{v}_{m+1}$ debe ser combinación lineal de éstos (Lema 4.1). Así, existen escalares c_1, c_2,..., c_m tales que

$$\mathbf{v}_{m+1} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_m\mathbf{v}_m. \tag{7.2}$$

Premultiplicando por la matriz A:

$$A\mathbf{v}_{m+1} = c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 + ... + c_mA\mathbf{v}_m,$$

y como $A\mathbf{v}_j = \lambda_j\mathbf{v}_j$ $(j = 1, 2, ..., k)$, entonces:

$$\lambda_{m+1}\mathbf{v}_{m+1} = c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + ... + c_m\lambda_m\mathbf{v}_m,$$

Reemplazando $\mathbf{v}_{m+1}$ en la última igualdad por la expresión dada en (7.2) y operando, se obtiene:

$$c_1\left(\lambda_1 - \lambda_{m+1}\right)\mathbf{v}_1 + c_2\left(\lambda_2 - \lambda_{m+1}\right)\mathbf{v}_2 + ... + c_m\left(\lambda_m - \lambda_{m+1}\right)\mathbf{v}_m = \mathbf{0}.$$

Dado que $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_m$ son linealmente independientes, entonces

$$c_j\,(\lambda_j - \lambda_{m+1}) = 0, \quad \text{para todo } j = 1, 2, ..., m,$$

y como todos los λ_j $(j = 1, 2, ..., k)$ son distintos, entonces $\lambda_j - \lambda_{m+1} \neq 0$ para todo $j = 1, 2, ..., m$. Así. necesariamente $c_1 = c_2 = ... = c_m = 0$ y, de acuerdo a (7.2) el vector $\mathbf{v}_{m+1}$ es el vector cero. Sin embargo, $\mathbf{v}_{m+1}$ es un autovector y por lo tanto es distinto de cero. Se obtuvo así una contradicción por suponer $m < k$, luego $m = k$ y entonces $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ es un conjunto linealmente independiente. □

Teorema 7.7. Si una matriz A cuadrada de orden n tiene n autovalores distintos entonces A es diagonalizable.

Demostración. Tomando para cada autovalor un autovector se obtienen n autovectores de A que corresponden a autovalores distintos. De acuerdo al lema anterior, estos n autovectores forman un conjunto linealmente independiente y por lo tanto A es una matriz diagonalizable. □

Ejemplo 7.12. Las siguientes dos matrices son diagonalizables:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 5 & 1 \\ 0 & -3 & 9 \\ 0 & 0 & 1 \end{bmatrix}.$$

La matriz A (dada en el Ejemplo 7.6 de la sección anterior) tiene como autovalores $\lambda_1 = 4$, $\lambda_2 = 2+\sqrt{3}$ y $\lambda_3 = 2-\sqrt{3}$. Ya que A es de orden 3 y tiene 3 autovalores distintos, A es diagonalizable.

La matriz B es triangular, sus autovalores son los elementos de la diagonal $\lambda_1 = 2$, $\lambda_2 = -3$ y $\lambda_3 = 1$. Como B es de 3×3 y tiene 3 autovalores distintos, entonces B es diagonalizable.◀

El teorema anterior establece una condición suficiente para que una matriz sea diagonalizable y es muy útil ya que identifica una clase amplia de matrices diagonalizables. Pero no es una condición necesaria que los autovalores sean distintos para que la matriz sea diagonalizable. Un caso extremo es la matriz identidad I_n que tiene solo el autovalor $\lambda = 1$, y claramente I_n es diagonalizable. Para que una matriz de $n \times n$ sea diagonalizable *debe tener n autovectores linealmente independientes*, que no significa tener n autovalores distintos.

El Lema 7.1 es un caso particular de uno más general: si λ_1, λ_2,..., λ_k son autovalores distintos y se elige un conjunto linealmente independiente de cada uno de los correspondientes autoespacios, el conjunto formado por todos estos autovectores es linealmente independiente. Por ejemplo, si se toman dos autovectores linealmente independientes de un autoespacio y tres autovectores linealmente independientes de otro autoespacio, el conjunto formado por los cinco vectores es linealmente independiente. El siguiente teorema prueba esta afirmación pero se omitirá en este momento su demostración.

Teorema 7.8. Sea A una matriz de $n \times n$ y λ_1, λ_2,..., λ_k autovalores distintos de A. Si B_1, B_2,..., B_k son subconjuntos linealmente independientes de los autoespacios correspondientes, entonces el conjunto $B = B_1 \cup B_2 \cup ... \cup B_k$ es linealmente independiente.

En base a este teorema, se tienen los dos siguientes criterios de diagonalización:

Corolario 7.5. Sea A una matriz de $n \times n$ y λ_1, λ_2,..., λ_k *todos* los autovalores distintos de A. Si B_1, B_2,..., B_k son bases para los autoespacios correspondientes, entonces:

La matriz A es diagonalizable si y solo si el conjunto $B = B_1 \cup B_2 \cup ... \cup B_k$ tiene exactamente n elementos.

Corolario 7.6. Sea A una matriz (real) de $n \times n$ que tiene n autovalores. Si λ_1, λ_2,..., λ_k son los autovalores distintos de A entonces:

La matriz A es diagonalizable si y solo si para cada λ_i la multiplicidades geométricas y algebraicas son iguales. Esto es, $m_g(\lambda_i) = m_a(\lambda_i)$ para cada $i = 1, 2, ..., k$.

Proceso de diagonalización

Los criterior anteriores nos permiten usar el proceso de diagonalización siguiente. Esto es, para una matriz dada A, encontrar un par de matrices P invertible y D diagonal tal que $P^{-1}AP = D$, o mostrar que A no es diagonalizable.

Sea A una matriz de $n \times n$.

1. Obtener el polinomio característico de A. Si alguna raíz es compleja, el proceso termina puesto que A no es diagonalizable. Caso contrario las n raíces (contando multiplicidades) son autovalores de A.

2. Si los n autovalores son distintos, A es diagonalizable y se prosigue con el paso 4.

3. Si hay autovalores repetidos, y λ_1, λ_2,..., λ_k son todos los autovalores distintos, obtener la dimensión de cada autoespacio $m_g(\lambda_i)$ y compararlo con $m_a(\lambda_i)$.

   Si para algún $i = 1, 2, ..., k$, $m_g(\lambda_i) \neq m_a(\lambda_i)$, el proceso se termina puesto que A no es diagonalizable.

   Si en cambio $m_g(\lambda_i) = m_a(\lambda_i)$ para todo $i = 1, 2, .., k$, entonces A es diagonalizable y se continúa con el paso 4.

4. Para cada autoespacio, $E_{\lambda_i} = N(\lambda_i I - A)$, obtener una base B_i. La unión de dichas bases forman un conjunto B que consta de n autovectores linealmente independientes de A.
   Una matriz P que tenga como columnas los n vectores de B y una matriz diagonal D que tenga en la posición (j, j) el autovalor de A correspondiente a la columna j de P, diagonalizan a A.

Observación 7.3. Claramente no hay un único par de matrices P y D que diagonalizan a A. Si se intercambian dos columnas de la matriz P y sus correspondientes autovalores en la matriz D, se obtiene un nuevo par de matrices P' y D' que también diagonalizan a A. Eventualmente $D' = D$ (si los autovectores que se intercambian corresponden al mismo autovalor). Véase Ejemplo 7.9.

Ejemplo 7.13. Se aplicará el proceso de diagonalización para la matriz $A = \begin{bmatrix} 2 & 3 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$. El polinomio característico de la matriz A es

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 2 & -3 & 0 \\ -4 & \lambda - 3 & 0 \\ 0 & 0 & \lambda - 6 \end{vmatrix} = (\lambda - 6) \begin{vmatrix} \lambda - 2 & -3 \\ -4 & \lambda - 3 \end{vmatrix} = (\lambda - 6)(\lambda^2 - 5\lambda - 6) = (\lambda - 6)^2(\lambda + 1).$$

Puede observarse que se calculó el determinante desarrollándolo según el último renglón, esto ayudó a la factorización del polinomio característico y por lo tanto a la búsqueda de sus raíces.

Las raíces del polinomio característico son: 6 (doble) y -1. Así, la matriz A tiene solo dos autovalores distintos $\lambda_1 = 6$ y $\lambda_2 = -1$ con $m_a(6) = 2$ y $m_a(-1) = 1$. Deben hallarse entonces las dimensiones de cada autoespacio E_6 y E_{-1}.

Para $\lambda_1 = 6$, es $E_6 = N(6I - A)$.

$$6I - A = \begin{bmatrix} 4 & -3 & 0 \\ -4 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \text{una forma escalonada es} \quad \begin{bmatrix} \boxed{4} & -3 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Puede observarse que $\rho(6I - A) = 1$ y luego $\nu(6I - A) = 2$. Así,

$$m_g(6) = 2 = m_a(6).$$

Para $\lambda_2 = -1$, es $E_{-1} = N(-I - A)$.

$$-I - A = \begin{bmatrix} -3 & -3 & 0 \\ -4 & -4 & 0 \\ 0 & 0 & -7 \end{bmatrix}, \quad \text{una forma escalonada es} \quad \begin{bmatrix} \boxed{-3} & -3 & 0 \\ 0 & 0 & \boxed{-7} \\ 0 & 0 & 0 \end{bmatrix}.$$

Puede observarse que $\rho(-I - A) = 2$ y luego $\nu(-I - A) = 1$. Así,

$$m_g(-1) = 1 = m_a(-1).$$

Por lo tanto, A es una matriz diagonalizable (que no tiene 3 autovalores distintos). Es claro que toda matriz diagonal semejante a A tiene en su diagonal a 6, 6 y -1.

Se hallarán un par de matrices P y D que tales que $P^{-1}AP = D$. Para ello es necesario encontrar una base para cada uno de los autoespacios.

Observando la forma escalonada de $6I - A$ obtenida, se encuentra que:

$$E_6 = N(6I - A) = \left\{ \begin{bmatrix} \frac{3}{4}s \\ s \\ t \end{bmatrix} : s, t \in \mathbb{R} \right\} = \text{gen} \left\{ \begin{bmatrix} \frac{3}{4} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\},$$

y dado que los vectores $\mathbf{v}_1 = \begin{bmatrix} \frac{3}{4} \\ 1 \\ 0 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ son autovectores linealmente independientes correspondientes a $\lambda_1 = 6$, éstos forman una base para E_6.

Por otro lado, observando la forma escalonada de $-I - A$, rápidamente se encuentra:

$$E_{-1} = N(-I - A) = \left\{ \begin{bmatrix} -t \\ t \\ 0 \end{bmatrix} : t \in \mathbb{R} \right\} = \text{gen} \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

El vector $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ es un autovector asociado a $\lambda_2 = -1$ y $\{\mathbf{v}_3\}$ es una base para E_{-1}.

Luego, las matrices

$$P = \begin{bmatrix} \frac{3}{4} & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

diagonalizan a la matriz A.

Solo como comprobación el lector puede verificar que $P^{-1}AP = D$, que es equivalente a probar que $AP = DP$.

Como se notó en la observación anterior, no existe preferencia respecto al orden de las columnas para la matriz P; solo debe tenerse en cuenta que un cambio en el orden de las columnas de P puede provocar un cambio en la matriz D. Por ejemplo, si se hubiese tomado

$$P = \begin{bmatrix} \frac{3}{4} & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

entonces la matriz diagonal correspondiente sería

$$D = \begin{bmatrix} 6 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 6 \end{bmatrix},$$

que es otro par de matrices que diagonalizan a A.◀

Ejemplo 7.14. Sea la matriz $A = \begin{bmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{bmatrix}$. El polinomio caracterísitco de A es

$$p_A(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & 1 & -4 \\ -3 & \lambda - 2 & 1 \\ -2 & -1 & \lambda + 1 \end{vmatrix} = \lambda^3 - 2\lambda^2 - 5\lambda + 6 = (\lambda - 1)(\lambda + 2)(\lambda - 3).$$

Como A tiene tres autovalores distintos $\lambda_1 = 1$, $\lambda_2 = -2$ y $\lambda_3 = 3$, A es diagonalizable.

Para $\lambda_1 = 1$, se tiene la matriz

$$I - A = \begin{bmatrix} 0 & 1 & -4 \\ -3 & -1 & 1 \\ -2 & -1 & 2 \end{bmatrix},$$

una forma escalonada equivalente es

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 4 \\ 0 & 0 & 0 \end{bmatrix},$$

y así el autoespacio $E_1 = N(I - A) = \text{gen}\left\{ \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix} \right\}$. El vector $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}$ es un autovector correspondiente a $\lambda_1 = 1$, y $B_1 = \{\mathbf{v}_1\}$ es una base de E_1.

Para $\lambda_2 = -2$, se tiene la matriz

$$-2I - A = \begin{bmatrix} -3 & 1 & -4 \\ -3 & -4 & 1 \\ -2 & -1 & -1 \end{bmatrix},$$

una forma escalonada es

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

y así el autoespacio $E_{-2} = N(-2I - A) = \text{gen}\left\{\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}\right\}$. El vector $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$ es un autovector correspondiente a $\lambda_2 = -2$, y $B_2 = \{\mathbf{v}_2\}$ es una base de E_{-2}.

Para $\lambda_3 = 3$, se tiene la matriz

$$3I - A = \begin{bmatrix} 2 & 1 & -4 \\ -3 & 1 & 1 \\ -2 & -1 & 4 \end{bmatrix},$$

una forma escalonada es

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

y así el autoespacio $E_3 = N(3I - A) = \text{gen}\left\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\right\}$. El vector $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ es un autovector correspondiente a $\lambda_3 = 3$, y $B_3 = \{\mathbf{v}_3\}$ es una base de E_3.

Luego, las matrices

$$P = \begin{bmatrix} -1 & 1 & 1 \\ 4 & -1 & 2 \\ 1 & -1 & 1 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

diagonalizan a la matriz A.◀

Ejemplo 7.15. Sea $A = \begin{bmatrix} 5 & 1 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$. El polinomio característico de A es

$$p_A(\lambda) = (\lambda - 5)^3(\lambda - 2),$$

cuyas raíces son 5 (triple) y 2. La matriz A tiene solo dos autovalores distintos $\lambda_1 = 5$ y $\lambda_2 = 2$ con $m_a(5) = 3$ y $m_a(2) = 1$.

Para $\lambda_1 = 5$, es $E_5 = N(5I - A)$. Entonces

$$5I - A = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

y claramente tiene rango 3 y por lo tanto la tanto su nulidad es 1. Así, $m_g(5) = 1 \neq m_a(5)$. Luego, la matriz A no es diagonalizable.◀

Potencia de matrices diagonalizables

El cálculo de las potencias, A^k, de una matriz A puede ser un trabajo muy tedioso. Pero esta cálculo se facilita si A es diagonalizable y conocemos dos matrices P y D que la diagonalizan.

En efecto, $D = P^{-1}AP$ y por lo tanto $A = PDP^{-1}$. Así,

$$A^2 = (PDP^{-1})(PDP^{-1}) = P(DID)P^{-1} = PD^2P^{-1}.$$

Generalizando, para todo $k \in \mathbb{Z}^+$,

$$A^k = PD^kP^{-1},$$

y el cálculo de D^k consiste solo en elevar a la potencia k los elementos en la diagonal de D.

Ejercicios 7.2

1. En cada ítem, se dan n autovectores linealmente independientes y sus correspondientes autovalores de una matriz A de $n \times n$.

   a) Encuentre una matriz P y una matriz diagonal D, que diagonalicen a A.

   b) Halle la matriz A.

   $\boxed{i.}$ $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\lambda_1 = 2$ $\quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\lambda_2 = -3$.

   ii. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\lambda_1 = 1$ $\quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\lambda_2 = 2$.

   $\boxed{iii.}$ $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\lambda_1 = 3$ $\quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\lambda_2 = -3$ $\quad \mathbf{u}_3 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$, $\lambda_3 = 3$.

$\boxed{2.}$ Use el criterio del Corolario 7.6 para probar que las siguientes matrices no son diagonalizables.

a) $\begin{bmatrix} 4 & 0 \\ 1 & 4 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$

c) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$

d) $\begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix}$

e) $\begin{bmatrix} 7 & 1 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 0 & 0 & 7 & 1 \\ 0 & 0 & 0 & 7 \end{bmatrix}$

f) $\begin{bmatrix} 5 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}$

$\boxed{3.}$ En cada ítem pruebe que A es diagonalizable y encuentre todas las matrices diagonales semejantes a A.

a) $A = \begin{bmatrix} 2 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

b) $A = \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}$

4. ¿Para qué valores de $a \in \mathbb{R}$ son diagonalizables las siguientes matrices?

a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & a \end{bmatrix}$ *b*) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & a & 0 \end{bmatrix}$ *c*) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ a & 0 & 0 \end{bmatrix}$

5. En cada ítem se dan una matriz P invertible y una matriz D diagonal. Sin calcular $A = PDP^{-1}$, encuentre otras matrices P_1 y D_1 tales que $A = P_1 D_1 P_1^{-1}$.

a) $P = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

b) $P = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$, $D = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}$

c) $P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

d) $P = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 0 \\ -1 & 4 & 0 \end{bmatrix}$, $D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

6. Para cada matriz del Ejercicio 6 de la sección anterior, indique si es diagonalizable. En caso afirmativo encuentre un par de matrices que la diagonalizan.

7. Para cada matriz A halle, si es posible, una matriz P no singular y una matriz D diagonal tal que $AP = PD$.

a) $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 2 & 1 & 2 \end{bmatrix}$

b) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

c) $A = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$

d) $A = \begin{bmatrix} 0 & -2 & 1 & 0 \\ 1 & 3 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

8. Encuentre A^9 para cada matriz A diagonalizable del ejercicio anterior.

9. Indique si las siguientes proposiciones son verdaderas o falsas. Justifique:

a) Si A es una matriz de 3×3 con dos autovalores distintos: 5 y 7, y $\rho(5I - A) = \rho(7I - A) = 1$, entonces A es diagonalizable.

b) Sea A una matriz de 3×3 con dos autovalores distintos, λ_1 y λ_2. Si $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ son autovectores correspondientes a λ_1 y $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ es un autovector correspondiente a λ_2, entonces A es diagonalizable.

c) Si A es una matriz de 3×3 y $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ es semejante a A, entonces el rango de A es 2.

d) Sea B una matriz semejante a A y $P^{-1}AP = B$. Si $\mathbf{x}$ es un autovector de A correspondiente al autovalor λ, entonces $P^{-1}\mathbf{x}$ es un autovector de B correspondiente a λ.

e) Si A es una matriz diagonalizable entonces A^t es diagonalizable.

f) Si A es una matriz invertible y diagonalizable, entonces A^{-1} es diagonalizable.

Respuestas a los Ejercicios

Capítulo 1

Sección 1.1

1. *a*) es lineal, no homogénea, coeficientes: 1, 3 y -1, término constante: 5, variable delantera: x, variables libres y y z, conjunto solución: $\{(5-3t+s, t, s) : t, s \in \mathbb{R}\}$.

   b) no lineal.

   c) es lineal, homogénea, coeficientes: 1, -3, -5 y 1, término constante: 0, variable delantera: x_1, variables libres x_2, x_3 y x_4, conjunto solución: $\{(3t+5s-r, t, s, r) : t, s, r \in \mathbb{R}\}$.

   d) es lineal, no homogénea, coeficientes: 0, 0, y 0, término constante: 9, variable delantera: no posee, variables libres: no posee, conjunto solución: $\emptyset$.

2. *a*) no tiene solución: $a = 4$, tiene infinitas soluciones: $a = -4$, tiene única solución $a \neq \pm 4$.

   b) infinitas soluciones: para todo $a \in \mathbb{R}$, no existe a tal que posee única solución o no posea solución.

   c) no existe solución: $a = -1$, infinitas soluciones: $a \neq -1$, no existe a tal que posee única solución.

   d) infinitas soluciones: para todo $a \in \mathbb{R}$, no existe a tal que posee única solución o no posea solución.

Sección 1.2

1. $$\begin{cases} -3x_1 & +x_2 & +x_3 & -x_4 & = & 3 \\ 2x_1 & & & -x_4 & = & 9 \\ & 2x_2 & +x_3 & -x_4 & = & -3 \\ 2x_1 & -x_2 & -x_3 & -x_4 & = & -6 \end{cases}$$

2. $a = \frac{3}{2}$, $b = \frac{5}{2}$.

5. 20 camisetas lisas y 25 estampadas.

6. *a*) $x = -57$, $y = -24$, $z = -8$.

   b) $x_1 = \frac{29}{2} - \frac{7}{2}t$, $x_2 = \frac{9}{2} + \frac{1}{2}t$, $x_3 = t$, $x_4 = 0$, $t \in \mathbb{R}$.

   c) $x = y = z = 0$.

   d) $x_1 = -3+s+2t$, $x_2 = s$, $x_3 = 8-t$, $x_4 = t$, $t, s \in \mathbb{R}$.

8. - $E_1 \leftrightarrow E_2$
   - $E_2 - 2E_1 \to E_2$, $E_3 + E_1 \to E_3$
   - $E_3 + \frac{2}{3}E_2 \to E_3$
   - $3E_4 \to E_4$

   Solución: $x = y = z = 2$.

Sección 1.3

1. matriz de coeficientes: $\begin{bmatrix} 1 & 1 & 1 & -1 \\ 2 & 1 & -1 & 0 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 2 & 0 & -1 & 1 \end{bmatrix}$, matriz aumentada: $\begin{bmatrix} 1 & 1 & 1 & -1 & -1 \\ 2 & 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 4 \\ 2 & 0 & -1 & 1 & 0 \end{bmatrix}$.

3. A, B, D, E, F, G y H están en forma escalonada; D, F, G y H están en forma escalonada reducida. Columnas pivote de A: 1era, 2da y 4ta; columnas pivote de B: 1era y 2da; columnas pivote de D: 1era, 3ra y 5ta; columnas pivote de E: 1era; columnas pivote de F: 1era; columnas pivote de G: 1era, 2da y 3era; columnas pivote de H: 1era, 2da y 6ta.

5. *a*) $A \sim B$.

   b) A y B no son equivalentes.

   c) $A \sim B$.

7. *a*) $\begin{bmatrix} 1 & 0 & 0 & \frac{3}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & -2 \end{bmatrix}$ es la forma escalonada reducida de A; columnas pivote de A: 1era, 2da y 3era.

b) $\begin{bmatrix} 1 & 0 & -\frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & -\frac{19}{3} & -\frac{2}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}$ es la forma escalonada reducida de A; columnas pivote de A: 1era y 2da.

c) $\begin{bmatrix} 1 & 3 & 2 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ es la forma escalonada reducida de A; columnas pivote de A: 1era.

8. No. No. Es 15.

Sección 1.4

1. *a*) $x_1 = 1$, $x_2 = 2$, $x_3 = -2$, $x_4 = -4$.
   b) $x = t - 2$, $y = 5 + z - t$, $w = -4 - z + t$, $z \in \mathbb{R}$, $t \in \mathbb{R}$.
   c) inconsistente.
   d) $x_1 = x_2 = x_3 = x_4 = 0$.
   e) $x_1 = \frac{8}{3} + 2t$, $x_2 = \frac{7}{3} + 2t$, $x_3 = 3 + 3t$, $t \in \mathbb{R}$.
   f) $x = -15 - \frac{14}{3}t + \frac{14}{3}s$, $y = -6 - \frac{5}{3}t + \frac{11}{3}s$, $z = t$, $w = s$, $s, t \in \mathbb{R}$.
2. 32 de AA1, 25 de AA2, 17 de AA3 y 20 de B1.
3. Sí, es posible.
4. *a*) $\begin{cases} x_1 + x_2 = 800 \\ x_2 - x_3 + x_4 = 300 \\ x_4 + x_5 = 500 \\ x_1 + x_5 = 600 \\ x_3 = 400 \end{cases}$
   Solución: $x_1 = 600 - x_5$, $x_2 = 200 + x_5$, $x_3 = 400$, $x_4 = 500 - x_5$, $0 \le x_5 \le 500$.
   b) $x_1 = 20$, $x_2 = 20 + x_4$, $x_3 = -80 + x_4$, $x_4 \ge 80$.
5. $2a - b + c = 0$.
6. *a*) no existen μ y λ
   b) $\mu = 2$, $\lambda \in \mathbb{R}$.
   c) $\mu \neq 2$, $\lambda \in \mathbb{R}$.
7. *a*) $\alpha \neq 2$ y $\alpha \neq 3$
   b) $\alpha = 2$. Solución: $x = \frac{5}{4} + \frac{5}{4}t$, $y = \frac{1}{4} - \frac{3}{4}t$, $z = t$, $t \in \mathbb{R}$.
   c) $\alpha = 3$.
8. Inconsistente: $b = 2$ y $a \neq 0$; Infinitas soluciones: $b = 2$ y $a = 0$; Única solución: $b \neq 2$ y $a \in \mathbb{R}$.
9. *a*) 4 y 6 *b*) 0 y -2
10. *a*) Sistema 1: $x = 2$, $y = 1$; Sistema 2: inconsistente; Sistema 3: $x = -1$, $y = 1$.
    b) Sistema 1: $x = 1$, $y = -\frac{1}{5}$, $z = \frac{2}{5}$; Sistema 2: $x = 1$, $y = \frac{1}{5}$, $z = -\frac{2}{5}$; Sistema 3: $x = -2$, $y = \frac{3}{5}$, $z = \frac{4}{5}$.
    c) Sistema 1: $x_1 = \frac{1}{3} + \frac{1}{3}t$, $x_2 = \frac{2}{3}t - \frac{1}{3}$, $x_3 = t$, $t \in \mathbb{R}$; Sistema 2: inconsistente.

Capítulo 2

Sección 2.1

1. $x = -\frac{2}{5}$, $y = -\frac{6}{5}$, $z = -\frac{3}{5}$
2. *a*) $\begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$
   b) $(-4, 4, -2, 4)$
   c) $\begin{bmatrix} -3 \\ -3 \\ 3 \\ 0 \end{bmatrix}$
   d) $(2, -\frac{7}{2}, \frac{3}{2}, \frac{11}{2})$
   e) $\begin{bmatrix} -8 \\ -6 \\ 32 \end{bmatrix}$
   f) no es posible
   g) $(0, 0, -3, 3, -15)$
4. *a*) $\begin{bmatrix} 0 \\ 5/2 \\ -13/2 \\ 1 \end{bmatrix}$
   b) $\begin{bmatrix} -3 \\ -5 \\ 8 \\ -4 \end{bmatrix}$
   c) $\begin{bmatrix} 3/4 \\ 1/4 \\ -1/2 \\ 1 \end{bmatrix}$
   d) no es posible.
5. *a*) Sí lo es, $\mathbf{b} = 2\mathbf{a}_1 + 3\mathbf{a}_2 + 0\mathbf{a}_3$.
   b) No lo es.
   c) Sí lo es, $\mathbf{b} = \left(\frac{4}{3} + \frac{1}{3}t + \frac{1}{3}s\right)\mathbf{a}_1 + \left(\frac{4}{3} - \frac{5}{3}t - \frac{2}{3}s\right)\mathbf{a}_2 + t\mathbf{a}_3 + s\mathbf{a}_4$, $s, t \in \mathbb{R}$.
   d) No lo es.
6. $a = \pm 1$
7. *a*) $\left(\frac{17}{14}, -\frac{17}{7}, \frac{8}{7}\right)$
   b) $m_1 = \frac{21}{2} - \frac{15}{4}t$, $m_2 = \frac{3}{2} - \frac{1}{4}t$, $m_3 = t$ con $0 < t < \frac{14}{5}$.

Sección 2.2

1. $x = 1$, $y = \frac{1}{3}$, $z = -\frac{2}{3}$
2. *a*) $\begin{bmatrix} -3 & 5 \\ -2 & -2 \end{bmatrix}$
   b) $\begin{bmatrix} -3 & -6 & -27 \\ -9 & -6 & 15 \end{bmatrix}$
   c) no es posible
   d) $\begin{bmatrix} -10 & 12 & 4 \\ -10 & -4 & -2 \\ 2 & 2 & 0 \end{bmatrix}$

4. *a*) $\begin{bmatrix} 1 & \frac{9}{2} & 1 \\ -4 & 0 & -3 \end{bmatrix}$

   b) $\begin{bmatrix} \frac{1}{10} & \frac{8}{5} & \frac{2}{5} \\ -\frac{8}{5} & -\frac{3}{10} & -\frac{13}{10} \end{bmatrix}$

   c) $\begin{bmatrix} -1 & -4 & 2 \\ 4 & 3 & 7 \end{bmatrix}$

   d) $\begin{bmatrix} -\frac{1}{6} & \frac{5}{6} & \frac{1}{6} \\ -1 & -\frac{1}{2} & -1 \end{bmatrix}$

5. *a*) -15 *c*) 7

   b) 35 *d*) -14

6. $a = -6$

7. *a*) $\begin{bmatrix} -8 & 0 & -4 \\ 4 & 0 & 2 \\ 12 & 0 & 6 \end{bmatrix}$

   b) $[0]$

   c) $\begin{bmatrix} 2 & 2 & -8 \\ -12 & 0 & -6 \end{bmatrix}$

   d) no es posible

   e) $\begin{bmatrix} -6 & 10 \\ -3 & -1 \end{bmatrix}$

   f) $\begin{bmatrix} 3 & -7 & 2 \\ 4 & -3 & 2 \\ -2 & -9 & -2 \end{bmatrix}$

8. 6

9. $\begin{bmatrix} 1 \\ 5 \\ -10 \end{bmatrix}$

11. *a*) $\begin{bmatrix} 5 \\ 0 \\ 10 \end{bmatrix}$

    b) $\begin{bmatrix} 6 \\ -3 \\ -1 \\ -3 \\ 7 \end{bmatrix}$

12. *a*) $\begin{bmatrix} 2 & -4 \\ 1 & 1 \\ -1 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}$

    b) $\begin{bmatrix} 1 & -4 & -1 \\ 2 & 1 & 3 \\ -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$

13. $\left\{ \begin{array}{rcl} -2x + 2z & = & 4 \\ 3x + y + z & = & 1 \\ -x + y - z & = & 2 \end{array} \right.$; $x = -\frac{5}{6}$, $y = \frac{7}{3}$, $z = \frac{7}{6}$.

14. *a*) sí *b*) sí *c*) no

16. $\begin{bmatrix} 0 & \frac{9}{2} & \frac{9}{2} \\ \frac{3}{2} & \frac{9}{2} & 3 \\ 3 & \frac{15}{2} & \frac{3}{2} \end{bmatrix}$

17. A, $\mathbf{O}_4$ y C.

18. A, $\mathbf{O}_4$ y C.

19. *a*) $\begin{bmatrix} 56 & 40 & 0 \\ 80 & 56 & 0 \\ 16 & 24 & -8 \end{bmatrix}$

    b) $\begin{bmatrix} 1 & 1 & 1 \\ 20 & 15 & 4 \\ 5 & 4 & 0 \end{bmatrix}$

    c) $\begin{bmatrix} -41 & -42 & -22 \\ -40 & -33 & -28 \\ -26 & -40 & -3 \end{bmatrix}$

    d) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

21. No, no es cierto.

22. *a*) Falso *d*) Verdadero

    b) Falso

    c) Verdadero *e*) Falso

Sección 2.3

1. *a*) $A^{-1} = \begin{bmatrix} -\frac{1}{14} & \frac{3}{14} \\ \frac{2}{7} & \frac{1}{7} \end{bmatrix}$

   b) A no es invertible

   c) $A^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix}$

   d) A no es invertible

   e) $A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{8} & 0 \\ 0 & 0 & 3 \end{bmatrix}$

   f) $A^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{5}{6} & -1 & -\frac{5}{6} \\ 0 & 1 & -1 & 0 \\ -\frac{1}{9} & -\frac{11}{18} & 1 & \frac{5}{18} \\ \frac{2}{9} & \frac{13}{18} & -1 & -\frac{1}{18} \end{bmatrix}$

   g) A no es invertible

   h) $A^{-1} = \begin{bmatrix} 6 & -5 & 6 & -3 \\ 6 & -6 & 7 & -3 \\ -2 & 2 & -2 & 1 \\ 1 & -1 & 1 & 0 \end{bmatrix}$

2. $\begin{bmatrix} \frac{1}{5} & \frac{1}{5} \\ -\frac{3}{10} & \frac{3}{10} \end{bmatrix}$

3. $A^2 = \begin{bmatrix} 4 & 1 & -2 \\ 1 & 9 & 1 \\ 1 & 3 & 0 \end{bmatrix}$,

$A^3 = \begin{bmatrix} -7 & 1 & 4 \\ 7 & 28 & 1 \\ 1 & 9 & 1 \end{bmatrix}$,

$A^{-2} = \begin{bmatrix} -3 & -6 & 19 \\ 1 & 2 & -6 \\ -6 & -11 & 35 \end{bmatrix}$,

$A^{-3} = \begin{bmatrix} 19 & 35 & -111 \\ -6 & -11 & 35 \\ 35 & 64 & -203 \end{bmatrix}$

4. $a \neq 4$

5. *a*) no es posible calcular la inversa de la matriz de coeficientes

   b) sí es posible calcular la inversa de la matriz de coeficientes; $x = b - 2b^2 - b^3$, $y = b - b^2$, $z = b - b^2 - b^3$

7. *a*) $\begin{bmatrix} \frac{1}{2} & -\frac{7}{2} \\ 0 & 1 \end{bmatrix}$

   b) $\begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}$

   c) $\begin{bmatrix} -3 & -3 & -7 \\ 0 & 0 & 20 \\ 0 & 0 & -6 \end{bmatrix}$

   d) $\begin{bmatrix} -5 & 23 & 4 \\ 5 & -20 & -3 \\ -3 & 14 & 2 \end{bmatrix}$

   e) $\begin{bmatrix} 1/10 & -7/10 \\ 7/20 & 11/20 \end{bmatrix}$

11. *a*) Falso   *d*) Falso

    b) Verdadero

    c) Verdadero   *e*) Verdadero

Capítulo 3

Sección 3.1

1. *a*) $M_{11} = -6$, $C_{11} = -6$, $M_{12} = 4$, $C_{12} = -4$, $M_{13} = 8$, $C_{13} = 8$, $M_{21} = -1$, $C_{21} = 1$, $M_{22} = 5$, $C_{22} = 5$, $M_{23} = -3$, $C_{23} = 3$, $M_{31} = -16$, $C_{31} = -16$, $M_{32} = 2$, $C_{32} = -2$, $M_{33} = 4$, $C_{33} = 4$.

   b) 26.

2. *a*) 26   *g*) 9

   b) $-2\sqrt{2}$ *h*) -3

   c) 0   *i*) 96

   d) 24   *j*) -30

   e) -14 *k*) 120

   f) -1 *l*) -243

3. *a*) 2.   *b*) 2, 3, $\frac{2}{3}$.   *c*) 6, 5.

Sección 3.2

1. *a*) la columna 3 es una columna cero.

   b) los renglones 1 y 3 son iguales.

   c) la columna 4 es -2 por la columna 2.

2. *a*) 8   *b*) 10   *c*) -2

4. $|B| = 3$, $|C| = -3$, $|D| = 6$, $|E| = -48$, $|F| = -243$, $|G| = -\frac{16}{27}$, $|H| = -81$.

5. *a*) -2 *d*) -2

   b) 4   *e*) -4

   c) -54 *f*) -2

6. A, B y C son invertibles; D es singular.

7. *a*) $a \neq -4 \pm \sqrt{6}$ *b*) $a \neq -1 \pm \frac{\sqrt{33}}{3}$

8. *a*) B es invertible.

   b) $|2B^t| = -960$, $|ABC| = 360$, $|DA + DB| = 0$, $\left|-A^4\right| = 16$, $\left|-(3A)^2\right| = 26244$

9. $\lambda = 2$ ó $\lambda = -1$.

Sección 3.3

1. *a*) $\begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ \frac{1}{21} & -\frac{2}{21} \end{bmatrix}$

   b) no es invertible.

   c) $\begin{bmatrix} \frac{2}{5} & -\frac{2}{5} & -\frac{1}{5} \\ -\frac{1}{10} & \frac{1}{10} & \frac{3}{10} \\ \frac{7}{10} & \frac{3}{10} & -\frac{1}{10} \end{bmatrix}$

   d) $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{5}{2} & 1 \\ -\frac{1}{2} & \frac{3}{2} & 1 \end{bmatrix}$

2. *a*) Es posible; $x = \frac{27}{5}$, $y = \frac{8}{15}$.

   b) Es posible; $x_1 = \frac{5}{4}$, $x_2 = \frac{5}{2}$, $x_3 = \frac{11}{4}$.

   c) No es posible; inconsistente.

   d) Es posible; $x_1 = -4$, $x_2 = -2$, $x_3 = 1$.

3. $x = a\cos\theta + b\,\mathrm{sen}\,\theta$, $y = b\cos\theta - a\,\mathrm{sen}\,\theta$.

4. $a \neq 0$

Capítulo 4

Sección 4.1

3. *a*) es un espacio vectorial.
   b) no es un espacio vectorial.
   c) no es un espacio vectorial.
   d) no es un espacio vectorial.
   e) es un espacio vectorial.

Sección 4.2

1. *a*) es un subespacio.
   b) no es un subespacio.
   c) es un subespacio.
   d) es un subespacio.
   e) no es un subespacio.
   f) es un subespacio.
   g) no es un subespacio.
   h) es un subespacio.
   i) no es un subespacio.
   j) es un subespacio.
   k) es un subespacio.
   l) es un subespacio.
   m) es un subespacio.
   n) es un subespacio.

3. *a*) no es un subespacio.
   b) es un subespacio.
   c) es un subespacio.
   d) no es un subespacio.
   e) es un subespacio.

Sección 4.3

1. *a*) Sí, $\mathbf{u} = \left(2+\frac{29}{2}t\right)\mathbf{u}_1 + \left(-2-\frac{79}{2}t\right)\mathbf{u}_2 + \left(3+\frac{51}{2}t\right)\mathbf{u}_3 + \left(1-\frac{33}{2}t\right)\mathbf{u}_4 + t\mathbf{u}_5,\ t\in\mathbb{R}$.
   b) Sí, $\mathbf{u} = \left(-\frac{2}{7}+2t\right)\mathbf{u}_1 + \left(-\frac{10}{7}-t\right)\mathbf{u}_2 + (4-2t)\,\mathbf{u}_3 - \frac{9}{7}\mathbf{u}_4 + t\mathbf{u}_5,\ t\in\mathbb{R}$.
   c) Sí, $\mathbf{u} = 2\mathbf{u}_1 + (2+t)\,\mathbf{u}_2 + (-3-2t)\,\mathbf{u}_3 + (-2+t)\,\mathbf{u}_4 + t\mathbf{u}_5,\ t\in\mathbb{R}$.

2. *a*) Sí, $p_4 = 4p_1 + 4p_2 - 5p_3$.
   b) No.
   c) No.

3. Para todo $a = -\frac{4}{3}$, $b = \frac{10}{3}$, $c = \frac{7}{6}$.

4. El plano en el espacio de ecuación $3x+y+z=0$.

5. El espacio generado por las matrices A y B es el conjunto de las matrices de la forma $\begin{bmatrix} -t-\frac{3}{4}s & -t-\frac{5}{4}s \\ t & s \end{bmatrix}$ con t y s en $\mathbb{R}$.

6. *a*) Sí.
   b) Sí.

7. *a*) Sí, lo genera.
   b) Sí, lo genera.
   c) No lo genera.
   d) Sí, lo genera.
   e) Sí, lo genera.
   f) No lo genera.

8. $a \neq 0$.

10. *a*) gen $\{(2,0,0,-1),(0,0,1,-3),(0,0,0,1)\}$
    b) gen $\left\{ \begin{bmatrix} 1\\1\\0\\0\\0 \end{bmatrix}, \begin{bmatrix} -1\\0\\1\\0\\0 \end{bmatrix}, \begin{bmatrix} 1\\0\\0\\1\\0 \end{bmatrix}, \begin{bmatrix} -3\\0\\0\\0\\1 \end{bmatrix} \right\}$
    c) gen $\left\{1+x^2,\ -x+x^2\right\}$
    d) gen $\{2,\ \cos x\}$
    e) gen $\left\{ \begin{bmatrix} 2 & 0\\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1\\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1\\ -1 & 0 \end{bmatrix} \right\}$

11. *a*) Falso
    b) Verdadero
    c) Falso
    d) Verdadero

Sección 4.4

1. *a*) linealmente dependiente.
   b) linealmente independiente.
   c) linealmente independiente.
   d) linealmente independiente.
   e) linealmente independiente.
   f) linealmente dependiente.
   g) linealmente independiente.
   h) linealmente dependiente.
   i) linealmente independiente.

2. a y b reales tales que $a \neq b$.

Sección 4.5

1. *a*) es base. *b*) no es base. *c*) no es base. *d*) no es base.

2. *a*) no es base. *b*) no es base. *c*) es base.

3. *a*) es base. *b*) no es base. *c*) no es base. *d*) no es base.

6. *a*) 2 *b*) 2 *c*) 2 *d*) 2 *e*) 3 *f*) 2 *g*) 3 *h*) 2 *i*) 4

7. $a = 3$

8. 2

9. Una base: $\left\{e^x,\ xe^x,\ x^2e^x\right\}$; dimensión 3.

10. 2

11. *a*) $\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}2\\-1\\3\end{bmatrix}, \begin{bmatrix}-1\\3\\4\end{bmatrix}\right\}$

 b) $\{1+x, -5x\}$

12. *a*) $\{(1,-1,3),(1,0,0),(0,1,0)\}$

 b) $\left\{\begin{bmatrix}1&2\\1&0\end{bmatrix}, \begin{bmatrix}-1&1\\3&0\end{bmatrix}, \begin{bmatrix}1&0\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix}\right\}$

13. *a*) Verdadero *b*) Verdadero *c*) Falso *d*) Falso *e*) Verdadero *f*) Verdadero *g*) Verdadero *h*) Verdadero *i*) Verdadero *j*) Verdadero *k*) Verdadero *l*) Verdadero

Sección 4.6

1. $\begin{bmatrix}2\\1\end{bmatrix}$

2. *a*) $\begin{bmatrix}-1\\3\\5\end{bmatrix}$

 b) $\begin{bmatrix}-11/7\\2/7\end{bmatrix}$

 c) $\begin{bmatrix}-3\\2\\-5\\4\end{bmatrix}$

3. *a*) $\mathbf{x} = (1,2,-5)$

 b) $\mathbf{x} = (6,7,5)$

 c) $\mathbf{x} = (8,-1,-6)$

4. *a*) $p = -8 - x$

 b) $p = -3 + x + 2x^2$

 c) $p = 6 - 11x - 3x^2 + 4x^3$

5. *a*) $\begin{bmatrix}20&16\\16&24\end{bmatrix}$

 b) $\begin{bmatrix}2\\1/2\\1/2\\9/4\end{bmatrix}$

6. *a*) $S = \{e^x, e^{-x}\}$

 b) $\begin{bmatrix}1/2\\1/2\end{bmatrix}$

 c) $g(x) = \operatorname{senh} x$

8. *a*) $\begin{bmatrix}-13/6&5/6\\-2/3&1/3\end{bmatrix}$

 b) $\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}$

 c) $\begin{bmatrix}4/11&5/11\\1/11&4/11\end{bmatrix}$

 d) $\begin{bmatrix}1&1&1&1\\0&1&1&1\\0&0&1&1\\0&0&0&1\end{bmatrix}$

9. *a*) $\begin{bmatrix}3&\frac{15}{2}&\frac{3}{2}\\-3&-5&-2\\-1&-3&0\end{bmatrix}$

 b) $\begin{bmatrix}-2&-\frac{3}{2}&-\frac{5}{2}\\\frac{2}{3}&\frac{1}{2}&\frac{1}{2}\\\frac{4}{3}&\frac{1}{2}&\frac{5}{2}\end{bmatrix}$

10. *a*) $\begin{bmatrix}3\sqrt{2}\\\sqrt{2}\end{bmatrix}$ *b*) $\begin{bmatrix}-3/2\\3/2\end{bmatrix}$

11. *a*) $\begin{bmatrix}-1/2\\1/4\\1/4\\-1/8\end{bmatrix}$

 b) $\begin{bmatrix}1&0&1/2&0\\0&1/2&0&3/4\\0&0&1/4&0\\0&0&0&1/8\end{bmatrix}$

Sección 4.7

1. *a*) una base para el espacio de renglones $\{(3,-2,8,2,1,4),(0,0,-8,5,0,-10),(0,0,0,0,0,2)\}$; una base para el espacio de columnas $\left\{\begin{bmatrix}3\\0\\0\end{bmatrix},\begin{bmatrix}8\\-8\\0\end{bmatrix},\begin{bmatrix}4\\-10\\2\end{bmatrix}\right\}$; una base para el espacio nulo $\left\{\begin{bmatrix}2/3\\1\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}-7/3\\0\\5/8\\1\\0\\0\end{bmatrix},\begin{bmatrix}-1/3\\0\\0\\0\\1\\0\end{bmatrix}\right\}$; rango 3; nulidad 3.

   b) una base para el espacio de renglones $\{(2,-3,5),(0,0,8)\}$; una base para el espacio de columnas $\left\{\begin{bmatrix}2\\0\\0\\0\end{bmatrix},\begin{bmatrix}5\\8\\0\\0\end{bmatrix}\right\}$; una base para el espacio nulo $\left\{\begin{bmatrix}3\\2\\0\end{bmatrix}\right\}$; rango 2; nulidad 1.

   c) una base para el espacio de columnas $\{(3,-1,2,1),(0,-4,3,5),(0,0,2,1),(0,0,0,10)\}$; una base para el espacio de columnas $\left\{\begin{bmatrix}3\\0\\0\\0\end{bmatrix},\begin{bmatrix}-1\\-4\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\2\\0\end{bmatrix},\begin{bmatrix}1\\5\\1\\10\end{bmatrix}\right\}$; el espacio nulo es el espacio cero; rango 4; nulidad 0.

2. *a*) una base para el espacio de renglones $\{(1,2,-3),(0,-3,-3),(0,0,14)\}$; una base para el espacio de columnas $\left\{\begin{bmatrix}1\\2\\-1\end{bmatrix},\begin{bmatrix}2\\4\\-5\end{bmatrix},\begin{bmatrix}-3\\8\\0\end{bmatrix}\right\}$; el espacio nulo es el espacio cero; rango 3; nulidad 0.

   b) una base para el espacio de renglones $\{(0,2,-3,1,2),(0,0,0,4,3)\}$; una base para el espacio de columnas $\left\{\begin{bmatrix}2\\-2\\4\end{bmatrix},\begin{bmatrix}1\\3\\6\end{bmatrix}\right\}$; una base para el espacio nulo $\left\{\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\3/2\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\-5/2\\0\\-3/4\\1\end{bmatrix}\right\}$; rango 2; nulidad 3.

   c) una base para el espacio de renglones $\{(3,2,-1),(0,-1,7)\}$; una base para el espacio de columnas $\left\{\begin{bmatrix}3\\6\\3\\0\end{bmatrix},\begin{bmatrix}2\\3\\1\\1\end{bmatrix}\right\}$; una base para el espacio nulo $\left\{\begin{bmatrix}-13\\21\\3\end{bmatrix}\right\}$; rango 2; nulidad 1.

3. *a*) $\operatorname{col}(A)=\operatorname{gen}\left\{\begin{bmatrix}-1\\2\\0\end{bmatrix},\begin{bmatrix}1\\2\\1\end{bmatrix},\begin{bmatrix}2\\4\\2\end{bmatrix}\right\}$
   $\operatorname{ren}(A)=\operatorname{gen}\{(-1,1,1,2),(2,2,2,4),(0,1,1,2)\}$

   b) $\{(-1,1,1,2),(0,1,1,2)\}$

   c) $\left\{\begin{bmatrix}-1\\2\\0\end{bmatrix},\begin{bmatrix}0\\4\\1\end{bmatrix}\right\}$

   d) $\left\{\begin{bmatrix}-1\\2\\0\end{bmatrix},\begin{bmatrix}1\\2\\1\end{bmatrix}\right\}$

   e) $N(A)=\left\{\begin{bmatrix}0\\-t-2s\\t\\s\end{bmatrix}: t,s\in\mathbb{R}\right\}$; una base para $N(A)$ es $\left\{\begin{bmatrix}0\\-1\\1\\0\end{bmatrix},\begin{bmatrix}0\\-2\\0\\1\end{bmatrix}\right\}$

   f) $\rho(A)=\nu(A)=2$

   g) **a** está en $\operatorname{col}(A)$ pero **b** y **c** no lo están; $\mathbf{a}=(-1)\mathbf{c}_1+(-t-2s)\mathbf{c}_2+t\mathbf{c}_3+s\mathbf{c}_4$, $t,s\in\mathbb{R}$, donde $\mathbf{c}_i$ $(i=1,2,3,4)$ son las columnas de A.

4. *a*) $\left\{\begin{bmatrix}1\\-1\end{bmatrix},\begin{bmatrix}3\\4\end{bmatrix}\right\}$

   b) $\{(1,-1,0),(1,2,1),(0,1,1)\}$

   c) $\left\{\begin{bmatrix}1\\2\\3\\1\end{bmatrix},\begin{bmatrix}3\\1\\4\\0\end{bmatrix},\begin{bmatrix}-5\\1\\-4\\0\end{bmatrix}\right\}$

5. *a*) $\{1-3x,\ 4x\}$

   b) $\{-1+2x^2,\ x-x^2\}$

   c) $\{1-2x-3x^2+x^3,\ x+2x^2,\ 3-2x-x^2+2x^3\}$

6. *a*) $\left\{\begin{bmatrix}-1&0\\1&0\end{bmatrix},\begin{bmatrix}1&-1\\0&0\end{bmatrix},\begin{bmatrix}-1&1\\1&0\end{bmatrix}\right\}$

   b) $\left\{\begin{bmatrix}1&1\\1&1\end{bmatrix},\begin{bmatrix}1&1\\1&0\end{bmatrix},\begin{bmatrix}1&1\\0&0\end{bmatrix},\begin{bmatrix}1&0\\0&0\end{bmatrix}\right\}$

c) $\left\{ \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -2 & 2 \\ 2 & 4 & 6 \end{bmatrix}, \begin{bmatrix} 0 & -2 & 2 \\ 2 & 6 & 7 \end{bmatrix} \right\}$

7. a) $\left\{ \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$

b) $\{1 + 3x - x^2 + x^3, 1 + x + x^3, 2 + x^2 + x^3, x + x^3\}$

c) $\left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

8. a) i. $\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$

ii. $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$

b) i. $\{1 + x + x^3,\ -2 + x + x^2,\ x^2 + x^3,\ 1\}$

ii. $\left\{ \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -2 \\ 2 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$

9. a) $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

b) $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$

c) No.

10. 235

13. Sí.

Capítulo 5

Sección 5.1

1. 3; $\sqrt{17}$; $\sqrt{7}$; -11; 8; $(0, -\frac{4}{3}, \frac{8}{3}, \frac{4}{3})$

3. $\arccos\left(-\frac{11}{315}\right) \approx 92°$; $\arccos\left(\frac{4}{51}\right) \approx 85,5°$

4. a) Sí b) No c) Sí

8. a) Verdadero b) Verdadero c) Verdadero d) Verdadero

Sección 5.2

5. a) $\sqrt{5}$; $\sqrt{77}$; 13; $\sqrt{56}$; $\arccos\left(\frac{13}{\sqrt{385}}\right) \approx 28,5°$

c) $\left(\frac{15}{2}t, t\right)$, $t \in \mathbb{R}$

d) $\left(\frac{65}{77}, -\frac{13}{77}\right)$

6. a) 10; $\sqrt{17}$; $\sqrt{14}$

b) $\arccos\left(\frac{10}{\sqrt{239}}\right) \approx 49,7°$

c) $p = p(x)$

d) $\frac{40}{17} + \frac{10}{17}x$

e) $\frac{40}{17} + \frac{10}{7}x$

7. b) i. 120; 19; $36\sqrt{2}$; $\sqrt{19}$

ii. $\arccos\left(\frac{10}{3\sqrt{38}}\right) \approx 57,3°$

iii. $-\frac{120}{19} + \frac{240}{19}x$

iv. $r(x) = a + bx + cx^2$ tal que $6a + 15b + 24c = 0$; por ejemplo: $r(x) = -4 + x^2$ $(a = 0, c = 1)$.

8. a) $\sqrt{\frac{\pi^3}{3}}$; $\sqrt{\frac{e^{2\pi}-1}{2}}$; $\sqrt{\frac{\pi}{2}}$

b) $\left(\frac{\pi^3}{3} - 2e^{\pi}(\pi - 1) + \frac{e^{2\pi}-5}{2}\right)^{1/2}$

c) $\frac{2e^{\pi}(\pi-1)+2}{e^{2\pi}-1}e^x$

d) No.

Sección 5.3

1. $[\mathbf{v}]_B = \begin{bmatrix} 27 \\ -1 \\ -9 \end{bmatrix}$

2. a) $(-1, 4, 0)$ b) $(6, 4, 1)$

3. a) $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$, $\mathbf{w}_1 = \left(-\frac{4}{3}, 2, -\frac{4}{3}, \frac{4}{3}\right) \in \mathbb{W}$ y $\mathbf{w}_2 = \left(\frac{16}{3}, 1, \frac{13}{3}, -\frac{7}{3}\right)$ ortogonal a $\mathbb{W}$; $d(\mathbf{v}, \mathbb{W}) = \frac{5\sqrt{19}}{3}$.

b) $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$, $\mathbf{w}_1 = (5, 2, 3, 6) \in \mathbb{W}$ y $\mathbf{w}_2 = (-2, 2, 2, 0)$ ortogonal a $\mathbb{W}$; $d(\mathbf{v}, \mathbb{W}) = 2\sqrt{3}$.

4. a) $(3, -1, 1, -1)$; $2\sqrt{6}$.

b) $(-1, -5, -3, 9)$; $\sqrt{170}$.

5. a) $\left\{ \left(0, \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right), \left(\frac{8\sqrt{5}}{\sqrt{609}}, \frac{17\sqrt{5}}{5\sqrt{609}}, -\frac{34\sqrt{5}}{5\sqrt{609}}\right) \right\}$

b) $\left\{ \left(\frac{3}{2\sqrt{5}}, \frac{1}{2\sqrt{5}}, -\frac{1}{2\sqrt{5}}, \frac{3}{2\sqrt{5}}\right), \left(\frac{1}{2\sqrt{5}}, \frac{3}{2\sqrt{5}}, \frac{3}{2\sqrt{5}}, -\frac{1}{2\sqrt{5}}\right), \left(-\frac{3}{2\sqrt{5}}, \frac{1}{2\sqrt{5}}, \frac{1}{2\sqrt{5}}, \frac{3}{2\sqrt{5}}\right) \right\}$

c) $\left\{ \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \sqrt{\frac{2}{3}}\right) \right\}$

d) $\left\{\begin{bmatrix} -3/\sqrt{10} \\ 1/\sqrt{10} \\ 0 \end{bmatrix}\right\}$

e) $\left\{\left(\frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}, 0\right), (0,0,1)\right\}$

f) $\left\{\frac{2}{\sqrt{41}}\begin{bmatrix} 1 \\ 0 \\ 3 \\ 1/2 \end{bmatrix}, \sqrt{\frac{41}{62}}\begin{bmatrix} -11/41 \\ 1 \\ 8/41 \\ -26/41 \end{bmatrix}\right\}$

g) $\left\{\frac{1}{\sqrt{5}}(2,1,0,0), \sqrt{\frac{5}{14}}(-\frac{3}{5}, \frac{6}{5}, 1, 0)\right.$ $\left.\frac{7}{\sqrt{2}}(\frac{1}{14}, -\frac{1}{7}, \frac{3}{14}, 0)\right\}$

6. *a)* $\{1 + x - x^2, \frac{2}{3} + \frac{2}{3}x + \frac{4}{3}x^2 - x^3,$ $\frac{8}{11} - \frac{3}{11}x + \frac{5}{11}x^2 + \frac{10}{11}x^3\}$

   b) $\{1 + x, \frac{1}{2} - \frac{1}{2}x + x^2\}$

7. *a)* $\left\{\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix},\right.$ $\left.\begin{bmatrix} -\frac{1}{4} & -\frac{1}{4} & -\frac{5}{4} \\ \frac{3}{4} & 1 & \frac{1}{2} \end{bmatrix}\right\}$

   b) $\left\{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \begin{bmatrix} -\frac{1}{3} & 1 & -1 \\ 0 & \frac{2}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}\right\}$

8. *a)* $\{x,\ e^x - 3x\}$

   b) $\{x,\ 1 - \frac{3}{2}x,\ \frac{4}{21} - \frac{5}{7}x + x^5\}$

   c) $\left\{\operatorname{sen} x,\ \ \cos x - \frac{2\operatorname{sen}^2 1}{2 - \operatorname{sen} 2}\operatorname{sen} x\right\}$

9. *a)* $\left(\frac{184}{609}, \frac{809}{609}, \frac{209}{609}\right)$; $\left(\frac{425}{609}, -\frac{200}{609}, \frac{400}{609}\right)$

   b) $\left(\frac{1}{2}, -\frac{1}{2}, \frac{3}{2}, \frac{1}{2}\right)$; $\left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)$

10. *a)* $\frac{13}{9} + \frac{1}{3}x + \frac{7}{9}x^2 + \frac{5}{9}x^3 \in \mathbb{W}$; $-\frac{4}{9} + \frac{2}{3}x + \frac{2}{9}x^2 + \frac{4}{9}x^3 \in \mathbb{W}^\perp$

    b) $-\frac{1}{2} + x - \frac{1}{2}x^2 \in \mathbb{W}$; $\frac{1}{2} - \frac{1}{2}x^2 \in \mathbb{W}^\perp$

11. *a)* $\begin{bmatrix} \frac{5}{14} & \frac{5}{14} & \frac{2}{7} \\ \frac{3}{7} & \frac{1}{14} & \frac{11}{14} \end{bmatrix}$

    b) $\begin{bmatrix} \frac{1}{4} & -\frac{3}{4} & \frac{3}{4} \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{34} \end{bmatrix}$

12. $\left(\frac{5}{2} - \frac{6e - 15}{e^2 - 7}\right)x + \frac{2e - 5}{e^2 - 7}e^x$

15. *c)* i. $\{(-2t, t, t) : t \in \mathbb{R}\}$; $\{(-2, 1, 1)\}$

    ii. $\{(-s, -t - s, t, s) : s, t \in \mathbb{R}\}$; $\{(-1, -1, 0, 1), (0, -1, 0, 1)\}$

    iii. $\{(0, -t, t, 0) : t \in \mathbb{R}\}$; $\{(0, -1, 1, 0)\}$

Capítulo 6

Sección 6.1

1. *a)* $D_T = \mathbb{R}^2$; $C_T = \mathbb{R}^3$

   b) $(-4, 3, 0)$, $(3, -1, 0)$ y $(0, 0, 0)$.

   c) $(0, 0)$

   d) No.

   e) $\{(a, b, 0) : a, b \in \mathbb{R}\}$

2. *a)* sí
   b) no
   c) sí
   d) sí
   e) no
   f) sí
   g) sí
   h) no
   i) sí
   j) sí
   k) sí
   l) sí

3. *b)* $\begin{bmatrix} -1 \\ -2 \\ 1 \\ 3 \end{bmatrix}$

4. *a)* $n = 2$, $m = 3$, $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 \\ -x_1 + 5x_2 \\ 4x_2 \end{bmatrix}$

   b) $n = 4$, $m = 1$, $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_1 - 2x_2 + 5x_3 + 3x_4$

5. *a)* $\left(\frac{5}{2}y - \frac{5}{2}x,\ \frac{3}{2}y + \frac{1}{2}x\right)$

   b) $\mathbb{R}^2$

6. $T(x_1, x_2, x_3) = \left(-x_1 - \frac{1}{2}x_2 + \frac{1}{2}x_3,\ x_1 + x_2 + x_3\right)$

7. $\frac{b - a}{4}(1 + x) + \frac{5a - b}{8}x^2$

8. $x + \frac{1}{2}x^2 + \frac{1}{3}x^3$, $-x + \frac{1}{4}x^4$ y x^2.

10. *a)* Falso.
    b) Falso.
    c) Verdadero.

Sección 6.2

1. *a)* $\{(0, 0)\}$

   c) gen $\{-1 + x\}$

   d) $\{(0, 0, 0)\}$

   f) gen $\{1\}$

   i) $\{O\}$, donde O es la matriz cero de $n \times n$

2. *a)* $\ker(T) = \{(0, 0)\}$; $R(T) = \text{gen}\{(1, 0, 0), (0, 0, 1)\}$; $\nu(T) = 0$; $\rho(T) = 2$.

b) $\ker(T) = \{(0,0,0)\}$; $R(T) = \mathbb{R}^3$; $\nu(T) = 0$; $\rho(T) = 3$.

c) $\ker(T) = \text{gen}\left\{-1-x,\ -1-x^2\right\}$; $R(T) = \text{gen}\{x\}$; $\nu(T) = 2$; $\rho(T) = 1$.

d) $\ker(T) = \text{gen}\left\{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right\}$; $R(T) = \left\{\begin{bmatrix} a' & b' \\ 0 & c' \end{bmatrix} \in \mathbb{M}_2 : a', b', c' \in \mathbb{R}\right\}$; $\nu(T) = 1$; $\rho(T) = 3$.

e) $\ker(T) = \text{gen}\left\{-1+x+x^2\right\}$; $R(T) = \mathbb{R}^2$; $\nu(T) = 1$; $\rho(T) = 2$.

f) $\ker(T) = \text{gen}\{(1,1,0)\}$; $R(T) = \left\{\begin{bmatrix} a' & 0 & 0 \\ b' & 0 & b' \end{bmatrix} \in \mathbb{M}_2 : a', b' \in \mathbb{R}\right\}$; $\nu(T) = 1$; $\rho(T) = 2$.

3. $\rho(T) = 2$ y $\nu(T) = 2$.

4. $\ker(T) = \{0\}$; $R(T) = \mathbb{P}_2$; $\nu(T) = 0$; $\rho(T) = 3$.

5. *b*) $\rho(T) = 2$ y $\nu(T) = 2$.

6. *a*) $\ker(T) = \left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}$; $R(T) = \mathbb{R}^2$; $\rho(T) = 2$; $\nu(T) = 0$.

b) $\ker(T) = \text{gen}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$; $R(T) = \text{gen}\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$; $\rho(T) = 1$; $\nu(T) = 1$.

c) $\ker(T) = \text{gen}\left\{\begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}\right\}$; $R(T) = \mathbb{R}^2$; $\rho(T) = 2$; $\nu(T) = 1$.

d) $\ker(T) = \left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}$; $R(T) = \text{gen}\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}\right\}$; $\rho(T) = 2$; $\nu(T) = 0$.

e) $\ker(T) = \text{gen}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$; $R(T) = \text{gen}\left\{\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}\right\}$; $\rho(T) = 2$; $\nu(T) = 1$.

f) $\ker(T) = \text{gen}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}\right\}$; $R(T) = \text{gen}\left\{\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}\right\}$; $\rho(T) = 1$; $\nu(T) = 2$.

g) $\ker(T) = \text{gen}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$; $R(T) = \mathbb{R}^2$; $\rho(T) = 2$; $\nu(T) = 2$.

8. *a*) 2 *b*) 1 *c*) 2

9. $\mathbb{R}^4$

11. $\{a : a \in \mathbb{R}\}$

12. $\{bx : b \in \mathbb{R}\}$

Sección 6.3

1. *a*) $n = 2$, $m = 3$, $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \\ 1 & 3 \end{bmatrix}$

b) $n = m = 2$, no es lineal

c) $n = 3$, $m = 2$, no es lineal

d) $n = 2$, $m = 1$, $A = \begin{bmatrix} 3 & 0 \end{bmatrix}$

e) $n = 4$, $m = 3$, $A = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}$

f) $n = 3$, $m = 4$, $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

2. *a*) $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ *c*) $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

b) $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ *d*) $A = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$

3. *a*) $A = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$ *c*) $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

b) $A = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$ *d*) $A = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$

4. *a*) $A = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}$ *b*) $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

5. *a*) expansión de un factor 2 en la dirección de la coordenada x.

b) reflexión con respecto a la recta $y = x$

c) reflexión con respecto a la recta $y = -x$

d) proyección ortogonal sobre el eje x

e) proyección ortogonal sobre el eje y

f) rotación alrededor del origen de coordenadas un ángulo $\theta = 45°$

Sección 6.4

1. a) $\begin{bmatrix} -1 & -1 \\ 1 & 1 \\ 0 & -1 \end{bmatrix}$

   b) $\begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$

   c) $\begin{bmatrix} 0 & -1 & 1 \\ \frac{1}{2} & 1 & -\frac{1}{2} \end{bmatrix}$

   d) $\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$

2. $T(a+bx+cx^2) = (-a+b+4c) + (a+3b+4c)x$

3. $T(a+bx+cx^2) = (c-a) + (a+2b+c)x$

4. $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) =$
   $(5a+3b-3d, -3a-2b+d, -2a-2b+c+d,$
   $2a+b+c-d)$

5. 2. $\rho(T) = 2$, $\nu(T) = 1$

   4. $\rho(T) = 4$, $\nu(T) = 0$

6. $\ker(T) = \text{gen}\,\{(1,1,3)\}$
   $R(T) = \text{gen}\,\{1+x, \ -1+3x\}$

7. a) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 4 \\ 0 & 2 & 5 \\ 1 & 3 & 1 \end{bmatrix}$

   b) $-3x^2 + 5x^3 - 2x^4$

8. a) $\begin{bmatrix} 3 \\ 1 \\ -3 \end{bmatrix}$, $\begin{bmatrix} -2 \\ 6 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 2 \\ 7 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

   b) $(11,5,22)$, $(-42,32,-10)$, $(-56,86,17)$ y $(-13,17,2)$.

   c) $T(x,y,z,w) = \big(\frac{1217}{60}z - \frac{373}{60}x - \frac{82}{15}y - \frac{229}{60}w,$
   $\frac{51}{4}w - \frac{25}{4}x + 12y - \frac{79}{4}z,\ \frac{91}{6}w - \frac{59}{6}x + \frac{2}{3}y + \frac{37}{6}z\big)$

10. I_3

12. a) $A' = \begin{bmatrix} 0 & -3 \\ -1 & -1 \end{bmatrix}$, $A = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$

    b) $A' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 3 \end{bmatrix}$

    c) $A' = \begin{bmatrix} \frac{5}{2} & \frac{3}{2} \\ \frac{5}{2} & \frac{3}{2} \end{bmatrix}$, $A = \begin{bmatrix} 4 & 1 \\ 0 & 0 \end{bmatrix}$

13. $B' = \{(2,1),(1,1)\}$, $[T]_{B'} = \begin{bmatrix} -5 & -6 \\ 2 & 2 \end{bmatrix}$

14. $B' = \big\{1+x+2x^2, 1+x^2, 1+x\big\}$,
    $[T]_{B'} = \begin{bmatrix} -\frac{5}{2} & -\frac{3}{2} & -\frac{3}{2} \\ 7 & 3 & 5 \\ \frac{3}{2} & \frac{5}{2} & -\frac{3}{2} \end{bmatrix}$

15. a) $P = \begin{bmatrix} \frac{4}{3} & -3 \\ \frac{2}{3} & -1 \end{bmatrix}$, $Q = \begin{bmatrix} -\frac{3}{2} & \frac{9}{2} \\ -1 & 2 \end{bmatrix}$

    b) $[T]_{B'} = \begin{bmatrix} 5 & -9 \\ \frac{4}{3} & -2 \end{bmatrix}$

    c) $[T(\mathbf{v}])]_B = \begin{bmatrix} -\frac{2}{3} \\ 0 \end{bmatrix}$, $[T(\mathbf{v}])]_{B'} = \begin{bmatrix} 1 \\ \frac{2}{3} \end{bmatrix}$

Capítulo 7

Sección 7.1

1. $\mathbf{u}$ es un autovector y 5 es su correspondiente autovalor; $\mathbf{v}$ es un autovector y 0 es su correspondiente autovalor; $\mathbf{w}$ no es autovector.

2. $\mathbf{u}+\mathbf{v}$ no; $c\mathbf{u}$ sí.

3. $\lambda_1 = 1$, $\lambda_2 = 2$;
   $\lambda_1 = 0$, $\lambda_2 = -1$;
   $\lambda_1 = 5$, $\lambda_2 = -5$.

4. Para $\lambda_1 = -6$, $\mathbf{u} = \begin{bmatrix} 1 \\ -t \\ t \end{bmatrix}$ con $t \neq 0$.

   Para $\lambda_2 = 4$, $\mathbf{v} = \begin{bmatrix} 0 \\ 2t \\ 3t \end{bmatrix}$ con $t \neq 0$.

   Para $\lambda_3 = 7$, $\mathbf{w} = \begin{bmatrix} t \\ 0 \\ 0 \end{bmatrix}$ con $t \neq 0$.

5. $E_2 = \text{gen}\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$.

6. a) $p(\lambda) = \lambda^2 - 5\lambda$; $\lambda_1 = 0$ y $\lambda_2 = 5$; base para E_0: $\left\{\begin{bmatrix} -2 \\ 3 \end{bmatrix}\right\}$; base para E_5: $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$; $m_g(0) = m_a(0) = 1$ y $m_g(5) = m_a(5) = 1$.

   b) $p(\lambda) = \lambda^2+6\lambda+9$; $\lambda_1 = -3$; base para E_{-3}: $\left\{\begin{bmatrix} 3 \\ 1 \end{bmatrix}\right\}$; $m_g(-3) = 1$ y $m_a(-3) = 2$.

   c) $p(\lambda) = \lambda^2 - (a+b)\lambda$; $\lambda_1 = 0$ y $\lambda_2 = a+b$; base para E_0: $\left\{\begin{bmatrix} -b \\ a \end{bmatrix}\right\}$; base para E_{a+b}: $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$; $m_g(0) = m_a(0) = 1$ y $m_g(a+b) = m_a(a+b) = 1$.

d) $p(\lambda) = \lambda^3 - 6\lambda^2 + 11\lambda - 6$; $\lambda_1 = 1$, $\lambda_2 = 2$ y $\lambda_3 = 3$; base para E_1: $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$; base para E_2: $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$; base para E_3: $\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$; $m_g(1) = m_a(1) = 1$, $m_g(2) = m_a(2) = 1$ y $m_g(3) = m_a(3) = 1$.

e) $p(\lambda) = \lambda^3 - 4\lambda^2 - 2\lambda + 8$; $\lambda_1 = \sqrt{2}$, $\lambda_2 = -\sqrt{2}$ y $\lambda_3 = 4$; base para $E_{\sqrt{2}}$: $\left\{ \begin{bmatrix} 1 \\ 0 \\ \sqrt{2} \end{bmatrix} \right\}$; base para $E_{-\sqrt{2}}$: $\left\{ \begin{bmatrix} -1 \\ 0 \\ \sqrt{2} \end{bmatrix} \right\}$; base para E_4: $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$; $m_g(\sqrt{2}) = m_a(\sqrt{2}) = 1$, $m_g(-\sqrt{2}) = m_a(-\sqrt{2}) = 1$ y $m_g(4) = m_a(4) = 1$.

f) $p(\lambda) = \lambda^3 - 2\lambda^2 - 4\lambda + 8$; $\lambda_1 = 2$, y $\lambda_2 = -2$; base para E_2: $\left\{ \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$; base para E_{-2}: $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\}$; $m_g(2) = m_a(2) = 2$ y $m_g(-2) = m_a(-2) = 1$.

g) $p(\lambda) = \lambda^4 - 4\lambda^3 + 3\lambda^2 + 4\lambda - 4$; $\lambda_1 = -1$, $\lambda_2 = 1$ y $\lambda_3 = 2$; base para E_{-1}: $\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$; base para E_1: $\left\{ \begin{bmatrix} 2 \\ 0 \\ 3 \\ 0 \end{bmatrix} \right\}$; base para E_2: $\left\{ \begin{bmatrix} 1 \\ -3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 1 \\ 0 \end{bmatrix} \right\}$; $m_g(-1) = m_a(-1) = 1$, $m_g(1) = m_a(1) = 1$ y $m_g(2) = m_a(2) = 2$.

7. $\begin{bmatrix} -2 \\ -10 \\ 5 \end{bmatrix}$

Sección 7.2

1. i. $P = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$, $A = \begin{bmatrix} -3 & \frac{5}{2} \\ 0 & 2 \end{bmatrix}$

   ii. $P = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, $A = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$

   iii. $P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -2 & 0 \end{bmatrix}$, $D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, $A = \begin{bmatrix} \frac{5}{3} & -\frac{2}{3} & 2 \\ -\frac{4}{3} & \frac{7}{3} & 2 \\ \frac{8}{3} & \frac{4}{3} & -1 \end{bmatrix}$

3. *a*) $D_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $D_2 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ y $D_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

   b) $D_1 = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}$ y $D_2 = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$

4. *a*) $a \neq 2$ *b*) $a > 0$ *c*) para todo a

6. *a*) $D = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix}$, $P = \begin{bmatrix} -2 & 1 \\ 3 & 1 \end{bmatrix}$

   b) no es posible

   c) $D = \begin{bmatrix} 0 & 0 \\ 0 & a+b \end{bmatrix}$, $P = \begin{bmatrix} -b & 1 \\ a & 1 \end{bmatrix}$

   d) $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, $P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

   e) $D = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & -\sqrt{2} & 0 \\ 0 & 0 & 4 \end{bmatrix}$, $P = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ \sqrt{2} & \sqrt{2} & 0 \end{bmatrix}$

   f) $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$, $P = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}$

   g) $D = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$, $P = \begin{bmatrix} 0 & 2 & 1 & 0 \\ 0 & 0 & -3 & 3 \\ 1 & 3 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

7. *a*) $P = \begin{bmatrix} 2 & -3 & 1 \\ 0 & 0 & -6 \\ 2 & 2 & 4 \end{bmatrix}$, $D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

b) $P = \begin{bmatrix} -1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

c) no es posible

d) $P = \begin{bmatrix} 1 & -2 & 1 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$,

$$D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

9. *a*) Falso
   b) Verdadero
   c) Verdadero
   d) Verdadero
   e) Verdadero
   f) Verdadero

Bibliografía

- Anton, H. (2011). *Introducción al Álgebra Lineal: con aplicaciones en negocios, economía, ingeniería, física, ciencias de la computación, teoría de aproximación, ecología, sociología, demografía y genética.* (5ta. Edición). México: Limusa Wiley.
- Grossman, S. & Flores Godoy, J. (2012). *Álgebra Lineal.* (7ma. Edición). México: McGraw-Hill.
- Kolman, B. & Hill, D. (2006). *Álgebra Lineal.* (8va. Edición). México: Pearson Educación.
- Lang, S. (1998). *Introducción al Álgebra Lineal.* México: Addison Wesley Iberoamericana.
- Larson, R. & Falvo, D. (2010). *Fundamentos de Álgebra Lineal.* (6ta. Edición). México: Cengage Learning.
- Lay, D. (2012). *Álgebra Lineal y sus aplicaciones.* (4ta. Edición). México: Pearson Educación.
- Lay, D. (2013). *Álgebra Lineal para cursos con enfoque por competencias.* México: Pearson Educación.
- Nakos, G. & Joyner, D. (1999). *Álgebra Lineal con Aplicaciones.* México: International Thomson Editores.
- Sanz, P., Vázquez, F. & Ortega, P. (1998). *Problemas de Álgebra Lineal. Cuestiones, ejercicios y tratamiento en Derive.* Madrid: Prentice Hall.