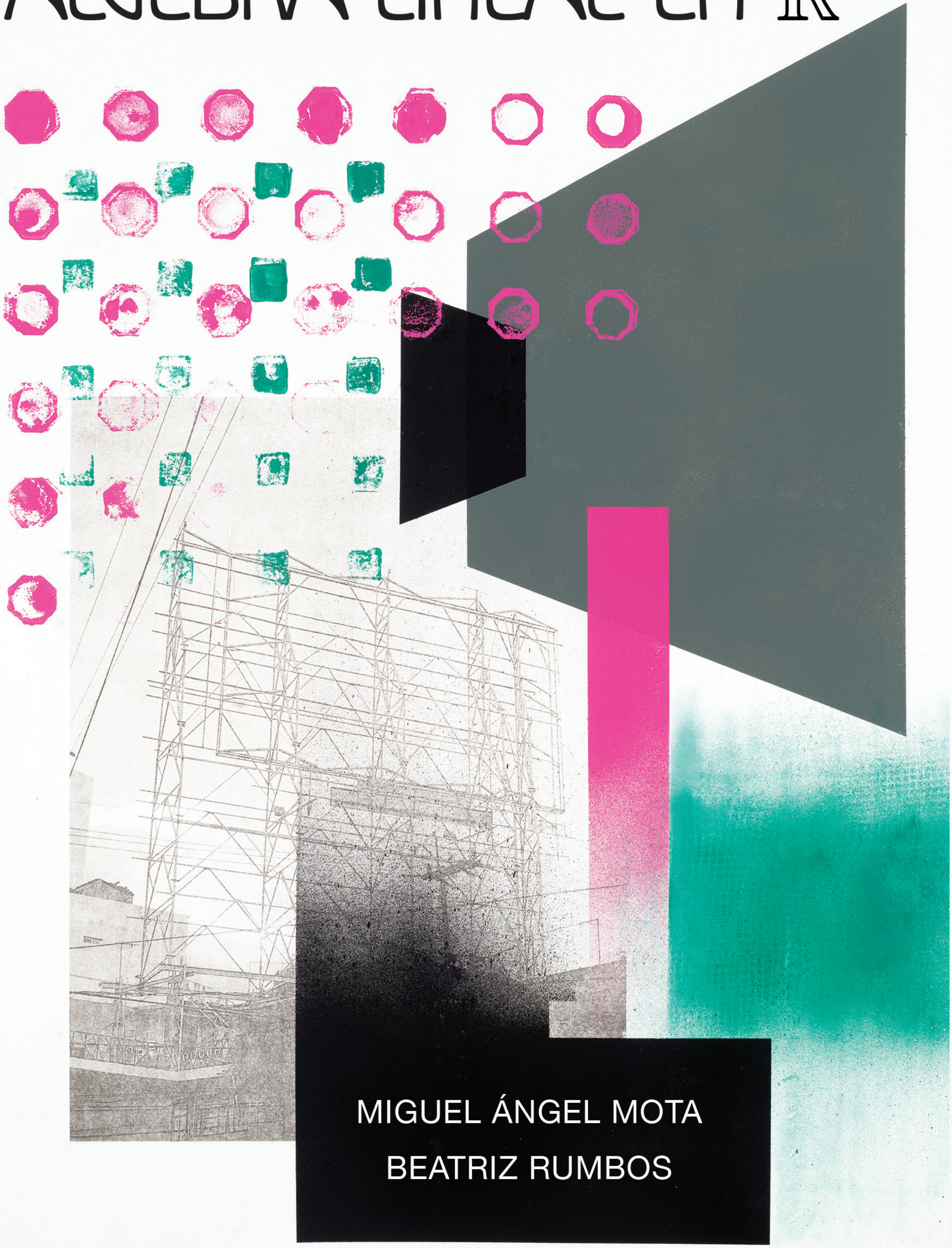
ÁLGEBRA LINEAL EN $\mathbb{R}^n$



MIGUEL ÁNGEL MOTA
BEATRIZ RUMBOS

Álgebra Lineal en $\mathbb{R}^n$

Miguel Ángel Mota
Beatriz Rumbos

Primera edición

A todos los espacios vectoriales, dondequiera que estén.
Beatriz Rumbos

A Yichun Shi y a Elvira Gaytán.
Miguel Ángel Mota

Índice general

Prefacio

Si realizamos una búsqueda de libros de álgebra lineal dentro del portal de *Amazon* aparecen miles de resultados por lo que, en primera instancia, exponemos las razones que nos llevaron a escribir el presente texto. Existe un consenso de que el lenguaje y los métodos del álgebra lineal permean el resto de las matemáticas y de que ésta tiene múltiples aplicaciones en disciplinas tan diversas como la economía, las ingenierías o la hoy naciente ciencia de datos. Sin embargo, no termina de ser claro cuál debe ser el nivel de exposición y el material a cubrirse dentro de un primer (y frecuentemente, único) curso en el área.

Nuestra propuesta consiste en ofrecer una perspectiva teórica, motivada por la intuición de lo que sucede en los espacios vectoriales reales más accesibles: el plano cartesiano, el espacio tridimensional y de forma más general, la estructura conocida como $\mathbb{R}^n$. A pesar de que nos restringimos al campo de los números reales y a los espacios $\mathbb{R}^n$, este texto permite que los estudiantes que así lo requieran transiten fácilmente a un segundo curso de álgebra lineal de carácter mucho más general.

A partir de conceptos y argumentos relativamente sencillos, buscamos que nuestros lectores no sólo manipulen correctamente los objetos clásicos como los sistemas de ecuaciones lineales, las matrices, los determinantes y otros más; sino que sean capaces de entender que éstos forman parte de un universo matemático mucho más amplio y profundo. En este sentido, buscamos desarrollar el pensamiento abstracto desde un ámbito donde la mayoría de nuestros estudiantes se sienten suficientemente cómodos. Por ejemplo, el capítulo de transformaciones lineales precede al de álgebra matricial, lo cual pudiera parecer extraño, sin embargo, el propósito es hacer el paralelismo entre la multiplicación de matrices y la composición de cierto tipo de funciones, operación con la cual el estudiante ya está familiarizado desde sus primeros cursos de precálculo.

De acuerdo a nuestra experiencia con estudiantes de distintas ingenierías y licenciaturas (incluida matemáticas), es posible cubrir la mayor parte del material aquí presentado en un curso de 56 horas de duración y también puede adaptarse fácilmente a formatos más cortos, omitiendo algunos temas o demostraciones, o bien simplificando la presentación de los mismos. Hemos introducido la función determinante utilizando el enfoque multilineal de Emil Artin con el fin de evitar los engorrosos argumentos combinatorios que surgen cuando se opta por una presentación vía permutaciones. El instructor puede ignorar este enfoque y pasar directamente a la parte meramente operativa del determinante, mencionando explícitamente aquellas propiedades cuya verificación se ha dejado de lado.

A diferencia de la mayoría de los libros existentes de álgebra lineal, en este texto no hacemos énfasis en las múltiples y variadas aplicaciones del material presentado. Esto no quiere decir que minimicemos su importancia, sino que asumimos que éstas serán abordadas en otros cursos o bien, el profesor podrá incorporar las aplicaciones prácticas que considere necesarias para complementar el curso. No obstante, y con el fin de dar una motivación adicional a todos aquellos que se inician en esta aventura, es muy deseable que se aborde al menos una de las dos aplicaciones que aparecen en este libro, mismas que son cruciales para realizar cierto tipo de predicciones. Nos referimos a Cadenas de Markov y a la aproximación lineal que se obtiene utilizando mímimos cuadrados. Esta última constituye un final apropiado para este libro al tratarse de un método que encuentra lo más parecido a una solución óptima para un sistema lineal inconsistente cuyo objetivo es obtener la recta que mejor aproxima a un conjunto de datos.

Quisieramos agradecer a todos los colegas que nos han hecho correcciones y sugerencias valiosas para enriquecer y mejorar este texto, en especial a Lucía Ramírez y a Conchita Ruiz por haber revisado cuidadosa y exhaustivamente este material. Finalmente, agradecemos también el apoyo de la Asociación Mexicana de Cultura A.C. y del Instituto Tecnológico Autónomo de México (ITAM).

Capítulo 1

Preliminares

Conjuntos

Un **conjunto**, en su concepción más elemental, es una colección de objetos denominados **elementos** del conjunto. Típicamente, los conjuntos suelen denotarse por letras mayúsculas: $A, B, C, \ldots$y sus elementos por letras minúsculas: $a, b, c, \ldots$ El símbolo "$\in$" denota la pertenencia de un elemento a un conjunto, por ejemplo

$$a \in A$$

significa que "a" es un elemento del conjunto A. La no pertenencia a un conjunto se denota por "$\notin$" de manera que

$$x \notin A$$

se lee como "x no es un elemento de A". Al conjunto sin elementos se le denomina **conjunto vacío** y se denota por $\emptyset$. Si dados dos conjuntos A y B se tiene que todos los elementos de A son elementos de B, entonces decimos que A es un **subconjunto** de B y esto se denota como[1]

$$A \subset B.$$

[1] También es común utilizar la notación

$$A \subseteq B.$$

En particular, $A \subset A$ y dado que el conjunto vacío no tiene elementos, por vacuidad, se tiene que $\emptyset \subset A$ para cualquier conjunto A.

Dos conjuntos A y B son iguales si ambos tienen los mismos elementos. En este caso se tiene que $A \subset B$ y $B \subset A$. La igualdad entre los conjuntos se denota simplemente por $A = B$. Si A es un subconjunto de B pero A no es igual a B (en este caso existirá al menos un elemento de B que no pertenece al conjunto A), decimos que A es un subconjunto propio de B (o está contenido propiamente en B). En este caso se utiliza la notación

$$A \subsetneq B.$$

Para especificar un conjunto es necesario saber quiénes son sus elementos. Para ello pueden listarse los elementos (esto es práctico si se trata de un conjunto pequeño o con cierta estructura recursiva) o bien describirse con una o varias propiedades que los caracterizen. Pueden escribirse así

$$\begin{aligned}
A &= \{a, b, 666, \#, \bigstar\},\\
B &= \{1, 2, 3, \ldots\},\\
C &= \{\text{anfibios de color verde}\},\\
D &= \{x \in \mathbb{R} \mid -1 < x < 1\}.
\end{aligned}$$

En el caso del conjunto D, la descripción de sus elementos se lee como: "el conjunto de todos los x que pertenecen a los números reales tales que[2] cumplen $-1 < x < 1$". Por supuesto, una forma compacta de denotar a este conjunto es mediante la simbología $(-1, 1)$ que hace referencia al correspondiente intervalo abierto de la recta real.

Dados dos conjuntos A y B, podemos "juntar sus elementos" en un conjunto llamado la **unión** de A y B, denotada por $A \cup B$ y definida como

$$A \cup B = \{x \mid x \in A \text{ o } x \in B\},$$

en donde el "o" indica que x puede ser un elemento de A, de B o de ambos. De forma análoga se define la **intersección** de los conjuntos A y B como el conjunto de sus elementos comunes y se denota por $A \cap B$. Formalmente,

$$A \cap B = \{x \mid x \in A \text{ y } x \in B\}.$$

Asimismo, también se define el **complemento relativo**[3] de B en A

[2] Es frecuente que la linea vertical "|" que significa "tales que cumplen" se sustituya por dos puntos ": ". Así, podríamos alternativamente denotar al conjunto C como

$$C = \{x \in \mathbb{R} \colon\ -1 < x < 1\}.$$

como

$$A \setminus B = \{x \mid x \in A \text{ y } x \notin B\}.$$

Así, $A \setminus B$ se refiere al conjunto de todos los elementos de A que no están en B.

Es usual trabajar con conjuntos cuyos elementos pertenecen a algún conjunto más grande (universal) que contiene a todos los conjuntos de interés. Si denotamos a este conjunto universal por X, para todo conjunto A tendremos que $A \subset X$. El complemento relativo de A con respecto a X se conoce simplemente como **complemento** y se denota por A^c, tenemos así que

$$A^c = X \setminus A = \{x \in X \mid x \notin A\}.$$

Ejemplo 1.1 *Para cualesquiera A, B conjuntos, se tiene que*

$$A \setminus B = A \cap B^c.$$

En efecto,

$$\begin{aligned} A \setminus B &= \{x \mid x \in A \ y \ x \notin B\} \\ &= \{x \mid x \in A \ y \ x \in B^c\} = A \cap B^c. \end{aligned}$$

Ejemplo 1.2 *Supongamos que nuestro universo X es el conjunto de letras del alfabeto griego. Si $A = \{\alpha, \beta, \gamma, \gamma, \gamma\}$ y $B = \{\beta, \alpha, \gamma\}$, entonces $A = B$. Para demostrar esto simplemente observamos que dado cualquier $x \in A$ se cumple que $x \in B$, por lo que $A \subset B$, y dado cualquier $x \in B$ se cumple que $x \in A$, con lo cual $B \subset A$; por lo tanto, $A = B$. Observemos que es irrelevante el orden en el que aparecen los elementos, así como también si éstos aparecen o no repetidos.*

Ejemplo 1.3 *Sea X el conjunto de números naturales $\{1, 2, \ldots\}$. Sean $A = \{x \in X \mid x \text{ es impar}\}$, $B = \{x \in X \mid x \text{ es par}\}$, $C = \{6, 66, 666\}$ y $D = \{1, 2, 5, 6, 7\}$. Es inmediato notar que:*

$$\begin{aligned} A &\subsetneq X, \\ A \cup B &= X, \\ A \cap B &= \varnothing \\ A^c &= B, \\ B^c &= A, \\ A \setminus B &= A, \\ C \cap D &= \{6\}, \\ C \setminus D &= \{66, 666\}. \end{aligned}$$

[3] En ocasiones llamada **diferencia de conjuntos** y denotada por $A - B$.

Conjuntos de números

Algunos conjuntos muy conocidos en matemáticas son:

El conjunto de los **números naturales** $\mathbb{N} = \{1, 2, 3, \ldots\}$, el conjunto de los **números enteros** $\mathbb{Z} = \{\ldots, -2-1, 0, 1, 2, \ldots\}$, el conjunto de los **números racionales** $\mathbb{Q} = \left\{ \frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0 \right\}$ y el conjunto de los **números reales** $\mathbb{R}$.

En todos ellos pueden definirse las operaciones de suma ($+$) y producto ($\times$) entre cada pareja de elementos, mismas que satisfacen una serie de propiedades sumamente "convenientes"[4]. El álgebra abstracta estudia la estructura algebraica de éstos y otros muchos conjuntos. En particular, los llamados espacios vectoriales son conjuntos que poseen lo que se conoce como una estructura *lineal.* El álgebra lineal estudia estos espacios y las funciones entre ellos que preservan esta estructura. En este texto nos concentraremos en el caso particular del espacio vectorial $\mathbb{R}^n$, mismo que se describirá más adelante.

Es más o menos claro como obtener $\mathbb{Q}$ a partir de $\mathbb{Z}$ y éste a partir de $\mathbb{N}$. Asimismo, no es muy difícil probar que estos tres conjuntos tienen el mismo número de elementos (o la misma cardinalidad) y, por lo tanto, son numerables. La construcción del conjunto de números reales es mucho más compleja de entender ya que algebraicamente posee la misma estructura que los números racionales, en el sentido de que ambos constituyen lo que se conoce como un **campo**; sin embargo, $\mathbb{Q}$ y $\mathbb{R}$ difieren sustancialmente en cuanto a su estructura cardinal y geométrica (o más bien topológica). El detalle de cómo obtener $\mathbb{R}$ a partir de $\mathbb{Q}$ va más allá de los alcances de este libro y pertenece más bien al ámbito de un curso de análisis matemático.

Recordemos que cualquier elemento de $\mathbb{Q}$ tiene una expresión decimal finita o bien infinita pero con repetición periódica a partir de algún punto. A diferencia de los racionales, los **números irracionales**, denotados por $\mathbb{I}$, poseen una expansión decimal infinita que no presenta regularidad alguna, por ejemplo, $\pi = 3.1416\ldots$, $e = 2.7182\ldots$, $\sqrt{2} = 1.4142, \ldots$, etcétera. Como conjunto, los reales son la unión de los racionales y los irracionales ($\mathbb{R} = \mathbb{Q} \cup \mathbb{I}$) y, desde una perspectiva más gráfica, son representables a través de los puntos de una recta

[4]Nos referimos a las propiedades de conmutatividad, asociatividad, distributividad y la existencia de elementos neutros e inversos aditivos y multiplicativos, según sea el caso.

infinita, llamada la "recta real" (Figura 1.1).

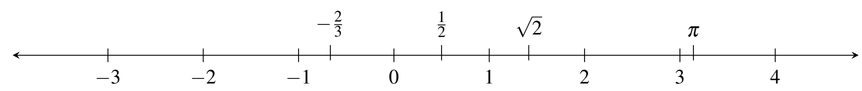


Figura 1.1: Recta Real.

La necesidad de introducir cada vez más números en el sentido de que

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

está motivada por el deseo natural de resolver cierto tipo ecuaciones. Por ejemplo, para encontrar la solución de las ecuaciones

$$\begin{aligned} x + 1 &= 0, \\ 2x &= 1, \\ x^2 &= 2, \end{aligned}$$

es necesario considerar a $\mathbb{Z}$, $\mathbb{Q}$ y $\mathbb{R}$, respectivamente. Observemos que la ecuación

$$x^2 = -1,$$

en apariencia inofensiva, no tiene solución en ninguno de estos conjuntos de números. Para resolver esta ecuación se requiere del conjunto de los **números complejos**, denotado por $\mathbb{C}$, mismos que extienden a los números reales de manera que $\mathbb{R} \subset \mathbb{C}$. En este curso no utilizaremos a los números complejos, todos los coeficientes y variables tomarán valores en el conjunto de números reales o bien, en alguno de los subconjuntos $\mathbb{N}, \mathbb{Z}$ o $\mathbb{Q}$.

Demostraciones para principiantes

En matemáticas existen enunciados fundamentales que llamamos axiomas y que presuponemos ciertos sin necesidad de demostrarlos. El resto de la disciplina consiste de resultados cuya validez debe demostrarse, ya sea a partir de los axiomas o de otros resultados que ya se han validado previamente. Cada enunciado matemático puede ser verdadero o falso (¡no se permite que a veces sea verdadero y a veces falso!) y en caso de que sea verdadero decimos que el enunciado es válido. Típicamente se tiene un enunciado compuesto del tipo: "Si X, entonces Y" o bien algo equivalente. Este enunciado compuesto es válido si cuando X es válido, entonces Y también lo es. Para demostrar que esta aseveración es cierta, suponemos, por hipótesis, que X es válido y mediante una secuencia de pasos que se siguen lógicamente el uno del otro, llegamos a que Y es válido. Se dice fácil, pero como siempre, ¡el diablo está en los detalles! Veamos algunos ejemplos sencillos.

Ejemplo 1.4 *Dados conjuntos A y B se cumple*

$$(A \cup B)^c = A^c \cap B^c.$$

Para demostrar la igualdad entre estos conjuntos primero tomamos un elemento arbitrario del conjunto $(A \cup B)^c$, digamos x. Por lo tanto $x \in (A \cup B)^c$, equivalentemente, $x \notin A \cup B$ por lo que x no es un elemento de A ni de B, es decir $x \notin A$ y $x \notin B$ o bien, $x \in A^c$ y $x \in B^c$ y, por lo tanto, $x \in A^c \cap B^c$ y como cualquier elemento del conjunto $(A \cup B)^c$ está en el conjunto $A^c \cap B^c$, hemos probado que

$$(A \cup B)^c \subset A^c \cap B^c. \tag{$\clubsuit$}$$

De forma similar, si comenzamos con x cualquier elemento de $A^c \cap B^c$, se tiene que $x \in A^c$ y $x \in B^c$ o bien $x \notin A$ y $x \notin B$, por lo que $x \notin A \cup B$, equivalentemente $x \in (A \cup B)^c$ y tenemos que

$$A^c \cap B^c \subset (A \cup B)^c. \tag{$\spadesuit$}$$

En virtud de (♣) y (♠) concluimos que

$$(A \cup B)^c = A^c \cap B^c.$$

Esta igualdad entre conjuntos es una de las llamadas Leyes de De Morgan.

Ejemplo 1.5 *En este ejemplo, la palabra número se refiere a un número entero. Se define el siguiente subconjunto de enteros:*

$$Pares = \{x \in \mathbb{Z} \mid x = 2n, \text{ para algún } n \in \mathbb{Z}\}.$$

Demostrar la validez del siguiente enunciado: "el producto de dos números pares es un número par". ¿Por dónde comenzamos? Primero, entendamos que significa la definición de enteros pares: si nos dicen que x es un entero par, entonces x puede escribirse como $2n$ para algún entero n, en otras palabras, x es un múltiplo de 2. Segundo, es conveniente reescribir el enunciado de la siguiente forma: "Para cualesquiera enteros a y b, si a y b son pares, entonces ab es par". Tomamos como punto de partida que tenemos dos números pares arbitrarios, a y b y procedemos como sigue: como a y b son pares, existen enteros m y n tales que $a = 2m$ y $b = 2n$, por lo tanto,

$$ab = (2m)(2n) = 2(2mn)$$

y tenemos que ab también es un múltiplo de 2, con lo cual es par y concluimos la demostración. ■ *(A partir de este punto utilizaremos este pequeño cuadrado para denotar el final de una demostración).*

Notemos que en la demostración anterior probamos que cuando el enunciado “a y b son enteros pares” es verdadero, entonces el enunciado “ab es un entero par” también lo es. Decimos que “ab es un entero par” es una condición necesaria para que a y b sean enteros pares, pues si ab fuese impar, a y b no podrían ser pares. No obstante, “ab es un entero par” no es una condición suficiente para garantizar que a y b sean pares, por ejemplo, podríamos tener: $a = 2$, $b = 3$ por lo que $ab = (2)(3) = 6$, que es par; sin embargo, $b = 3$ no lo es. Equivalentemente, hemos probado que el enunciado “si a y b son enteros tales que su producto ab es un número par, entonces a y b son pares” es falso, ya que proporcionamos un contraejemplo para dicho enunciado.

En general, dados dos enunciados, digamos X y Y, el enunciado condicional “si X entonces Y” es válido, si la suposición de que X es verdadero fuerza a que Y también lo sea. Esto último puede interpretarse como “X es suficiente para Y” o como “Y es necesario para X” (ya que si Y fuese falso, entonces X también tendría que serlo).

En ocasiones se requiere demostrar la equivalencia de dos enunciados matemáticos, digamos X y Y. Para este propósito se utiliza el siguiente enunciado compuesto:

$$X \text{ si y sólo si } Y.$$

Esto es simplemente una forma abreviada de poner

$$\text{Si } X, \text{ entonces } Y \text{ y si } Y, \text{ entonces } X.$$

Para que este enunciado compuesto sea válido, bajo cualquier escenario los enunciados X y Y deben ser ambos verdaderos o ambos falsos. En esta situación, la validez de X es una condición necesaria y suficiente para la validez de Y.

Ejemplo 1.6 *Para este ejemplo se define el conjunto de los enteros impares (o nones) como sigue:*

$$\mathit{Impares} = \{x \in \mathbb{Z} \mid x = 2n + 1, \text{ para algún } n \in \mathbb{Z}\}.$$

Demostrar que para todo entero x se tiene que $x + 2$ es impar si y sólo si $x + 5$ es par. Probemos primero “Si $x + 2$ es impar, entonces $x + 5$ es par. Como punto de partida, tenemos un número entero arbitrario x tal que $x + 2$ es impar, es decir, podemos escribir a $x + 2$ como

$$x + 2 = 2n + 1$$

para algún entero n. Nuestra meta es demostrar que $x + 5$ es par. Para este efecto, sumamos 3 en ambos lados de la expresión anterior para obtener

$$x + 2 + 3 = 2n + 1 + 3,$$

por lo tanto

$$x + 5 = 2n + 4 = 2(n + 2).$$

Esto nos dice que $x + 5$ *pertenece al conjunto de enteros pares y concluimos esta parte de la demostración. Falta demostrar que "Si* $x + 5$ *es par, entonces* $x + 2$ *es impar". Procedemos de forma análoga: como* $x + 5$ *es par podemos escribirlo como*

$$x + 5 = 2m$$

para algún entero m. *Restando* 3 *en ambos lados de esta expresión obtenemos*

$$x + 2 = 2m - 3 = 2m - 4 + 1 = 2(m - 2) + 1.$$

De esta última expresión concluimos que $x+2$ *pertenece al conjunto de enteros impares, concluyendo así la demostración.*■

Para concluir con esta sección veremos como demostrar que un enunciado es verdadero, ¡asumiendo que es falso y llegando a una contradicción! Por ejemplo, podríamos demostrar la validez del enunciado compuesto "si X, entonces Y", asumiendo que Y es falso y X verdadero y, a partir de ahí llegar a una contradicción, concluyendo que no podemos asumir que Y sea falso. Por ejemplo, supongamos que queremos demostrar la validez del siguiente enunciado para cualesquiera q y r números reales: "si q es un número racional y r es un irracional, entonces $q - r$ es irracional". Como $q \in \mathbb{Q}$, se tiene que $q = \frac{a}{b}$ con $a, b \in \mathbb{Z}$ y $b \neq 0$. Si asumimos que el enunciado "$q - r$ es irracional" es falso, se tendrá que $q - r \in \mathbb{Q}$ y existirán enteros c y d con $d \neq 0$ tales que,

$$q - r = \frac{c}{d}.$$

Sustituyendo para q se tiene

$$\frac{a}{b} - r = \frac{c}{d}$$

o bien

$$r = \frac{a}{b} - \frac{c}{d} = \frac{ad - bc}{bd} \in \mathbb{Q}.$$

Esto último contradice el supuesto "r es irracional", por lo tanto, el enunciado "$q - r$ es irracional" debe ser verdadero. ■

La validez de la demostración anterior se debe a que el enunciado "si X, entonces Y" es equivalente al enunciado "si no Y, entonces no X", en donde "no X" se refiere a la negación del enunciado X. Es común realizar demostraciones por contradicción cuando se trata de probar la existencia de alguna propiedad u objeto. Dos enunciados clásicos (conocidos y demostrados en la antigua Grecia) cuyas demostraciones se efectúan por contradicción son:

- El número de enteros primos es infinito.
- $\sqrt{2}$ es un número irracional.

El lector interesado puede encontrar las demostraciones en múltiples textos o páginas electrónicas[5]. La idea general de la demostración del primer enunciado es suponer que el conjunto de todos los números primos es finito, digamos N. A partir de ahí se construye un número primo que no pertenece a este conjunto, lo cual es una contradicción. Claramente, dado que no existe una forma recursiva de construir al conjunto de números primos, no puede probarse en forma directa que es infinito.

Para la demostración del segundo enunciado, se parte del hecho de que los números reales son racionales o irracionales y no pueden ser ambos simultáneamente. Asumimos que $\sqrt{2}$ no es irracional, por lo tanto es racional y puede expresarse como

$$\sqrt{2} = \frac{a}{b},$$

para a, b enteros, $b \neq 0$. A partir de aquí se llega a una contradicción y concluimos que $\sqrt{2}$ debe ser irracional. Una vez más, sería imposible hacer la demostración en forma directa ya que tendríamos que escribir a $\sqrt{2}$ con una expresión decimal infinita sin patrón alguno.

Principio de Inducción

En matemáticas surge con frecuencia la necesidad de demostrar que una propiedad es válida para todos los números naturales (o para todos los naturales mayores o iguales que un cierto número). En este caso puede utilizarse el llamado **Principio de Inducción** para demostrar la propiedad en cuestión.

Principio de Inducción Supongamos que S es el subconjunto de números naturales que satisface la propiedad $\mathcal{P}$ y S satisface las siguientes dos características:

a) $1 \in S$.

b) Para cualquier número natural de la forma $n-1$ con $n \geq 2$, si $n-1 \in S$, entonces $n \in S$.

[5] Por ejemplo en:
http://artofproblemsolving.com/wiki/index.php?title=Proof_by_contradiction

Entonces, $S = \mathbb{N}$, es decir, todos los números naturales satisfacen la propiedad $\mathcal{P}$.

La conclusión de este principio es bastante intuitiva en el sentido de que cada número natural, comenzando con el 1, "transfiere"la propiedad $\mathcal{P}$ a su sucesor. Ahora utilizaremos esta técnica para verificar una fórmula que, cuenta la leyenda, fue descubierta por K.F. Gauss cuando era niño.

Ejemplo 1.7 *Mostraremos por inducción que para toda $n \in \mathbb{N}$ se cumple la siguiente igualdad:*

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

De acuerdo al Principio de Inducción, primero debemos verificar que la fórmula se cumple para $n = 1$, lo cual es cierto pues

$$1 = \frac{1(1+1)}{2}.$$

Supongamos ahora que la fómula es válida para $n-1$ y mostraremos que lo mismo sucede para n. Ahora bien, gracias a nuestra hipótesis inductiva tenemos que

$$1 + 2 + \cdots + (n-1) = \frac{(n-1)((n-1)+1)}{2} = \frac{(n-1)(n)}{2}.$$

Sumando n a ambos lados de esta igualdad llegamos a que

$$\begin{aligned} 1 + 2 + \cdots + (n-1) + n &= \frac{(n-1)(n)}{2} + n \\ &= \frac{(n-1)(n) + 2n}{2} \\ &= \frac{n((n-1)+2)}{2} \\ &= \frac{n(n+1)}{2}. \end{aligned}$$

Concluimos pues, que para todo $n \in \mathbb{N}$ la suma de los primeros n naturales es igual a $\frac{n(n+1)}{2}$. ■

El principio de inducción admite muchas variantes. Por ejemplo, si quisieramos probar que una propiedad P es válida para todo número natural $n \geq 2$ (como sucede al final del Ejercicio 1.5), por supuesto sería suficiente verificar que 2 satisface P y la implicación de que si $n-1$ tiene la propiedad P $(n \geq 3)$, entonces n también la tiene. Aprender a realizar demostraciones es algo que requiere de mucha práctica, así que la mejor forma de dominar el tema es “aprendiendo sobre la marcha”.

Ejercicios

Ejercicio 1.1 *Sean A, B y C conjuntos, demostrar que se cumplen las siguientes:*

a. $A \cap B \subset A$ *y* $A \cap B \subset B$.

b. $A \subset A \cup B$ *y* $B \subset A \cup B$.

c. $A \cup (A \cap B) = A$.

d. $(A^c)^c = A$.

e. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

f. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

g. $(A \cup B)^c = A^c \cap B^c$, *ésta es la otra Ley de De Morgan.*

Ejercicio 1.2 *Demostrar que* $\mathbb{I} = \mathbb{R} \setminus \mathbb{Q}$.

Ejercicio 1.3 *Demostrar que para cualquier $a \in \mathbb{Z}$ se cumple*

a. "a es par si y sólo si $a + 1$ es impar".

b. "a es par si y sólo si a^2 es par"

c. El entero $a(a + 1)$ es par.

Ejercicio 1.4 *Demostrar que si $r \in \mathbb{Q}$, $r \neq 0$ y $s \in \mathbb{I}$, entonces $rs \in \mathbb{I}$.*

Ejercicio 1.5 *Utilizar el Principio de Inducción para demostrar las siguientes igualdades.*

a. Para todo $n \in \mathbb{N}$ se cumple

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2.$$

b. Sea $x \in \mathbb{R}$, $x \neq 1$, para todo $n \in \mathbb{N}$ se cumple

$$1 + x + x^2 + \cdots + x^n = \frac{x^{n+1} - 1}{x - 1}.$$

c. Para todo $n \in \mathbb{N}$, $n \geq 2$, se cumple

$$\left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{9}\right)\cdots\left(1 - \frac{1}{n^2}\right) = \frac{n + 1}{2n}.$$

Capítulo 2

Sistemas de ecuaciones lineales

Introducción

Como es bien sabido, una ecuación lineal con coeficientes reales en una variable es una igualdad de la forma

$$ax = b,$$

donde a es el *coeficiente de la variable* x y b es el *término independiente*, ambos números reales. La calificación de "lineal" se refiere a que la variable "x" aparece con exponente igual a uno. Resolver la ecuación se refiere a encontrar los valores reales de x, si es que existen, para los cuales se satisface la igualdad y el conjunto de soluciones se denota comúnmente por S. Para resolver esta ecuación se tienen dos casos posibles.

Caso 1: $a \neq 0$**.** En esta situación es posible multiplicar ambos lados de la ecuación $ax = b$ por $a^{-1} = \frac{1}{a}$, de manera que la igualdad se conserva, obteniendose[1]

$$x = \frac{1}{a}(ax) = \frac{b}{a}.$$

[1]Notemos que a^{-1} existe ya que $a \neq 0$.

Por ejemplo, la solución de la ecuación $4x = 2$ es $x = \frac{2}{4} = \frac{1}{2}$. Así, nuestro conjunto solución es

$$S = \left\{ x \in \mathbb{R} \mid x = \frac{1}{2} \right\},$$

que es un conjunto con un único elemento que puede expresarse simplemente como $S = \left\{ \frac{1}{2} \right\}$.

Caso 2: $a = 0$**.** Aquí existen dos subcasos posibles.

Subcaso 2.1: $b \neq 0$**.** Como ningún número real multiplicado por cero es igual a algo distinto de cero, el conjunto solución asociado es vacío, es decir, $S = \emptyset$. Por ejemplo, este es el caso de la ecuación $0x = 7$.

Subcaso 2.2: $b = 0$**.** Como cualquier número real multiplicado por cero es igual a cero, el conjunto solución de la ecuación $0x = 0$ es el conjunto de todos los números reales, es decir, $S = \mathbb{R}$.

Consideremos ahora ecuaciones lineales con dos variables (al igual que antes, las variables aparecen con exponente uno), esto es, ecuaciones de la forma

$$ax + by = c,$$

donde a es el *coeficiente de la variable* x, b es el *coeficiente de la variable* y y c es el *término independiente*. La resolución de esta ecuación tiene cuatro casos posibles.

Caso 1: $a \neq 0$ y $b \neq 0$. Aquí, al despejar y obtenemos que $y = \frac{c}{b} - \frac{a}{b}x$ y el correspondiente conjunto solución puede visualizarse en el plano cartesiano como una recta (con pendiente $-\frac{a}{b}$ y que corta al eje vertical en el punto $(0, \frac{c}{b})$) que atraviesa ambos ejes del plano.

Caso 2: $a = 0$ y $b \neq 0$. En esta situación obtenemos $0x + by = c$ de manera que $y = \frac{c}{b}$. Por ejemplo, la pareja (x, y) es solución de $0x + 4y = 2$ siempre que $y = \frac{1}{2}$ y sin ninguna restricción sobre x. Un momento de reflexión nos permite percatarnos que, de hecho, x puede tomar cualquier valor por lo que el conjunto solución es

$$S = \left\{ (x, \frac{1}{2}) \mid x \in \mathbb{R} \right\}.$$

Este conjunto puede visualizarse en el plano cartesiano como la recta horizontal $y = \frac{1}{2}$. Es claro que este caso podría pensarse como un caso particular del anterior cuando la pendiente de la recta es igual a cero.

Caso 3: $a \neq 0$ y $b = 0$. Aquí obtenemos $ax + 0y = c$. Despejando x obtenemos que $x = \frac{c}{a}$, por lo que el correspondiente conjunto solución es $S = \{(\frac{c}{a}, y) \mid y \in \mathbb{R}\}$. Este conjunto puede visualizarse en el plano cartesiano como la recta vertical $x = \frac{c}{a}$.

Caso 4: $a = 0$ y $b = 0$**.** Existen dos subcasos posibles.

Subcaso 4.1: $c \neq 0$**.** La ecuación correspondiente es $0x + 0y = c$. y el conjunto solución asociado es $S = \emptyset$.

Subcaso 4.2: $c = 0$**.** Como cualquier pareja ordenada de números reales (x, y) satisface la ecuación $0x + 0y = 0$, el conjunto solución de esta ecuación es $S = \{(x, y) \mid x, y \in \mathbb{R}\}$.

En la Figura 2.1 se representan los conjuntos solución de los casos 1 (si $c/a, c/b > 0$) y 2 (si $c/b > 0$)

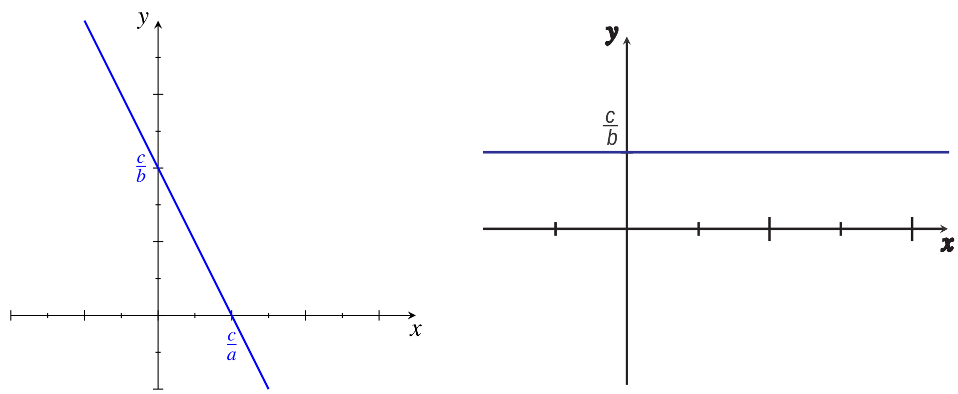


Figura 2.1: Conjuntos solución.

Recordemos ahora que

$$\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$$

es el conjunto de todas las parejas ordenadas cuyas componentes son números reales de manera que, en el Subcaso 4.2, $S = \mathbb{R}^2$. Análogamente,

$$\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$$

se define como el conjunto de todas las tripletas ordenadas cuyas componentes son números reales. En general, tenemos la siguiente definición para cualquier número natural n.

Definición 2.1 $\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_1, x_2, \ldots, x_n \in \mathbb{R}\}$ *es el conjunto de todas las n-tuplas o $n-$adas ordenadas cuyas componentes son números reales. A un elemento de $\mathbb{R}^n$ también se le llama* ***vector*** *de $\mathbb{R}^n$.*

Sistemas de dos por dos

Un sistema con dos ecuaciones lineales y dos variables x, y es un sistema de la forma

$$\begin{cases} ax + by = c, \\ a'x + b'y = c', \end{cases}$$

donde a, b, c, a', b' y c' son números reales. Decimos que un par ordenado (t_1, t_2) es solución del sistema si y sólo si ("**syss**" de aquí en adelante) es solución de cada una de las dos ecuaciones que componen al sistema.

Ejemplo 2.2 *Encontrar el conjunto solución del sistema*

$$\begin{cases} 2x - 2y = 8, & \textit{(I)} \\ -4x + 2y = -6. & \textit{(II)} \end{cases} \tag{A}$$

Si multiplicamos ambos lados de la ecuación I por $\frac{1}{2}$, obtenemos el nuevo sistema

$$\begin{cases} x - y = 4, & \textit{(III)} \\ -4x + 2y = -6. & \textit{(II)} \end{cases} \tag{B}$$

Notemos que los sistemas A y B son equivalentes *en el sentido de que tienen el mismo conjunto solución. De hecho, por la forma en que se obtuvo B, es claro que toda solución de A también satisface B. Para probar que toda solución de B también es solución de A basta multiplicar ambos lados de (III) por $2 = \left(\frac{1}{2}\right)^{-1}$.*

Ahora, comenzando con el sistema B y sumando 4 veces la ecuación (III) a la ecuación (II) obtenemos el sistema

$$\begin{cases} x - y = 4, & \textit{(III)} \\ \quad - 2y = 10. & \textit{(IV)} \end{cases} \tag{C}$$

Dado que

$$(IV) = 4(III) + (II),$$

se sigue que

$$(II) = (IV) - 4(III),$$

de manera que los sistemas B y C también son equivalentes. Pero como A es equivalente a B concluimos que los tres sistemas tienen exactamente el mismo conjunto solución.

Como siguiente paso multiplicamos ambos lados de la ecuación (IV) por $-\frac{1}{2}$ obteniendo el sistema equivalente

$$\begin{cases} x - y = 4, & \textit{(III)} \\ \quad y = -5. & \textit{(V)} \end{cases} \tag{D}$$

Finalmente, sumando una vez [2] *la ecuación (V) a la (III) llegamos a*

$$\begin{cases} x = -1, & \textit{(VI)} \\ y = -5. & \textit{(V)} \end{cases} \tag{E}$$

Este último sistema, cuya solución se lee en forma directa, es equivalente a todos los anteriores. De esta forma, el conjunto solución de todos estos sistemas está dado por

$$S = \{(x, y) \in \mathbb{R}^2 | x = -1, y = -5\}$$

o simplemente, $S = \{(-1, -5)\}$.

Los pasos que hemos empleado en el ejemplo anterior pueden sistematizarse de una forma más efectiva observando que lo relevante no son las variables sino sus coeficientes. Para ello, basta notar que cada ecuación puede representarse por el renglón de sus coeficientes y su término independiente. De esta forma, las operaciones realizadas arriba entre los sistemas pueden visualizarse como sigue:

Ejemplo 2.3

$$\left[\begin{array}{rr|r} 2 & -2 & 8 \\ -4 & 2 & -6 \end{array}\right] \overset{\frac{1}{2}R_1}{\sim} \left[\begin{array}{rr|r} 1 & -1 & 4 \\ -4 & 2 & -6 \end{array}\right] \overset{4R_1+R_2}{\sim} \left[\begin{array}{rr|r} 1 & -1 & 4 \\ 0 & -2 & 10 \end{array}\right] \overset{-\frac{1}{2}R_2}{\sim}$$

$$\left[\begin{array}{rr|r} 1 & -1 & 4 \\ 0 & 1 & -5 \end{array}\right] \overset{1R_2+R_1}{\sim} \left[\begin{array}{rr|r} 1 & 0 & -1 \\ 0 & 1 & -5 \end{array}\right].$$

Cada sistema queda representado por un arreglo con dos renglones de números denotados por R_1 y R_2. El símbolo $\sim$ significa que hay equivalencia entre sistemas, es decir, ambos comparten el mismo conjunto solución. Adicionalmente, kR_i (donde k es un número real distinto del cero) significa que multiplicamos al renglón i por k mientras que $kR_j + R_i$ indica que la equivalencia se obtiene sumando k veces el renglón R_j al renglón R_i. Esta última operación entre renglones se realiza con el fin de llevar al sistema de ecuaciones original a otro equivalente cuya solución es más simple de obtener[3]. Más adelante, en el Ejemplo 2.17, consideraremos una última operación entre renglones de tal forma que con estas tres operaciones seremos capaces de resolver

[2] Evidentemente, esto es lo mismo que sumar (V) a (III), pero queremos presentar esta operación como un caso particular de un método más general.

[3] Notemos que sumar un múltiplo de un renglón a otro es equivalente a sumar el múltiplo de una ecuación a otra. Esto se conoce como **método de eliminación** ya que su propósito es convertir en cero a algún coeficiente, para así "aniquilar" la presencia de una variable dentro de la correspondiente ecuación.

cualquier sistema de ecuaciones lineales (sin importar que tan grande sea éste).

Geométricamente, cada una de las ecuaciones del sistema original (A) representa una recta en $\mathbb{R}^2$. Las dos rectas se cortan en el punto $(-1, -5)$, como se ilustra en la Figura 2.2.

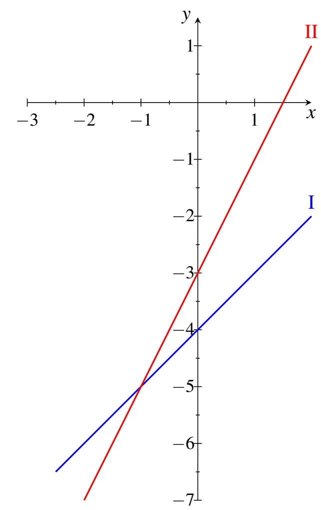


Figura 2.2: Solución del sistema (A).

El lector se preguntará el porqué utilizar el método anterior para resolver el sistema de dos ecuaciones con dos variables. Después de todo, podríamos utilizar el llamado método de sustitución (despejar una variable de una ecuación, sustituirla en la otra y resolver). Sin embargo, la sustitución se vuelve inmanejable cuando se tiene un mayor número de ecuaciones y variables. Afortunadamente, el método que acabamos de introducir se extiende de forma inmediata a los casos más complejos. Para poder formalizarlo, introducimos el concepto de "matriz" en la siguiente sección.

Matrices

Una **matriz** A de $m \times n$ es un arreglo rectangular de mn números reales colocados en m renglones horizontales y n columnas verticales. De esta forma, A puede visualizarse como

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

La notación $A \in M_{m\times n}(\mathbb{R})$ es una abreviatura para decir que A es una matriz de $m \times n$ con entradas o componentes reales. Dado que en este texto dichas entradas siempre serán números reales, utilizaremos la notación simplificada $A \in M_{m\times n}$. En general, las matrices se denotarán por letras mayúsculas: $A, B, C, \ldots$

Definición 2.4 *Dada $A \in M_{m\times n}$, $1 \leq i \leq m$ y $1 \leq j \leq n$, A_i denota al renglón i de A, A^j denota a la columna j de A y a_{ij} denota a la componente que se localiza en el renglón i y en la columna j de A (obsérvese que se utiliza la letra minúscula "a" en concordancia con la letra "A" mayúscula que denota a la matriz). De esta forma,*

$$A_i = \begin{bmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{bmatrix},$$
$$A^j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Formalmente, y para evitar posibles confusiones cuando m o n son números mayores a 9, la componente a_{ij} debería denotarse por $a_{i,j}$. Sin embargo, aquí omitiremos el uso de comas para simplificar la notación.

Ejemplo 2.5 *Sean*

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}, B = \begin{bmatrix} 2 & 6 & 3 \end{bmatrix} \text{ y } C = \begin{bmatrix} 1 & 0 & 6 \\ 7 & 8 & 3 \\ 9 & 6 & 4 \end{bmatrix}.$$

Utilizando la notación que acabamos de introducir tenemos que

$$A \in M_{3\times 2},\ B \in M_{1\times 3} \text{ y } C \in M_{3\times 3}.$$

Además,

$$A_1 = \begin{bmatrix} 1 & 2 \end{bmatrix}, a_{32} = 6, C_2 = \begin{bmatrix} 7 & 8 & 3 \end{bmatrix} \text{ y } C^2 = \begin{bmatrix} 0 \\ 8 \\ 6 \end{bmatrix}.$$

Definición 2.6 *Sean $A, B \in M_{m\times n}$. Decimos que $A = B$ syss ambas matrices coinciden entrada por entrada. Esto es, syss $a_{ij} = b_{ij}$ para toda i y toda j.*

Ejemplo 2.7 *Si*

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

entonces $A \neq B$ (pues $a_{22} \neq b_{22}$) y $C \neq D$ (pues $c_{11} = 2 \neq 3 = d_{11}$).

En lo que sigue haremos un recuento taxonómico de algunas matrices relevantes.

Definición 2.8 *La **matriz cero** $\Theta \in M_{m\times n}$ es la matriz de $m \times n$ cuyas entradas son todas iguales a cero. Por ejemplo, si $\Theta \in M_{2\times 2}$ y $\Theta' \in M_{2\times 3}$ son las correspondientes matrices cero, entonces*

$$\Theta = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ y } \Theta' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Definición 2.9 *Sea $A \in M_{m\times n}$. Decimos que A es una **matriz cuadrada** de orden n syss $m = n$.*

Ejemplo 2.10 *La matriz $\begin{bmatrix} 1 & 0 \\ 7 & 8 \end{bmatrix}$ es cuadrada de orden 2.*

Definición 2.11 *Si*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

*es una matriz cuadrada de orden n, entonces la **diagonal principal** se define como la sucesión $a_{11}, a_{22}, \dots, a_{nn}$.*

Ejemplo 2.12 *La diagonal principal de $\begin{bmatrix} 1 & 0 \\ 7 & 8 \end{bmatrix}$ es 1, 8.*

Definición 2.13 *La **matriz identidad** $I \in M_{n\times n}$ es la matriz cuadrada de orden n tal que todos sus elementos en la diagonal principal son 1 y todos sus elementos fuera de dicha diagonal son cero. Alternativamente, dada $A \in M_{n\times n}$,*

$$A = I \quad \textit{syss} \quad a_{ij} = \begin{cases} 1 & \textit{si } i = j, \\ 0 & \textit{si } i \neq j. \end{cases}$$

Así, la matriz identidad $I \in M_{n\times n}$ se visualiza como

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Definición 2.14 *Decimos que $\vec{v}$ es un **vector** de $\mathbb{R}^n$ syss $\vec{v} \in M_{n\times 1}$ o $\vec{v} \in M_{1\times n}$ es decir, syss $\vec{v}$ es una columna o un renglón con n componentes.*

Aunque $\begin{bmatrix} \pi \\ -3 \end{bmatrix}$ y $\begin{bmatrix} \pi & -3 \end{bmatrix}$ son formalmente objetos distintos (pues el primero es una matriz con dos renglones y una columna mientras que el segundo es una matriz con un renglón y dos columnas), resulta natural decir que ambos coinciden en la primera y segunda componente (tales componentes son respectivamente π y -3). Por ello, los vectores de $\mathbb{R}^n$ pueden representarse como columnas o como renglones y la Definición 2.14 no entra en conflicto con la Definición 2.1.

Representación matricial

Un **sistema lineal** de m ecuaciones y n variables es un conjunto con m igualdades de la forma

$$\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{array}\right. \tag{SL}$$

Decimos que el vector $(t_1, t_2, \ldots, t_n) \in \mathbb{R}^n$ es una solución del sistema SL syss al reemplazar cada variable x_i por el valor t_i se satisfacen las m igualdades que lo componen. Por lo tanto, el conjunto solución es siempre un subconjunto de $\mathbb{R}^n$ (aunque este conjunto puede ser vacío).

Utilizando la notación matricial representamos al sistema anterior como

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

Esta matriz es conocida como **la matriz aumentada del sistema**. El vector

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

es decir, la última columna de la matriz aumentada se denomina el **vector de términos independientes.** Finalmente, la matriz

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

es decir, la submatriz que se obtiene al omitir la última columna de la matriz aumentada se conoce como **la matriz de coeficientes**. Obsérvese que la columna i consiste de los coeficientes de la variable x_i, de manera que debe evitarse cambiar el orden de las columnas para no confundir las variables. La matriz aumentada del sistema puede representarse como

$$(A \mid \vec{b}).$$

Ejemplo 2.15 *La matriz aumentada del sistema*

$$\begin{aligned} w + 3x + 2y + 6z &= 7, \\ w \;+\; 8x \;+\; z &= -1, \\ w + x + y + z &= 1, \end{aligned}$$

es la matriz

$$\left[\begin{array}{cccc|c} 1 & 3 & 2 & 6 & 7 \\ 1 & 8 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{array}\right].$$

Además,

$$\begin{bmatrix} 7 \\ -1 \\ 1 \end{bmatrix} \quad y \quad \begin{bmatrix} 1 & 3 & 2 & 6 \\ 1 & 8 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

son, respectivamente, el vector de términos independientes y la matriz de los coeficientes de este sistema.

Operaciones elementales

Observemos que los sistemas

$$\left\{\begin{aligned} w + 3x + 2y + 6z &= 7, \\ w \;+\; 8x \;+\; z &= -1, \\ w + x + y + z &= 1, \end{aligned}\right. \qquad \text{y} \qquad \left\{\begin{aligned} w + 3x + 2y + 6z &= 7, \\ w + x + y + z &= 1, \\ w \;+\; 8x \;+\; z &= -1, \end{aligned}\right.$$

son equivalentes, por lo que tiene sentido escribir

$$\left[\begin{array}{cccc|c} 1 & 3 & 2 & 6 & 7 \\ 1 & 8 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{array}\right] \overset{R_2 \leftrightarrow R_3}{\sim} \left[\begin{array}{cccc|c} 1 & 3 & 2 & 6 & 7 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 8 & 0 & 1 & -1 \end{array}\right].$$

Así pues, remitiéndonos a la discusión que sigue al Ejemplo 2.2, contamos con tres diferentes tipos de **operaciones elementales** sobre matrices que preservan el conjunto solución entre los sistemas de ecuaciones subyacentes:

Definición 2.16 (operaciones elementales)

1. Una operación elemental de **tipo 1** consiste en a intercambiar dos renglones.

2. Una operación elemental de **tipo 2** consiste en multiplicar un renglón por una constante distinta de cero.

3. Una operación elemental de **tipo 3** consiste en realizar una instancia del **método de eliminación** esto es, en sumar un múltiplo de un renglón a otro.

Ejemplo 2.17 *Utilizando operaciones elementales encontraremos el conjunto solución del sistema*

$$\begin{aligned} 2x + y + 3z &= 0, \\ x + y + z &= 1, \\ 2x + y - z &= 4. \end{aligned}$$

Para ello notamos que

$$\left[\begin{array}{ccc|c} 2 & 1 & 3 & 0 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & 4 \end{array}\right] \overset{R_1 \leftrightarrow R_2}{\sim} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 1 & 3 & 0 \\ 2 & 1 & -1 & 4 \end{array}\right] \overset{-2R_1+R_2}{\sim}$$
$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -1 & 1 & -2 \\ 2 & 1 & -1 & 4 \end{array}\right] \overset{-2R_1+R_3}{\sim} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -1 & 1 & -2 \\ 0 & -1 & -3 & 2 \end{array}\right] \overset{-R_2}{\sim}$$
$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & -1 & -3 & 2 \end{array}\right] \overset{1R_2+R_3}{\sim} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & -4 & 4 \end{array}\right] \overset{-\frac{1}{4}R_3}{\sim}$$
$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & -1 \end{array}\right] \overset{1R_3+R_2}{\sim} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right] \overset{-1R_3+R_1}{\sim}$$
$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right] \overset{-1R_2+R_1}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right].$$

En particular, esta última matriz nos dice que

$$\begin{aligned} 1x + 0y + 0z &= 1, \\ 0x + 1y + 0z &= 1, \\ 0x + 0y + 1z &= -1. \end{aligned}$$

Por lo tanto, nuestro conjunto solución es $S = \{(1, 1, -1)\}$.

El orden y la elección de las operaciones elementales que hemos escogido en el ejemplo anterior no son únicos pero sí tienen una cierta lógica. Esencialmente, intentamos llevar a la matriz a lo que se conoce como su **forma escalonada reducida**. Este concepto será definido formalmente en la Definición 2.23, pero por lo pronto, adelantamos que, para llegar a ella buscamos crear (mediante operaciones elementales de los tres tipos) una escalera descendente que comienza en el primer renglón y cuyos escalones no tienen saltos entre renglones, aunque sí los pueden tener entre columnas, en cuyo caso sólo se exige que la escalera sea lo más estrecha posible.

Utilizamos primero el método de eliminación de arriba hacia abajo y de izquierda a derecha para aniquilar a todas las componentes que están por debajo del inicio de cada uno de los futuros escalones. Las operaciones elementales de tipo dos son empleadas a conveniencia y sirven, entre otras cosas, para asegurarnos que si un renglón tiene al menos una componenete distinta de cero, entonces su primera componente distinta de cero es igual a uno. Este uno es llamado un **uno principal**[4].

Nuestra escalera estará conformada por esta sucesión de unos principales y, con el fin de evitar saltos entre renglones, posiblemente será necesario intercambiarlos. Finalmente, recurrimos al método de eliminación - esta vez de abajo hacia arriba y de derecha a izquiera - con el fin de garantizar que si una columna tiene un uno principal, entonces todas las otras entradas de esa columna son cero.

Concretamente, para una matriz no nula podemos realizar este proceso mediante el siguiente algoritmo.

Algoritmo de Gauss Jordan

1. Localizar la primera columna no nula (de izquierda a derecha) y elegir alguna de sus entradas distinta de 0. Supongamos que dicha entrada está en el $i-$ ésimo renglón.

2. Utilizar la operación elemental de intercambio de renglones para intercambiar el primer renglón con el $i-$ ésimo; de esta forma, la columna del paso anterior se transforma en otra con una entrada diferente de 0 en el primer renglón.

3. Utilizar la operación elemental de multiplicar un renglón por un número distinto de 0 para transformar en 1 a la primera entrada de la columna obtenida en el paso anterior. Este 1 es un uno principal y es el primer elemento (de izquierda a derecha) distinto de 0 que aparece en el primer renglón.

[4]En general, un **elemento principal** se refiere a la primera componente distinta de cero de un renglón, independientemente de si es o no igual a uno.

4. Utilizar el método de eliminación (sumando múltiplos del primer renglón a otros renglones) para aniquilar todas las componentes que están por debajo del elemento principal del primer renglón[5].

5. Repetir los pasos 1 a 4 para la submatriz que se obtiene eliminando el primer renglón de la matriz obtenida en el paso anterior.

6. Repetir el paso anterior hasta que ya no sea posible obtener unos principales.

7. Finalizar el proceso utilizando el método de eliminación para aniquilar todas las entradas por arriba de los unos principales de la matriz.

Consistencia y sistemas homogéneos

Definición 2.18 *Decimos que un sistema lineal con m ecuaciones y n variables es* ***homogéneo*** *syss su vector correspondiente de términos independientes es igual al vector cero de $\mathbb{R}^m$.*

Definición 2.19 *Un sistema lineal es* ***consistente*** *syss tiene al menos una solución. Si el conjunto solución es vacío, entonces diremos que el sistema es* ***inconsistente****.*

Proposición 2.20 *Todo sistema lineal homogéneo es consistente.*

Demostración. Basta notar que si un sistema homogéneo tiene m ecuaciones y n variables, entonces el vector cero de $\mathbb{R}^n$ siempre pertenece al conjunto solución de dicho sistema de manera que este conjunto nunca es vacío. ■

Como el vector cero de $\mathbb{R}^n$ siempre es solución de un sistema homogéneo con n variables, decimos que el vector cero es **la solución trivial** de cualesquiera de estos sistemas. Por supuesto que algunos sistemas homogéneos tienen muchas soluciones (por ejemplo, este es el caso del sistema cuya única ecuación es $x - y = 0$ y cuyo conjunto solución es $\{(x, x) \mid x \in \mathbb{R}\}$) pero algunos otros sólo tienen la solución trivial.

Ejemplo 2.21 *Utilizando la heurística previa a la Definición 2.18, resolveremos el sistema homogéneo*

$$\begin{aligned} 2y + 3z &= 0, \\ 2x - 3y + 2z &= 0, \\ 3x + \frac{1}{2}y - z &= 0. \end{aligned}$$

[5] Los pasos 3 y 4 son intercambiables, de acuerdo a lo que parezca más simple.

Para ello notamos que

$$\left[\begin{array}{rrr|r} 0 & 2 & 3 & 0 \\ 2 & -3 & 2 & 0 \\ 3 & \frac{1}{2} & -1 & 0 \end{array}\right] \overset{R_1 \leftrightarrow R_2}{\sim} \left[\begin{array}{rrr|r} 2 & -3 & 2 & 0 \\ 0 & 2 & 3 & 0 \\ 3 & \frac{1}{2} & -1 & 0 \end{array}\right] \overset{-\frac{3}{2}R_1+R_3}{\sim}$$
$$\left[\begin{array}{rrr|r} 2 & -3 & 2 & 0 \\ 0 & 2 & 3 & 0 \\ 0 & 5 & -4 & 0 \end{array}\right] \overset{-\frac{5}{2}R_2+R_3}{\sim} \left[\begin{array}{rrr|r} 2 & -3 & 2 & 0 \\ 0 & 2 & 3 & 0 \\ 0 & 0 & -\frac{23}{2} & 0 \end{array}\right] \overset{-\frac{2}{23}R_3}{\sim}$$
$$\left[\begin{array}{rrr|r} 2 & -3 & 2 & 0 \\ 0 & 2 & 3 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \overset{-3R_3+R_2}{\sim} \left[\begin{array}{rrr|r} 2 & -3 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \overset{-2R_3+R_1}{\sim}$$
$$\left[\begin{array}{rrr|r} 2 & -3 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \overset{\frac{1}{2}R_2}{\sim} \left[\begin{array}{rrr|r} 2 & -3 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \overset{3R_2+R_1}{\sim}$$
$$\left[\begin{array}{rrr|r} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \overset{\frac{1}{2}R_1}{\sim} \left[\begin{array}{rrr|r} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

Por lo tanto, $x = 0$, $y = 0$, $z = 0$. *Dicho de otra forma, nuestro conjunto solución es trivial ya que*

$$S = \{(0, 0, 0)\}.$$

Forma escalonada reducida

El lector posiblemente ya ha notado que en todos los ejemplos anteriores, la estructura de la última matriz obtenida mediante operaciones elementales, misma que permite inferir fácilmente el correspondiente conjunto solución, tiene ciertas características que resumimos a continuación.

Definición 2.22 *Una matriz está en* ***forma escalonada (FE)*** *syss se cumplen las siguientes tres condiciones:*

1. *Si la matriz tiene renglones que consisten únicamente de ceros, esos renglones están en su parte inferior.*

2. *Si un renglón tiene componentes distintas de cero, entonces la primera (de izquierda a derecha) de esas componentes distintas de cero es igual a uno. Ese uno es llamado un* ***uno principal****.*

3. *Cada uno principal aparece a la derecha y por debajo de los unos principales de los renglones precedentes.*

Definición 2.23 *Una matriz está en* ***forma escalonada reducida (FER)*** *syss satisface las tres condiciones de la Definición 2.22 junto con la siguiente condición adicional:*

4. *Cada uno principal es el único elemento diferente de cero dentro de la columna en la cual se encuentra.*

Observación 2.24 *La tercera condición de la Definición 2.22 implica que si $A \in M_{m\times n}$ está en forma escalonada (reducida), entonces su número de unos principales es menor o igual al mínimo entre m y n.*

Observación 2.25 *Los pasos 1 a 6 del Algoritmo de Gauss - Jordan nos llevan a la forma escalonada de cualquier matriz, adicionalmente, el paso 7 nos lleva a la forma escalonada reducida.*

Ejemplo 2.26 *Las matrices*

$$\begin{bmatrix} 0 & 1 & \pi & 2 & 7 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

están en forma escalonada, sin embargo no están en forma escalonada reducida pues no cumplen la propiedad (4). Las matrices

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

no están en forma escalonada, en el primer caso no se satisface (2) *y en el segundo no se cumple* (1).

Ejemplo 2.27 *Las matrices*

$$A = \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 7 \end{bmatrix}, \; B = \begin{bmatrix} 0 & 1 & 5 & 0 & 7 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \; I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; \Theta = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

están en forma escalonada reducida. Sin embargo, las siguientes matrices no están en forma escalonada reducida:

$$C = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \; D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \; E = \begin{bmatrix} 0 & 1 & 5 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \; F = \begin{bmatrix} 0 & 1 & 8 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

En efecto, C no satisface (1)*, D no satisface* (2)*, E no satisface* (3) *y F no satisface* (4) *(aunque F sí está en forma escalonada).*

Teorema de Gauss - Jordan

El **método de Gauss - Jordan** consiste en transformar (o reducir), mediante los tres tipos de operaciones elementales, una matriz aumentada a su FER. La idea de tener un arreglo rectangular de los coeficientes (tipo matriz) y llegar a la forma escalonada para encontrar el conjunto solución de un sistema era ya conocido en la antigua China desde 179 dC. En el siglo XIX, K. F. Gauss utilizó esta técnica, por lo cual se denomina *eliminación gaussiana* a la obtención de la forma escalonada de una matriz. En 1887, W. Jordan extendió este procedimiento para llegar a la forma escalonada reducida. La eliminación gaussiana es más eficiente en el sentido que contiene aproximadamente la mitad de los pasos que el método de Gauss-Jordan; no obstante, para el desarrollo teórico es conveniente llegar a la FER.

Teorema 2.28 (Gauss-Jordan) *Para cualquier matriz $A \in M_{m\times n}$, existe una única matriz B tal que B está en FER y B se obtiene a partir de A aplicando operaciones elementales. La matriz B se conoce como "la forma escalonada reducida de A".*

Demostración. Se proporciona un esbozo de la demostración para la existencia de la FER, la unicidad se demostrará más adelante en el Capítulo 6.

Sea $A \in M_{m\times n}$, si $A = \Theta$ (la matriz cero), el resultado es obvio pues A es su propia FER. Si $A \neq \Theta$, se hace inducción sobre el número de renglones de A : Si $m = 1$, la FER de A se obtiene multiplicando al único renglón de A por el inverso multiplicativo de su primera componente distinta de cero. Si $m > 1$, tomamos la primera columna (de izquierda a derecha) distinta de cero. Supongamos que se trata de la columna A^j, y sea $a_{ij} \neq 0$ cualquier componente (no nula) de A^j. Realizando la operación elemental $R_1 \longleftrightarrow R_i$ transformamos A^j en la columna

$$\begin{bmatrix} a_{ij} \\ * \\ \vdots \\ * \end{bmatrix}.$$

Realizando la operación elemental $\frac{1}{a_{ij}}R_1$ sobre el primer renglón de A, a_{ij} se transforma en "1". Las componentes $a_{lj} \neq 0$ con $l > 1$ se aniquilan realizando operaciones elementales del tipo $-a_{lj}R_1 + R_l$. En este punto, A^j se ha transformado en la columna

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

Sea $B \in M_{m\times n}$, la matriz que se ha obtenido a partir de A transformando la columna A^j como arriba. Ahora se considera la matriz $\hat{B} \in M_{(m-1)\times n}$ que se obtiene eliminando el primer renglón de B. Por hipótesis de inducción existe la FER de la matriz $\hat{B}$ que puede visualizarse como

$$Fer(\hat{B}) = \begin{array}{c} \qquad\qquad\qquad j+1 \\ \qquad\qquad\qquad \downarrow \\ \begin{bmatrix} 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{bmatrix} \end{array} \in M_{(m-1)\times n}.$$

Realizando las mismas operaciones elementales sobre B que se utilizaron para obtener la FER de $\hat{B}$ obtenemos que la matriz A se ha transformado en una matriz C que se visualiza como sigue:

$$C = \begin{array}{c} \qquad\qquad\qquad j+1 \\ \qquad\qquad\qquad \downarrow \\ \begin{bmatrix} 0 & \cdots & 0 & 1 & c_{1(j+1)} & \cdots & c_{1n} \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{bmatrix} \end{array},$$

en donde la matriz que se obtiene eliminando el primer renglón de C es la FER de $\hat{B}$. Si la $k-\acute{e}sima$ columna de esta FER tiene un uno (principal) en el renglón l (es decir, en el renglón $l+1$ de C) y $c_{1k} \neq 0$, se realiza la operación elemental $-c_{1k}R_{l+1} + R_1$ sobre la matriz C. Una vez realizado esto para todos los casos posibles, la matriz resultante será la FER de A. ∎

Ejemplo 2.29 *Utilizando el método de Gauss-Jordan verificaremos que el sistema*

$$\begin{aligned} x + 2y &= 0, \\ -x - 2y &= 1, \end{aligned}$$

es inconsistente. Para ello, basta notar que

$$\left[\begin{array}{rr|r} 1 & 2 & 0 \\ -1 & -2 & 1 \end{array}\right] \overset{1R_1+R_2}{\sim} \left[\begin{array}{rr|r} 1 & 2 & 0 \\ 0 & 0 & 1 \end{array}\right].$$

El segundo renglón de la matriz nos dice que $0x + 0y = 1$, por lo que este sistema no tiene solución.

Ejemplo 2.30 *Consideremos el sistema*

$$\begin{aligned} 3x - y &= a, \\ x - y &= b, \\ 2x + y &= c. \end{aligned}$$

y supongamos que queremos encontrar todos los valores de a, b y c para los cuales el sistema es consistente. Para ello, obtenemos la FER de la matriz aumentada como sigue:

$$\left[\begin{array}{cc|c} 3 & -1 & a \\ 1 & -1 & b \\ 2 & 1 & c \end{array}\right] \overset{R_1 \leftrightarrow R_2}{\sim} \left[\begin{array}{cc|c} 1 & -1 & b \\ 3 & -1 & a \\ 2 & 1 & c \end{array}\right] \overset{-3R_1+R_2}{\sim}$$
$$\left[\begin{array}{cc|c} 1 & -1 & b \\ 0 & 2 & a-3b \\ 2 & 1 & c \end{array}\right] \overset{-2R_1+R_3}{\sim} \left[\begin{array}{cc|c} 1 & -1 & b \\ 0 & 2 & a-3b \\ 0 & 3 & c-2b \end{array}\right] \overset{\frac{R_2}{2};\frac{R_3}{3}}{\sim}$$
$$\left[\begin{array}{cc|c} 1 & -1 & b \\ 0 & 1 & \frac{a-3b}{2} \\ 0 & 1 & \frac{c-2b}{3} \end{array}\right] \overset{1R_2+R_1;\,-1R_2+R_3}{\sim} \left[\begin{array}{cc|c} 1 & 0 & b+\frac{a-3b}{2} \\ 0 & 1 & \frac{a-3b}{2} \\ 0 & 0 & \frac{c-2b}{3}-\frac{a-3b}{2} \end{array}\right].$$

Observemos que este sistema será consistente syss

$$\frac{c-2b}{3} - \frac{a-3b}{2} = 0$$

o bien syss

$$3a - 5b - 2c = 0.$$

En este caso se tendrá que

$$x = b + \frac{a-3b}{2}, \;\; y = \frac{a-3b}{2}.$$

El ejemplo anterior nos sugiere el siguiente resultado.

Proposición 2.31 *Si $A \in M_{m\times n}$ es la matriz de coeficientes de un sistema de ecuaciones y $m > n$, entonces la FER de A tiene al menos un renglón de ceros.*

Demostración. El número de unos principales en la FER de A es a lo más igual al número de columnas, es decir n, y como $m > n$ los últimos $m - n$ renglones de la FER son renglones de ceros. ■

En algunas ocasiones los sistemas de ecuaciones no sólo son consistentes (i.e., su conjunto solución no es vacío), sino que tienen además un número infinito de soluciones. El método de Gauss Jordan también es capaz de detectar este fenómeno.

Ejemplo 2.32 *Si deseamos encontrar el conjunto solución del sistema*

$$\begin{aligned} x + 2y - 3z &= -4, \\ 2x + y - 3z &= 4, \end{aligned}$$

basta notar que

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & -4 \\ 2 & 1 & -3 & 4 \end{array}\right] \overset{-2R_1+R_2}{\sim} \left[\begin{array}{ccc|c} 1 & 2 & -3 & -4 \\ 0 & -3 & 3 & 12 \end{array}\right] \overset{-\frac{1}{3}R_2}{\sim}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & -4 \\ 0 & 1 & -1 & -4 \end{array}\right] \overset{-2R_2+R_1}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 4 \\ 0 & 1 & -1 & -4 \end{array}\right].$$

De esta forma, obtenemos que $x - z = 4$, $y - z = -4$ *y que* z *es una variable libre (i.e., puede tomar cualquier valor real). Como* $x = 4 + z$ *y* $y = -4 + z$, *obtenemos que el conjunto solución del sistema puede expresarse como*

$$S = \{(4 + z, -4 + z, z) \mid z \in \mathbb{R}\},$$

por lo que existe un número infinito de soluciones (tantas como los números reales). Por ejemplo, haciendo $z = -1, 0, 1$ *se tienen, respectivamente, las siguientes soluciones:*

$$(3, -5, -1),\ (4, -4, 0),\ (5, -3, 1).$$

Definición 2.33 *Supongamos que* L *es un sistema con* m *ecuaciones y* n *variables y que* x *es una de estas variables. Decimos que*

- x *es una* ***variable restringida*** *de* L *syss la columna correspondiente a* x *en la FER de la matriz aumentada contiene un uno principal.*
- x *es una* ***variable libre*** *de* L *syss* x *no es una variable restringida de* L.

A lo largo de este libro y en caso de que existan variables libres, escriberemos al conjunto solución en términos de éstas.

Ejemplo 2.34 *Considérese el sistema*

$$x - y = 0$$

cuya matriz aumentada y FER es

$$\left[\begin{array}{cc|c} 1 & -1 & 0 \end{array}\right].$$

Por lo tanto, x es una variable restringida y y es una variable libre de manera que

$$\begin{aligned} x &= y \\ y &\in \mathbb{R}. \end{aligned}$$

Equivalentemente, el conjunto solución es el siguiente subconjunto de $\mathbb{R}^2$:

$$S = \{(y, y) \mid y \in \mathbb{R}\},$$

mismo que representa una recta. Ahora bien, el sistema

$$x - y + 0z = 0$$

tiene como matriz aumentada y FER a la matriz

$$\left[\begin{array}{ccc|c} 1 & -1 & 0 & 0 \end{array}\right].$$

Por lo tanto, x es una variable restringida y y, z son variables libres de manera que

$$\begin{aligned} x &= y, \\ y, z &\in \mathbb{R}. \end{aligned}$$

El conjunto solución es el subconjunto de $\mathbb{R}^3$ dado por

$$S = \{(y, y, z) \mid y, z \in \mathbb{R}\},$$

mismo que representa un plano[6].

Cardinalidad del conjunto solución

Notamos ahora que el número infinito de soluciones para el sistema

$$\begin{aligned} x + 2y - 3z &= -4, \\ 2x + y - 3z &= 4, \end{aligned}$$

que consideramos en el ejemplo anterior implica un número infinito de soluciones para el **sistema homogéneo asociado** (el que se obtiene a partir del sistema original igualando todos los términos independientes a cero)

$$\begin{aligned} x + 2y - 3z &= 0, \\ 2x + y - 3z &= 0. \end{aligned}$$

[6]En el siguiente capítulo veremos que S puede escribirse como

$$S = \{y(1, 1, 0) + z(0, 0, 1) \mid y, z \in \mathbb{R}\}.$$

En efecto, la columna (o vector) $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ no cambia bajo operaciones elementales; entonces, sin necesidad de volver a realizar las operaciones del ejemplo 2.32, obtenemos que

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 0 \\ 2 & 1 & -3 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right].$$

Observamos que la variable z, misma que estaba libre en el conjunto solución S del sistema original, vuelve a estar libre en el conjunto solución S_H del sistema homogéneo asociado, de hecho

$$S_H = \{(z, z, z) \mid z \in \mathbb{R}\}.$$

Proposición 2.35 *Si un sistema lineal L tiene un número infinito de soluciones, entonces su sistema homogéneo asociado H también tiene un número infinito de soluciones.*

Demostración. Supongamos que nuestro sistema original L tiene m ecuaciones y n variables. Sea $(A \mid \vec{a})$ la matriz aumentada asociada al sistema en donde A y $\vec{a}$ son, respectivamente, la matriz de coeficientes y el vector de términos independientes de L. Supongamos, además, que $(B \mid \vec{b})$ es la FER de $(A \mid \vec{a})$. Por la hipótesis de infinitud, existe al menos una variable libre en esta FER. Notemos ahora que la matriz aumentada del sistema homogéneo asociado H es $(A \mid \vec{0})$, donde $\vec{0}$ es el vector cero de $\mathbb{R}^m$ (o columna con m ceros). Como $\vec{0}$ es invariante bajo operaciones elementales, la FER de $(A \mid \vec{0})$ es $(B \mid \vec{0})$. Por lo tanto, esta segunda FER es consistente y también tiene al menos una variable libre. Esto da lugar a un número infinito de soluciones para H. ■

Dado un sistema lineal, éste es inconsistente o bien tiene al menos una solución. En este segundo caso cabe la posibilidad de que existan o no variables libres. Este sencillo razonamiento justifica el siguiente resultado.

Proposición 2.36 *Si L es un sistema lineal, entonces ocurre una y sólo una de las siguientes alternativas:*

(a) L es inconsistente. Esto es, la FER de L tiene al menos un renglón del tipo

$$[0 \;\; 0 \ldots 0 \mid c],$$

donde c es un real distinto de cero.

(b) L tiene exactamente una solución. Esto es, no sucede (a) y la FER de L no tiene variables libres.

(c) *L tiene un número infinito de soluciones. Esto es, no suceden ni (a) ni (b) y la FER de L tiene al menos una variable libre.*

Ejemplo 2.37 *Encontrar todos los valores h y k tales que el sistema*

$$\begin{aligned} x_1 - 3x_2 &= 1, \\ 2x_1 + hx_2 &= k. \end{aligned}$$

(a) *No tiene solución.*

(b) *Tiene solución única.*

(c) *Tiene un número infinito de soluciones.*

Para resolver los tres casos primero notamos que

$$\left[\begin{array}{cc|c} 1 & -3 & 1 \\ 2 & h & k \end{array}\right] \overset{-2R_1+R_2}{\sim} \left[\begin{array}{cc|c} 1 & -3 & 1 \\ 0 & h+6 & k-2 \end{array}\right].$$

De acuerdo a la Proposición 2.36, este sistema es inconsistente syss

$$h + 6 = 0 \;\; y \;\; k - 2 \neq 0.$$

Así, la solución para (a) es $h = -6$, $k \neq 2$. Con respecto a (b), tenemos que el sistema es consistente syss sus dos variables están restringidas. Pero, para que $h+6$ pueda "convertirse en un uno" deberíamos de poder utilizar la operación elemental $\frac{1}{h+6}R_2$. Para ello, basta que $h + 6 \neq 0$ (i.e, $h \neq -6$)). Finalmente, notamos que (c) ocurre syss (b) y (a) no ocurren. De esta forma, la solución para (c) es $h = -6$ y $k = 2$.

Proposición 2.38 *Si H es un sistema lineal homogéneo con más variables que ecuaciones, entonces H tiene un número infinito de soluciones.*

Demostración. Supongamos que H tiene m ecuaciones y n variables. Por hipótesis $m < n$. Como todo sistema homogéneo es consistente, H debe satisfacer alguna de las dos últimas opciones citadas en la proposición anterior. Ahora bien, como

$$\#\text{variables restringidas de } H = \#\text{unos principales} \leq m < n,$$

H tiene al menos una variable libre y existe un número infinito de soluciones. ■

Observación 2.39 *Si en la proposición anterior el sistema no fuese homogéneo, podría suceder que éste fuese inconsistente de manera que no tendría solución alguna.*

Ejemplo 2.40 *Sin necesidad de cálculos engorrosos, la proposición anterior nos asegura que el sistema homogéneo*

$$\begin{aligned} \pi x + 8.9y + 99z &= 0, \\ 26x + 666y - 3.9z &= 0, \end{aligned}$$

tiene un número infinito de soluciones.

La demostración de la siguiente proposición está autocontenida en el enunciado de la misma

Proposición 2.41 *Supongamos que H es un sistema lineal homogéneo con n ecuaciones y n variables. Las siguientes afirmaciones son equivalentes:*

(a) H tiene solución única.

(b) H no tiene variables libres.[7]

(c) La FER de la matriz de coeficientes de H (la cual es una matriz cuadrada) tiene n unos principales.

(d) La FER de la matriz de coeficientes de H es la matriz identidad $I \in M_{n\times n}$.

Ejemplo 2.42 *Es sencillo notar que el sistema homogéneo*

$$\begin{aligned} x - y + z &= 0, \\ 2x - 2y + z &= 0, \\ 3x - 3y + z &= 0, \end{aligned}$$

tiene un número infinito de soluciones (tómese, por ejemplo, $z = 0$ y $x = y$). Así pues, la proposición anterior garantiza que la FER de la matriz

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

no es la matriz identidad.

[7] Recuérdese que todo sistema homogéneo tiene al menos una solución.

Ejercicios

Ejercicio 2.1 *Encontrar la matriz aumentada, la matriz de coeficientes y el vector de términos independientes de los siguientes sistemas de ecuaciones.*

$$a.\ \begin{cases} 6w + x = 8, \\ w + \pi x + y = -1, \\ w + 2x + y + z = 1. \end{cases}$$

$$b.\ \begin{cases} 3x + 2y = 7, \\ -x + y = -1, \\ 3x - \sqrt{2}y = 6.6. \end{cases}$$

Ejercicio 2.2 *Determinar si las siguientes matrices están en forma escalonada (FE). En caso afirmativo, determinar si también están en forma escalonada reducida (FER).*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ B = \begin{bmatrix} 1 & 0 & -7 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},\ C = \begin{bmatrix} 0 & 1 & 5 & 0 & -9 \\ 0 & 0 & 0 & 1 & 6.2 \end{bmatrix},$$

$$D = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},\ E = \begin{bmatrix} 1 & 0 & 9 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},\ F = \begin{bmatrix} 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

Ejercicio 2.3 *Encontrar la FER de las siguientes matrices:*

$$A = \begin{bmatrix} 1 & 1 & 8 \\ 0 & 2 & -2 \\ 4 & 4 & 0 \end{bmatrix},\ B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 3 & 3 & 6 & 9 \end{bmatrix},\ C = \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 & 0 & 3 \\ 2 & 4 & 0 & 1 & 0 & 1 \\ 1 & 2 & -1 & 1 & 1 & 6 \\ 2 & 4 & 1 & 2 & 1 & 5 \end{bmatrix}.$$

Ejercicio 2.4 *Utilizar el ejercicio anterior para identificar las variables libres y las variables restringidas de los siguientes sistemas homogéneos.*

$$a.\ \begin{cases} x + y + 8z = 0, \\ 2y - 2z = 0, \\ 4x + 4y = 0. \end{cases}$$

$$b.\ \begin{cases} w + 2x + 3y + 4z = 0, \\ 5w + 6x + 7y + 8z = 0, \\ 3w + 3x + 6y + 9z = 0. \end{cases}$$

c. $\begin{cases} u + 2v + 3w + x = -1, \\ w + x = 3, \\ 2u + 4v + x = 1, \\ u + 2v - w + x + y = 6, \\ 2u + 4v + w + 2x + y = 5. \end{cases}$

Ejercicio 2.5 *Utilizando el método de Gauss-Jordan encontrar el conjunto solución de los siguientes sistemas de ecuaciones.*

a. $\begin{cases} -2x + 6y = -22, \\ 3x + 9y = -21. \end{cases}$

b. $\begin{cases} -2x + 6y = 0, \\ 3x + 9y = 0. \end{cases}$

c. $\begin{cases} 2x + y + 3z = 21, \\ x + y + z = 8, \\ 2x + y - z = 5. \end{cases}$

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

e. $\begin{cases} -5x + 10y = 10, \\ 2x - 4y = 6. \end{cases}$

f. $\begin{cases} -5x + 10y = 0, \\ 2x - 4y = 0. \end{cases}$

g. $\begin{cases} -5x + 10y = 15, \\ 2x - 4y = -6. \end{cases}$

h. $\begin{cases} w + 2x + 3y + 4z = 1, \\ 5w + 6x + 7y + 8z = 2, \\ 3w + 3x + 6y + 9z = 3. \end{cases}$

i. $\begin{cases} w + 2x + 3y + 4z = 0, \\ 5w + 6x + 7y + 8z = 0, \\ 3w + 3x + 6y + 9z = 0. \end{cases}$

Obsérvese que h) e i) están relacionados con los Ejercicios 2.3 y 2.4.

Ejercicio 2.6 *Determinar si los sistemas de ecuaciones, cuya FER se proporciona a continuación, son consistentes. En dicho caso, encontrar* *las variables libres, las variables restringidas y el conjunto solución si las variables son denotadas por* $x_1, x_2, \ldots$

$$A = \left[\begin{array}{cccc|c} 1 & -2 & 0 & -1 & 1 \\ 0 & 0 & 1 & 2 & \pi \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \quad B = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 3 \end{array}\right],$$

$$C = \left[\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & -2 \end{array}\right], \quad D = \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Ejercicio 2.7 *Probar (en general) o refutar (con un contraejemplo) los siguientes enunciados (en cada caso se debe indicar claramente "verdadero" o "falso").*

a. Si un sistema lineal tiene más variables que ecuaciones, entonces el sistema es consistente.

b. Si un sistema lineal tiene más ecuaciones que variables, entonces el sistema es consistente.

c. Si un sistema lineal tiene más variables que ecuaciones, entonces el sistema es inconsistente.

d. Si un sistema lineal tiene más ecuaciones que variables, entonces el sistema es inconsistente.

e. Si un sistema lineal es consistente y tiene más variables que ecuaciones, entonces el sistema tiene infinitas soluciones.

f. Si un sistema lineal es consistente y tiene más ecuaciones que variables, entonces el sistema tiene infinitas soluciones.

g. Si la FER de la matriz de coeficientes de un sistema lineal es la matriz identidad, entonces ese sistema tiene solución única.

h. Si un sistema lineal tiene solución única, entonces la FER de la matriz de coeficientes de ese sistema es la matriz identidad.

i. Sean $A \in M_{m\times n}$ *y* $\vec{b} \in \mathbb{R}^n$*. Si* $[A \mid \vec{b}]$ *está en forma escalonada reducida, entonces* A *también lo está.*

j. Un sistema lineal homogéneo siempre es consistente.

Ejercicio 2.8 *Encontrar todos los valores de* c *y* k *tales que cada uno de los sistemas:*

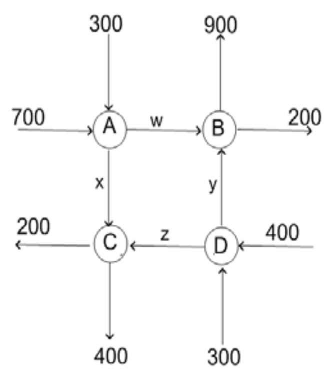
$$i)\ \begin{cases} 2x - cy = 8, \\ 3x + 9y = k. \end{cases} \qquad ii)\ \begin{cases} x + 2y + 3z = k, \\ x + 3y + 4z = -2, \\ x + 4y + cz = -3. \end{cases}$$

a. No tiene solución.

b. Tiene solución única.

c. Tiene infinitas soluciones.

Ejercicio 2.9 *Considerar el diagrama de abajo, en el cual los números representan la cantidad de vehículos por hora que llegan y salen de los cruceros A,B,C y D. Suponer que el tráfico es fluido, de manera que cada hora la cantidad de vehículos que entra a cada crucero es igual al número de vehículos que sale del mismo. Plantear un sistema lineal de ecuaciones para calcular el número de vehículos por hora que transitan entre los cruceros (denotados por w, x, y y z).*



Ejercicio 2.10 *Al caminar por las calles de la ciudad de México es inevitable observar los numerosos puestos de comercio informal: comida, ropa, electrónicos, música, películas de dudosa procedencia y muchos otros objetos diversos. Veamos la contabilidad semanal de José, uno de estos comerciantes informales.*

Para poder instalar su puesto en las inmediaciones de una conocida estación del metro, José permite que diariamente un elemento de la policía revise sus ingresos recaudados y tome una cuarta parte de los mismos; acto seguido, la lideresa del ambulantaje de la zona recauda de José la octava parte del ingreso remanente. Asimismo, al final de cada semana el policía le paga 500 pesos a la lideresa y de esta forma todos conviven tranquilamente.

En una semana dada el ingreso neto (después del pago de cuotas) de José fue de \$2 100. Para esta misma semana, encontrar el ingreso bruto de José y el ingreso neto que obtuvieron el policía y la lideresa (a costa de José).

Capítulo 3

El espacio vectorial $\mathbb{R}^n$

Introducción

El método de Gauss Jordan visto en el capítulo anterior proporciona un algoritmo para la resolución de sistemas de ecuaciones lineales. Sin embargo, una pregunta natural es si se trata simplemente de un algoritmo aislado y puramente mecánico o bien si estos sistemas y sus soluciones forman parte de un universo más amplio. La respuesta va en la segunda dirección y es por eso que ahora se introducirá un caso particular del concepto de espacio vectorial, mismo que constituye uno de los objetos fundamentales en el desarrollo de la matemática.

En general, a los elementos de cualquier espacio vectorial se les llama vectores, en particular $\mathbb{R}^n$ es un espacio vectorial y la Definición 2.1 en la cual $(x_1, x_{2,} \ldots, x_n) \in \mathbb{R}^n$ se denomina un vector, es correcta. La estructura algebraica relevante para los vectores es que éstos pueden sumarse entre sí y multiplicarse por elementos de cierto conjunto de números (los números reales en el caso de $\mathbb{R}^n$). Adicionalmente, estas operaciones cumplen varias propiedades que daremos más adelante.

Como ya hemos visto antes, un vector $\vec{x} \in \mathbb{R}^n$, coincide con una matriz de $n \times 1$ y puede representarse mediante el arreglo

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

En los casos cuando $n = 1, 2$ o 3, también puede visualizarse geométricamente como una "flecha" que va del origen al punto $(x_1, x_{2,} \ldots, x_n)$. La Figura 3.1 muestra esta visualización para los vectores $\vec{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $\vec{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ en $\mathbb{R}^2$.

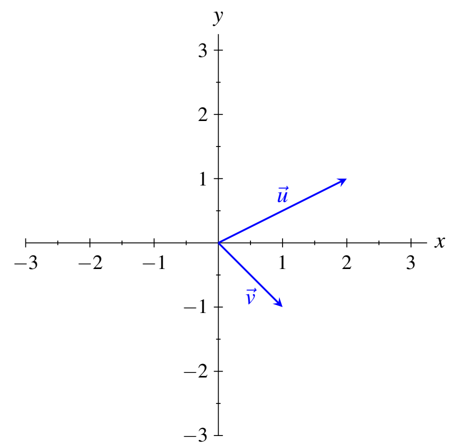


Figura 3.1: Vectores en $\mathbb{R}^2$.

Asimismo, en la Figura 3.2 se muestra el vector vector $\vec{u} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ en $\mathbb{R}^3$.

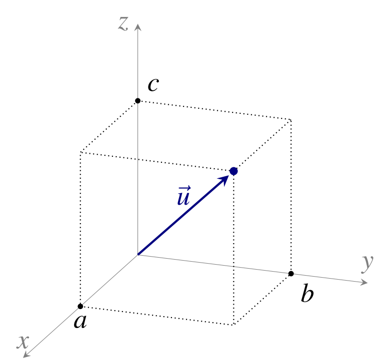


Figura 3.2: Vector en $\mathbb{R}^3$.

A pesar de que no hemos definido formalmente la suma de vectores y la multiplicación de un vector por un número real, la intuición nos dice que estas operaciones corresponden a las nociones de superposición y reescalamiento obvias. Así pues, si $\vec{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $\vec{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, entonces $\vec{u} + \vec{v} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ y $2\vec{v} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ tienen la siguiente interpretación geométrica.

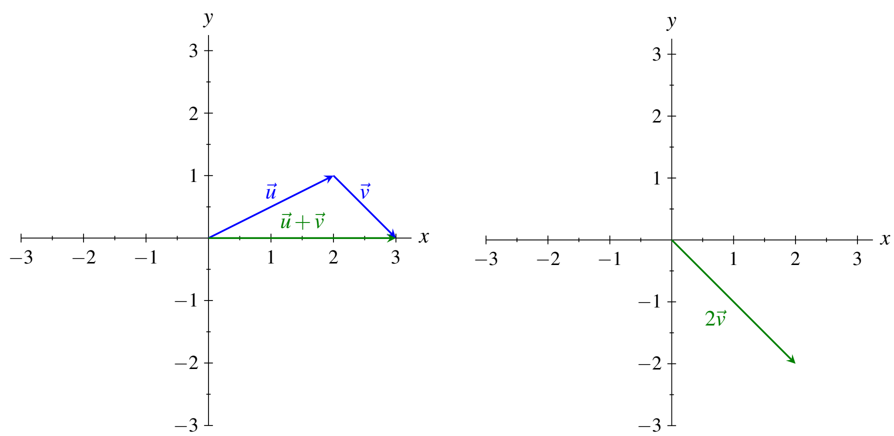


Figura 3.3: Superposición y reescalamiento.

Estructura algebraica de $\mathbb{R}^n$

De acuerdo a la Definición 2.6, dados $\vec{u}$ y $\vec{v}$ vectores de $\mathbb{R}^n$, diremos que $\vec{u}$ es igual a $\vec{v}$ syss coinciden entrada a entrada. Esto es, $\vec{u} = \vec{v}$ syss $u_i = v_i$ para $i = 1, 2, \ldots, n$. En lo que sigue introducimos dos operaciones sobre $\mathbb{R}^n$.

Definición 3.1 *Sean*

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad y \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

*vectores de $\mathbb{R}^n$ y sea c un número real. La **suma** de $\vec{a}$ y $\vec{b}$ (denotada por $\vec{a}+\vec{b}$) y el **producto** de c con $\vec{a}$ (denotado por $c\vec{a}$) se definen como*

$$\vec{a}+\vec{b} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix} \quad y \quad c\vec{a} = \begin{bmatrix} ca_1 \\ ca_2 \\ \vdots \\ ca_n \end{bmatrix},$$

es decir, las operaciones se realizan "coordenada a coordenada".

Observación 3.2 *Se hace énfasis en que el producto que se ha definido, no es entre dos vectores, sino entre un número real c y un vector $\vec{a} \in \mathbb{R}^n$. A partir de ahora, la palabra* ***escalar*** *será sinónimo de número real, por lo que $c\vec{a}$ puede denominarse como el producto escalar de c por el vector $\vec{a}$.*

Notación: en el contexto de $\mathbb{R}^n$ utilizaremos las siguientes convenciones.

- *El vector cero (aquel cuyas componentes son todas cero) es denotado por $\vec{0}$.*
- *$-\vec{a}$ denota al vector $(-1)\vec{a}$.*
- *$\vec{a} - \vec{b}$ significa $\vec{a} + (-\vec{b})$.*

Ejemplo 3.3 *Si $\vec{a} = \begin{bmatrix} 3 \\ 4 \\ 1 \\ -1 \end{bmatrix}$ y $\vec{b} = \begin{bmatrix} 3 \\ -5 \\ 0 \\ 1 \end{bmatrix}$ son vectores en $\mathbb{R}^4$, entonces*

$$2\vec{a} + 3\vec{b} = \begin{bmatrix} 6 \\ 8 \\ 2 \\ -2 \end{bmatrix} + \begin{bmatrix} 9 \\ -15 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 15 \\ -7 \\ 2 \\ 1 \end{bmatrix}.$$

Ejemplo 3.4 *Si $\vec{a} = \begin{bmatrix} -1 \\ 7 \end{bmatrix}, \vec{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ son vectores en $\mathbb{R}^2$, entonces $\vec{a} + \vec{0} = \vec{a}$.*

Ejemplo 3.5 *En el caso de $\mathbb{R}^2$, la suma de vectores puede visualizarse con el llamado método del paralelogramo como se muestra en la Figura 3.4.*

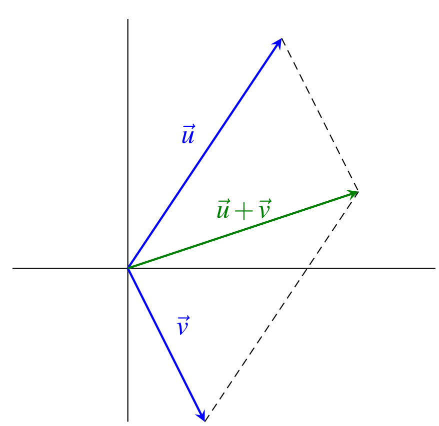


Figura 3.4: Suma de vectores

La siguiente proposición es sencilla (pero tediosa) de demostrar apelando a las propiedades básicas de $\mathbb{R}$.

Proposición 3.6 $\mathbb{R}^n$ *junto con las operaciones de suma entre vectores y de producto escalar que acabamos de definir satisface las siguientes propiedades.*

V1 *Si* $\vec{u}, \vec{v} \in \mathbb{R}^n$, *entonces* $\vec{u}+\vec{v} = \vec{v}+\vec{u}$ *(conmutatividad de la suma).*

V2 *Si* $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$, *entonces* $\vec{u}+(\vec{v}+\vec{w}) = (\vec{u}+\vec{v})+\vec{w}$ *(asociatividad de la suma).*

V3 *Si* $\vec{u} \in \mathbb{R}^n$, *entonces* $\vec{u}+\vec{0} = \vec{u}$ *(existencia de neutro aditivo).*

V4 *Si* $\vec{u} \in \mathbb{R}^n$, *entonces* $\vec{u}-\vec{u} = \vec{0}$ *(existencia de inversos aditivos).*

V5 *Si* $\vec{u} \in \mathbb{R}^n$, *entonces* $1\vec{u} = \vec{u}$.

V6 *Si* $\vec{u} \in \mathbb{R}^n$ *y* β *y* γ *son escalares, entonces* $(\beta\gamma)\vec{u} = \beta(\gamma\vec{u})$.

V7 *Si* $\vec{u} \in \mathbb{R}^n$ *y* β *y* γ *son escalares, entonces* $(\beta+\gamma)\vec{u} = \beta\vec{u}+\gamma\vec{u}$ *(primera forma de distributividad).*

V8 *Si* $\vec{u}, \vec{v} \in \mathbb{R}^n$ *y* β *es un escalar, entonces* $\beta(\vec{u}+\vec{v}) = \beta\vec{u}+\beta\vec{v}$ *(segunda forma de distributividad).*

Demostración. A manera de ejemplo probamos $(V3)$ y $(V7)$. En el caso de $(V3)$ notamos que

$$\vec{u}+\vec{0} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} u_1+0 \\ u_2+0 \\ \vdots \\ u_n+0 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \vec{u}$$

De forma similar, en el caso de $(V7)$ tenemos que

$$(\beta+\gamma)\vec{u} = \begin{bmatrix} (\beta+\gamma)u_1 \\ (\beta+\gamma)u_2 \\ \vdots \\ (\beta+\gamma)u_n \end{bmatrix} = \begin{bmatrix} \beta u_1+\gamma u_1 \\ \beta u_2+\gamma u_2 \\ \vdots \\ \beta u_n+\gamma u_n \end{bmatrix}$$

$$= \begin{bmatrix} \beta u_1 \\ \beta u_2 \\ \vdots \\ \beta u_n \end{bmatrix} + \begin{bmatrix} \gamma u_1 \\ \gamma u_2 \\ \vdots \\ \gamma u_n \end{bmatrix} = \beta \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \gamma \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \beta\vec{u}+\gamma\vec{u},$$

con lo cual $(V7)$ queda demostrada. ■

A partir de estas ocho propiedades, conocidas como **propiedades de espacio vectorial**, se deducen otros resultados cada vez más complejos. A continuación se proporcionan algunos ejemplos.

Proposición 3.7 (Ley de la cancelación) *Si $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^n$ y $\vec{a} + \vec{b} = \vec{a} + \vec{c}$, entonces $\vec{b} = \vec{c}$.*

Demostración. Supongamos $\vec{a} + \vec{b} = \vec{a} + \vec{c}$. Sumando $-\vec{a}$ en ambos lados de esta igualdad y aplicando sucesivamente las propiedades de asociatividad, del inverso y del neutro aditivo, obtenemos

$$\begin{aligned} -\vec{a} + (\vec{a} + \vec{b}) &= -\vec{a} + (\vec{a} + \vec{c}), \\ (-\vec{a} + \vec{a}) + \vec{b} &= (-\vec{a} + \vec{a}) + \vec{c}, \\ \vec{0} + \vec{b} &= \vec{0} + \vec{c}, \\ \vec{b} &= \vec{c}, \end{aligned}$$

concluyendo así la demostración. ■

Corolario 3.8 *Si $\vec{a}, \vec{b} \in \mathbb{R}^n$ y $\vec{a} + \vec{b} = \vec{a}$, entonces $\vec{b} = \vec{0}$.*

Demostración. Por hipótesis tenemos que $\vec{a} + \vec{b} = \vec{a}$, equivalentemente $\vec{a} + \vec{b} = \vec{a} + \vec{0}$. Por la ley de la cancelación concluimos que $\vec{b} = \vec{0}$. ■

Proposición 3.9 *Si $\vec{u} \in \mathbb{R}^n$ y β es un escalar, entonces $\beta\vec{u} = \vec{0}$ syss $\beta = 0$ o $\vec{u} = \vec{0}$.*

Demostración. Si $\beta = 0$ o $\vec{u} = \vec{0}$, entonces es claro que $\beta\vec{u} = \vec{0}$. Supongamos ahora que

$$\beta\vec{u} = \vec{0}; \tag{$\bigstar$}$$

entonces debemos demostar que $\beta = 0$ o $\vec{u} = \vec{0}$. Si $\beta = 0$, la demostración queda terminada. Supongamos que $\beta \neq 0$ (y, por lo tanto, tiene sentido hablar de $\frac{1}{\beta}$). Multiplicando ambos lados de ($\bigstar$) por $\frac{1}{\beta}$ obtenemos que

$$\vec{u} = 1\vec{u} = (\frac{1}{\beta}\beta)\vec{u} = \frac{1}{\beta}(\beta\vec{u}) = \frac{1}{\beta}\vec{0} = \vec{0}$$

y concluimos la demostración. ■

Combinaciones lineales

Definición 3.10 *Sean $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m, \vec{v}$ vectores de $\mathbb{R}^n$. Decimos que el vector $\vec{v}$ es una **combinación lineal** de los vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$, syss existen escalares $\beta_1, \beta_2, \ldots, \beta_m$ tales que*

$$\vec{v} = \beta_1\vec{u}_1 + \beta_2\vec{u}_2 + \cdots + \beta_m\vec{u}_m.$$

La suma anterior puede expresarse en forma compacta como

$$\vec{v} = \sum_{i=1}^{m}\beta_i\vec{u}_i;$$

sin embargo, para mayor claridad en este libro se favorecerá el uso de la notación extendida.

Ejemplo 3.11 *Consideremos los vectores*

$$\vec{v} = \begin{bmatrix}1\\2\\0\end{bmatrix},\ \vec{u_1} = \begin{bmatrix}0\\-1\\0\end{bmatrix},\ \vec{u_2} = \begin{bmatrix}\frac{1}{2}\\0\\0\end{bmatrix}$$

en $\mathbb{R}^3$. *Claramente,*

$$\vec{v} = \begin{bmatrix}1\\2\\0\end{bmatrix} = -2\begin{bmatrix}0\\-1\\0\end{bmatrix} + 2\begin{bmatrix}\frac{1}{2}\\0\\0\end{bmatrix} = -2\vec{u}_1 + 2\vec{u}_2,$$

por lo que $\vec{v}$ es una combinación lineal de $\vec{u}_1$ y $\vec{u}_2$.

Ejemplo 3.12 *Consideremos a los vectores $\vec{v}, \vec{u}_1, \vec{u}_2$ del ejemplo anterior junto con el vector*

$$\vec{u}_3 = \begin{bmatrix}666\\\pi\\-765.8\end{bmatrix}$$

Como $v = -2\vec{u}_1 + 2\vec{u}_2$, es claro que v es combinación lineal de $\vec{u}_1$, $\vec{u}_2$ y $\vec{u}_3$ (pues $v = -2\vec{u}_1 + 2\vec{u}_2 + 0\vec{u}_3$).

Definición 3.13 *Sean $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ vectores de $\mathbb{R}^n$. Definimos al* ***conjunto generado por*** *$\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$, denotado por span$\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m\}$, como el conjunto de todas las combinaciones lineales de estos m vectores. Esto es,*

$$span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m\} = \{\beta_1\vec{u}_1 + \beta_2\vec{u}_2 + \cdots + \beta_m\vec{u}_m \mid \beta_1, \beta_2, \ldots, \beta_m \in \mathbb{R}\}.$$

Ejemplo 3.14 *Sea $\vec{v} \in \mathbb{R}^n$, entonces, span$\{\vec{v}\} = \{\beta\vec{v} \in \mathbb{R}^n \mid \beta \in \mathbb{R}\}$. Si $\vec{v} = \vec{0}$, este conjunto contiene únicamente al vector $\vec{0}$. Si $\vec{v} \neq \vec{0}$, podemos visualizar a span$\{\vec{v}\}$ como la recta infinita en $\mathbb{R}^n$ que pasa por el origen en la dirección de $\vec{v}$ (ver Figura 3.5).*

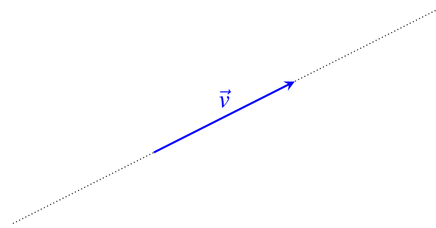


Figura 3.5: $span\{\vec{v}\}$.

Ejemplo 3.15 *Consideremos, de nuevo, a los vectores $\vec{v}, \vec{u}_1, \vec{u}_2 \in \mathbb{R}^3$ del Ejemplo 3.11. Ya sabemos que $\vec{v} \in span\{\vec{u}_1, \vec{u}_2\}$ (pues $v = -2\vec{u}_1 + 2\vec{u}_2$) y que $\vec{0} \in span\{\vec{u}_1, \vec{u}_2\}$ (pues $\vec{0} = 0\vec{u}_1 + 0\vec{u}_2$). Ahora nos gustaría describir, en general, al conjunto $span\{\vec{u}_1, \vec{u}_2\}$. Para ello, notamos que $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in span\{\vec{u}_1, \vec{u}_2\}$ syss existen reales β_1, β_2 tales que*

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \beta_1\vec{u}_1 + \beta_2\vec{u}_2 = \begin{bmatrix} 0 \\ -\beta_1 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{\beta_2}{2} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 + \frac{1}{2}\beta_2 \\ -\beta_1 + 0 \\ 0 + 0 \end{bmatrix}.$$

Como dos vectores son iguales syss coinciden componente a componente, se tiene un sistema de ecuaciones en el cual las variables son los escalares β_1 y β_2. La FER de este sistema se obtiene a partir de su matriz aumentada como sigue.

$$\left[\begin{array}{rr|r} 0 & \frac{1}{2} & a \\ -1 & 0 & b \\ 0 & 0 & c \end{array}\right] \overset{R_1 \leftrightarrow R_2}{\sim} \left[\begin{array}{rr|r} -1 & 0 & b \\ 0 & \frac{1}{2} & a \\ 0 & 0 & c \end{array}\right] \overset{-R_1}{\sim}$$
$$\left[\begin{array}{rr|r} 1 & 0 & -b \\ 0 & \frac{1}{2} & a \\ 0 & 0 & c \end{array}\right] \overset{2R_2}{\sim} \left[\begin{array}{rr|r} 1 & 0 & -b \\ 0 & 1 & 2a \\ 0 & 0 & c \end{array}\right].$$

Por lo tanto, el sistema es consistente (i.e., existen tales β_1 y β_2) syss $c = 0$ y

$$span\{\vec{u}_1, \vec{u}_2\} = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3 \mid c = 0 \right\} = \left\{ \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} \in \mathbb{R}^3 \mid a, b \in \mathbb{R} \right\}.$$

Además,

$$\begin{bmatrix} a \\ b \\ 0 \end{bmatrix} = \beta_1\vec{u}_1 + \beta_2\vec{u}_2 \quad syss \quad \beta_1 = -b, \beta_2 = 2a.$$

Así, en el caso concreto del vector $\begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix}$ *tenemos que* $\begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix} = -4u_1 - 6u_2$.

Ejemplo 3.16 *Para describir span* $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ *en* $\mathbb{R}^2$*, notemos que*

$$\begin{aligned} span\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} &= \left\{ \beta_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \middle| \beta_1, \beta_2 \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \middle| \beta_1, \beta_2 \in \mathbb{R} \right\} = \mathbb{R}^2. \end{aligned}$$

Ejemplo 3.17 *En general, dados dos vectores* $\vec{u}$ *y* $\vec{v}$ *(digamos en* $\mathbb{R}^2$ *o* $\mathbb{R}^3$*), tenemos que span*$\{\vec{u}, \vec{v}\}$ *puede ser una recta o un plano, dependiendo de si los vectores son o no colineales. Esto se ilustra en la Figura 3.6.*

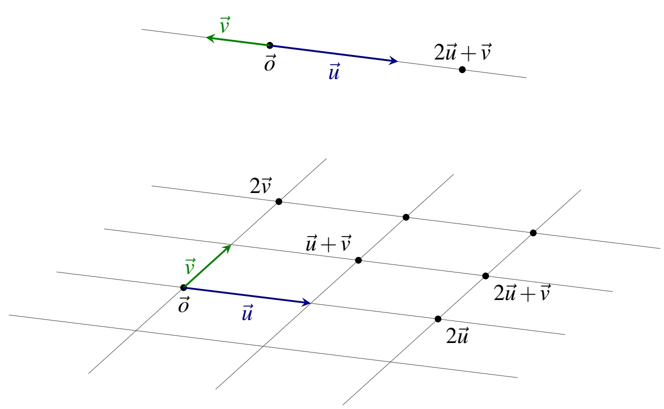


Figura 3.6: Conjuntos generados por dos vectores colineales y no colineales.

Ejemplo 3.18 *Consideremos ahora a los vectores*

$$\vec{v_1} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$$

y supongamos que deseamos describir a los vectores del conjunto span$\{\vec{v}_1, \vec{v}_2\}$*. Para ello notamos que, por definición,*

$$\begin{aligned} span\{\vec{v_1}, \vec{v_2}\} &= \left\{ x_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} x_1 \\ 2x_2 \\ x_1 - x_2 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R} \right\}. \end{aligned}$$

Por lo tanto,

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in span\{\vec{v}_1, \vec{v}_2\}$$

syss existen $x_1, x_2 \in \mathbb{R}$ tales que

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_2 \\ x_1 - x_2 \end{bmatrix}.$$

Como dos vectores son iguales syss coinciden componente a componente se tiene el siguiente sistema de ecuaciones,

$$\begin{aligned} x_1 &= a, \\ 2x_2 &= b, \\ x_1 - x_2 &= c, \end{aligned}$$

cuya matriz aumentada es

$$\left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & 2 & b \\ 1 & -1 & c \end{array}\right].$$

Obtenemos su FER como sigue,

$$\left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & 2 & b \\ 1 & -1 & c \end{array}\right] \overset{-R_1+R_3}{\sim} \left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & 2 & b \\ 0 & -1 & c-a \end{array}\right]$$

$$\overset{R_2 \leftrightarrow R_3}{\sim} \left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & -1 & c-a \\ 0 & 2 & b \end{array}\right] \overset{2R_2+R_3}{\sim} \left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & -1 & c-a \\ 0 & 0 & b+2c-2a \end{array}\right] \overset{-R_2}{\sim}$$
$$\left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & a-c \\ 0 & 0 & b+2c-2a \end{array}\right].$$

Por lo tanto, el sistema es consistente (i.e., existen tales x_1 y x_2) syss

$$b + 2c - 2a = 0.$$

Equivalentemente,

$$b = 2a - 2c.$$

En tal caso, tenemos que la matriz anterior está dada como

$$\left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & a-c \\ 0 & 0 & 0 \end{array}\right].$$

De aquí se obtiene que $x_1 = a, x_2 = a - c$. *Así, en el caso concreto del vector*

$$\vec{w} = \begin{bmatrix} 8 \\ 14 \\ 1 \end{bmatrix},$$

tenemos que $\vec{w} \in span\{\vec{v}_1, \vec{v}_2\}$ *puesto que* $2(8) - 2(1) = 14$, *por lo tanto podemos expresar a* $\vec{w}$ *como*

$$\vec{w} = 8\vec{v}_1 + (8-1)\vec{v}_2 = 8\vec{v}_1 + 7\vec{v}_2.$$

Finalizamos esta sección con tres resultados bastante intuitivos acerca del comportamiento del *span* de un conjunto de vectores.

Proposición 3.19 *Si* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ *son vectores de* $\mathbb{R}^n$*, entonces*

$$\vec{0} \in span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m\}.$$

Demostración. Basta notar que $\vec{0} = 0\vec{u}_1 + 0\vec{u}_2 + \cdots + 0\vec{u}_m$. ■

Proposición 3.20 *Si* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}$ *son vectores de* $\mathbb{R}^n$*, entonces*

$$span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\} \subset span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}\}.$$

Demostración. Sea $\vec{v} \in span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\}$; debemos probar que

$$\vec{v} \in span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}\}.$$

Por definición, existen escalares $\beta_1, \beta_2, \ldots, \beta_p$ tales que

$$v = \beta_1\vec{u}_1 + \beta_2\vec{u}_2 + \cdots + \beta_p\vec{u}_p,$$

por lo tanto,

$$v = \beta_1\vec{u}_1 + \beta_2\vec{u}_2 + \cdots + \beta_p\vec{u}_p + 0\vec{u}_{p+1} \in span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}\}$$

y se concluye la demostración. ■

Antes de pasar al siguiente resultado recordemos que dos conjuntos M y N son iguales syss tienen exactamente los mismos elementos. Así, para demostrar que $M = N$, basta demostrar que todos los elementos de M son elementos de N ($M \subset N$) y que todos los elementos de N son elementos de M ($N \subset M$).

Ejemplo 3.21 *Sean* $\vec{v}_1, \vec{v}_2$ *y* $\vec{v}_3$ *vectores en* $\mathbb{R}^n$ *y* $\vec{v}_3 \in span\{\vec{v}_1, \vec{v}_2\}$. *Entonces se tiene que*

$$span\{\vec{v}_1, \vec{v}_2\} = span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}.$$

En este ejemplo hay que probar la igualdad entre dos conjuntos. La contención $span\{\vec{v}_1, \vec{v}_2\} \subset span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ *es inmediata por la Proposición 3.20. Para probar la otra contención, tomamos cualquier elemento* $\vec{w} \in span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. *Por definición, tenemos que existen escalares* $\gamma_1, \gamma_2, \gamma_3$ *tales que*

$$\vec{w} = \gamma_1\vec{v}_1 + \gamma_2\vec{v}_2 + \gamma_3\vec{v}_3. \qquad (\bigstar)$$

Por hipótesis, $\vec{v}_3 \in span\{\vec{v}_1, \vec{v}_2\}$ *y existen* $\beta_1, \beta_2 \in \mathbb{R}$ *tales que,* $v_3 = \beta_1\vec{v}_1 + \beta_2\vec{v}_2$. *Sustituyendo esta expresión para* v_3 *en* ($\bigstar$) *se obtiene*

$$\vec{w} = \gamma_1\vec{v}_1 + \gamma_2\vec{v}_2 + \gamma_3\left(\beta_1\vec{v}_1 + \beta_2\vec{v}_2\right)$$

y podemos reescribir esta igualdad como

$$\vec{w} = \left(\gamma_1 + \gamma_3\beta_1\right)\vec{v}_1 + \left(\gamma_2 + \gamma_3\beta_2\right)\vec{v}_2,$$

es decir, $\vec{w} \in span\{\vec{v}_1, \vec{v}_2\}$ *y se tiene que* $span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} \subset span\{\vec{v}_1, \vec{v}_2\}$, *con lo cual demostrada la igualdad entre ambos conjuntos.*

El ejemplo anterior genera múltiples aplicaciones como, por ejemplo, el asegurar que

$$span\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\} = span\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}\sqrt{2}\\0\\-\pi\end{bmatrix}\right\}.$$

A continuación proporcionamos una generalización de este tipo de situaciones.

Proposición 3.22 *Si* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}$ *son vectores de* $\mathbb{R}^n$ *y*

$$\vec{u}_{p+1} \in span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\},$$

entonces

$$span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\} = span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}\}.$$

Demostración. Gracias a la Proposición 3.20 sabemos que

$$span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\} \subset span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}\}.$$

Por lo tanto, resta verificar que

$$span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}\} \subset span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\}.$$

Por hipótesis, existen escalares $\beta_1, \beta_2, \ldots, \beta_p$ tales que

$$\vec{u}_{p+1} = \beta_1\vec{u}_1 + \beta_2\vec{u}_2 + \cdots + \beta_p\vec{u}_p.$$

Sea ahora $\vec{v} \in span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{u}_{p+1}\}$, entonces debemos probar que

$$\vec{v} \in span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\}.$$

Por definición, existen escalares $\gamma_1, \gamma_2, \ldots, \gamma_{p+1}$ tales que

$$\vec{v} = \gamma_1\vec{u}_1 + \gamma_2\vec{u}_2 + \cdots + \gamma_{p+1}\vec{u}_{p+1}$$

y por lo tanto,

$$\vec{v} = \gamma_1\vec{u}_1 + \gamma_2\vec{u}_2 + \cdots + \gamma_{p+1}\underbrace{(\beta_1\vec{u}_1 + \beta_2\vec{u}_2 + \cdots + \beta_p\vec{u}_p)}_{\vec{u}_{p+1}} =$$

$$(\gamma_1 + \gamma_{p+1}\beta_1)\vec{u}_1 + (\gamma_2 + \gamma_{p+1}\beta_2)\vec{u}_2 + \cdots + (\gamma_p + \gamma_{p+1}\beta_p)\vec{u}_p$$

y $\vec{v} \in span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\}$. ■

Algunas representaciones útiles

Consideremos el siguiente sistema con m ecuaciones y n variables.

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

Como siempre, denotamos a la matriz correspondiente de coeficientes por A, al vector de términos independientes por $\vec{b}$ y por S al conjunto solución del sistema lineal. Notamos que $(t_1, t_2, \ldots, t_n) \in S$ syss

$$\begin{bmatrix} a_{11}t_1 + a_{12}t_2 + \ldots + a_{1n}t_n \\ a_{21}t_1 + a_{22}t_2 + \ldots + a_{2n}t_n \\ \vdots \\ a_{m1}t_1 + a_{m2}t_2 + \ldots + a_{mn}t_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Esto es, syss

$$\begin{bmatrix} t_1a_{11} \\ t_1a_{21} \\ \vdots \\ t_1a_{m1} \end{bmatrix} + \begin{bmatrix} t_2a_{12} \\ t_2a_{22} \\ \vdots \\ t_2a_{m2} \end{bmatrix} + \ldots + \begin{bmatrix} t_na_{1n} \\ t_na_{2n} \\ \vdots \\ t_na_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

o bien,

$$t_1A^1 + t_2A^2 + \ldots t_nA^n = \vec{b}.$$

De esta forma, hemos probado el siguiente resultado.

Proposición 3.23 *$(t_1, t_2, \ldots, t_n) \in \mathbb{R}^n$ es solución del sistema cuya matriz aumentada es $(A \mid \vec{b})$ syss*

$$t_1A^1 + t_2A^2 + \ldots + t_nA^n = \vec{b}.$$

En particular, dicho sistema es consistente syss $\vec{b} \in span\{A^1, A^2, \ldots, A^n\}$.

Las consideraciones anteriores motivan la siguiente definición.

Definición 3.24 (Producto de una matriz por un vector) *Sea $A \in M_{m\times n}$ y sea*

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

un vector de $\mathbb{R}^n$, en donde n es el número de columnas de A. Definimos el producto de A y $\vec{x}$, denotado por $A\vec{x}$ como

$$A\vec{x} = x_1A^1 + x_2A^2 + \ldots + x_nA^n.$$

Ejemplo 3.25 *Si*

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 6 \\ 4 & 0 \end{bmatrix}, \ \vec{x} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \ \vec{y} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad y \ \ \vec{z} = \begin{bmatrix} 7 \\ 0 \\ 5 \end{bmatrix},$$

entonces

$$A\vec{x} = -1\begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 10 \\ -4 \end{bmatrix},$$
$$A\vec{y} = 0\begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} + 0\begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

y $A\vec{z}$ no está definido.

Utilizando esta nueva operación, podemos traducir la Proposición 3.23 como sigue.

Corolario 3.26 *Supongamos que*

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

es un vector de $\mathbb{R}^n$, pensado como una matriz de $n \times 1$. Las siguientes afirmaciones son equivalentes:

1. $\vec{x}$ *es solución del sistema lineal cuya matriz aumentada es* $(A \mid \vec{b})$.

2. $A\vec{x} = \vec{b}$.

3. $x_1A^1 + x_2A^2 + \ldots + x_nA^n = \vec{b}$.

Así pues, el sistema de ecuaciones

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

también puede reescribirse en **forma matricial** como

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

o alternativamente en **forma vectorial** como

$$x_1A^1 + x_2A^2 + \cdots + x_nA^n = \vec{b},$$

en donde $A^j \in \mathbb{R}^m$.

Ejemplo 3.27 *El sistema de ecuaciones*

$$\begin{aligned} 3x_2 + 6x_3 &= 3, \\ x_1 + x_2 + x_3 &= 0, \end{aligned}$$

puede representarse como

$$x_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

o bien, como

$$\begin{bmatrix} 0 & 3 & 6 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}.$$

El siguiente corolario es una reinterpretación del Corolario 3.26 que sistematiza algunos de los pasos empleados en los Ejemplos 3.15 y 3.18.

Corolario 3.28 *Sean*

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

un vector de $\mathbb{R}^n$ *(pensado como una matriz de* $n \times 1$*),* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n, \vec{v}$ *vectores de* $\mathbb{R}^m$ *y* $B \in M_{m\times n}$ *la matriz cuya columna* B^j *coincide con* $\vec{u}_j$ *para toda* $j = 1, 2, \ldots, n$*. Las siguientes afirmaciones son equivalentes:*

(a) $\vec{x}$ *es solución del sistema cuya matriz aumentada es* $(B \mid \vec{v})$*.*

(b) $x_1\vec{u}_1 + x_2\vec{u}_2 + \ldots + x_n\vec{u}_n = \vec{v}$*.*

Demostración. Gracias al Corolario 3.26 tenemos que $\vec{x}$ es solución del sistema cuya matriz aumentada es $(B \mid \vec{v})$ syss

$$x_1\vec{u}_1 + x_2\vec{u}_2 + \ldots + x_n\vec{u}_n = x_1B^1 + x_2B^2 + \ldots + x_nB^n = \vec{v}$$

con lo cual se concluye la demostración. ■

Ejemplo 3.29 *Encontrar todos los pares ordenados* $(c_1, c_2) \in \mathbb{R}^2$ *tales que*

$$c_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \end{bmatrix}.$$

Para este propósito debemos encontrar todos los pares $(c_1, c_2) \in \mathbb{R}^2$ *que satisfacen*

$$\begin{bmatrix} 1 & 0 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$$

y como

$$\left[\begin{array}{cc|c} 1 & 0 & 1 \\ 4 & 3 & 7 \end{array}\right] \overset{-4R_1+R_2}{\sim} \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 3 & 3 \end{array}\right] \overset{\frac{1}{3}R_2}{\sim} \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right],$$

el conjunto solución es $S = \{(1, 1)\}$*.*

Ejemplo 3.30 *Supongamos que se nos pide determinar si* $\vec{w}$ *es combinación lineal de* $\vec{v}_1, \vec{v}_2, \vec{v}_3$*, donde*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \vec{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Una forma de proceder es analizando el siguiente sistema,

$$x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Para este propósito encontramos la FER de su matriz aumentada como sigue,

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 1 & -1 & 2 & 2 \end{array}\right] \overset{-R_1+R_2}{\sim} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

y observamos que este sistema es inconsistente. Por lo tanto,

$$\vec{w} \notin span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}.$$

Alternativamente, podríamos notar que $\vec{v}_2, \vec{v}_3 \in span\{\vec{v}_1\}$ y por la Proposición 3.22 tenemos que

$$span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = span\{\vec{v}_1\}.$$

Sin embargo,

$$span\{\vec{v}_1\} = \left\{\beta \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbb{R}^2 \mid \beta \in \mathbb{R}\right\} \subsetneq \mathbb{R}^2$$

y claramente,

$$\vec{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \notin span\{\vec{v}_1\} = span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}.$$

Observemos que, geométricamente, $span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es simplemente la recta $y = x$ en el plano xy.

Recordemos que el símbolo $\subsetneq$ significa "contención propia"; esto es, $A \subsetneq B$ quiere decir que A está contenido en B, pero que B contiene elementos que no están en A. El ejemplo anterior muestra que si $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ son vectores de $\mathbb{R}^n$ con $p \geq n$, entonces no es necesariamente cierto que $span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} = \mathbb{R}^n$ (hay algunos casos donde esta igualdad es verdadera y otros, como el que acabamos de ver, donde es falsa). Dos preguntas naturales que abordaremos más adelante son ¿Cuántos vectores son necesarios para generar a todo $\mathbb{R}^n$? y ¿qué deben satisfacer estos vectores?

Ejercicios

Ejercicio 3.1 *Completar la demostración de la Proposición 3.6.*

Ejercicio 3.2 *Probar que si $\vec{v}$ es un vector distinto del vector cero y c_1 y c_2 son dos escalares con $c_1 \neq c_2$, entonces $c_1\vec{v} \neq c_2\vec{v}$.*

Ejercicio 3.3 *Demostrar que si β es un escalar, A es una matriz de $m \times n$ y $\vec{v}$ es un vector de $\mathbb{R}^n$, entonces $\beta(A\vec{v}) = A(\beta\vec{v})$.*

Ejercicio 3.4 *Considerar los vectores*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \; \vec{v}_2 = \begin{bmatrix} -3 \\ 3 \end{bmatrix}, \; \vec{v}_3 = \begin{bmatrix} 5 \\ 10 \end{bmatrix}, \; \vec{v}_4 = \begin{bmatrix} -7 \\ 13 \end{bmatrix}.$$

a. *Determinar si $\vec{v}_4 \in span\{\vec{v}_1, \vec{v}_2\}$. En caso afirmativo, encontrar todos los valores de x_1 y x_2 tales que $\vec{v}_4 = x_1\vec{v}_1 + x_2\vec{v}_2$.*

b. *Determinar si $\vec{v}_4 \in span\{\vec{v}_1, \vec{v}_3\}$. En caso afirmativo, encontrar todos los valores de x_1 y x_2 tales que $\vec{v}_4 = x_1\vec{v}_1 + x_2\vec{v}_3$.*

c. *Determinar si $\vec{v}_4 \in span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. En caso afirmativo, encontrar todos los valores de x_1, x_2 y x_3 tales que $\vec{v}_4 = x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3$.*

d. *Determinar si $\vec{v}_2 \in span\{\vec{v}_1\}$. En caso afirmativo, encontrar todos los valores de x_1 tales que $\vec{v}_2 = x_1\vec{v}_1$.*

e. *Determinar si $\vec{v}_3 \in span\{\vec{v}_1\}$. En caso afirmativo, encontrar todos los valores de x_1 tales que $\vec{v}_3 = x_1\vec{v}_1$.*

Ejercicio 3.5 *Considerar los vectores*

$$\vec{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix}, \; \vec{v}_2 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \; \vec{v}_3 = \begin{bmatrix} 4 \\ -3 \\ 5 \end{bmatrix}, \; \vec{v}_4 = \begin{bmatrix} -5 \\ -3 \\ 10 \end{bmatrix}, \; \vec{v}_5 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

a. *Determinar si $\vec{v}_3 \in span\{\vec{v}_1, \vec{v}_2\}$. En caso afirmativo, encontrar todos los valores de x_1 y x_2 tales que*

$$\vec{v}_3 = x_1\vec{v}_1 + x_2\vec{v}_2.$$

b. *Determinar si $\vec{v}_4 \in span\{\vec{v}_1, \vec{v}_2\}$. En caso afirmativo, encontrar todos los valores de x_1 y x_2 tales que*

$$\vec{v}_4 = x_1\vec{v}_1 + x_2\vec{v}_2.$$

c. Determinar si $\vec{v}_5 \in span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$*. En caso afirmativo, encontrar todos los valores de* x_1*,* x_2 *y* x_3 *tales que*

$$\vec{v}_5 = x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3.$$

d. Determinar si $\vec{v}_5 \in span\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$*. En caso afirmativo, encontrar todos los valores de* x_1*,* x_2 *y* x_3 *tales que*

$$\vec{v}_5 = x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_4.$$

Ejercicio 3.6 *Encontrar todos los valores de* a*,* b *y* c *tales que el vector*

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in span\left\{ \begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -3 \\ 5 \end{bmatrix} \right\}.$$

Ejercicio 3.7 *Encontrar todos los valores de* a*,* b *y* c *tales que el vector*

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in span\left\{ \begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Ejercicio 3.8 *Utilizando el ejercicio anterior y el inciso b) del Ejercicio 3.5 encontrar todos los valores de* a*,* b *y* c *tales que el vector*

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in span\left\{ \begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ -3 \\ 10 \end{bmatrix} \right\}.$$

Ejercicio 3.9 *Calcular* $A\vec{t_1}$ *y* $A\vec{t_2}$*, donde*

$$A = \begin{bmatrix} 2 & 1 & 5 & -7 \\ 4 & 1 & 10 & 2 \\ 6 & 1 & 15 & 7 \end{bmatrix}, \ \vec{t_1} = \begin{bmatrix} \frac{1}{2} \\ -1 \\ \frac{1}{5} \\ 0 \end{bmatrix}, \ \vec{t_2} = \begin{bmatrix} -1 \\ 2 \\ -2 \\ 3 \end{bmatrix}.$$

Ejercicio 3.10 *Representar a los siguientes sistemas mediante una ecuación matricial del tipo* $A\vec{x} = \vec{b}$*.*

a. $\begin{cases} -2x_1 + 6x_2 = -22, \\ 3x_1 + 9x_2 = -21. \end{cases}$

b. $\begin{cases} 2x_1 + x_2 + 3x_3 = 21, \\ x_1 + x_2 + x_3 = 8, \\ 2x_1 + x_2 - x_3 = 5. \end{cases}$

c. $\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 1, \\ 5x_1 + 6x_2 + 7x_3 + 8x_4 = 2, \\ 3x_1 + 3x_2 + 6x_3 + 9x_4 = 3. \end{cases}$

Ejercicio 3.11 *Representar a los sistemas del ejercicio anterior mediante una ecuación vectorial del tipo*

$$x_1 A^1 + \ldots + x_n A^n = \vec{b}.$$

Ejercicio 3.12 *Considerar las matrices*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -5 \end{bmatrix}, \; B = \begin{bmatrix} 2.5 & 15 & -7 \\ \frac{1}{3} & 2 & 0 \end{bmatrix},$$
$$C = \begin{bmatrix} 2 & -7 & -5 \\ 4 & -14 & -10 \end{bmatrix}, \; D = \begin{bmatrix} 2 \\ -\pi \end{bmatrix}.$$

a. Determinar si las columnas de A generan todo $\mathbb{R}^2$. Esto es, determinar si span$\{A^1, A^2\} = \mathbb{R}^2$.

b. Determinar si las columnas de B generan todo $\mathbb{R}^2$. Esto es, determinar si span$\{B^1, B^2, B^3\} = \mathbb{R}^2$.

c. Determinar si las columnas de C generan todo $\mathbb{R}^2$. Esto es, determinar si span$\{C^1, C^2, C^3\} = \mathbb{R}^2$.

d. Determinar si las columnas de D generan todo $\mathbb{R}^2$. Esto es, determinar si span$\{D^1\} = \mathbb{R}^2$.

Ejercicio 3.13 *Considerar las matrices*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 11 & 10 & 9 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 6 & 10 & 9 \end{bmatrix} \; C = \begin{bmatrix} 6 & -4 \\ 7 & -\pi \\ \sqrt{2} & 1 \end{bmatrix}.$$

a. Determinar si las columnas de A generan todo $\mathbb{R}^3$. Esto es, determinar si span$\{A^1, A^2, A^3\} = \mathbb{R}^3$.

b. Determinar si las columnas de B generan todo $\mathbb{R}^3$. Esto es, determinar si span$\{B^1, B^2, B^3\} = \mathbb{R}^3$.

c. Determinar si las columnas de C generan todo $\mathbb{R}^3$. Esto es, determinar si span$\{C^1, C^2\} = \mathbb{R}^3$.

Ejercicio 3.14 *Probar que si $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m \in span\{\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_n\}$, entonces*

$$span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\} \subset span\{\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_n\}.$$

Ejercicio 3.15 *Demostrar que*

$$span\{\vec{v}_1, \vec{v}_2\} = span\{\vec{v}_1 + \vec{v}_2, \vec{v}_1 - \vec{v}_2\}.$$

Ejercicio 3.16 *Probar que si* $c_1, c_2, \ldots, c_m$ *son escalares distintos de cero, entonces*

$$span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\} = span\{c_1\vec{v}_1, c_2\vec{v}_2, \ldots, c_m\vec{v}_m\}.$$

Capítulo 4

Independencia lineal

Introducción

En el capítulo anterior definimos el producto de una matriz $A \in M_{m\times n}$ y de un vector $\vec{t} \in \mathbb{R}^n$ (donde $\vec{t}$ puede pensarse como una matriz de $n \times 1$) como

$$A\vec{t} = t_1A^1 + t_2A^2 + \ldots + t_nA^n.$$

Este producto es, por definición, una combinación lineal de las columnas A^j de la matriz A y sirve, entre otras cosas, para proporcionar una forma alternativa de representar sistemas de ecuaciones lineales.

En este capítulo responderemos a la siguiente pregunta: Dado un sistema lineal consistente, ¿bajo qué condiciones podemos asegurar que hay exactamente una solución? Antes de intentar contestarla analicemos los siguientes ejemplos.

Ejemplo 4.1 *Consideremos al sistema lineal L dado por*

$$\begin{aligned} 3x_1 - x_2 &= 6, \\ x_1 + x_2 + x_3 &= 1, \end{aligned}$$

mismo que puede expresarse en forma matricial como

$$\begin{bmatrix} 3 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}.$$

Supongamos que deseamos, primero encontrar su conjunto solución y segundo, determinar de cuántas formas distintas podemos expresar al vector $\begin{bmatrix}0\\0\end{bmatrix}$ *como combinación lineal de los vectores*

$$\begin{bmatrix}3\\1\end{bmatrix}, \begin{bmatrix}-1\\1\end{bmatrix} \quad y \quad \begin{bmatrix}0\\1\end{bmatrix};$$

es decir, de los vectores columna de la matriz del sistema.

Obtenemos la FER de la matriz aumentada del sistema como sigue:

$$\left[\begin{array}{rrr|r} 3 & -1 & 0 & 6 \\ 1 & 1 & 1 & 1 \end{array}\right] \overset{R_1\leftrightarrow R_2}{\sim} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 1 \\ 3 & -1 & 0 & 6 \end{array}\right] \overset{-3R_1+R_2}{\sim} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 1 \\ 0 & -4 & -3 & 3 \end{array}\right]$$

$$\overset{-\frac{1}{4}R_2}{\sim} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 1 \\ 0 & 1 & \frac{3}{4} & -\frac{3}{4} \end{array}\right] \overset{-R_2+R_1}{\sim} \left[\begin{array}{rrr|r} 1 & 0 & \frac{1}{4} & \frac{7}{4} \\ 0 & 1 & \frac{3}{4} & -\frac{3}{4} \end{array}\right].$$

De aquí que el conjunto solución está dado por

$$S_L = \left\{ \begin{bmatrix} \frac{7}{4} - \frac{1}{4}x_3 \\ -\frac{3}{4} - \frac{3}{4}x_3 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_3 \in \mathbb{R} \right\}.$$

Observemos que este conjunto puede escribirse alternativamente como

$$S_L = \left\{ \begin{bmatrix} \frac{7}{4} \\ -\frac{3}{4} \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -\frac{1}{4} \\ -\frac{3}{4} \\ 1 \end{bmatrix} \in \mathbb{R}^3 \mid x_3 \in \mathbb{R} \right\},$$

mismo que puede visualizarse como una recta en dirección del vector

$$\begin{bmatrix} -\frac{1}{4} \\ -\frac{3}{4} \\ 1 \end{bmatrix}$$

y que pasa por el punto $(\frac{7}{4}, -\frac{3}{4}, 0)$ *en* $\mathbb{R}^3$.

Finalmente, queremos determinar las soluciones para el sistema homogéneo de ecuaciones

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

Este sistema puede reescribirse en forma matricial como

$$\begin{bmatrix} 3 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix},$$

que es simplemente el sistema homogéneo H *asociado a nuestro sistema original. Por la Proposicón 2.35 sabemos que, dado que* L *tiene un* *número infinito de soluciones, H también tiene un número infinito de soluciones. De hecho, el conjunto solución de este sistema homogéneo es*

$$S_H = \left\{ \begin{bmatrix} -\frac{1}{4}x_3 \\ -\frac{3}{4}x_3 \\ x_3 \end{bmatrix} \mid x_3 \in \mathbb{R} \right\},$$

que puede escribirse alternativamente como

$$S_H = \left\{ x_3 \begin{bmatrix} -\frac{1}{4} \\ -\frac{3}{4} \\ 1 \end{bmatrix} \in \mathbb{R}^3 \mid x_3 \in \mathbb{R} \right\}$$

y que representa la recta en dirección de $\begin{bmatrix} -\frac{1}{4} \\ -\frac{3}{4} \\ 1 \end{bmatrix}$ *que pasa por el origen, es decir, S_L y S_H son rectas paralelas en $\mathbb{R}^3$ (ver Figura 4.1). La conclusión es que hay un número infinito de formas distintas de expresar al vector cero como combinación lineal de los tres vectores en cuestión.*

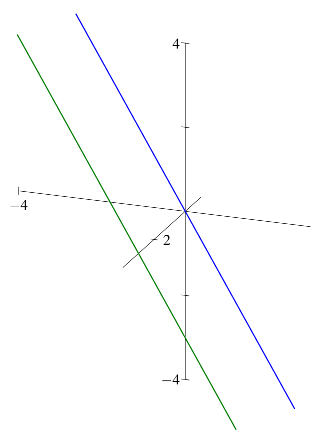


Figura 4.1: S_L y S_H.

El resultado, aparentemente inocuo, de que el vector cero puede ser expresado como combinación lineal de un conjunto de vectores de varias formas, nos lleva a uno de los conceptos más importantes del álgebra lineal: el de independencia (y dependencia) lineal, que abordamos a continuación.

Dependencia e independencia lineal

Definición 4.2 *Sean $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ vectores de $\mathbb{R}^n$. Decimos que el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es **linealmente dependiente (LD)** syss existen escalares $\beta_1, \beta_2, \ldots, \beta_p$, no todos cero**,** tales que*

$$\beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \ldots + \beta_p\vec{v}_p = \vec{0}.$$

*Esto es, $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es linealmente dependiente syss el vector cero de $\mathbb{R}^n$ puede expresarse como una combinación lineal **no trivial** (con algún coeficiente diferente del cero) de estos p vectores.*

Ejemplo 4.3 *Consideremos a los vectores de $\mathbb{R}^3$*

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} -2 \\ -4 \\ 0 \end{bmatrix} \text{ y } \vec{u}_3 = \begin{bmatrix} 666 \\ \pi \\ -1 \end{bmatrix}.$$

Entonces el conjunto $\{\vec{u}_1, \vec{u}_2\}$ es LD pues $2\vec{u}_1 + 1\vec{u}_2 = \vec{0}$. Asimismo, el conjunto $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ también es linealmente dependiente puesto que $2\vec{u}_1 + 1\vec{u}_2 + 0\vec{u}_3 = \vec{0}$.

Ejemplo 4.4 *Como ya hemos visto en el Ejemplo 4.1, existe un número infinito de soluciones para la ecuación*

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

por lo que el conjunto

$$\left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

es LD.

Ejemplo 4.5 *Expresar al vector $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ como combinación lineal de los vectores $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Para este efecto, observamos que la ecuación*

$$x \begin{bmatrix} 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

tiene como única solución $x = 0 = y$ ya que la FER de la matriz

$$\left[\begin{array}{cc|c} 1 & -1 & 0 \\ 1 & 1 & 0 \end{array}\right]$$

está dada por

$$\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right].$$

Concluimos que la combinación lineal trivial

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

es la única posible. En particular, el conjunto de vectores

$$\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}-1\\1\end{bmatrix}\right\}$$

no es LD.

Definición 4.6 *Sean $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ vectores de $\mathbb{R}^n$. Decimos que el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es **linealmente independiente (LI)** syss no es linealmente dependiente, es decir, si $\beta_1, \beta_2, \ldots, \beta_p$ son escalares tales que*

$$\beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \ldots + \beta_p\vec{v}_p = \vec{0},$$

entonces $\beta_1 = \beta_2 = \cdots = \beta_p = 0$.

Proposición 4.7 *Si $\vec{v}$ es un vector de $\mathbb{R}^n$ distinto de cero, entonces el conjunto $\{\vec{v}\}$ es LI.*

Demostración. Supongamos que existe $c \in \mathbb{R}$ tal que $c\vec{v} = \vec{0}$; debemos comprobar que $c = 0$. Pero $c\vec{v} = \vec{0}$ syss $c = 0$ o $\vec{v} = \vec{0}$ y como $\vec{v} \neq \vec{0}$, concluimos que $c = 0$. ■

Ejemplo 4.8 *El conjunto* $\left\{\begin{bmatrix}1\\-2\\3\\-4\end{bmatrix}\right\}$ *en $\mathbb{R}^4$ es LI.*

Observación 4.9 *Si $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ son vectores de $\mathbb{R}^n$ y B es la matriz de $n \times p$ definida de tal forma que la columna B^j coincide con v_j, es decir,*

$$B = [\vec{v}_1\ \vec{v}_2 \cdots \vec{v}_p],$$

entonces las siguientes afirmaciones son equivalentes:

1. *$\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es un conjunto linealmente independiente.*

2. *La única forma de expresar al vector cero como una combinación lineal de $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ es la trivial.*

3. $x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_p\vec{v}_p = \vec{0}$ *implica que el vector* $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}$ *está forzado a ser el vector cero de* $\mathbb{R}^p$.

4. *La ecuación matricial* $B\vec{x} = \vec{0}$ *tiene exactamente una solución (la solución trivial)*

Ejemplo 4.10 *Considérense los vectores de* $\mathbb{R}^3$

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \ y \ \vec{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

y supongamos que deseamos determinar si el conjunto $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ *es linealmente independiente. Para ello, consideramos la matriz de* 3×3, *definida como* $B = [\vec{u}_1 \ \vec{u}_2 \ \vec{u}_3]$ *y al sistema homogéneo* $B\vec{x} = \vec{0}$. *Notamos que la matriz aumentada de este sistema está dada por*

$$\left[\begin{array}{rrr|r} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 1 & 1 & 0 \end{array}\right]$$

y tiene como FER a la matriz

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] = (I \mid \vec{0}).$$

Por lo tanto hay exactamente una solución (la solución trivial) y el conjunto $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ *es LI.*

Observemos que el conjunto de vectores

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

en $\mathbb{R}^2$ y el conjunto análogo,

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

en $\mathbb{R}^3$, tienen, entre otras, las siguientes carácteristicas:

1. Ambos son linealmente independientes.

2. Cualquier vector en $\mathbb{R}^2$ o, alternativamente, en $\mathbb{R}^3$ puede expresarse trivialmente como combinación lineal de ellos.

En efecto, dado cualquier vector $\begin{bmatrix} a \\ b \end{bmatrix}$ en $\mathbb{R}^2$, se tiene que

$$\begin{bmatrix} a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Asimismo si $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$, entonces

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Evidentemente lo anterior puede generalizarse a cualquier $\mathbb{R}^n$ y esto nos lleva a otro concepto muy importante, el de base, mismo que se desarrollará formalmente en el Capítulo 11. Por lo pronto consideramos un caso particular, el de base estándar que se define a continuación.

Definición 4.11 *La **base estándar** para $\mathbb{R}^n$ es el conjunto $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$, donde $\vec{e}_i$ es el vector de $\mathbb{R}^n$ cuya i-ésima entrada es igual a uno y cuyas otras entradas son todas cero. Es decir, podemos visualizar a $\vec{e}_i$ como*

$$i \longrightarrow \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

en donde el "1" aparece en la i-ésima coordenada.

Observación 4.12 *Si $A \in M_{m\times n}$, dado $\vec{e}_i \in \mathbb{R}^n$, entonces $A\ \vec{e}_i = A^i$.*

Proposición 4.13 *La base estándar para $\mathbb{R}^n$ es LI.*

Demostración. Notemos que el conjunto

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\}$$

es LI syss el sistema homogéneo

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

cuya FER es $(I \mid \vec{0})$, tiene solución única, lo cual es obvio. ■

Proposición 4.14 *Sea* $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$, *entonces*

$$\vec{v} \in span\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}.$$

Demostración. En virtud de que

$$\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + \cdots + v_n\vec{e}_n,$$

el resultado es inmediato. ■

Algunos resultados

Observemos que dado el conjunto $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$ de vectores en $\mathbb{R}^3$, se tiene que

$$1 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

de manera que el conjunto es LD, lo cual nos lleva a la siguiente proposición.

Proposición 4.15 *Si uno de los vectores* $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p \in \mathbb{R}^n$ *es el vector cero, entonces el conjunto* $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$ *es LD.*

Demostración. Por conmutatividad de la suma entre vectores podemos suponer que $\vec{v}_1 = \vec{0}$. Entonces

$$1\vec{v}_1 + 0\vec{v}_2 + \dots + 0\vec{v}_p = \vec{0},$$

por lo que es posible expresar al vector cero como una combinación lineal no trivial de $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$. Concluimos pues que el conjunto formado por estos p vectores es LD. ■

Ejemplo 4.16 *Consideremos los vectores*

$$\vec{u}_1 = \begin{bmatrix} 1 \\ \pi \\ 7 \\ -1 \end{bmatrix}, \ \vec{u}_2 = \begin{bmatrix} -7 \\ -1 \\ 0 \\ 6 \end{bmatrix}, \ \vec{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

en $\mathbb{R}^4$*. Como*

$$0\vec{u}_1 + 0\vec{u}_2 + 666\vec{u}_3 = \vec{0},$$

tenemos que el conjunto $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ *es LD. Por supuesto, la elección del escalar distinto del cero que multiplica a* $\vec{u}_3$ *en la igualdad anterior es completamente irrelevante.*

Proposición 4.17 *Si el conjunto* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ *de vectores de* $\mathbb{R}^n$ *es LD y* $\vec{w} \in \mathbb{R}^n$ *es cualquier otro vector, entonces el conjunto*

$$\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p, \vec{w}\}$$

también es LD.

Demostración. Por hipótesis, existen escalares $\beta_1, \beta_2, \ldots, \beta_p$ no todos cero tales que

$$\beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \ldots + \beta_p\vec{v}_p = \vec{0}.$$

Por lo tanto,

$$\beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \ldots + \beta_p\vec{v}_p + 0\vec{w} = \vec{0}$$

es una combinación lineal no trivial (pues al menos uno de los escalares allí involucrados es distinto de cero) y el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p, \vec{w}\}$ es LD. ■

El resultado anterior simplemente nos dice que si un conjunto de vectores es LD, entonces cualquier conjunto que lo contenga también lo será.

Ejemplo 4.18 *Consideremos a los vectores de* $\mathbb{R}^4$,

$$\vec{w}_1 = \begin{bmatrix} 0 \\ 2 \\ 3 \\ 0 \end{bmatrix}, \vec{w}_2 = \begin{bmatrix} 3 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \vec{w}_3 = \begin{bmatrix} 9 \\ 4 \\ 6 \\ 0 \end{bmatrix} \ y \ \vec{w}_4 = \begin{bmatrix} 0.89 \\ 3\pi \\ 7 \\ -79 \end{bmatrix}$$

y supongamos que deseamos determinar si el conjunto formado por estos cuatro vectores es LD o LI. Para ello, primero notamos que $2\vec{w}_1 + 3\vec{w}_2 - \vec{w}_3 = \vec{0}$*, por lo que el conjunto* $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ *es LD. Esto último, junto con la proposición anterior, implican que* $\{\vec{w}_1, \vec{w}_2, \vec{w}_3, \vec{w}_4\}$ *también es LD (de hecho,* $2\vec{w}_1 + 3\vec{w}_2 - \vec{w}_3 + 0\vec{w}_4 = \vec{0}$*).*

Observamos que, en el ejemplo anterior, el conjunto $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ es LD y a partir de la igualdad

$$2\vec{w}_1 + 3\vec{w}_2 - \vec{w}_3 = \vec{0},$$

se sigue que $\vec{w}_3 \in span\{\vec{w}_1, \vec{w}_2\}$ (pues $\vec{w}_3 = 2\vec{w}_1 + 3\vec{w}_2$). De igual manera, $\vec{w}_1 \in span\{\vec{w}_2, \vec{w}_3\}$ (pues $\vec{w}_1 = -\frac{3}{2}\vec{w}_2 + \frac{1}{2}\vec{w}_3$) y $\vec{w}_2 \in span\{\vec{w}_1, \vec{w}_3\}$ (pues $\vec{w}_2 = -\frac{2}{3}\vec{w}_1 + \frac{1}{3}\vec{w}_3$). La siguiente proposición generaliza esta observación.

Proposición 4.19 *Sean $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p, \vec{v}_{p+1}$ vectores de $\mathbb{R}^n$. El conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p, \vec{v}_{p+1}\}$ es LD syss al menos uno de estos vectores es combinación lineal de los otros p vectores.*

Demostración. Si el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p, \vec{v}_{p+1}\}$ es LD. Por definición, existen escalares $\beta_1, \beta_2, \ldots, \beta_p, \beta_{p+1}$, no todos cero, tales que

$$\vec{0} = \beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \ldots + \beta_p\vec{v}_p + \beta_{p+1}\vec{v}_{p+1}.$$

Por conmutatividad de la suma entre vectores podemos suponer que $\beta_{p+1} \neq 0$ y, por lo tanto, existe $\frac{1}{\beta_{p+1}}$y $\vec{v}_{p+1}$ puede despejarse como

$$\vec{v}_{p+1} = -\frac{\beta_1}{\beta_{p+1}}\vec{v}_1 - \frac{\beta_2}{\beta_{p+1}}\vec{v}_2 + \ldots - \frac{\beta_p}{\beta_{p+1}}\vec{v}_p,$$

es decir, $\vec{v}_{p+1} \in span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$.

Supongamos ahora que al menos uno de los $p+1$ vectores del conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p, \vec{v}_{p+1}\}$ es combinación lineal de los otros p vectores. Sin pérdida de generalidad, asumimos que $\vec{v}_{p+1} \in span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$. Entonces existen escalares $\alpha_1, \ldots, \alpha_p$ tales que

$$\alpha_1\vec{v}_1 + \ldots + \alpha_p\vec{v}_p = \vec{v}_{p+1}$$

o lo que es lo mismo,

$$\alpha_1\vec{v}_1 + \ldots + \alpha_p\vec{v}_p + (-1)\vec{v}_{p+1} = \vec{0}.$$

Por lo tanto, el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p, \vec{v}_{p+1}\}$ es LD. ■

Observación 4.20 *La Proposición 4.19 dice que en un conjunto LD, al menos uno de sus elementos es combinación lineal los otros, pero no afirma que éste sea el caso para todos ellos. En el caso del conjunto $\{\vec{w}_1, \vec{w}_2, \vec{w}_3, \vec{w}_4\}$ del Ejemplo 4.18 (mismo que es LD), $\vec{w}_1 \in span\{\vec{w}_2, \vec{w}_3, \vec{w}_4\}$, $\vec{w}_2 \in span\{\vec{w}_1, \vec{w}_3, \vec{w}_4\}$, $\vec{w}_3 \in span\{\vec{w}_1, \vec{w}_2, \vec{w}_4\}$ y sin embargo, $\vec{w}_4 \notin span\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$.*

Corolario 4.21 *Sean $\vec{v}_1, \vec{v}_2$ vectores de $\mathbb{R}^n$. El conjunto $\{\vec{v}_1, \vec{v}_2\}$ es LD syss uno de los vectores es multiplo escalar del otro (combinación lineal del otro). Más aún, si $\vec{v}_1$ y $\vec{v}_2$ son distintos del vector cero, entonces $\{\vec{v}_1, \vec{v}_2\}$ es LD syss cada uno de estos vectores es múltiplo escalar del otro.*

Demostración. La primera parte de este corolario se sigue directamente de la Proposición 4.19. Para la segunda parte supongamos que $\vec{v}_1$ y $\vec{v}_2$ son distintos del vector cero. Entonces $\{\vec{v}_1, \vec{v}_2\}$ es LD syss existe un escalar α tal que

$$\vec{v}_1 = \alpha\vec{v}_2 \;\text{ o }\; \vec{v}_2 = \alpha\vec{v}_1.$$

En cualquier caso α debe ser distinto de cero ya que $\vec{v}_1$ y $\vec{v}_2$ son distintos del vector cero. Se concluye que $\{\vec{v}_1, \vec{v}_2\}$ es LD syss existe un escalar $\alpha \neq 0$ tal que

$$\vec{v}_1 = \alpha\vec{v}_2 \;\text{ y }\; \vec{v}_2 = \frac{1}{\alpha}\vec{v}_1$$

con lo cual concluye la demostración. ■

Ejemplo 4.22 *Consideremos a los vectores*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -2 \\ -4 \end{bmatrix}$$

en $\mathbb{R}^2$. Claramente se tiene que $\vec{v}_2 = -2\vec{v}_1$, por lo que el conjunto $\{\vec{v}_1, \vec{v}_2\}$ es LD.

Ejemplo 4.23 *Consideremos los vectores*

$$\vec{z}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \vec{z}_2 = \begin{bmatrix} -\pi \\ 0 \end{bmatrix}$$

en $\mathbb{R}^2$. Entonces $\{\vec{z}_1, \vec{z}_2\}$ es un conjunto LI ya que no existe un escalar c tal que $c\vec{z}_1 = \vec{z}_2$. En efecto, si tal escalar existiese, tendríamos simultáneamente que $1c = -\pi$ y que $2c = 0$, lo cual es imposible.

Ejemplo 4.24 *Consideremos a los vectores de $\mathbb{R}^4$*

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 0 \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} -2 \\ 56 \\ 8.8 \\ 9 \end{bmatrix} \text{ y } \vec{u}_3 = \begin{bmatrix} 9 \\ 18 \\ 36 \\ 0 \end{bmatrix}.$$

Como $\vec{u}_3$ es claramente un multiplo de $\vec{u}_1$, $\{\vec{u}_1, \vec{u}_3\}$ es LD, por lo que $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ también es LD en virtud de la Proposición 4.17.

Proposición 4.25 *Si el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ en $\mathbb{R}^n$ es LI y $\vec{v}_{p+1} \in \mathbb{R}^n$ es cualquier otro vector con la propiedad de que*

$$\vec{v}_{p+1} \notin span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\},$$

entonces el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p, \vec{v}_{p+1}\}$ también es LI.

Demostración. Supongamos que existen escalares $\beta_1, \beta_2, \ldots, \beta_{p+1}$ tales que

$$\beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \cdots + \beta_p\vec{v}_p + \beta_{p+1}\vec{v}_{p+1} = \vec{0}. \quad (\bigstar)$$

Debemos verificar que $\beta_1 = \cdots = \beta_p = \beta_{p+1} = 0$.

Para ello, comenzamos afirmando que $\beta_{p+1} = 0$. En efecto, si $\beta_{p+1} \neq 0$, entonces $\frac{1}{\beta_{p+1}}$ es un número real y, utilizando la igualdad ($\bigstar$), se sigue que

$$\vec{v}_{p+1} = -\frac{\beta_1}{\beta_{p+1}}\vec{v}_1 - \frac{\beta_2}{\beta_{p+1}}\vec{v}_2 - \cdots - \frac{\beta_p}{\beta_{p+1}}\vec{v}_p,$$

lo cual contradice la hipótesis de que $v_{p+1} \notin span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$. Por lo tanto $\beta_{p+1} = 0$. Esto último junto con la igualdad ($\bigstar$) implican que

$$\beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \cdots + \beta_p\vec{v}_p = \vec{0}.$$

Finalmente, por hipótesis, el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es LI. y concluimos que $\beta_1 = \beta_2 = \cdots = \beta_p = 0$. ∎

Ejemplo 4.26 *Sean*

$$\vec{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} \pi \\ 0 \\ 6 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 1 \\ \sqrt{2} \\ 6 \end{bmatrix}$$

vectores en $\mathbb{R}^3$ y supongamos que deseamos comprobar si el conjunto $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es LI. Existen al menos dos métodos posibles para efectuar tal comprobación.

El primero consiste en verificar que el sistema homogéneo

$$\left[\begin{array}{ccc|c} 2 & \pi & 1 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 6 & 6 & 0 \end{array}\right]$$

tiene solución única, lo cual no haremos aquí.

El segundo utiliza algunas de las propiedades que ya hemos demostrado: es claro que $\vec{v}_1$ no es un multiplo escalar de $\vec{v}_2$, por lo que el conjunto $\{\vec{v}_1, \vec{v}_2\}$ es LI. También es sencillo (puesto que $\vec{v}_1$ y $\vec{v}_2$ tienen ceros en su segunda coordenada y este no es el caso para $\vec{v}_3$) notar que no existen escalares x_1 y x_2 tales que $x_1\vec{v}_1 + x_2\vec{v}_2 = \vec{v}_3$. El resultado anterior asegura que el conjunto $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es LI.

Ahora demostraremos que tres vectores en $\mathbb{R}^2$ siempre forman un conjunto LD. A manera de ejemplo supongamos que los vectores en cuestión son

$$\vec{v_1} = \begin{bmatrix} 1 \\ 77 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 66 \\ 2 \end{bmatrix}, \vec{v_3} = \begin{bmatrix} 96 \\ 2.8 \end{bmatrix}.$$

Para comprobar dicha afirmación debemos verificar que el sistema homogéneo

$$\left[\begin{array}{ccc|c} 1 & 66 & 96 & 0 \\ 77 & 2 & 2.8 & 0 \end{array}\right]$$

tiene una infinidad de soluciones, lo cual se sigue de la Proposición 2.38, pues se trata de un sistema homogéneo con más variables que ecuaciones que tiene al menos una variable libre. Este es el argumento central detrás del siguiente resultado.

Proposición 4.27 *Sean $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ vectores de $\mathbb{R}^n$. Si $m > n$, entonces el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ es LD.*

Demostración. Basta notar que el sistema homogéneo $A\vec{x} = \vec{0}$, en donde la matriz A tiene como columnas a estos m vectores, tiene una infinidad de soluciones. ■

Corolario 4.28 *Sean $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ vectores de $\mathbb{R}^n$.*

- *Si $m > n$, entonces el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ es LD.*
- *Si $m < n$, entonces span$\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\} \subsetneq \mathbb{R}^n$.*

Demostración. La primera parte del corolario es simplemente la Proposición 4.27. Para probar la segunda parte, tomamos la matriz

$$A = [\vec{v}_1 \ \vec{v}_2 \cdots \vec{v}_m]$$

y suponemos que su FER está dada por $\tilde{A}$. Por lo tanto, la Proposición 2.31 nos dice que, como $m < n$, el $n-$ ésimo renglón de $\tilde{A}$ es un renglón de ceros. Si tomamos

$$\vec{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^n,$$

se tiene que $(\tilde{A} \mid \vec{e}_n)$ corresponde a un sistema inconsistente. Sea $\vec{b} \in \mathbb{R}^n$ tal que[1]

$$(A \mid \vec{b}) \sim (\tilde{A} \mid \vec{e}_n);$$

[1] La matriz aumentada $(A \mid \vec{b})$ se obtiene a partir de $(B \mid \vec{e}_n)$ realizando las operaciones elementales inversas que se utilizaron para obtener B a partir de A.

entonces, $\vec{b} \notin span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ y $span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\} \subsetneq \mathbb{R}^n$. ■

De manera metafórica podríamos decir que en el primer caso pecamos por exceso (de vectores) y la penitencia es la dependencia lineal, mientras que en el segundo pecamos por defecto y la penitencia se traduce en no poder generar a todo $\mathbb{R}^n$. En particular contestamos parcialmente la pregunta formulada al final de capítulo anterior: ¿cuántos vectores se necesitan para generar $\mathbb{R}^n$? La segunda parte del Corolario 4.28 nos dice que al menos se requieren n vectores.

Ejemplo 4.29 *En el espíritu del Corolario 4.28 mostraremos que el conjunto* $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ *de dos vectores en* $\mathbb{R}^3$ *no puede generar todo* $\mathbb{R}^3$. *Para este efecto es necesario encontrar* $\vec{b} \in \mathbb{R}^3$ *tal que*

$$\vec{b} \notin span \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Consideramos la matriz cuyas columnas son los vectores del conjunto, es decir

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

Obtenemos su FER como sigue:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \overset{R_2 \leftrightarrow R_3}{\sim} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \overset{-R_1 + R_3}{\sim} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \overset{-R_2 + R_3}{\sim} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \tilde{A}.$$

Notamos que para $\vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^3$ *el sistema cuya matriz aumentada es* $(\tilde{A} \mid \vec{e}_3)$ *es inconsistente. Podemos obtener un sistema equivalente* $(A \mid \vec{b})$, *para algún* $\vec{b} \in \mathbb{R}^3$, *simplemente revirtiendo las operaciones elementales utilizadas para obtener la equivalencia* $A \sim \tilde{A}$. *En efecto,*

$$(\tilde{A} \mid \vec{e}_3) = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \overset{R_2 + R_3}{\sim} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right] \overset{R_1 + R_3}{\sim}$$

$$\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{array}\right] \overset{R_3 \leftrightarrow R_2}{\sim} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{array}\right] = (A \mid \vec{b}).$$

De aquí obtenemos que

$$\vec{b} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \notin span \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

y por lo tanto,

$$span \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\} \subsetneq \mathbb{R}^3.$$

Ser o no ser la matriz identidad, ésa es la cuestión

En el capítulo anterior vimos que el poder establecer si el sistema lineal $A\vec{x} = \vec{b}$ tiene o no solución equivale a determinar si el vector $\vec{b}$ es combinación lineal de las columnas de A. En el caso particular en el cual $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ son vectores de $\mathbb{R}^n$ y

$$A = [\vec{v}_1 \ \vec{v}_2 \cdots \vec{v}_n]$$

es la matriz de $n \times n$ cuyas columnas son estos vectores, existen dos escenarios posibles para la FER de A :

Escenario 1 *La FER de A es la matriz identidad $I \in M_{n\times n}$. Entonces para cualquier vector $\vec{b} \in \mathbb{R}^n$, la matriz aumentada $(A \mid \vec{b})$ del sistema $A\vec{x} = \vec{b}$ es equivalente a $(I \mid \vec{c})$ para algún vector $\vec{c} \in \mathbb{R}^n$. Por lo tanto, este sistema tiene solución única. En particular,*

$$(A \mid \vec{0}) \sim (I \mid \vec{0})$$

y el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ es LI.

Escenario 2 *La FER de A (llamémosle C)* ***no*** *es la matriz identidad. Entonces $(A \mid \vec{0})$ es equivalente a $(C \mid \vec{0})$ y el sistema homogéneo asociado tiene al menos una variable libre (pues C* tiene al menos un renglón de ceros)*. Por lo tanto, el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ es LD.*

En este escenario, podemos apelar a la Proposición 4.19 para concluir que al menos uno de los n vectores en cuestión debe ser combinación lineal de los otros $n-1$. Por ejemplo, si

$$\vec{v}_1 \in span\{\vec{v}_2, \ldots, \vec{v}_n\},$$

entonces la Proposición 3.22 y el Corolario 4.28 implican que

$$span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} = span\{\vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n\} \subsetneq \mathbb{R}^n.$$

Estos dos escenarios se resumen en los siguientes dos corolarios:

Corolario 4.30 *Sean $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ vectores de $\mathbb{R}^n$. Las siguientes afirmaciones son equivalentes.*

1. *El conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ es LI.*
2. $span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} = \mathbb{R}^n$.
3. *Si $A = [\vec{v}_1\ \vec{v}_2 \cdots \vec{v}_n] \in M_{n\times n}$, entonces $A\vec{x} = \vec{b}$ tiene una solución única para cualquier $\vec{b} \in \mathbb{R}^n$.*

Corolario 4.31 *Sean $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ vectores de $\mathbb{R}^n$ y*

$$V = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}.$$

Las siguientes afirmaciones son equivalentes.

1. *El conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ es LD.*
2. $V \subsetneq \mathbb{R}^n$.
3. *Si $A = [\vec{v}_1\ \vec{v}_2 \cdots \vec{v}_n] \in M_{n\times n}$, entonces $A\vec{x} = \vec{b}$ es un sistema inconsistente si $\vec{b} \notin V$ y consistente cuando $\vec{b} \in V$. En este último caso el sistema tiene una infinidad de soluciones tal y como veremos en el Corolario 4.40.*

Ejemplo 4.32 *Supongamos que deseamos determinar si el sistema*

$$\begin{aligned} x &= b_1, \\ 2x + y &= b_2 \end{aligned}$$

tiene solución para cualquier elección de b_1 y b_2. Esto es, nos preguntamos si span $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\}$ es o no todo $\mathbb{R}^2$. Para ello, basta observar que el vector $\begin{bmatrix}1\\2\end{bmatrix}$ no es multiplo escalar de $\begin{bmatrix}0\\1\end{bmatrix}$, por lo tanto, el conjunto $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\}$ es LI y por el Corolario 4.30, el sistema tiene solución única.

Ejemplo 4.33 *Sean*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \ y \ \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

vectores en $\mathbb{R}^3$*. Nos preguntamos si* $span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ *es o no todo* $\mathbb{R}^3$*. Para ello, basta notar que la matriz*

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

cuyas columnas son estos vectores, tiene como FER a la matriz identidad. Por lo tanto, $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ *es un conjunto LI* [2] *y* $span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \mathbb{R}^3$*.*

Ejemplo 4.34 *Consideremos a los vectores* $\vec{v}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ *y* $\vec{v}_2 = \begin{bmatrix} -3 \\ 6 \end{bmatrix}$ *de* $\mathbb{R}^2$*, mismos que son LD pues* $\vec{v}_2 = -3\ \vec{v}_1$*. Si*

$$A = \begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix} \in M_{2\times 2}$$

y $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \in \mathbb{R}^2$*, es fácil verificar que la FER del sistema* $A\vec{x} = \vec{b}$ *está dada por*

$$\left[\begin{array}{cc|c} 1 & -3 & b_1 \\ 0 & 0 & b_2 + 2b_1 \end{array}\right];$$

con lo cual, el sistema es inconsistente siempre y cuando $b_2 + 2b_1 \neq 0$*. En cambio, si* $b_2 + 2b_1 = 0$*, el sistema tiene una infinidad de soluciones. Concretamente, en este caso el conjunto solución está dado por*

$$S = \left\{ \begin{bmatrix} b_1 + 3y \\ y \end{bmatrix} \in \mathbb{R}^2 \mid y \in \mathbb{R} \right\}.$$

Finalmente,

$$\begin{aligned} span\{\vec{v}_1, \vec{v}_2\} &= \left\{ \begin{bmatrix} b_1 \\ -2b_1 \end{bmatrix} \in \mathbb{R}^2 \mid b_1 \in \mathbb{R} \right\} \\ &= \left\{ b_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} \in \mathbb{R}^2 \mid b_1 \in \mathbb{R} \right\}. \end{aligned}$$

---

[2] Una forma alternativa de determinar la independencia lineal del este conjunto es notando que $\vec{v}_2$ no es multiplo escalar de $\vec{v}_1$ y que $\vec{v}_3 \notin span\{\vec{v}_1, \vec{v}_2\}$.

El caso general

En esta última sección probaremos que si un sistema lineal con m ecuaciones y n variables es consistente, entonces el sistema tiene solución única syss las columnas de la matriz de coeficientes forman un conjunto LI. Para este fin, se encontrará una relación entre el conjunto solución del sistema lineal y el conjunto solución del sistema homogéneo asociado.

En el Ejemplo 2.32 del Capítulo 2 observamos que el conjunto solución del sistema

$$\begin{aligned} x + 2y - 3z &= -4, \\ 2x + y - 3z &= 4 \end{aligned}$$

está dado por

$$S = \left\{ \begin{bmatrix} 4+z \\ -4+z \\ z \end{bmatrix} \mid z \in \mathbb{R} \right\},$$

que puede escribirse alternativamente como

$$S = \left\{ \begin{bmatrix} 4 \\ -4 \\ 0 \end{bmatrix} + z \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \mid z \in \mathbb{R} \right\}.$$

Podemos obtener elementos específicos o soluciones particulares de este conjunto solución tomando distintos valores de z. Por ejemplo, si $z = 0$, se obtiene la solución particular

$$\vec{s}_p = \begin{bmatrix} 4 \\ -4 \\ 0 \end{bmatrix}.$$

También notamos que la solución del sistema homogéneo asociado es el conjunto

$$S_H = \left\{ \begin{bmatrix} z \\ z \\ z \end{bmatrix} \mid z \in \mathbb{R} \right\} = \left\{ z \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \mid z \in \mathbb{R} \right\}.$$

Concluimos que cualquier elemento del conjunto solución S puede expresarse como la suma de una solución particular de dicho sistema y una solución del sistema homogéneo asociado. La formalización de este resultado requiere de una nueva noción de suma y de un resultado de distributividad que presentamos a cotinuación.

Definición 4.35 *Sea $A\vec{x} = \vec{b}$ un sistema lineal con $A \in M_{m\times n}$ y $\vec{b} \in \mathbb{R}^m$. Supongamos que el sistema es consistente, de manera que tiene al menos una solución y su conjunto solución S no es vacío. Sea $\vec{s}_p \in S$, cualquier elemento de este conjunto (al cual llamamos una **solución particular**). Si $A\vec{x} = \vec{0}$ es el sistema homogéneo asociado (que siempre es consistente) y S_H su conjunto solución, se define el conjunto $\vec{s}_p + S_H$ como sigue:*

$$\vec{s}_p + S_H = \{\vec{s}_p + \vec{h} \mid \vec{h} \in S_H\},$$

es decir, $\vec{s} \in \vec{s}_p + S_H$ syss existe una solución $\vec{h}$ del sistema homogéneo tal que $\vec{s} = \vec{s}_p + \vec{h}$.

Lema 4.36 *Sean $A \in M_{m\times n}$ y sean $\vec{u}$ y $\vec{z}$ vectores de $\mathbb{R}^n$ (vistos como matrices de $n \times 1$. Entonces:*

(a) $A(\vec{u} + \vec{z}) = A\vec{u} + A\vec{z}$.

(b) $A(\vec{u} - \vec{z}) = A\vec{u} - A\vec{z}$.

Demostración. Para (a) notamos que:

$$\begin{aligned} A(\vec{u}+\vec{z}) &= A \begin{bmatrix} u_1 + z_1 \\ \vdots \\ u_n + z_n \end{bmatrix} \\ &= (u_1 + z_1)A^1 + \cdots + (u_n + z_n)A^n \\ &= (u_1A^1 + \cdots + u_nA^n) + (z_1A^1 + \cdots + z_nA^n) \\ &= A \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + A \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = A\vec{u} + A\vec{z}. \end{aligned}$$

La prueba de (b) es análoga y por lo tanto la omitimos. ■

Proposición 4.37 *Utilizando los mismos supuestos de la definición anterior se cumple la siguiente igualdad entre conjuntos:*

$$S = \vec{s}_p + S_H.$$

Demostración. Sean $\vec{s}_p \in S$ una solución particular y $\vec{s} \in S$ cualquier elemento de S. Notemos que, utilizando el Lema 4.36,

$$A(\vec{s} - \vec{s}_p) = A\vec{s} - A\vec{s}_p = \vec{b} - \vec{b} = \vec{0},$$

por lo tanto $\vec{s} - \vec{s}_p \in S_H$ y existe $\vec{h} \in S_H$ tal que $\vec{s} - \vec{s}_p = \vec{h}$. Equivalentemente,

$$\vec{s} = \vec{s}_p + \vec{h} \in \vec{s}_p + S_H$$

y hemos probado que $S \subset \vec{s}_p + S_H$. En forma análoga, si $\vec{s}_p + \vec{h}$ es cualquier elemento de $\vec{s}_p + S_H$, entonces, una vez más apelando al Lema 4.36,

$$A(\vec{s}_p + \vec{h}) = A\vec{s}_p + A\vec{h} = \vec{b} + \vec{0} = \vec{b},$$

concluyendo que $\vec{s}_p + \vec{h} \in S$ y $\vec{s}_p + S_H \subset S$. Por lo tanto, hemos demostrado ambas inclusiones y tenemos que $S = \vec{s}_p + S_H$. ∎

Observación 4.38 *Es importante hacer notar que si el sistema*

$$A\vec{x} = \vec{b}$$

es inconsistente, la proposición anterior es evidentemente falsa. En efecto, en dicha situación $S = \varnothing$ *pero* $S_H = \{\vec{0}\}$.

Observación 4.39 *Si* $\vec{h}_1$ *y* $\vec{h}_2$ *son dos elementos distintos de* S_H*, entonces* $\vec{s}_p + \vec{h}_1$ *y* $\vec{s}_p + \vec{h}_2$ *son dos elementos distintos de* $\vec{s}_p + S_H$*. Esto es una consecuencia directa de la ley de cancelación.*

Corolario 4.40 *Sea* $A \in M_{m \times n}$*, sea* $\vec{b}$ *un vector de* $\mathbb{R}^m$ *y supongamos que el sistema* $A\vec{x} = \vec{b}$ *es consistente (es decir,* $S \neq \varnothing$*). Entonces se tiene que:*

1. *La solución del sistema es única syss las columnas de* A *forman un conjunto LI.*

2. *El sistema tiene un número infinito de soluciones syss las columnas de* A *forman un conjunto LD.*

Demostración. Es suficiente probar la primera parte ya que la segunda es un enunciado equivalente. Supongamos que $\vec{s}_p$ es la solución única del sistema $A\vec{x} = \vec{b}$, entonces

$$\{\vec{s}_p\} = S = \vec{s}_p + S_H.$$

Por la Proposición 4.37 se tiene que dado $\vec{h} \in S_H$ debe cumplirse la igualdad $\vec{s}_p = \vec{s}_p + \vec{h}$. Aplicando la ley de la cancelación llegamos a que $\vec{h} = \vec{0}$ de manera que $S_H = \{\vec{0}\}$. De aquí que el sistema homogéneo $A\vec{x} = \vec{0}$ tiene una única solución y por la Observación 4.9, las columnas de A forman un conjunto LI.

Ahora bien, suponiendo que el conjunto de las columnas de A es LI, una vez más, por la Observación 4.9 el sistema homogéneo asociado tiene como conjunto solución a $S_H = \{\vec{0}\}$. Por la Proposición 4.37

$$S = \vec{s}_p + S_H = \left\{\vec{s}_p + \vec{0}\right\} = \{\vec{s}_p\},$$

de manera que $S = \{\vec{s}_p\}$ es un conjunto con un único elemento. ■

La demostración directa del segundo enunciado del corolario anterior resulta ser interesante debido a argumentos de finitud e infinitud. Veamos, si el sistema $A\vec{x} = \vec{b}$ tiene un número infinito de soluciones, entonces S es un conjunto infinito y por lo tanto, también lo será S_H. De aquí que $\vec{0}$ puede escribirse como combinación lineal no trivial de las columnas de A y éstas forman un conjunto LD. Ahora bien, si partimos de que las columnas de A son LD, entonces S_H cuenta con un número infinito de elementos. Se antoja concluir que $\vec{s}_p + S_H$ y en consecuencia S, son infinitos. En efecto lo serán pero tenemos que apelar a la ley de la cancelación para garantizar que al sumar $\vec{s}_p$ al conjunto infinito S_H, el conjunto resultante sigue siendo infinito. Ahora bien, la Observación 4.39 dice que dados dos vectores distintos $\vec{h}_1$ y $\vec{h}_2$ en S_H, los vectores $\vec{s}_p + \vec{h}_1$ y $\vec{s}_p + \vec{h}_2$ serán elementos distintos en S. Entonces el sistema $A\vec{x} = \vec{b}$ tiene más de una solución y S es infinito.

Ejemplo 4.41 *Sean*

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -3 & 0 \\ 1 & 1 & -1 \\ 5 & -1 & 1 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 2 \\ 3 \\ -2 \\ 2 \end{bmatrix} \quad y \ \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Es fácil ver que $\vec{x}_p = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$ *es una solución particular del sistema* $A\vec{x} = \vec{b}$. *Consideremos el sistema homogéneo* $A\vec{x} = \vec{0}$. *La FER de la matriz aumentada correspondiente está dada como*

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

por lo que la única solución es $\vec{x} = \vec{0}$. *De aquí que las columnas de* A *forman un conjunto LI y* $\vec{x}_p$ *es la única solución del sistema* $A\vec{x} = \vec{b}$.

Ejemplo 4.42 *Para resolver el sistema*

$$\begin{aligned} 4x + 2z &= -4, \\ -3x - 3y &= 3, \\ x - y + z &= -1 \end{aligned} \qquad (\spadesuit)$$

procedemos como sigue. Puede verificarse fácilmente que $\vec{s}_p = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$ *es una solución particular de (♠). Adicionalmente, la matriz aumentada del sistema homogéneo asociado está dada por*

$$\left[\begin{array}{rrr|r} 4 & 0 & 2 & 0 \\ -3 & -3 & 0 & 0 \\ 1 & -1 & 1 & 0 \end{array}\right],$$

cuya FER es

$$\left[\begin{array}{rrr|r} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Entonces, el conjunto solución del sistema homogéneo es

$$\begin{aligned} S_H &= \left\{ \begin{bmatrix} -\frac{z}{2} \\ \frac{z}{2} \\ z \end{bmatrix} \in \mathbb{R}^3 \mid z \in \mathbb{R} \right\} \\ &= \left\{ \frac{z}{2} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \in \mathbb{R}^3 \mid z \in \mathbb{R} \right\} \\ &= \left\{ s \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \in \mathbb{R}^3 \mid s \in \mathbb{R} \right\}. \end{aligned}$$

Dado que el sistema homogéneo asociado tiene un número infinito de soluciones, las columnas de la matriz de coeficientes del sistema forman un conjunto LD. Asimismo, el sistema original tiene una infinidad de soluciones y el conjunto solución puede expresarse como

$$\begin{aligned} S &= \left\{ \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \in \mathbb{R}^3 \mid s \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} -1 - s \\ s \\ 2s \end{bmatrix} \in \mathbb{R}^3 \mid s \in \mathbb{R} \right\}. \end{aligned}$$

Ejercicios

Ejercicio 4.1 *Considerar los vectores*

$$\vec{v}_1 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 6 \\ -15 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 6 \\ 15 \end{bmatrix}, \vec{v}_4 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{v}_5 = \begin{bmatrix} -\frac{1}{5} \\ \frac{1}{2} \end{bmatrix}.$$

Determinar si cada uno de los siguientes seis conjuntos es LI o LD y encontrar cada uno de los seis conjuntos correpondientes generados (este ejercicio sólo requiere cálculos muy sencillos).

a. $\{\vec{v}_1\}$.

b. $\{\vec{v}_1, \vec{v}_2\}$.

c. $\{\vec{v}_1, \vec{v}_3\}$.

d. $\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$.

e. $\{\vec{v}_1, \vec{v}_3, \vec{v}_4\}$.

f. $\{\vec{v}_1, \vec{v}_2, \vec{v}_5\}$.

Ejercicio 4.2 *Considerar los vectores*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \ \vec{v}_2 = \begin{bmatrix} \frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}, \vec{v}_4 = \begin{bmatrix} -5 \\ -16 \\ -43 \end{bmatrix},$$
$$\vec{v}_5 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \ \vec{v}_6 = \begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}, \ \vec{v}_7 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

Determinar si cada uno de los siguientes siete conjuntos es LI o LD y encontrar cada uno de los cinco conjuntos correspondientes generados.

a. $\{\vec{v}_1, \vec{v}_2\}$.

b. $\{\vec{v}_1, \vec{v}_3\}$.

c. $\{\vec{v}_1, \vec{v}_3, \vec{v}_4\}$.

d. $\{\vec{v}_1, \vec{v}_3, \vec{v}_5\}$.

e. $\{\vec{v}_1, \vec{v}_3, \vec{v}_6\}$.

f. $\{\vec{v}_1, \vec{v}_3, \ \vec{v}_7\}$.

g. $\{\vec{v}_1, \vec{v}_3, \vec{v}_5, \vec{v}_6\}$.

Ejercicio 4.3 *Considerar los vectores*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} \pi \\ 2\pi \\ \pi \\ 4.6 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ 666 \\ 0 \\ 0 \end{bmatrix}, \vec{v}_4 = \begin{bmatrix} 0 \\ 111 \\ 999 \\ 0 \end{bmatrix},$$
$$\vec{v}_5 = \begin{bmatrix} 0 \\ 777 \\ 999 \\ 0 \end{bmatrix}, \vec{v}_6 = \begin{bmatrix} 9 \\ 8 \\ 3 \\ 9 \end{bmatrix}.$$

Determinar si cada uno de los siguientes tres conjuntos es LI o LD (este ejercicio sólo requiere cálculos muy sencillos).

a. $\{\vec{v}_2, \vec{v}_3, \vec{v}_4, \vec{v}_5\}$.

b. $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_5, \vec{v}_6\}$.

c. $\{\vec{v}_1, \vec{v}_2, \vec{v}_4, \vec{v}_5\}$.

Ejercicio 4.4 *Considerar las matrices*

$$A = \begin{bmatrix} 1.5 & 6 \\ 5 & h \end{bmatrix}, \ B = \begin{bmatrix} 1 & 2 & 3 \\ -4 & -5 & 6 \\ 6 & i & 9 \end{bmatrix} \ C = \begin{bmatrix} 1 & 2 & 0 \\ 5 & 1 & j \\ -2 & 5 & 9 \\ 3 & 4 & 1 \end{bmatrix}.$$

a. Encontrar todos los valores de h tales que las columnas de A forman un conjunto LI.

b. Encontrar todos los valores de i tales que las columnas de B forman un conjunto LD.

c. Encontrar todos los valores de j tales que las columnas de C forman un conjunto LI.

Ejercicio 4.5 *Probar (en general) o refutar (con un contraejemplo) los siguientes enunciados (en cada caso se debe indicar claramente "verdadero" o "falso").*

a. Si $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\} \subset \mathbb{R}^n$ es LD, entonces $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es LD.

b. Si $\{\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4\} \subset \mathbb{R}^n$ es LI, entonces $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es LI.

c. Suponga que $\vec{w}_1, \vec{w}_2$ y $\vec{w}_3$ son tres vectores de $\mathbb{R}^3$ distintos entre sí. Si $\{\vec{w}_1, \vec{w}_2\}$, $\{\vec{w}_1, \vec{w}_3\}$ y $\{\vec{w}_2, \vec{w}_3\}$ son tres conjuntos LI, entonces $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ es LI.

d. *Si* $\{\vec{z}_1, \vec{z}_2, \vec{z}_3\} \subset \mathbb{R}^n$ *es LD, entonces* $\vec{z}_1 \in span\{\vec{z}_2, \vec{z}_3\}$.

e. *Sea* A *una matriz de* $m \times n$ *y sean* $\vec{b}$ *y* $\vec{d}$ *dos vectores de* $\mathbb{R}^m$. *Si el sistema* $A\vec{x} = \vec{b}$ *tiene una infinidad de soluciones, entonces el sistema* $A\vec{x} = \vec{d}$ *también tiene una infinidad de soluciones.*

f. *Sea* A *una matriz de* $m \times n$ *y sean* $\vec{b}$ *y* $\vec{d}$ *dos vectores de* $\mathbb{R}^m$ *tales que los sistemas* $A\vec{x} = \vec{b}$ *y* $A\vec{x} = \vec{d}$ *son ambos consistentes. Si el sistema* $A\vec{x} = \vec{b}$ *tiene una infinidad de soluciones, entonces el sistema* $A\vec{x} = \vec{d}$ *también tiene una infinidad de soluciones.*

g. *Sea* A *una matriz de* $m \times n$ *y sea* $\vec{b}$ *un vector de* $\mathbb{R}^m$. *Si las columnas de* A *forman un conjunto LI y* $\{A^1, \ldots, A^n\} \cup \{\vec{b}\}$ *es un conjunto LD, entonces el sistema* $A\vec{x} = \vec{b}$ *tiene solución única.*

Ejercicio 4.6 *Probar que si el conjunto* $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ *es LI, entonces* $\{\vec{v}_1 + \vec{v}_2 + \vec{v}_3, \vec{v}_2 + \vec{v}_3, \vec{v}_3\}$ *es LI.*

Ejercicio 4.7 *Demostrar que si el conjunto* $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ *es LI, entonces* $\{\vec{v}_1 - \vec{v}_2, \vec{v}_1 + \vec{v}_3, \vec{v}_2 - \vec{v}_3\}$ *es LI.*

Ejercicio 4.8 *Probar que si* $c_1, c_2, \ldots, c_m$ *son escalares distintos de cero y* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ *es LI, entonces* $\{c_1\vec{v}_1, c_2\vec{v}_2, \ldots, c_m\vec{v}_m\}$ *es LI.*

Ejercicio 4.9 *El conjunto solución del sistema homogéneo*

$$\begin{aligned} x + 2y - z &= 0, \\ 2x + y + z &= 0 \end{aligned}$$

es $S_H = \left\{ s \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \mid s \in \mathbb{R} \right\}$. *Utilizando este hecho, encontrar la solución del sistema*

$$\begin{aligned} x + 2y - z &= 3, \\ 2x + y + z &= 3. \end{aligned}$$

Capítulo 5

Transformaciones lineales

Introducción

Cuando se tienen objetos matemáticos es natural querer estudiar las relaciones que existen entre ellos. Para este efecto, se utilizan las funciones entre estos objetos que preservan las propiedades que los caracterizan. Por ejemplo, si se tienen conjuntos ordenados A y B, entonces las funciones

$$f : A \longrightarrow B$$

que se desean analizar son aquellas que preservan el orden, es decir, si $x, y \in A$ y $x < y$, entonces queremos que se cumpla $f(x) < f(y)$. Asimismo, en Cálculo nos interesan ciertas propiedades geométricas (o topológicas) del conjunto de números reales (o de $\mathbb{R}^n$ en general) y las funciones de interés son aquellas que preservan sucesiones convergentes de puntos.

En los ejemplos anteriores se desea que las funciones preserven una relación de orden, en el primer caso, y alguna relación de cercanía entre elementos, en el segundo caso. Para los espacios vectoriales nos interesa preservar la estructura algebraica lineal, es decir, la suma entre vectores y la multiplicación por escalares. Geométricamente, las estructuras lineales originadas por las combinaciones lineales de vectores en un espacio vectorial (rectas y planos, por ejemplo), se transformarán en estructuras lineales de otro espacio vectorial. Por lo tanto, las funciones relevantes entre espacios vectoriales (en particular entre los espacios

$\mathbb{R}^n$) son aquellas que preservan estas operaciones de suma y producto por escalares.

Transformaciones lineales

Definición 5.1 *Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una función. Decimos que T es una* ***transformación lineal*** *syss se satisfacen las siguientes dos condiciones:*

1. *T preserva sumas. Esto es, si $\vec{x}, \vec{y} \in \mathbb{R}^n$, entonces*
$$T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y}).$$

2. *T preserva multiplicación escalar. Esto es, si $\vec{x} \in \mathbb{R}^n$ y c es un escalar, entonces*
$$T(c\vec{x}) = cT(\vec{x}).$$

Ejemplo 5.2 *Sea $T : \mathbb{R} \longrightarrow \mathbb{R}$ la función dada por $T(x) = 5x$. Notemos que dados $x, y \in \mathbb{R}$ se tiene que*

$$T(x + y) = 5(x + y) = 5x + 5y = T(x) + T(y).$$

Adicionalmente, si $c \in \mathbb{R}$,

$$T(cx) = 5(cx) = c(5x) = cT(x),$$

por lo tanto T es una transformación lineal.

Ejemplo 5.3 *Sea $S : \mathbb{R} \longrightarrow \mathbb{R}$ dada por $S(x) = x^2$ y supongamos que deseamos determinar si S es o no una transformación lineal. Para ello observamos que S no preserva sumas. En efecto, notemos que*

$$S(1 + 1) = S(2) = 4 \neq 2 = S(1) + S(1).$$

Por lo tanto, S no es una transformación lineal.

Ejemplo 5.4 *Sea $T_1 : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ la función dada por*

$$T_1\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 3x_1 \end{bmatrix}.$$

Para familiarizarnos con esta función notemos, por ejemplo, que

$$T_1\left(\begin{bmatrix}2\\4\\5\end{bmatrix}\right)=\begin{bmatrix}0\\6\end{bmatrix}.$$

Supongamos que deseamos determinar si T_1 es o no una transformación lineal. Para ello, veamos primero que T_1 preserva sumas:

$$T_1\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}+\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}\right)=T_1\left(\begin{bmatrix}x_1+y_1\\x_2+y_2\\x_3+y_3\end{bmatrix}\right)=\begin{bmatrix}0\\3(x_1+y_1)\end{bmatrix}=\begin{bmatrix}0\\3x_1+3y_1\end{bmatrix}$$

$$=\begin{bmatrix}0\\3x_1\end{bmatrix}+\begin{bmatrix}0\\3y_1\end{bmatrix}=T_1\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right)+T_1\left(\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}\right).$$

El segundo paso consiste en comprobar que T_1 preserva multiplicación escalar. En efecto,

$$\begin{aligned}T_1\left(c\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right)&=T_1\left(\begin{bmatrix}cx_1\\cx_2\\cx_3\end{bmatrix}\right)=\begin{bmatrix}0\\3(cx_1)\end{bmatrix}\\&=c\begin{bmatrix}0\\3x_1\end{bmatrix}=cT_1\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right).\end{aligned}$$

Por lo tanto, T_1 sí es una transformación lineal.

Ejemplo 5.5 *Sea $T_2:\mathbb{R}^2\longrightarrow\mathbb{R}^4$ la función dada por*

$$T_2\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}x\\0\\y\\2x+3y\end{bmatrix}.$$

Notemos, por ejemplo, que

$$T_2\left(\begin{bmatrix}-5\\4\end{bmatrix}\right)=\begin{bmatrix}-5\\0\\4\\2\end{bmatrix}$$

Supongamos ahora que deseamos determinar si T_2 es o no una transformación lineal. Para ello, comprobemos primero que T_2 preserva *sumas:*

$$T_2\left(\begin{bmatrix}a\\b\end{bmatrix}+\begin{bmatrix}c\\d\end{bmatrix}\right)=T_2\left(\begin{bmatrix}a+c\\b+d\end{bmatrix}\right)=$$

$$\begin{bmatrix}a+c\\0\\b+d\\2(a+c)+3(b+d)\end{bmatrix}=\begin{bmatrix}a+c\\0\\b+d\\(2a+3b)+(2c+3d)\end{bmatrix}=$$

$$\begin{bmatrix}a\\0\\b\\2a+3b\end{bmatrix}+\begin{bmatrix}c\\0\\d\\2c+3d\end{bmatrix}=T_2\left(\begin{bmatrix}a\\b\end{bmatrix}\right)+T_2\left(\begin{bmatrix}c\\d\end{bmatrix}\right).$$

El segundo paso consiste en comprobar que T_2 preserva multiplicación escalar:

$$T_2\left(c\begin{bmatrix}a\\b\end{bmatrix}\right)=T_2\left(\begin{bmatrix}ca\\cb\end{bmatrix}\right)=\begin{bmatrix}ca\\0\\cb\\2(ca)+3(cb)\end{bmatrix}$$
$$=c\begin{bmatrix}a\\0\\b\\2a+3b\end{bmatrix}=cT_2\left(\begin{bmatrix}a\\b\end{bmatrix}\right).$$

Por lo tanto, T_2 sí es una transformación lineal.

Ejemplo 5.6 *Sea $S:\mathbb{R}^2\longrightarrow\mathbb{R}^4$ la función dada por*

$$S\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}x\\1\\y\\2x+3y\end{bmatrix}.$$

y supongamos que deseamos determinar si S es o no una transformación lineal. Para ello, observamos que S no preserva multiplicación escalar. En efecto,

$$S\left(2\begin{bmatrix}1\\0\end{bmatrix}\right)=S\left(\begin{bmatrix}2\\0\end{bmatrix}\right)=\begin{bmatrix}2\\1\\0\\4\end{bmatrix}\neq\begin{bmatrix}2\\2\\0\\4\end{bmatrix}=2\begin{bmatrix}1\\1\\0\\2\end{bmatrix}=2\left(S\begin{bmatrix}1\\0\end{bmatrix}\right).$$

Por lo tanto, S no es una transformación lineal.

Ejemplo 5.7 *La función identidad $I : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ (dada por $I(\vec{x}) = \vec{x}$, para todo $\vec{x} \in \mathbb{R}^n$) es una transformación lineal. En efecto, si $\vec{x}, \vec{y} \in \mathbb{R}^n$ y c es un escalar, entonces $I(\vec{x} + \vec{y}) = \vec{x} + \vec{y} = I(\vec{x}) + I(\vec{y})$ (por lo que I preserva sumas) e $I(c\vec{x}) = c\vec{x} = cI(\vec{x})$ (por lo que I preserva multiplicación escalar).*

Ejemplo 5.8 *La función constante cero $\Theta : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ (dada por $\Theta(\vec{x}) = \vec{0}$, para todo $\vec{x} \in \mathbb{R}^n$) es una transformación lineal. En efecto, si $\vec{x}, \vec{y} \in \mathbb{R}^n$ y c es un escalar, entonces $\Theta(\vec{x}+\vec{y}) = \vec{0} = \vec{0}+\vec{0} = \Theta(\vec{x})+\Theta(\vec{y})$ (por lo que Θ preserva sumas) y $\Theta(c\vec{x}) = \vec{0} = c\vec{0} = c\Theta(\vec{x})$ (por lo que Θ preserva multiplicación escalar).*

El siguiente ejemplo es de naturaleza más práctica.

Ejemplo 5.9 *Una empresa polaca produce tres tipos de pelucas: A, B y C. Los costos (en złotys) de producir cada unidad (de tipo A, B y C) aparecen en la siguiente tabla:*

	A	*B*	*C*
MO	*3*	*4*	*5*
MT	*6*	*2*	*1*

donde MO y MT hacen referencia, respectivamente, a mano de obra y a materiales. Así, si esta empresa desea producir 5 *pelucas de tipo A, 3 de tipo B y 6 de tipo C, el vector de $\mathbb{R}^2$ cuya primera y segunda entrada son los costos totales de MO y de MT asociados a estas 14 pelucas está dado por*

$$\begin{bmatrix} 57 \\ 42 \end{bmatrix} = 5 \begin{bmatrix} 3 \\ 6 \end{bmatrix} + 3 \begin{bmatrix} 4 \\ 2 \end{bmatrix} + 6 \begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 & 4 & 5 \\ 6 & 2 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 6 \end{bmatrix}.$$

Supongamos ahora que la empresa desea producir x pelucas de tipo A, y de tipo B y z de tipo C. En general, el vector de $\mathbb{R}^2$ cuya primera y segunda entrada son los costos totales de MO y de MT asociados a estas $x + y + z$ pelucas está dado por

$$\begin{bmatrix} 3 & 4 & 5 \\ 6 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

De forma natural esta multiplicación define una función de $\mathbb{R}^3$ en $\mathbb{R}^2$. La siguiente proposición asegura que dicha función es, de hecho, una transformación lineal.

Proposición 5.10 *Sea A una matriz de $m \times n$. Si $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es la función dada por $T(\vec{x}) = A\vec{x}$, para todo $\vec{x} \in \mathbb{R}^n$, entonces T es una transformación lineal.*

Demostración. Para verificar que T preserva sumas simplemente notamos que si $\vec{x}, \vec{y} \in \mathbb{R}^n$, entonces

$$T(\vec{x} + \vec{y}) = A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y} = T(\vec{x}) + T(\vec{y}).$$

Por otro lado, para verificar que T preserva multiplicación escalar simplemente notamos que si $\vec{x} \in \mathbb{R}^n$ y c es un escalar, entonces

$$T(c\vec{x}) = A(c\vec{x}) = A \begin{bmatrix} cx_1 \\ \vdots \\ cx_n \end{bmatrix} = cx_1A^1 + \cdots + cx_nA^n$$

$$= c(x_1A^1 + \cdots + x_nA^n) = c\,(A\vec{x}) = cT(\vec{x})$$

y con esto concluimos que T es una transformación lineal. ■

La proposición anterior dice que toda matriz determina una transformación lineal. El resultado inverso, es decir que toda transformación lineal da lugar a una matriz, también es válido. Antes de enunciarlo formalmente se proporcionan algunos conceptos preliminares.

Recordemos que la base estándar para $\mathbb{R}^n$ es el conjunto

$$\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\},$$

en donde e_i es el vector de $\mathbb{R}^n$ cuya i-ésima entrada es igual a uno y cuyas otras entradas son todas cero. Adicionalmente, si $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ es un vector de $\mathbb{R}^n$, entonces $\vec{x}$ puede representarse como combinación lineal de $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ de la siguiente forma:

$$\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n.$$

Lema 5.11 *Sean $A, B \in M_{m\times n}$. Entonces se cumple que $A = B$ syss $A\vec{x} = B\vec{x}$ para todo $\vec{x} \in \mathbb{R}^n$.*

Demostración. Evidentemente $A = B$ implica que $A\vec{x} = B\vec{x}$ para todo $\vec{x} \in \mathbb{R}^n$. Para la implicación inversa se considera la base estándar de $\mathbb{R}^n$ dada por $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$. Por hipótesis, para todo $i = 1, 2, \ldots, n$, se cumple

$$A^i = A\vec{e}_i = B\vec{e}_i = B^i$$

y, por lo tanto $A = B$ puesto que ambas matrices poseen las mismas columnas. ■

Proposición 5.12 *Si $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es una transformación lineal, entonces existe una única matriz $A \in M_{m\times n}$, tal que $T(\vec{x}) = A\vec{x}$ para todo $\vec{x} \in \mathbb{R}^n$. Dicha matriz A, es conocida como la* ***matriz estándar que representa a*** *T y sus columnas están dadas por $A^i = T(\vec{e}_i)$, $i = 1, 2, \ldots, n$, donde $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ es la base estándar de $\mathbb{R}^n$.*

Demostración. Como mencionamos anteriormente, para cualquier vector $\vec{x}$ de $\mathbb{R}^n$ ocurre que

$$\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n.$$

Aplicando T a ambos lados de esta igualdad se tiene

$$\begin{aligned} T(\vec{x}) &= T(x_1\vec{e}_1 + \ldots + x_n\vec{e}_n) \\ &= T(x_1\vec{e}_1) + \cdots + T(x_n\vec{e}_n) \\ &= x_1T(\vec{e}_1) + \cdots + x_nT(\vec{e}_n) \\ &= x_1A^1 + \cdots + x_nA^n \\ &= A\vec{x}, \end{aligned}$$

justo como se quería probar. Aquí es importante mencionar que la segunda y tercera igualdad se obtuvieron utilizando que T preserva sumas y multiplicación escalar, respectivamente. Por lo tanto, la hipótesis de linealidad es fundamental en esta demostración.

Finalmente, por el Lema 5.12 concluimos que si A y B son dos matrices que representan a T, entonces $A = B$ lo cual prueba la unicidad de A. ■

Ejemplo 5.13 *Consideremos la transformación lineal* $T_1 : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *dada por*

$$T_1\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 3x_1 \end{bmatrix}$$

e introducida previamente en el Ejemplo 5.4. Como

$$T_1\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, T_1\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, T_1\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

se tiene que la matriz estándar que representa a T_1 *es*

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 3 & 0 & 0 \end{bmatrix}.$$

Por lo tanto, para todo vector $\vec{x} \in \mathbb{R}^3$ *se tiene que* $T(\vec{x}) = A\vec{x}$*. En particular,*

$$\begin{aligned} T_1\left(\begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}\right) &= \begin{bmatrix} 0 & 0 & 0 \\ 3 & 0 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix} \\ &= 2\begin{bmatrix} 0 \\ 3 \end{bmatrix} + 4\begin{bmatrix} 0 \\ 0 \end{bmatrix} + 5\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}, \end{aligned}$$

lo cual coincide con lo que mencionamos en el Ejemplo 5.4.

Ejemplo 5.14 *La matriz estándar que representa la transformación lineal identidad* $I : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *es la matriz cuya columna* j *es el vector* $I(\vec{e}_j) = \vec{e}_j$*. Dicha matriz es, por supuesto, la matriz identidad de* $n \times n$*.*

Ejemplo 5.15 *Consideremos la transformación lineal* $T_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$ *dada por*

$$T_2\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ 0 \\ y \\ 2x+3y \end{bmatrix}$$

e introducida previamente en el Ejemplo 5.5. Como

$$T_2\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix} \text{ y } T_2\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 3 \end{bmatrix},$$

se tiene que la matriz estándar que representa a T_2 *está dada como*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 2 & 3 \end{bmatrix}.$$

Por lo tanto, para todo vector $\vec{x} \in \mathbb{R}^2$ *se tiene que* $T_2(\vec{x}) = A\vec{x}$*. En particular,*

$$T_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

y también

$$T_2\left(\begin{bmatrix} -5 \\ 4 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -5 \\ 4 \end{bmatrix} = -5 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -5 \\ 0 \\ 4 \\ 2 \end{bmatrix},$$

lo cual coincide con lo que mencionamos en el Ejemplo 5.5.

Proposición 5.16 *Si* $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *es una transformación lineal,* $\vec{0}_n$ *denota al vector cero de* $\mathbb{R}^n$ *y* $\vec{0}_m$ *denota al vector cero de* $\mathbb{R}^m$*, entonces* $T(\vec{0}_n) = \vec{0}_m$*.*

Demostración. Supongamos que $A \in M_{m\times n}$ es la matriz estándar que representa a T. Entonces, $T(\vec{x}) = A\vec{x}$, para todo vector $\vec{x}$ de $\mathbb{R}^n$. En particular,

$$T(\vec{0}_n) = A\vec{0}_n = 0A^1 + \cdots + 0A^n = \vec{0}_m.$$

Una demostración alternativa consiste en observar que, gracias al hecho de que T preserva sumas, se tiene

$$T(\vec{0}_n) + T(\vec{0}_n) = T(\vec{0}_n + \vec{0}_n) = T(\vec{0}_n) = T(\vec{0}_n) + \vec{0}_m.$$

Esto último y la ley de la cancelación implican que $T(\vec{0}_n) = \vec{0}_m$. ∎

Rango y suprayectividad

Definición 5.17 *Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una transformación lineal. Definimos el **rango** de T, denotado por Rango(T), como el conjunto de todos los vectores $\vec{b} \in \mathbb{R}^m$ tales que existe $\vec{x} \in \mathbb{R}^n$ con la propiedad de que $T(\vec{x}) = \vec{b}$. Esto es, el rango de T es el conjunto de las imágenes, bajo de T, de los vectores de $\mathbb{R}^n$. En símbolos,*

$$Rango(T) = \{\vec{b} \mid T(\vec{x}) = \vec{b} \text{ para algún } \vec{x} \in \mathbb{R}^n\}$$

o alternativamente,

$$Rango(T) = \{T(\vec{x}) \mid \vec{x} \in \mathbb{R}^n\}.$$

Observación 5.18 *Gracias a la Proposición 5.16 es claro que si*

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

es una transformación lineal, entonces $\vec{0}_m \in$ Rango(T).

Ejemplo 5.19 *El rango de la transformación identidad*

$$I : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

es igual al conjunto

$$\{I(\vec{x}) \mid \vec{x} \in \mathbb{R}^n\} = \{\vec{x} \mid \vec{x} \in \mathbb{R}^n\} = \mathbb{R}^n.$$

Ejemplo 5.20 *El rango de la transformación constante cero*

$$\Theta : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

es igual al conjunto

$$\{\Theta(\vec{x}) \mid \vec{x} \in \mathbb{R}^n\} = \{\vec{0}_m\}.$$

Ejemplo 5.21 *Consideremos la transformación lineal* $T_3 : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$ *dada por*

$$T_3\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Como $T_3\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$*, es claro que* $\begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} \in Rango(T_3)$*. Supongamos ahora que deseamos describir el rango de* T_3*. Observando que la imagen de cualquier vector bajo* T_3 *es tal que su segunda, tercera y cuarta componente son iguales a cero, es claro que*

$$Rango(T_3) \subset \left\{ \begin{bmatrix} a \\ 0 \\ 0 \\ 0 \end{bmatrix} \mid a \in \mathbb{R} \right\}.$$

Afirmamos que, de hecho,

$$Rango(T_3) = \left\{ \begin{bmatrix} a \\ 0 \\ 0 \\ 0 \end{bmatrix} \mid a \in \mathbb{R} \right\}.$$

En efecto,

$$T_3\left(\begin{bmatrix} a \\ 0 \end{bmatrix}\right) = \begin{bmatrix} a+0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \in Rango(T_3).$$

Como el lector se podrá imaginar, no siempre es obvio determinar a simple vista el rango de una transformación lineal. Por ello, la siguiente observación (cuya demostración está autocontenida en el enunciado de la misma) es particularmente útil.

Observación 5.22 *Sean* $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *una transformación lineal,* $A \in M_{m\times n}$ *la matriz estándar que representa a* T *y* $\vec{b} \in \mathbb{R}^m$*. Las siguientes afirmaciones son equivalentes.*

1. $\vec{b} \in Rango(T)$.
2. *Existe* $\vec{x} \in \mathbb{R}^n$ *con la propiedad de que* $T(\vec{x}) = \vec{b}$.
3. *Existe* $\vec{x} \in \mathbb{R}^n$ *con la propiedad de que* $A\vec{x} = \vec{b}$.

4. *El sistema lineal cuya matriz aumentada es* $(A \mid \vec{b})$ *es consistente.*
5. $\vec{b} \in span\{A^1, A^2, \ldots A^n\}$.

En particular, se tiene que $Rango(T) = span\{A^1, A^2, \ldots A^n\}$.

Ejemplo 5.23 *Sea* $T_4 : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ *la transformación lineal dada por*

$$T_4\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} y \\ x \\ 0 \end{bmatrix}$$

y supongamos que deseamos, primero, encontrar la matriz estándar que representa a T_4 *y después, encontrar el rango de* T_4. *En el primer caso recordamos que la matriz buscada es* $[T_4(\vec{e}_1)\ T_4(\vec{e}_2)]$. *Utilizando la definición de* T_4 *se trata de la matriz*

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

Nótese que el rango de esta transformación es igual al conjunto de todos los vectores de $\mathbb{R}^3$ *cuya tercera componente es igual a cero. Este resultado también puede comprobarse observando que el rango de* T_4 *es el conjunto de todos los vectores* $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ *tales que la matriz aumentada*

$$\left[\begin{array}{cc|c} 0 & 1 & a \\ 1 & 0 & b \\ 0 & 0 & c \end{array}\right]$$

representa un sistema consistente. Obviamente, tal consistencia se obtiene syss $c = 0$.

Ejemplo 5.24 *Sea* $T_5 : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *la transformación lineal que satisface las siguientes igualdades:*

$$T_5\left(\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 4 \end{bmatrix}\ ,\ T_5\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\ \text{ y } T_5\left(\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

y supongamos que deseamos, primero, encontrar la matriz estándar que representa a T_5 *y después encontrar el rango de* T_5.

Para la primera tarea recordamos que la matriz que representa a T_5 *es la matriz* $A = [T_5(\vec{e}_1)\ T_5(\vec{e}_2)\ T_5(\vec{e}_3)]$. *Ahora bien, para calcular* $T_5(\vec{e}_1)$ *notamos que*

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

Por lo tanto,

$$T_5\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = T_5\left(\begin{bmatrix}1\\2\\0\end{bmatrix} - 2\begin{bmatrix}0\\1\\0\end{bmatrix}\right).$$

Pero T_5 *preserva sumas y multiplicación escalar, por lo que*

$$\begin{aligned} T_5\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) &= T_5\left(\begin{bmatrix}1\\2\\0\end{bmatrix}\right) - 2T_5\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right)\\ &= \begin{bmatrix}0\\4\end{bmatrix} - 2\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}-2\\4\end{bmatrix}. \end{aligned}$$

De manera análoga, y con el fin de calcular $T_5(\vec{e}_3)$, *notamos que*

$$\begin{bmatrix}0\\0\\1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}0\\0\\2\end{bmatrix}.$$

Por lo tanto,

$$T_5\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = T_5\left(\frac{1}{2}\begin{bmatrix}0\\0\\2\end{bmatrix}\right) = \frac{1}{2}T_5\left(\begin{bmatrix}0\\0\\2\end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix}6\\0\end{bmatrix} = \begin{bmatrix}3\\0\end{bmatrix}$$

(la segunda igualdad se sigue porque T_5 *preserva multiplicación escalar). En conclusión, la matriz que representa a la transformación* T_5 *es igual a la matriz*

$$A = \begin{bmatrix}-2 & 1 & 3\\ 4 & 0 & 0\end{bmatrix}.$$

Finalmente, el rango de T_5 *es el conjunto de todos los vectores* $\begin{bmatrix}a\\b\end{bmatrix}$ *tales que la matriz*

$$\left[\begin{array}{ccc|c}-2 & 1 & 3 & a\\ 4 & 0 & 0 & b\end{array}\right]$$

representa a un sistema consistente. Otra forma más práctica de calcular este rango consiste en notar que

$$\begin{aligned} Rango(T_5) &= span\left\{\begin{bmatrix}-2\\4\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}3\\0\end{bmatrix}\right\}\\ &= span\left\{\begin{bmatrix}-2\\4\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}\right\} = \mathbb{R}^2. \end{aligned}$$

Esta última igualdad se sigue del Corolario 4.30.

Ejemplo 5.25 *Sean T_5 y A como en el ejemplo anterior y supongamos que buscamos el conjunto de todos los vectores $\vec{x} \in \mathbb{R}^3$ tales que*

$$T_5(\vec{x}) = \begin{bmatrix} 4 \\ 12 \end{bmatrix}.$$

Ahora bien, como $T_5(\vec{x}) = A\vec{x}$, nuestro problema se reduce a resolver el sistema $A\vec{x} = \begin{bmatrix} 4 \\ 12 \end{bmatrix}$ cuya matriz aumentada es

$$\left[\begin{array}{rrr|r} -2 & 1 & 3 & 4 \\ 4 & 0 & 0 & 12 \end{array}\right]$$

y que tiene como FER a la matriz

$$\left[\begin{array}{rrr|r} 1 & 0 & 0 & 3 \\ 0 & 1 & 3 & 10 \end{array}\right].$$

Por lo tanto, el conjunto buscado es

$$\left\{ \begin{bmatrix} 3 \\ 10 - 3x_3 \\ x_3 \end{bmatrix} \mid x_3 \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} 3 \\ 10 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix} \mid x_3 \in \mathbb{R} \right\}.$$

Definición 5.26 *Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una transformación lineal. Decimos que T es **suprayectiva** (o "sobre") syss $Rango(T) = \mathbb{R}^m$.*

Como ya hemos visto, en el caso de las transformaciones lineales

$$T_1 : \mathbb{R}^3 \longrightarrow \mathbb{R}^2 \text{ y } T_5 : \mathbb{R}^3 \longrightarrow \mathbb{R}^2,$$

hay transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$ con $n > m$ que no son suprayectivas y otras que sí lo son. En el caso de las transformaciones T_2, T_3 y T_4 tenemos que éstas van de $\mathbb{R}^n$ en $\mathbb{R}^m$ con $n < m$, por lo que, utilizando el siguiente resultado, podemos afirmar que ninguna de las tres es suprayectiva (¡tal afirmación puede realizarse incluso sin conocer el comportamiento específico de cada una de ellas!).

Proposición 5.27 *Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una transformación lineal. Si $n < m$, entonces T no es suprayectiva.*

Demostración. Sea A la matriz estándar que representa a T. Gracias a la Observación 5.22 sabemos que $Rango(T) = span\{A^1, \ldots, A^n\}$. Este conjunto es, por supuesto, un subconjunto de $\mathbb{R}^m$. Como $n < m$

y gracias a la máxima de "pecado por defecto" al final del Corolario 4.28, podemos asegurar que

$$Rango(T) = span\{A^1, \ldots, A^n\} \subsetneq \mathbb{R}^m,$$

por lo que T no es suprayectiva. ■

Terminamos este análisis de suprayectividad para transformaciones lineales $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ considerando el caso $n = m$.

Proposición 5.28 *Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una transformación lineal y supongamos que A es la matriz estándar que representa a T. Entonces T es suprayectiva syss el conjunto $\{A^1, \ldots A^n\}$ es LI (o lo que es lo mismo, syss la FER de A es la matriz identidad).*

Demostración. Gracias a la Observación 5.22 sabemos que

$$Rango(T) = span\{A^1, \ldots, A^n\}.$$

Este conjunto es, por supuesto, un subconjunto de $\mathbb{R}^n$. Ahora bien, por el Corolario 4.30 sabemos que el conjunto generado por n vectores de $\mathbb{R}^n$ es igual a todo $\mathbb{R}^n$ syss esos n vectores forman un conjunto LI. Por lo tanto, T es suprayectiva (i.e., $Rango(T) = \mathbb{R}^n$) syss el conjunto $\{A^1, \ldots A^n\}$ es LI. ■

Ejemplo 5.29 *Sea $T_6 : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ la transformación lineal dada por*

$$T_6\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ x - y \\ z \end{bmatrix}$$

y supongamos que deseamos determinar si T_6 es suprayectiva. Para ello notamos, primero, que la matriz estándar que representa a T_6 es la matriz

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

cuya forma escalonada reducida es la matriz identidad de 3×3. Por la proposición anterior concluimos que T_6 es suprayectiva.

Espacio nulo e inyectividad

Definición 5.30 *Sea $T : \mathbb{R}^n \longrightarrow R^m$ una transformación lineal. Definimos el **espacio nulo** de T, denotado por $Null(T)$, como el conjunto de elementos de $\mathbb{R}^n$ que bajo la transformación T son enviados al vector $\vec{0}_m$. En símbolos,*

$$Null(T) = \{\vec{x} \in \mathbb{R}^n \mid T(\vec{x}) = \vec{0}_m\}.$$

Observación 5.31 *La Proposición 5.16 asegura que si*

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

es una transformación lineal, entonces $\vec{0}_n \in Null(T)$.

Ejemplo 5.32 *Dada la transformación identidad*

$$I : \mathbb{R}^n \longrightarrow \mathbb{R}^n,$$

$$Null(I) = \{\vec{x} \in \mathbb{R}^n \mid I(\vec{x}) = \vec{0}\} = \{\vec{0}_n\}.$$

Ejemplo 5.33 *Dada la transformación constante cero*

$$\Theta : \mathbb{R}^n \longrightarrow \mathbb{R}^m,$$

$$Null(\Theta) = \{\vec{x} \in \mathbb{R}^n \mid \Theta(\vec{x}) = \vec{0}\} = \mathbb{R}^n.$$

Ejemplo 5.34 *Consideremos la transformación lineal*

$$T_3 : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$$

del Ejemplo 5.21 dada por

$$T_3\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Para obtener $Null(T_3)$ *notamos que*

$$T_3\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{syss} \quad x = -y,$$

de manera que

$$Null(T_3) = \left\{ \begin{bmatrix} -y \\ y \end{bmatrix} \mid y \in \mathbb{R} \right\}.$$

La siguiente observación nos dice que $Null(T)$ puede calcularse como el conjunto solución de un sistema lineal homogéneo.

Observación 5.35 *Sean $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una transformación lineal, $A \in M_{m\times n}$ la matriz estándar que representa a T y $\vec{x} \in \mathbb{R}^n$. Las siguientes afirmaciones son equivalentes:*

1. $\vec{x} \in Null(T)$.
2. $T(\vec{x}) = \vec{0}_m$.
3. $A\vec{x} = \vec{0}_m$.

Ejemplo 5.36 *Queremos encontrar $Null(T_7)$, en donde $T_7 : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ es la transformación lineal definida por*

$$T_7\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + 4x_2 + 6x_3 + 8x_4 \\ x_1 + 3x_2 + 5x_4 \\ x_1 + x_2 + 6x_3 + 3x_4 \end{bmatrix}.$$

La matriz estándar que representa a T_7 está dada como

$$A = \begin{bmatrix} 2 & 4 & 6 & 8 \\ 1 & 3 & 0 & 5 \\ 1 & 1 & 6 & 3 \end{bmatrix}.$$

Como la FER de $[A \mid \vec{0}]$ es igual a

$$\left[\begin{array}{cccc|c} 1 & 0 & 9 & 2 & 0 \\ 0 & 1 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

tenemos que $A\vec{x} = \vec{0}$ syss

$$\begin{aligned} x_1 &= -9x_3 - 2x_4, \\ x_2 &= 3x_3 - x_4, \\ x_3, x_4 &\in \mathbb{R}. \end{aligned}$$

Por lo tanto, la Observación 5.35 asegura que

$$\begin{aligned} Null(T_7) &= \left\{ \begin{bmatrix} -9x_3 - 2x_4 \\ 3x_3 - x_4 \\ x_3 \\ x_4 \end{bmatrix} \mid x_3, x_4 \in \mathbb{R} \right\} \\ &= \left\{ x_3 \begin{bmatrix} -9 \\ 3 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \end{bmatrix} \mid x_3, x_4 \in \mathbb{R} \right\} \\ &= span \left\{ \begin{bmatrix} -9 \\ 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}. \end{aligned}$$

Definición 5.37 *Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una transformación lineal. Decimos que T es* ***inyectiva*** *(o* ***uno-a-uno****) syss para cualesquiera dos vectores $\vec{x}, \vec{y} \in \mathbb{R}^n$ se cumple que*

$$\text{si } \vec{x} \neq \vec{y}, \text{ entonces } T(\vec{x}) \neq T(\vec{y})$$

(lo que se lee como que dos vectores distintos tienen imágenes distintas). Equivalentemente, T es inyectiva syss para cualesquiera dos vectores $\vec{x}, \vec{y} \in \mathbb{R}^n$ se cumple que

$$\text{si } T(\vec{x}) = T(\vec{y}), \text{ entonces } \vec{x} = \vec{y}.$$

Ejemplo 5.38 *Queremos determinar si la transformación lineal*

$$T_5 : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$$

introducida previamente en el Ejemplo 5.24 es inyectiva. Para este efecto, notamos que

$$T_5\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} -2 & 1 & 3 \\ 4 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} =$$

$$x \begin{bmatrix} -2 \\ 4 \end{bmatrix} + y \begin{bmatrix} 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} -2x + y + 3z \\ 4x \end{bmatrix},$$

por lo que

$$T_5\left(\begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = T_5\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\right).$$

Dado que T_5 envía dos vectores distintos a la misma imagen, concluimos que T_5 no es inyectiva.

Ejemplo 5.39 *Consideremos la transformación lineal $T_3 : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$ dada por*

$$T_3\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + y \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

introducida previamente en el Ejemplo 5.21. Como

$$T_3\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} = T_3\left(\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right),$$

T_3 no es inyectiva.

Ejemplo 5.40 *Consideremos la transformación lineal* $T_8 : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$ *dada por*

$$T_8\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ 2x \\ 0 \\ 0 \end{bmatrix}$$

y supongamos que deseamos determinar si T_8 *es o no inyectiva. Si*

$$T_8\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = T_8\left(\begin{bmatrix} c \\ d \end{bmatrix}\right),$$

entonces

$$\begin{bmatrix} a+b \\ 2a \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} c+d \\ 2c \\ 0 \\ 0 \end{bmatrix}$$

y se obtienen las igualdades,

$$\begin{aligned} a+b &= c+d, \\ 2a &= 2c. \end{aligned}$$

Por lo tanto, $a = c$ *y* $b = d$*. Concluímos pues que*

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix}$$

y que T_8 *es inyectiva.*

Ahora daremos una serie de caracterizaciones que permiten decidir fácilmente si una transformación lineal es o no inyectiva.

Proposición 5.41 *Si* $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *es una transformación lineal y* $A \in M_{m\times n}$ *es su matriz estándar asociada, entonces las siguientes afirmaciones son equivalentes:*

(a) T *es inyectiva.*

(b) Si $\vec{x}$ *es tal que* $T(\vec{x}) = \vec{0}_m$*, entonces* $\vec{x} = \vec{0}_n$*.*

(c) $Null(T) = \{\vec{0}_n\}$*.*

(d) El sistema lineal $A\vec{x} = \vec{0}_m$ *tiene solución única.*

(e) El conjunto $\{A^1, \ldots, A^n\}$ *es LI.*

Demostración. Mostremos primero que (a) implica (b). Supongamos que T es inyectiva y que $T(\vec{x}) = \vec{0}_m$. Debemos verificar que $\vec{x}$ está forzado a ser el vector cero de $\mathbb{R}^n$. Ahora bien, como $T(\vec{0}_n) = \vec{0}_m = T(\vec{x})$ y T es inyectiva, debe cumplirse que $\vec{x} = \vec{0}_n$.

Supongamos ahora que T satisface (b) y que $\vec{a}$ y $\vec{b}$ son tales que

$$T(\vec{a}) = T(\vec{b}); \tag{$\bigstar$}$$

debemos probar que $\vec{a} = \vec{b}$. Por linealidad de T y gracias a ($\bigstar$), se tiene que

$$T(\vec{a} - \vec{b}) = T(\vec{a}) - T(\vec{b}) = \vec{0}_m$$

y utilizando (b) concluimos que $\vec{a} - \vec{b} = \vec{0}_n$. Esto es, $\vec{a} = \vec{b}$. La equivalencia entre (b), (c) y (d) es inmediata a partir de la definición de $Null(T)$ y de la Observación 5.35. Finalmente, la equivalencia entre (d) y (e) se sigue de la Observación 4.9. ■

Ejemplo 5.42 *En el Ejemplo 5.39 ya verificamos que la transformación lineal $T_3 : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$, dada por*

$$T_3\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

no es inyectiva. Esta conclusión puede también obtenerse recordando que en el Ejemplo 5.34 obtuvimos que

$$Null(T_3) = \left\{ \begin{bmatrix} -y \\ y \end{bmatrix} \mid y \in \mathbb{R} \right\},$$

por lo que $Null(T_3) \neq \{\vec{0}\}$ y por la Proposición 5.41, T_3 no es inyectiva. Alternativamente, puede utilizarse la misma proposición, observando que la matriz estándar que representa a T_3 está dada como

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

y notando que sus columnas forman un conjunto LD, por lo que T_3 no es inyectiva.

Ejemplo 5.43 *En el Ejemplo 5.40 ya verificamos que la transformación lineal $T_8 : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$, dada por*

$$T_8\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ 2x \\ 0 \\ 0 \end{bmatrix},$$

es inyectiva. Utilizando la Proposición 5.41 llegaremos a la misma conclusión. Para ello, notamos que la matriz estándar que representa a T_8 *está dada por*

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

y sus columnas forman un conjunto LI, con lo cual T_8 *es inyectiva.*

Ejemplo 5.44 *En el Ejemplo 5.38 ya verificamos que la transformación lineal* $T_5 : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$*, cuya matriz estándar es*

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

no es inyectiva. Utilizando la Proposición 5.41 llegamos a la misma conclusión: el sistema homogéneo dado por

$$\begin{bmatrix} -2 & 1 & 3 \\ 4 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

tiene una infinidad de soluciones (nótese que hay más variables que ecuaciones), por lo que T_5 *no es inyectiva.*

El ejemplo anterior puede generalizarse como sigue:

Proposición 5.45 *Si* $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *es una transformación lineal y* $n > m$*, entonces* T *no es inyectiva.*

Demostración. Sea $A \in M_{m\times n}$ la matriz estándar que representa a T y consideremos al sistema homogéneo $A\vec{x} = \vec{0}_m$. Como este sistema es consistente y tiene una infinidad de soluciones (pues tiene más variables que ecuaciones), concluimos que T no es inyectiva. De hecho, esta infinidad de soluciones nos dice que existe un número infinito vectores en $\mathbb{R}^n$ que, bajo T, tienen como imagen al vector $\vec{0}_m$. ■

El siguiente resultado (basado esencialmente en las Proposiciones 5.28 y 5.41 y la discusión que precede a los Corolarios 4.30 y 4.31.) nos dice que, en el caso de transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^n$, la inyectividad es sinónimo de otras muchas propiedades.

Teorema 5.46 *Si* $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *es una transformación lineal y* $A \in M_{n\times n}$ *es la matriz estándar que representa a* T*, entonces las siguientes afirmaciones son equivalentes.*

1. $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *es suprayectiva.*

2. *Para cualquier vector* $\vec{b} \in \mathbb{R}^n$*, el sistema* $A\vec{x} = b$ *tiene solución (única).*

3. *La FER de* A *es la matriz identidad de* $n \times n$*.*

4. *Las columnas de* A *forman un conjunto linealmente independiente.*

5. *El sistema homogéneo* $A\vec{x} = \vec{0}_n$ *tiene exactamente una solución (a saber, la trivial).*

6. T *es inyectiva.*

7. $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *es biyectiva.*

Como es de esperar, decimos que una transformación lineal

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

es **biyectiva** syss T es suprayectiva e inyectiva.

Ejemplo 5.47 *Consideremos la transformación lineal*

$$T_6 : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

dada por

$$T_6\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ x - y \\ z \end{bmatrix},$$

introducida previamente en el Ejemplo 5.29. Ya hemos notado que la matriz estándar que representa a T_6 *es la matriz*

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

Claramente, sus columnas forman un conjunto LI por lo que T_6 *y su matriz asociada satisfacen todas las propiedades enunciadas en el teorema anterior. En particular,* T_6 *no sólo es suprayectiva (como habíamos determinado anteriormente), sino que además es inyectiva y en consecuencia, biyectiva.*

Finalizamos esta sección mostrando que las transformaciones lineales inyectivas preservan independencia lineal en el siguiente sentido.

Definición 5.48 *Dada una transformación lineal* $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$*, decimos que* T *preserva independencia lineal si dados cualesquiera vectores* $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p \in \mathbb{R}^n$ *con* $p \geq 1$ *se tiene que si* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ *es LI, entonces* $\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_p)\}$ *es LI.*

Proposición 5.49 *Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una transformación lineal, entonces T es inyectiva syss T preserva independencia lineal.*

Demostración. Supongamos que T es inyectiva y sean $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ un conjunto LI de vectores de $\mathbb{R}^n$ y $c_1, \ldots . c_p$ escalares, tales que

$$c_1 T(\vec{v}_1) + c_2 T(\vec{v}_2) + \cdots + c_p T(\vec{v}_p) = \vec{0}_m.$$

Por linealidad, tenemos que

$$\begin{aligned} c_1 T(\vec{v}_1) + c_2 T(\vec{v}_2) + \cdots + c_p T(\vec{v}_p) = \\ T(c_1 \vec{v}_1 + c_2 \vec{v}_2 + \cdots + c_p \vec{v}_p) = \vec{0}_m \end{aligned}$$

y como T es inyectiva, se tiene que

$$c_1 \vec{v}_1 + c_2 \vec{v}_2 + \cdots + c_p \vec{v}_p = \vec{0}_n.$$

Esta igualdad implica que

$$c_1 = c_2 = \cdots = c_p = 0,$$

puesto que $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es un conjunto LI. Concluimos así que el conjunto $\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_p)\}$ también es LI.

Supongamos ahora que T preserva la independencia lineal y sean $\vec{u}$ y $\vec{v}$ dos vectores distintos de $\mathbb{R}^n$. Entonces, $\vec{u} - \vec{v} \neq \vec{0}$ y $\{\vec{u} - \vec{v}\}$ es LI. Por hipótesis, el conjunto

$$\{T(\vec{u} - \vec{v})\}$$

es LI y por lo tanto,

$$\vec{0} \neq T(\vec{u} - \vec{v}) = T(\vec{u}) - T(\vec{v})$$

y $T(\vec{u}) \neq T(\vec{v})$. Concluimos así que T es inyectiva. ■

Consideraciones geométricas

Las propiedades 1 y 2 de la Definición 5.1 que caracterizan las transformaciones lineales tienen implicaciones geométricas naturales. Concretamente, el mote de "lineal" es porque envían líneas en líneas y de forma más general, paralelogramos en paralelogramos, lo cual ilustraremos en el contexto de $\mathbb{R}^2$.

Dados una transformación lineal $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ y vectores $\vec{u}, \vec{w} \in \mathbb{R}^2$, sea $\mathcal{L}$ la recta en la dirección de $\vec{u}$ que pasa por el punto final del vector $\vec{w}$. Esto es,

$$\mathcal{L} = \{\vec{w} + a\vec{u},\ a \in \mathbb{R}\}.$$

En virtud de que $T(\vec{w} + a\vec{u}) = T(\vec{w}) + aT(\vec{u})$, la imagen de $\mathcal{L}$ bajo T es igual al conjunto

$$T(\mathcal{L}) = \{\vec{v} \in \mathbb{R}^2 \mid \vec{v} = T(\vec{w}) + aT(\vec{u}),\ a \in \mathbb{R}\}.$$

Es decir, esta imagen no es sino la recta en dirección del vector $T(\vec{u})$ que pasa por el punto final de $T(\vec{w})$. Estas consideraciones se ilustran en la Figura 5.1 donde $\vec{v} \in \mathcal{L}$ se transforma, bajo T, en $T(\vec{v}) \in T(\mathcal{L})$.

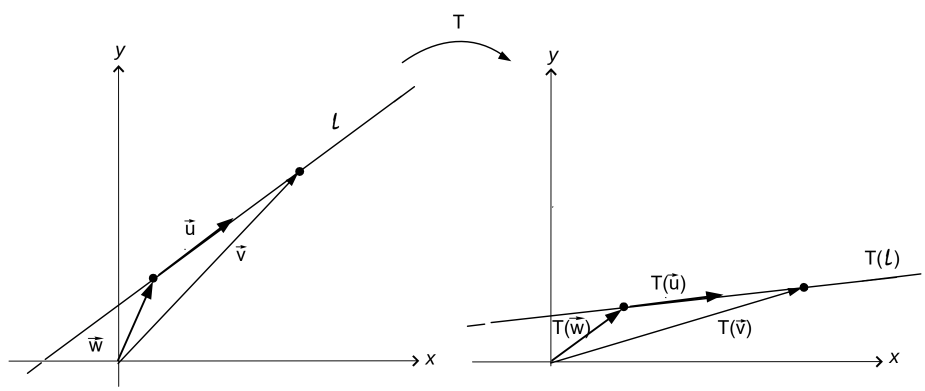


Figura 5.1: Imagen de una recta.

Observación 5.50 *Si $\vec{u} = \vec{0}$, entonces $\mathcal{L}$ y $T(\mathcal{L})$ consisten simplemente de un punto que puede considerarse como un caso degenerado de una recta. Asimismo, si $\vec{u} \neq \vec{0}$ y $T(\vec{u}) = \vec{0}$ (esto podría suceder si T no es inyectiva), entonces T transforma la recta $\mathcal{L}$ en un punto.*

Ahora supongamos que $\vec{u}$, $\vec{v}$ y $\vec{w}$ son tres vectores distintos en $\mathbb{R}^2$ con $\vec{u}$, $\vec{v} \neq \vec{0}$ (para evitar el caso trivial) y sea $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ una transformación lineal. A continuación, formamos un paralelogramo $\mathcal{P}$ con vértices dados por $\vec{w}$, $\vec{w} + \vec{u}, \vec{w} + \vec{u} + \vec{v}$ y $\vec{w} + \vec{v}$ como se muestra en la Figura 5.2.

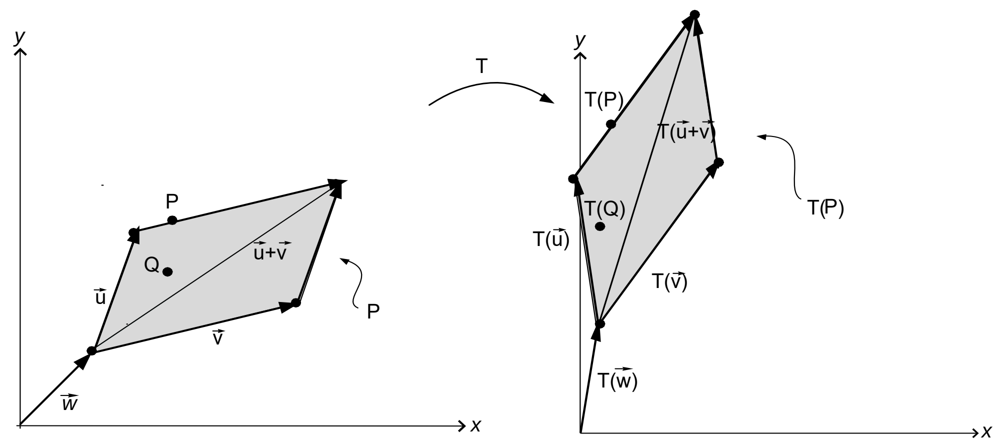


Figura 5.2: Imagen de un paralelogramo.

Entonces, $\mathcal{P}$ puede expresarse como el siguiente subconjunto de $\mathbb{R}^2$.

$$\mathcal{P} = \{\vec{w} + a\vec{u} + b\vec{v} \in \mathbb{R}^2 \mid a, b \in [0,1]\}.$$

Dado que $T(\vec{w} + a\vec{u} + b\vec{v}) = T(\vec{w}) + aT(\vec{u}) + bT(\vec{v})$, la imagen de $\mathcal{P}$ bajo la transformación T está dada como

$$T(\mathcal{P}) = \{T(\vec{w}) + aT(\vec{u}) + bT(\vec{v}) \in \mathbb{R}^2 \mid a, b \in [0,1]\},$$

conjunto que, claramente, también forma un paralelogramo.

Notemos que cualquier punto P en el perímetro de $\mathcal{P}$ es tal que $a \in \{0,1\}$ o $b \in \{0,1\}$ y que cualquier punto Q en el interior satisface $a, b \in (0,1)$. Entonces $T(P)$ y $T(Q)$ son, respectivamente, puntos en el perímetro y en el interior de $T(\mathcal{P})$ (ver Figura 5.2).

Observación 5.51 *Si $T(\vec{u}) = T(\vec{v})$, entonces T transformará el paralelogramo original en una recta que puede pensarse como un caso extremo de paralelogramo. Esta situación puede suceder cuando T no es inyectiva.*

Ejemplo 5.52 *Si el conjunto $\mathcal{C}$ en $\mathbb{R}^2$ es el cuadrado cuyos vértices están en los puntos $(0,0)$, $(1,0)$, $(1,1)$ y $(0,1)$, entonces para la transformación lineal dada por*

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 3x \\ 2y \end{bmatrix},$$

la imagen del cuadrado, $T(\mathcal{C})$, se ilustra en la Figura 5.3. Observamos que todos los puntos intermedios entre los vértices van a dar a puntos intermedios entre las imágenes de los mismos vértices, por ejemplo, la imagen del punto $(\frac{1}{2}, 1)$ es el punto $(\frac{3}{2}, 2)$ (ver Figura 5.3).

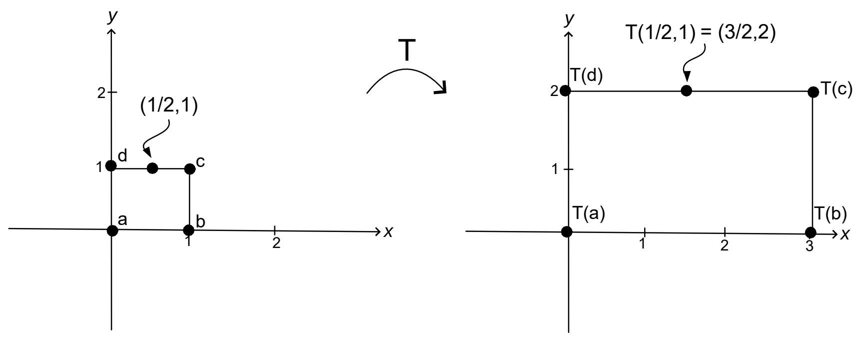


Figura 5.3: Imagen bajo T del cuadrado $\mathcal{C}$.

Ejemplo 5.53 *Sea $\mathcal{C}$ el mismo cuadrado del ejemplo anterior y considérese la transformación lineal (no inyectiva) dada por*

$$S\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x+y \end{bmatrix}.$$

La imagen del cuadrado es ahora el segmento de recta que va del origen al punto $(2,2)$ *y los cuatro vértices originales se colapsan en tres puntos:* $(0,0), (1,1)$ *y* $(2,2)$. *Esto se ilustra en la Figura 5.4.*

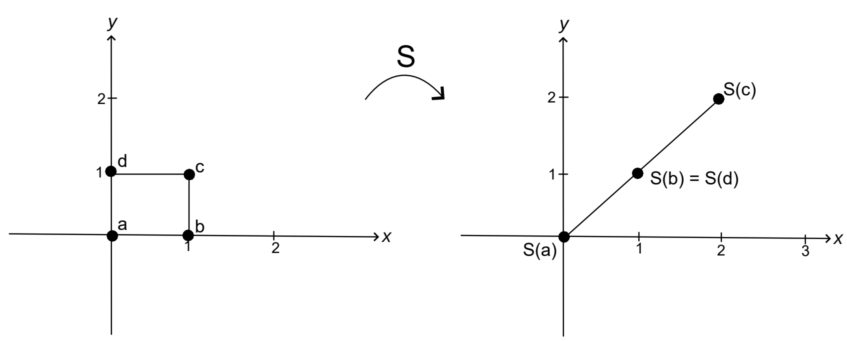


Figura 5.4: Imagen bajo S del cuadrado $\mathcal{C}$.

La acción de algunas transformaciones lineales (inyectivas) puede interpretarse en términos geométricos simples como reflexiones, contracciones, expansiones, rotaciones, deslizamientos, etcétera. El Ejemplo 5.52, visto anteriormente, representa una expansión del triple a lo largo del eje horizontal y del doble en el sentido vertical.

Ejemplo 5.54 *Consideremos la transformación lineal de* $\mathbb{R}^2$ *en* $\mathbb{R}^2$ *cuya matriz estándar está dada por*

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

Dado que

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix},$$

dicha transformación es una reflexión sobre el eje horizontal tal y como se ilustra en la Figura 5.5.

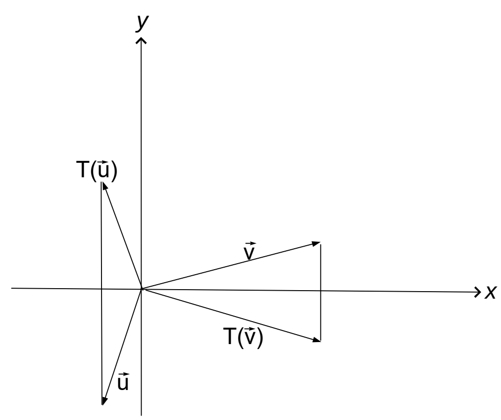


Figura 5.5.

Figura 5.5: Reflexión sobre el eje horizontal.

Ejemplo 5.55 *Si ahora deseamos una transformación lineal, digamos* R*, de* $\mathbb{R}^2$ *en* $\mathbb{R}^2$*, que rote a cada vector un ángulo* θ*, entonces necesitamos describir la acción de* R *sobre los vectores* $\vec{e}_1$ *y* $\vec{e}_2$*. Esto se muestra en la Figura 5.6 en la cual ilustramos la rotación sobre el círculo unitario. La matriz estándar asociada a esta transformación es simplemente*

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

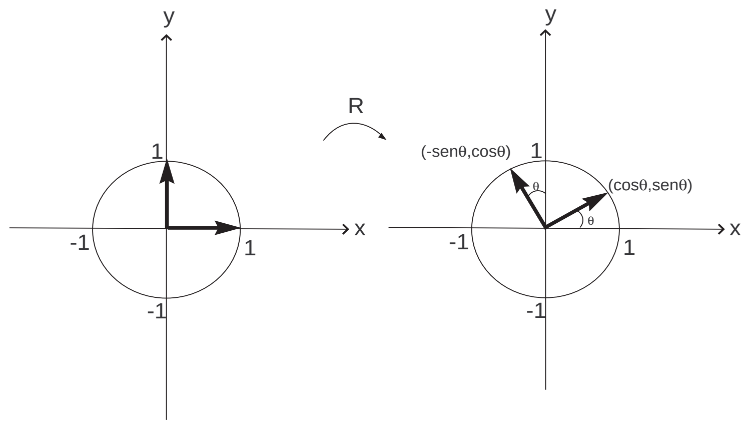


Figura 5.6: Rotación por un ángulo θ.

Ejemplo 5.56 *Consideremos la transformación lineal*

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ y \end{bmatrix}$$

cuya matriz estándar está dada por

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

Esta transformación representa lo que se conoce como un deslizamiento horizontal, lo cual se ilustra en la Figura 5.7.

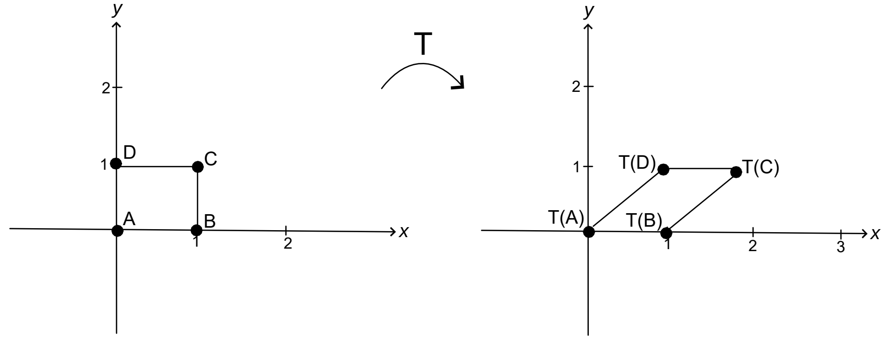


Figura 5.7: Deslizamiento horizontal.

Ejercicios

Ejercicio 5.1 *Considerar las siguientes funciones de $\mathbb{R}^2$ en $\mathbb{R}^2$ dadas por*

$$T_1\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x^2 \\ 7y \end{bmatrix},\ T_2\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x-y \\ x+y \end{bmatrix}, T_3\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y+1 \end{bmatrix}.$$

Determine si son transformaciones lineales. En caso afirmativo, encuentre la matriz estándar que las representa.

Ejercicio 5.2 *Sean S y U las funciones de $\mathbb{R}^3$ en $\mathbb{R}^3$ dadas por*

$$S\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 3x-\pi y+z \\ 2y+6z \\ x-3z \end{bmatrix}, U\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x-y \\ x+y \\ |z| \end{bmatrix}.$$

Determinar si son transformaciones lineales. En caso afirmativo, encontrar la matriz estándar que las representa.

Ejercicio 5.3 *Considerar las siguientes funciones:*

$$F:\mathbb{R}^3 \longrightarrow \mathbb{R}^2,\ G:\mathbb{R}^2 \longrightarrow \mathbb{R},\ H:\mathbb{R} \longrightarrow \mathbb{R}^4$$

dadas por

$$F\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} y \\ y \end{bmatrix},\ G\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = x+y,\ H(x) = \begin{bmatrix} x \\ 0 \\ 0 \\ x \end{bmatrix}.$$

Demostrar que son transformaciones lineales y encontrar las matrices estándar que las representan.

Ejercicio 5.4 *Determinar si existe una transformación lineal*

$$T:\mathbb{R}^2 \to \mathbb{R}^2$$

tal que

$$T\left(\begin{bmatrix} 3 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ T\left(\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}\ y\ T\left(\begin{bmatrix} 4 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Ejercicio 5.5 *La **inmersión natural** de $\mathbb{R}^2$ en $\mathbb{R}^3$ es la función $\Psi : \mathbb{R}^2 \to \mathbb{R}^3$ dada por*

$$\Psi\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}.$$

Probar que esta función es una transformación lineal y encontrar la matriz estándar que la representa.

Ejercicio 5.6 *La* ***proyección natural*** *de* $\mathbb{R}^3$ *sobre* $\mathbb{R}^2$ *es la función* $P : \mathbb{R}^3 \to \mathbb{R}^2$ *dada por*

$$P\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Probar que esta función es una transformación lineal y encontrar la matriz estándar que la representa.

Ejercicio 5.7 *Encontrar la matriz estándar que representa a la transformación lineal* $T : \mathbb{R}^3 \to \mathbb{R}^3$ *que satisface las siguientes tres igualdades:*

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix},\ T\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 4 \\ 4 \end{bmatrix},\ T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Ejercicio 5.8 *Encontrar la matriz estándar que representa a la transformación lineal* $U : \mathbb{R}^2 \to \mathbb{R}^2$ *que satisface las siguientes dos igualdades:*

$$U\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 4 \end{bmatrix},\ U\left(\begin{bmatrix} 0 \\ 0.5 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

Ejercicio 5.9 *Encontrar la matriz estándar que representa a la transformación lineal* $S : \mathbb{R}^2 \to \mathbb{R}^2$ *que satisface las siguientes dos igualdades:*

$$S\left(\begin{bmatrix} 3 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 9 \\ 17 \end{bmatrix},\ S\left(\begin{bmatrix} -2 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} -17 \\ -4 \end{bmatrix}.$$

Ejercicio 5.10 *Sean* T_1 *y* T_2 *transformaciones lineales de* $\mathbb{R}^n$ *en* $\mathbb{R}^m$. *Probar que si* $T_1 \neq T_2$, *entonces existe un elemento* $\vec{e}_j$ *de la base estándar tal que* $T_1(\vec{e}_j) \neq T_2(\ \vec{e}_j)$.

Ejercicio 5.11 *Encontrar mil transformaciones lineales,*

$$T : \mathbb{R}^2 \to \mathbb{R}^2,$$

distintas entre sí y tales que

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \pi \\ 666 \end{bmatrix}.$$

Para cada una de estas transformaciones especifique quién es $T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$.

Ejercicio 5.12 *Encontrar una transformación lineal $L : \mathbb{R}^2 \to \mathbb{R}^4$ tal que*

$$L\left(\begin{bmatrix}3\\6\end{bmatrix}\right) = \begin{bmatrix}81\\\sqrt{2}\\0\\-9\end{bmatrix}.$$

Ejercicio 5.13 *Encontrar el rango de las siguientes transformaciones y determine si cada una de ellas es o no suprayectiva.*

a. $L_1 : \mathbb{R}^2 \to \mathbb{R}^2$, *donde*

$$L_1\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \begin{bmatrix}\pi & -7\\666 & \sqrt{2}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}.$$

b. $L_2 : \mathbb{R}^2 \to \mathbb{R}^2$, *donde*

$$L_2\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \begin{bmatrix}18 & -3\\-3 & \frac{1}{2}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}.$$

c. $L_3 : \mathbb{R}^3 \to \mathbb{R}^2$, *donde*

$$L_3\left(\begin{bmatrix}x\\y\\z\end{bmatrix}\right) = \begin{bmatrix}18 & -3 & -12\\-3 & \frac{1}{2} & 2\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}.$$

d. $L_4 : \mathbb{R}^3 \to \mathbb{R}^3$, *donde*

$$L_4\left(\begin{bmatrix}x\\y\\z\end{bmatrix}\right) = \begin{bmatrix}2 & 5 & 7\\1 & 2 & 0\\0 & 0 & 6\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}.$$

e. $L_5 : \mathbb{R}^2 \to \mathbb{R}^3$, *donde*

$$L_5\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \begin{bmatrix}5 & -4\\-2 & 2\\-3 & 6\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}.$$

f. $L_6 : \mathbb{R}^3 \to \mathbb{R}^3$, *donde*

$$L_6\left(\begin{bmatrix}x\\y\\z\end{bmatrix}\right) = \begin{bmatrix}5 & -4 & 1\\-2 & 2 & 0\\-3 & 6 & 3\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}.$$

g. $L_7 : \mathbb{R}^3 \to \mathbb{R}^4$, *donde*

$$L_7\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 5 & -4 & 1 \\ -2 & 2 & 0 \\ -3 & 6 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Ejercicio 5.14 *Determinar si cada una de las transformaciones del ejercicio anterior es o no inyectiva.*

Ejercicio 5.15 *Demostrar que si* $T : \mathbb{R}^n \to \mathbb{R}^m$ *es una transformación lineal no inyectiva y* $\vec{y}_0$ *está en el rango de* T*, entonces existe un número infinito de vectores* $\vec{x}$ *tales que* $T(\vec{x}) = \vec{y}_0$*.*

Ejercicio 5.16 *Sean* A *y* B *matrices de* $m \times n$ *y sean* T *y* T' *las transformaciones de* $\mathbb{R}^n$ *en* $\mathbb{R}^m$ *dadas, respectivamente, por* $T(\vec{x}) = A\vec{x}$ *y* $T'(\vec{x}) = B\vec{x}$*. Pruebe que si* $A \neq B$*, entonces existe un vector* $\vec{x}_0$ *de* $\mathbb{R}^n$ *tal que* $T(\vec{x}_0) \neq T'(\vec{x}_0)$*.*

Ejercicio 5.17 *Sea* $T : \mathbb{R}^n \to \mathbb{R}^m$ *una transformación lineal y sean* $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ *vectores de* $\mathbb{R}^n$*. Probar que si el conjunto* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ *es LD, entonces el conjunto* $\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_p)\}$ *también lo es.*

Ejercicio 5.18 *Probar (en general) o refutar (con un contraejemplo) los siguientes enunciados (en cada caso se debe indicar claramente "verdadero" o "falso").*

a. *Sea* A *una matriz de* $m \times n$*, sea* $\vec{b}$ *un vector de* $\mathbb{R}^m$ *y sea*

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

la función dada por $T(\vec{x}) = A\vec{x} + \vec{b}$*. Si* $\vec{b} \neq \vec{0}$*, entonces* T *no es una transformación lineal.*

b. *Si* $T : \mathbb{R}^n \to \mathbb{R}^m$ *es una transformación lineal y* $n > m$*, entonces* $T : \mathbb{R}^n \to \mathbb{R}^m$ *es suprayectiva.*

c. *Si* $T : \mathbb{R}^n \to \mathbb{R}^m$ *es una transformación lineal y* $n < m$*, entonces* T *es inyectiva.*

d. *Si* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *es una transformación lineal distinta de la función constante cero, entonces* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *es biyectiva.*

e. *Si* $T : \mathbb{R}^n \to \mathbb{R}^m$ *es una transformación lineal y* $n \neq m$*, entonces* $T : \mathbb{R}^n \to \mathbb{R}^m$ *no es biyectiva.*

f. *Sea* $T : \mathbb{R}^n \to \mathbb{R}^m$ *una transformación lineal y sean* $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ *vectores de* $\mathbb{R}^n$. *Si el conjunto* $\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_p)\}$ *es LD, entonces el conjunto* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ *también lo es.*

g. *Sea* $T : \mathbb{R}^n \to \mathbb{R}^m$ *una transformación lineal y sean* $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ *vectores de* $\mathbb{R}^n$. *Si* T *es inyectiva y* $\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_p)\}$ *es LD, entonces el conjunto* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ *es LD.*

Ejercicio 5.19 *Sean* T_1 *y* T_2 *las transformaciones de* $\mathbb{R}^2$ *en* $\mathbb{R}^2$ *dadas por*

$$T_1\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -y \\ x \end{bmatrix}, \; T_2\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x - 2y \\ 3x + y \end{bmatrix}.$$

a. *Dibujar la imagen bajo* T_1 *del cuadrado cuyos vertices están en los puntos* $(0,0)$, $(1,0), (0,1)$ *y* $(1,1)$.

b. *Dibujar la imagen bajo* T_2 *del cuadrado cuyos vertices están en los puntos* $(0,0)$, $(1,0), (1,-1)$ *y* $(0,-1)$.

Ejercicio 5.20 *Encontrar la matriz estándar de las transformaciones lineales de* $\mathbb{R}^2$ *en* $\mathbb{R}^2$ *que representen las siguientes operaciones geométricas e ilustra la imagen del cuadrado con vértices* $(0,0), (1,0), (1,1)$ *y* $(0,1)$ *bajo las mismas.*

a. *Una reflexión sobre la recta* $y = x$.

b. *Una rotación por un ángulo* $\frac{\pi}{4}$.

c. *Una contracción del 50 % a lo largo del eje horizontal.*

Ejercicio 5.21 *Considera el cuadrado con vértices* $(0,0), (1,0), (0,1)$ *y* $(1,1)$. *Ilustra como se transforma este cuadrado bajo las transformaciones lineales representadas por las siguientes matrices.*

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

b. $\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$.

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

d. $\begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$.

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

Capítulo 6

Álgebra matricial

Introducción

Recordemos que en los cursos elementales de cálculo se estudian funciones (continuas o diferenciables, por ejemplo), cuyo dominio es algún subconjunto Ω de $\mathbb{R}^n$ y que toman valores en los números reales, es decir, funciones del tipo

$$f : \Omega \longrightarrow \mathbb{R},$$

las cuales pueden sumarse entre sí y multiplicarse por escalares simplemente definiendo, para toda pareja de funciones f, g de Ω en $\mathbb{R}$ y para todo número real c, las correspondientes funciones de suma y producto como sigue:

$$\begin{aligned}(f+g) &: \Omega \longrightarrow \mathbb{R},\\ (f+g)(x) &= f(x)+g(x),\\ (cf) &: \Omega \longrightarrow \mathbb{R},\\ (cf)(x) &= cf(x).\end{aligned}$$

Con estas definiciones, $f+g$ y cf siguen teniendo las mismas propiedades que las funciones originales. Por ejemplo, si f y g son continuas (o diferenciables o doblemente diferenciables, por ejemplo), $f+g$ y cf también lo serán. Asimismo, puede demostrarse que el conjunto de funciones

$$\{f : \Omega \longrightarrow \mathbb{R} \,|f \text{ es continua}\},$$

en donde la palabra “continua” puede sustituirse por “diferenciable” o “doblemente diferenciable” o alguna característica similar, satisface el símil de todas las propiedades del espacio vectorial $\mathbb{R}^n$ dadas en la Proposición 3.6.

Una pregunta natural es si estas operaciones pueden definirse para las transformaciones lineales de manera que la suma y el producto por un escalar también sean transformaciones lineales. Lo más natural es utilizar la misma definición que se utiliza en cálculo. Concretamente, si

$$\begin{aligned} T &: \mathbb{R}^n \longrightarrow \mathbb{R}^m \ \text{ y} \\ S &: \mathbb{R}^n \longrightarrow \mathbb{R}^m \end{aligned}$$

son dos transformaciones lineales y $c \in \mathbb{R}$, definimos

$$(T+S) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

como

$$(T+S)(\vec{x}) = T(\vec{x}) + S(\vec{x}).$$

Similarmente,

$$(cT) : \mathbb{R}^n \longrightarrow \mathbb{R}^m \ ,$$

está dada por

$$(cT)(\vec{x}) = cT(\vec{x}).$$

Es fácil verificar que las transformaciones así definidas también son lineales, los cual se deja como ejercicio.

Ahora bien, si A y B son las matrices estándar asociadas a T y S, la matriz estándar asociada a $T+S$ debería ser $A+B$ y la matriz estándar asociada a cT debería corresponder a cA ¿Cómo son estas matrices $A+B$ y cA? En este capítulo se estudian éstas y otras operaciones para matrices.

$M_{m\times n}$ como espacio vectorial

En este capítulo, dada una matriz $A \in M_{m\times n}$, utilizaremos de forma intercambiable los símbolos a_{ij} y A_{ij} para denotar a la componente que está en el renglón i y en la columna j de A.

Definición 6.1 *Sean $A, B \in M_{m\times n}$ y sea c un número real.*

*(i) La **suma de** A **y** B se define como la matriz $A+B \in M_{m\times n}$ cuyas entradas están dadas por*

$$(A+B)_{ij} = A_{ij} + B_{ij},$$

para toda $i = 1, 2, \ldots, m$ y para toda $j = 1, 2, \ldots, n$.

(ii) *El* ***producto del escalar*** c ***por la matriz*** A *se define como la matriz* $cA \in M_{m\times n}$ *cuyas entradas están dadas por*

$$(cA)_{ij} = cA_{ij},$$

para toda $i = 1, 2, \ldots, m$ *y para toda* $j = 1, 2, \ldots, n$.

Notación: En el contexto de $M_{m\times n}$ utilizaremos las siguientes convenciones.

- *Recordemos que la matriz cero (aquella cuyas componentes son todas cero) se denota por* Θ.
- *Si* $A \in M_{m\times n}$, *entonces* $-A$ *denota a la matriz* $(-1)A$.
- *Si* $A, B \in M_{m\times n}$, *entonces* $A - B$ *significa* $A + (-B)$.
- *Como siempre, la palabra escalar es sinónimo de número real.*

Ejemplo 6.2 *Consideremos a las matrices*

$$A = \begin{bmatrix} 3 & 4 & 5 \\ -4 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & -4 & 9 \\ 6 & 0 & -12 \end{bmatrix}.$$

Entonces

$$A + B = \begin{bmatrix} 10 & 0 & 14 \\ 2 & 2 & -11 \end{bmatrix}, 2A = \begin{bmatrix} 6 & 8 & 10 \\ -8 & 4 & 2 \end{bmatrix},$$
$$2A - 3B = \begin{bmatrix} -15 & 20 & -17 \\ -26 & 4 & 38 \end{bmatrix}.$$

Hasta ahora hemos definido las operaciones de suma y producto por escalares en dos contextos aparentemente distintos: el de matrices y el de transformaciones lineales. Sin embargo, el siguiente resultado establece una íntima conexión entre ambos.

Proposición 6.3 *Sean* $S, T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *dos transformaciones lineales con matrices estándar asociadas* A *y* B, *respectivamente. Sea* c *un escalar, entonces* $A + B$ *es la matriz estándar asociada a* $S + T$ *y* cA *es la matriz estándar asociada a* cS.

Demostración. Por definición, la matriz estándar asociada a $S + T$ está dada por

$$\begin{aligned}
&[(S+T)(\vec{e}_1)\ (S+T)(\vec{e}_1) \cdots (S+T)(\vec{e}_n)] \\
&= [S(\vec{e}_1) + T(\vec{e}_1)\ S(\vec{e}_2) + T(\vec{e}_2) \cdots S(\vec{e}_n) + T(\vec{e}_n)] \\
&= [S(\vec{e}_1)\ S(\vec{e}_2) \cdots S(\vec{e}_n)] + [T(\vec{e}_1)\ T(\vec{e}_2) \cdots T(\vec{e}_n)] \\
&= A + B.
\end{aligned}$$

Similarmente, la matriz asociada a cS se calcula como

$$\begin{aligned}[(cS)(\vec{e}_1)\ (cS)(\vec{e}_2)\cdots(cS)(\vec{e}_n)] &= [cS(\vec{e}_1)\ cS(\vec{e}_2)\cdots cS(\vec{e}_n)] \\ &= c\,[S(\vec{e}_1)\ S(\vec{e}_2)\cdots S(\vec{e}_n)] = cA,\end{aligned}$$

con lo cual se concluye la demostración. ■

Ahora veremos que $\mathbb{R}^n$ también tiene profundas similitudes algebraicas con $M_{m\times n}$. Concretamente ambos son casos particulares de lo que se conoce como un espacio vectorial.

Proposición 6.4 *El conjunto $M_{m\times n}$, junto con las operaciones que acabamos de definir de suma entre matrices y de producto por un escalar, satisface las siguientes propiedades:*

- *V1 Si $A, B \in M_{m\times n}$, entonces $A + B = B + A$ (conmutatividad de la suma).*
- *V2 Si $A, B, C \in M_{m\times n}$, entonces $(A + B) + C = A + (B + C)$ (asociatividad de la suma).*
- *V3 Si $A \in M_{m\times n}$, entonces $A+\Theta = A$ (existencia de neutro aditivo).*
- *V4 Si $A \in M_{m\times n}$, entonces $A - A = \Theta$ (existencia de inversos aditivos).*
- *V5 Si $A \in M_{m\times n}$, entonces $1A = A$.*
- *V6 Si $A \in M_{m\times n}$ y β y γ son escalares, entonces $(\beta\gamma)A = \beta(\gamma A)$.*
- *V7 Si $A \in M_{m\times n}$ y β y γ son escalares, entonces $(\beta+\gamma)A = \beta A+\gamma A$ (primera forma de distributividad).*
- *V8 Si $A, B \in M_{m\times n}$ y β es un escalar, entonces $\beta(A+B) = \beta A+\beta B$ (segunda forma de distributividad).*

Demostración. Con el fin de sugerir paralelismos con la Proposición 3.6, probaremos $(V3)$ y $(V7)$. En el caso de $(V3)$ notamos que

$$(A + \Theta)_{ij} = A_{ij} + \Theta_{ij} = A_{ij},$$

por lo que las matrices $A + \Theta$ y A son iguales (pues coinciden entrada por entrada). De forma similar, en el caso de $(V7)$ tenemos que

$$((\beta+\gamma)A)_{ij} = (\beta+\gamma)A_{ij} = \beta A_{ij}+\gamma A_{ij} = (\beta A)_{ij}+(\gamma A)_{ij} = (\beta A+\gamma A)_{ij},$$

por lo que las matrices $(\beta + \gamma)A$ y $\beta A + \gamma A$ son iguales (pues tienen las mismas componentes). ■

Estas ocho propiedades son idénticas a las dadas en la Proposición 3.6 para los espacios vectoriales $\mathbb{R}^n$ y como ya hemos dicho anteriormente, se denominan **propiedades de espacio vectorial**. En general, cualquier conjunto cuyos objetos pueden sumarse y multiplicarse por escalares, de tal manera que satisfacen los análogos de V1 a V8, es llamado un espacio vectorial. De esta forma, se generalizan las propiedades fundamentales de los espacios $\mathbb{R}^n$ a estructuras más abstractas, que nada tienen que ver con conjuntos de "flechas" en $\mathbb{R}^n$. Por ejemplo, el conjunto $M_{m\times n}$ o equivalentemente el conjunto

$$\{T : \mathbb{R}^n \longrightarrow \mathbb{R}^m \mid T \text{ es una transformación lineal}\}$$

tienen la estructura de espacio vectorial[1].

Se tiene así que las Proposiciones 3.7 y 3.9, junto con el Corolario 3.8, también son válidas en el nuevo entorno de matrices (o de transformaciones lineales). A continuación enunciamos estos resultados. Se proporciona la demostración para la ley de la cancelación para que el lector se de cuenta que es idéntica al caso de $\mathbb{R}^n$ dado en el Capítulo 3.

Proposición 6.5 (Ley de la cancelación) *Si $A, B, C \in M_{m\times n}$, y $A + B = A + C$, entonces $B = C$.*

Demostración. Supongamos $A + B = A + C$. Sumando $-A$ a ambos lados de esta igualdad y aplicando las propiedades de asociatividad del inverso y del neutro aditivo, obtenemos

$$\begin{aligned} B &= \Theta + B = (-A + A) + B = -A + (A + B) = -A + (A + C) \\ &= (-A + A) + C = \Theta + C = C, \end{aligned}$$

concluyendo así la demostración. ■

Corolario 6.6 *Si $A, B \in M_{m\times n}$ y $A + B = A$, entonces $B = \Theta$.*

Proposición 6.7 *Si $A \in M_{m\times n}$ y β es un escalar, entonces $\beta A = \Theta$ syss $\beta = 0$ o $A = \Theta$.*

Terminamos esta sección mostrando que la multiplicación de una matriz por un vector es siempre consistente con las operaciones de suma y multiplicación escalar.

[1] Así como también conjuntos de funciones del tipo

$$\{f : \Omega \longrightarrow \mathbb{R} \mid f \text{ es continua}\}.$$

Proposición 6.8 *Si A, B son matrices de $m \times n$, $\vec{x}$ y $\vec{y}$ son vectores de $\mathbb{R}^n$ y r es un escalar, entonces:*

1. $A(\vec{x}+\vec{y}) = A\vec{x} + A\vec{y}$.
2. $(A+B)\vec{x} = A\vec{x} + B\vec{x}$.
3. $(rA)\vec{x} = r(A\vec{x}) = A(r\vec{x})$.

Demostración. La demostración de (1) se realizó en el Lema 4.36 del Capítulo 4. En el caso de (2), definiendo las transformaciones lineales

$$S, T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

como $S(\vec{x}) = A\vec{x}$ y $T(\vec{x}) = B\vec{x}$, se tiene que por la Proposición 6.3

$$(A+B)\vec{x} = (S+T)\vec{x} = S(\vec{x}) + T(\vec{x}) = A\vec{x} + B\vec{x}.$$

Finalmente, para demostrar la primera igualdad de (3) - la segunda se realizó en el Ejercicio 3.3 -, notamos que si $S : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es como antes, una vez más por la Proposición 6.3 se obtiene

$$(rA)\vec{x} = (rS)(\vec{x}) = r(S(\vec{x})) = r(A\vec{x}),$$

lo cual concluye la demostración. ■

La transpuesta de una matriz

Definición 6.9 *Si $A \in M_{m\times n}$, entonces* ***la matriz transpuesta de** A, denotada por A^T, se define como la matriz $A^T \in M_{n\times m}$, donde*

$$(A^T)_{ij} = A_{ji},$$

para toda $i \leq n$ y para toda $j \leq m$. En otras palabras, la matriz transpuesta de A es la matriz que se obtiene a partir de A convirtiendo (en el orden natural) sus renglones en columnas (o lo que es lo mismo sus columnas en renglones).

Notemos que si $A \in M_{m\times n}$, entonces A^T puede visualizarse como

$$A^T = \begin{bmatrix} (A^1)^T \\ (A^2)^T \\ \vdots \\ (A^n)^T \end{bmatrix} = \begin{bmatrix} (A_1)^T & (A_2)^T & \cdots & (A_m)^T \end{bmatrix},$$

pues
$$\left(A^T\right)_j = \left(A^j\right)^T \text{ y } \left(A^T\right)^i = (A_i)^T.$$

Por ejemplo, si
$$A = \begin{bmatrix} 3 & 4 & 5 \\ -4 & 2 & 1 \end{bmatrix},$$
entonces
$$A^T = \begin{bmatrix} 3 & -4 \\ 4 & 2 \\ 5 & 1 \end{bmatrix}.$$
En particular,
$$(A_1)^T = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}, \ (A_2)^T = \begin{bmatrix} -4 \\ 2 \\ 1 \end{bmatrix},$$
$$\left(A^1\right)^T = \begin{bmatrix} 3 & -4 \end{bmatrix}, \ \left(A^2\right)^T = \begin{bmatrix} 4 & 2 \end{bmatrix}.$$
El siguiente resultado enlista algunas de las propiedades más utilizadas de la matriz transpuesta. La demostración de todas ellas es inmediata.

Proposición 6.10 *Las siguientes afirmaciones son verdaderas*

1. *Si $A \in M_{m\times n}$, entonces $(A^T)^T = A$.*
2. *Si $A, B \in M_{m\times n}$, entonces $(A+B)^T = A^T + B^T$.*
3. *Si $A \in M_{m\times n}$ y r es un escalar, entonces $(rA)^T = rA^T$.*

Definición 6.11 *Sea $A \in M_{n\times n}$, decimos que A es una* ***matriz simétrica*** *syss $A = A^T$.*

Ejemplo 6.12 *Si $A \in M_{n\times n}$, entonces la matriz $A + A^T$ es simétrica. En efecto, utilizando los incisos (1) y (2) de la proposición anterior tenemos que:*
$$\left(A + A^T\right)^T = A^T + (A^T)^T = A^T + A = A + A^T.$$

Observemos que si $A \in M_{n\times n}$ es una matriz simétrica, entonces para todo $i, j = 1, \ldots, n$ debe cumplirse $A_{ij} = \left(A^T\right)_{ij} = A_{ji}$, es decir, la matriz tiene simetría con respecto a su diagonal principal.

Ejemplo 6.13 *Las siguientes matrices son simétricas:*
$$\begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 4 & -5 \\ 4 & -1 & 3 \\ -5 & 3 & 666 \end{bmatrix}.$$

Ejemplo 6.14 *En cálculo vectorial se utiliza la matriz hessiana asociada a ciertas funciones reales. Ésta es la matriz simétrica de las segundas derivadas como se ve a continuación. Si*

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

es una función doblemente diferenciable y con segundas derivadas continuas, se define la ***matriz hessiana*** *de f como*

$$Hf(x,y) = \begin{bmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{bmatrix}.$$

Por continuidad de las derivadas tenemos que $f_{xy}(x,y) = f_{yx}(x,y)$, por lo que esta matriz es simétrica. Por ejemplo, si $f(x,y) = x^2 + xy + y^3$, entonces

$$Hf(x,y) = \begin{bmatrix} 2 & 1 \\ 1 & 6y \end{bmatrix}.$$

En general, si

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

es doblemente diferenciable con segundas derivadas continuas, la matriz hessiana de f (evaluada en $\vec{x} \in \mathbb{R}^n$) es la matriz simétrica dada por

$$Hf(\vec{x}) = \begin{bmatrix} f_{x_1x_1}(\vec{x}) & \cdots & f_{x_1x_n}(\vec{x}) \\ \vdots & \ddots & \vdots \\ f_{x_nx_1}(\vec{x}) & \cdots & f_{x_nx_n}(\vec{x}) \end{bmatrix}.$$

Producto punto o escalar

En esta sección se define una nueva operación entre vectores de $\mathbb{R}^n$.

Definición 6.15 *Sean*

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \ \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n.$$

El ***producto punto (o producto escalar)*** *entre $\vec{x}$ y $\vec{y}$ se define como*

$$\vec{x} \cdot \vec{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{i=1}^{n} x_iy_i.$$

Observemos que este producto asocia a cada pareja de vectores de $\mathbb{R}^n$ un escalar, es decir, un número real.

Ejemplo 6.16 *Dados*

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix} \in \mathbb{R}^2,$$

se tiene que

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix} = 1 \times 3 + (-1) \times 2 = 1.$$

La siguiente proposición proporciona las propiedades básicas del producto punto y su demostración se deja como ejercicio al lector.

Proposición 6.17 *Si $\vec{x}, \vec{y}, \vec{z}$ son vectores de $\mathbb{R}^n$ y r es un escalar, entonces se satisfacen las siguientes propiedades.*

P1 $\vec{x} \cdot \vec{x} = 0$ syss $\vec{x} = \vec{0}$.

P2 $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$.

P3 $\vec{x} \cdot (\vec{y} + \vec{z}\ ) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$.

P4 $(r\vec{x}) \cdot \vec{y} = r(\vec{x} \cdot \vec{y}) = \vec{x} \cdot (r\vec{y})$.

Ejemplo 6.18

$$\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \in \mathbb{R}^3,$$

se tiene que

$$\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix} = 5 - 5 + 0 = 0,$$

$$\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = 1 + 1 + 2 = 4,$$

por lo tanto

$$\begin{aligned} &\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \cdot \left(2 \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right) \\ &= \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \cdot \left(2 \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix} \right) + \left(\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right) \\ &= 2 \left(\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix} \right) + \left(\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right) \\ &= 2 \times 0 + 4 = 4. \end{aligned}$$

El producto punto tiene interpretaciones geométricas muy interesantes que exploraremos más adelante en los capítulos 14 y 15, pero por el momento nos limitaremos a su utilidad notacional para el desarrollo de operaciones puramente algebraicas. Por ejemplo, en ocasiones es necesario identificar coordenadas específicas de algún vector

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n.$$

Para este propósito, la siguiente relación, que es inmediata a partir de la definición del producto punto, es sumamente útil.

$$\vec{e}_i \cdot \vec{v} = v_i.$$

Por otro lado, dadas las matrices

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \in M_{1\times n} \text{ y } B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in M_{n\times 1};$$

notamos que las columnas de A consisten de un único elemento, a_j, y el producto AB, definido como en 3.24 del capítulo 3, es un escalar dado por[2],

$$AB = a_1b_1 + a_2b_2 + \cdots + a_nb_n \in \mathbb{R}.$$

El lector observador notará la similitud con el producto punto (o escalar) entre vectores que acabamos de definir. Esto no es casualidad puesto que se cumple

$$AB = a_1b_1 + a_2b_2 + \cdots + a_nb_n = A \cdot B,$$

en donde A y B se piensan como vectores[3] en $\mathbb{R}^n$.

Ejemplo 6.19 *Sean* $A = \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$ *y* $B = \begin{bmatrix} 1 \\ 0 \\ 666 \\ 7 \end{bmatrix}$ *vectores en* $\mathbb{R}^4$. *El producto punto* $A \cdot B$ *puede también escribirse como sigue.*

$$A \cdot B = (-1)(1) + (2)(0) + (0)(666) + (1)(7) = 8 = A^T B.$$

[2] En realidad, este producto es un vector con una única entrada (o una matriz de 1×1) que puede identificarse con un número real.

[3] Independientemente de que se trate de columnas o renglones, A y B pueden identificarse con los vectores $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_n)$ en $\mathbb{R}^n$.

Producto de matrices

En esta sección se introduce una operación nueva que llamamos producto de matrices. Para este fin, utilizamos como punto de partida la operación de composición entre transformaciones lineales. Sean

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^p, \quad S : \mathbb{R}^p \longrightarrow \mathbb{R}^m$$

dos transformaciones lineales y sea

$$S \circ T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

su composición, definida como

$$(S \circ T)(\vec{x}) = S(T(\vec{x})).$$

Es inmediato verificar que la función $S \circ T$ es también una transformación lineal. Por ejemplo, $S \circ T$ preserva sumas ya que dados $\vec{x}$, $\vec{y}$ vectores en $\mathbb{R}^n$, se tiene que

$$\begin{aligned}
(S \circ T)(\vec{x} + \vec{y}) &= S(T(\vec{x} + \vec{y})) \\
&= S(T(\vec{x}) + T(\vec{y})) \\
&= S(T(\vec{x})) + S(T(\vec{y})) \\
&= (S \circ T)(\vec{x}) + (S \circ T)(\vec{y}).
\end{aligned}$$

Análogamente, se demuestra que para cualquier escalar c y cualquier vector $\vec{x} \in \mathbb{R}^n$ se cumple

$$(S \circ T)(c\vec{x}) = c\,(S \circ T)(\vec{x}).$$

Sean $A \in M_{m\times p}$ y $B \in M_{p\times n}$ las matrices estándar asociadas a S y T, respectivamente. Las siguientes igualdades son inmediatas para todo $\vec{x} \in \mathbb{R}^n$:

$$(S \circ T)(\vec{x}) = S(T(\vec{x})) = S(B\vec{x}) = A\,(B\vec{x}).$$

De lo anterior, es natural definir el producto de estas matrices de tal forma que

$$(AB)\vec{x} = A\,(B\vec{x}),$$

de manera que la matriz estándar asociada a la composición $S \circ T$ sea, precisamente, la matriz producto AB. Se tiene así la siguiente definición.

Definición 6.20 *Sean $A \in M_{m\times p}$, $B \in M_{p\times n}$ matrices y sean*

$$S : \mathbb{R}^p \longrightarrow \mathbb{R}^m,\ T : \mathbb{R}^n \longrightarrow \mathbb{R}^p$$

las respectivas transformaciones asociadas. Esto es,

$$S(\vec{y}) = A\vec{y},\ T(\vec{x}) = B\vec{x},$$

para cualesquiera $\vec{x} \in \mathbb{R}^n$ y $\vec{y} \in \mathbb{R}^p$. Entonces se define el ***producto*** *AB como la matriz estándar de $m \times n$ asociada a la composición*

$$S \circ T : \mathbb{R}^n \longrightarrow \mathbb{R}^m.$$

Es decir, si $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ es la base estándar de $\mathbb{R}^n$, por construcción

$$AB = \left[\ S(T(\vec{e}_1)) \quad S(T(\vec{e}_2)) \quad \cdots \quad S(T(\vec{e}_n))\ \right]. \qquad (\bigstar)$$

La definición anterior nos dice que es necesario evaluar $S(T(\vec{e}_j))$ para toda $j = 1, 2, \ldots, n$. Esto puede ser laborioso ya que hay que encontrar las transformaciones S, T y $S \circ T$. Lo ideal sería poder calcular el producto AB a partir de las matrices A y B. Para este efecto, supongamos que

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mp} \end{bmatrix},\ B = \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{p1} & b_{p2} & \ldots & b_{pn} \end{bmatrix}.$$

De acuerdo a la definición dada por $(\bigstar)$, la $j-$*ésima* columna del producto AB está dada por

$$(AB)^j = (S \circ T)(\vec{e}_j) = S(T(\vec{e}_j)) = S(B\vec{e}_j) = S\left(B^j\right) = AB^j,$$

en donde la penúltima igualdad se cumple en virtud de que $B^j = B\vec{e}_j$. Por lo tanto, observemos que

$$AB = \left[AB^1\ AB^2\ \cdots\ AB^n\right].$$

Hemos demostrado así la siguiente proposición, misma que proporciona una definición alternativa del producto AB, sin apelar a las transformaciones lineales asociadas.

Proposición 6.21 *Si $A \in M_{m\times p}$ y $B \in M_{p\times n}$, entonces el producto $AB \in M_{m\times p}$ satisface que, para cada $j = 1, \ldots, n$, la $j-$ésima columna de AB es una combinación lineal de las columnas de A. Concretamente,*

$$\begin{aligned}(AB)^j &= b_{1j}A^1 + b_{2j}A^2 + \cdots + b_{pj}A^p \\ &= AB^j.\end{aligned}$$

Podemos refinar aún más el procedimiento anterior para obtener explícitamente cada entrada de la columna $(AB)^j$. En efecto, como

$$(AB)^j = AB^j = b_{1j}A^1 + b_{2j}A^2 + \cdots + b_{pj}A^p$$

y las columnas de A están dadas por

$$A^k = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix},$$

utilizando las dos igualdades que preceden al Ejemplo 6.19, tenemos que

$$\begin{aligned}
(AB)^j &= b_{1j} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + b_{2j} \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + b_{pj} \begin{bmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{mp} \end{bmatrix} \\
&= \begin{bmatrix} b_{1j}a_{11} + b_{2j}a_{12} + \cdots + b_{pj}a_{1p} \\ b_{1j}a_{21} + b_{2j}a_{22} + \cdots + b_{pj}a_{2p} \\ \vdots \\ b_{1j}a_{m1} + b_{2j}a_{m2} + \cdots + b_{pj}a_{mp} \end{bmatrix} \\
&= \begin{bmatrix} A_1B^j \\ A_2B^j \\ \vdots \\ A_mB^j \end{bmatrix} = \begin{bmatrix} A_1 \cdot B^j \\ A_2 \cdot B^j \\ \vdots \\ A_m \cdot B^j \end{bmatrix}.
\end{aligned}$$

Observamos que la entrada en el $i-$*ésimo* renglón de la columna $(AB)^j$ es simplemente

$$(AB)_{ij} = A_iB^j = A_i \cdot B^j.$$

Por lo tanto, disponemos de una fórmula sencilla para calcular cada entrada de la matriz producto de dos matrices y la siguiente proposición es inmediata.

Proposición 6.22 *Si $A \in M_{m\times p}$ y $B \in M_{p\times n}$, entonces*

$$(AB)_{ij} = A_iB^j = A_i \cdot B^j,$$

para toda $i \leq m$ y para toda $j \leq n$.

Observación 6.23 *Por construcción, el producto AB que acabamos de definir generaliza el producto $A\vec{x}$ de una matriz por un vector* [4]*.*

[4] Simplemente se piensa al vector $\vec{x}$ como una matriz con una sóla columna.

Observación 6.24 *El producto matricial AB está definido syss el número de columnas de A es igual al número de renglones de B. En tal caso, el número de renglones de AB es igual al número de renglones de A, mientras que el número de columnas de AB es igual al número de columnas de B.*

La Proposición 6.21 proporciona una descripción de las columnas de la matriz producto. Una pregunta natural es si puede hacerse algo similar con los renglones de la matriz producto. La siguiente proposición responde a esta pregunta en sentido afirmativo.

Proposición 6.25 *Si $A \in M_{m\times n}$ y $B \in M_{n\times p}$, entonces el producto $AB \in M_{m\times p}$ satisface que, para cada $i = 1, \ldots, m$, el $i-$ésimo renglón de AB está dado como*

$$(AB)_i = A_i B.$$

Demostración. Sea $i = 1, \ldots m$. La Proposición 6.22 nos dice que la entrada en la $j-$*ésima* columna del renglón $(AB)_i$ está dada por

$$A_i B^j.$$

Entonces,

$$(AB)_i = \begin{bmatrix} A_i B^1 \; A_i B^2 \cdots A_i B^p \end{bmatrix} = A_i B$$

y concluimos la demostración. ■

Ejemplo 6.26 *Consideremos las matrices*

$$A = \begin{bmatrix} 2 & 1 & -1 \\ 3 & 0 & -2 \end{bmatrix}, \; B = \begin{bmatrix} 2 & 0 \\ -1 & 0 \\ 4 & 1 \end{bmatrix}, \; C = \begin{bmatrix} 6 & 7 \\ -1 & 8 \end{bmatrix},$$
$$D = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}, \; E = \begin{bmatrix} 1 & -1 \end{bmatrix}.$$

y supongamos que deseamos calcular los productos AB, AC, AD, DE y CA. En el primer caso notamos que $A \in M_{2\times 3}$ y $B \in M_{3\times 2}$, por lo que $AB \in M_{2\times 2}$. Además,

$$(AB)_{11} = A_1 B^1 = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} = 4 - 1 - 4 = -1,$$

$$(AB)_{12} = A_1 B^2 = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = -1,$$

$$(AB)_{21} = A_2B^1 = \begin{bmatrix} 3 & 0 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} = 6 - 8 = -2,$$

$$(AB)_{22} = A_2B^2 = \begin{bmatrix} 3 & 0 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = -2,$$

Por lo tanto,

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

En el segundo caso, notamos que el número de columnas de A es distinto al número de renglones de C, por lo que el producto AC no está definido. En el tercer caso notamos que $A \in M_{2\times3}$ y que $D \in M_{3\times1}$, por lo que $AD \in M_{2\times1}$. Además,

$$(AD)_{11} = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} = 8 - 1 = 7,$$

$$(AD)_{21} = \begin{bmatrix} 3 & 0 & -2 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} = 12 - 2 = 10.$$

Por lo tanto,

$$AD = \begin{bmatrix} 7 \\ 10 \end{bmatrix}.$$

Alternativamente, como D es una matriz de 3×1 y puede pensarse como un vector en $\mathbb{R}^3$, AD también puede calcularse como

$$AD = 4 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 7 \\ 10 \end{bmatrix}.$$

Finalmente, el producto DE está definido pues $D \in M_{3\times1}$ y $E \in M_{1\times2}$. Obsérvese que los renglones de D y las columnas de E son simplemente números reales. Las entradas del producto $DE \in M_{3\times2}$ se calculan como sigue:

$$\begin{aligned} (DE)_{11} &= D_1E^1 = 4 \times 1 = 4, \; (DE)_{12} = D_1E^2 = 4 \times (-1) = -4, \\ (DE)_{21} &= D_2E^1 = 0 \times 1 = 0, \; (DE)_{22} = D_2E^2 = 0 \times (-1) = 0, \\ (DE)_{31} &= D_3E^1 = 1 \times 1 = 1, \; (DE)_{32} = D_3E^2 = 1 \times (-1) = -1, \end{aligned}$$

de manera que

$$DE = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} = \begin{bmatrix} 4 & -4 \\ 0 & 0 \\ 1 & -1 \end{bmatrix}.$$

Como $C \in M_{2\times 2}$ *y* $A \in M_{2\times 3}$ *tendremos que* $CA \in M_{2\times 3}$, *en efecto,*

$$\begin{aligned} CA &= \begin{bmatrix} 6 & 7 \\ -1 & 8 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 3 & 0 & -2 \end{bmatrix} \\ &= \begin{bmatrix} (6)(2)+(7)(3) & (6)(1)+(7)(0) & (6)(-1)+(7)(-2) \\ (-1)(2)+(8)(3) & (-1)(1)+(8)(0) & (-1)(-1)+(8)(-2) \end{bmatrix} \\ &= \begin{bmatrix} 33 & 6 & -20 \\ 22 & -1 & -15 \end{bmatrix}. \end{aligned}$$

Recordemos que el producto AC *no estaba definido (el número de columnas de* A *no coincide con el número de renglones de* C*); sin embargo, como acabamos de ver,* CA *sí lo está. Esto es una muestra de que el producto de matrices no es conmutativo.*

Ejemplo 6.27 *Cuando* A *y* B *son matrices cuadradas de* $n \times n$, *ambos productos,* AB *y* BA, *están definidos y son matrices de* $n \times n$. *Por ejemplo, sean*

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

entonces

$$AB = \begin{bmatrix} -2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} =$$

$$\begin{bmatrix} (-2)(3)+(-1)(-1) & (-2)(-1)+(-1)(3) \\ (1)(3)+(0)(-1) & (1)(-1)+(0)(3) \end{bmatrix} = \begin{bmatrix} -5 & -1 \\ 3 & -1 \end{bmatrix},$$

$$BA = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} -2 & -1 \\ 1 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} (3)(-2)+(-1)(1) & (3)(-1)+(-1)(0) \\ (-1)(-2)+(3)(1) & (-1)(-1)+(3)(0) \end{bmatrix} = \begin{bmatrix} -7 & -3 \\ 5 & 1 \end{bmatrix}.$$

Observemos que ambos productos AB *y* BA *existen pero* $AB \neq BA$, *de manera que el producto de matrices no es conmutativo aún cuando las matrices son cuadradas y del mismo tamaño.*

Ejemplo 6.28 *Consideremos las matrices*

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix},\ B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

y supongamos que deseamos calcular la tercera columna del producto AB. *Una forma eficiente de hacerlo (en lugar de calcular toda la matriz* AB*) es simplemente utilizando la Proposición 6.21. De esta forma,*

$$(AB)^3 = AB^3 = 3 \begin{bmatrix} 2 \\ 5 \end{bmatrix} + 6 \begin{bmatrix} 3 \\ 6 \end{bmatrix} + 9 \begin{bmatrix} 4 \\ 7 \end{bmatrix}.$$

Ejemplo 6.29 *Sean* $S : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *y* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *las transformaciones lineales dadas por*

$$S\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x - y \\ x + 2y \end{bmatrix}, T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + y - z \\ 2x + z \end{bmatrix}.$$

Las matrices estándar asociadas son, respectivamente,

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

Podemos calcular la composición

$$S \circ T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$$

de dos formas: la primera es directamente, sin utilizar las matrices. En este caso se tiene

$$\begin{aligned} (S \circ T)\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) &= S\left(T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right)\right) = S\left(\begin{bmatrix} x + y - z \\ 2x + z \end{bmatrix}\right) \\ &= \begin{bmatrix} 2(x + y - z) - (2x + z) \\ (x + y - z) + 2(2x + z) \end{bmatrix} = \begin{bmatrix} 2y - 3z \\ 5x + y + z \end{bmatrix}. \end{aligned}$$

La segunda forma de encontrar $S \circ T$ *es utilizando el producto* AB *como sigue.*

$$\begin{aligned} (S \circ T)\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) &= \underbrace{\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \end{bmatrix}}_{B}\begin{bmatrix} x \\ y \\ z \end{bmatrix} \\ &= \begin{bmatrix} 0 & 2 & -3 \\ 5 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2y - 3z \\ 5x + y + z \end{bmatrix}. \end{aligned}$$

Ejemplo 6.30 *Sean* $T_1 : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *y* $T_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$ *las transformaciones dadas por*

$$T_1\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x - z \\ y + z \end{bmatrix}, \quad T_2\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ 2x \\ 0 \end{bmatrix}.$$

Entonces $T_2 \circ T_1 : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ *está dada por*

$$(T_2 \circ T_1)\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = T_2\left(T_1\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right)\right) = T_2\left(\begin{bmatrix} x - z \\ y + z \end{bmatrix}\right) = \begin{bmatrix} x - z \\ y + z \\ 2(x - z) \\ 0 \end{bmatrix}.$$

Por lo tanto, la matriz estándar que representa a $T_2 \circ T_1$ es la matriz

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

Observamos que como

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

son las matrices que representan a T_2 y a T_1, respectivamente, se cumple

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

El siguiente resultado enlista algunas propiedades del producto de matrices.

Proposición 6.31 *Las siguientes afirmaciones son verdaderas*

1. *El producto de matrices no es conmutativo.*

2. *El producto de matrices es asociativo. Esto es, si $D \in M_{m\times n}$, $E \in M_{n\times p}$ y $F \in M_{p\times q}$, entonces $D(EF) = (DE)F$.*

3. *Si $A, B \in M_{m\times n}$ y $C \in M_{n\times p}$, entonces $(A+B)C = AC + BC$.*

4. *Si $A \in M_{m\times n}$ y $B, C \in M_{n\times p}$, entonces $A(B+C) = AB + AC$.*

5. *Si $A \in M_{m\times n}$ e I_m denota a la matriz identidad de $m \times m$, entonces $I_m A = A$.*

6. *Si $A \in M_{m\times n}$ e I_n denota a la matriz identidad de $n \times n$, entonces $AI_n = A$.*

7. *Si $A \in M_{m\times n}$ y $B \in M_{n\times p}$, entonces $(AB)^T = B^T A^T$.*

8. *Si $A \in M_{m\times n}$ y $B \in M_{n\times p}$ y $\alpha, \beta \in \mathbb{R}$, entonces $(\alpha A)(\beta B) = (\alpha\beta)AB$.*

Demostración. La demostración de (1) se sigue del Ejemplo 6.27. Los incisos (2), (5) y (6) son inmediatos utilizando propiedades básicas de la composición de funciones junto con el hecho de que el producto de las matrices corresponde a la composición de las transformaciones lineales asociadas. Por ejemplo, para demostrar (2), supongamos que R, S y T son las transformaciones lineales asociadas a las matrices D, E y F, respectivamente. Se tiene así la siguiente composición entre ellas

$$\mathbb{R}^q \xrightarrow{T} \mathbb{R}^p \xrightarrow{S} \mathbb{R}^n \xrightarrow{R} \mathbb{R}^m,$$

de manera que como la composición de funciones es asociativa, se tiene que

$$R \circ (S \circ T) = (R \circ S) \circ T$$

y por lo tanto, la matriz estándar que representa a $R \circ (S \circ T)$ es igual a la matriz estándar que representa a $(R \circ S) \circ T$. Esto es,

$$D(EF) = (DE)F.$$

Los incisos restantes pueden verificarse utilizando que dos matrices de $n \times m$ son iguales syss coinciden entrada a entrada. En el caso de (7), por ejemplo, utilizamos las propiedades de matrices transpuestas, la Proposición 6.22 y la conmutatividad del producto punto para obtener:

$$\begin{aligned} ((AB)^T)_{ij} &= (AB)_{ji} = A_j \cdot B^i = B^i \cdot A_j \\ &= \left(B^T\right)_i \cdot (A^T)^j = (B^T A^T)_{ij}, \end{aligned}$$

por lo que las matrices $(AB)^T$ y $B^T A^T$ son iguales. ■

Ejemplo 6.32 *Sean*

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

entonces

$$AB = \begin{bmatrix} 4 & -3 \\ 10 & -4 \\ 16 & -5 \end{bmatrix}.$$

Se tiene así que

$$\begin{aligned} (AB)^T &= \begin{bmatrix} 4 & 10 & 16 \\ -3 & -4 & -5 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \end{bmatrix} \\ &= B^T A^T. \end{aligned}$$

Otras matrices de interés

Matrices diagonales

Consideremos el conjunto de matrices cuadradas $M_{n\times n}$. Un subconjunto que se utiliza recurrentemente es el conjunto de las matrices diagonales definido como sigue.

Definición 6.33 *Una matriz $D \in M_{n\times n}$ es* ***diagonal*** *syss $D_{ij} = 0$ para toda $i \neq j$. Es decir, D es diagonal syss todos las entradas fuera de su diagonal son nulas.*

Ejemplo 6.34 *Las matrices $I, \Theta \in M_{n\times n}$ (identidad y cero, respectivamente) son diagonales.*

Ejemplo 6.35 *Las matrices*

$$\begin{bmatrix} 666 & 0 \\ 0 & \pi \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

son diagonales.

Observación 6.36 *Las matrices diagonales son simétricas.*

Observación 6.37 *Si $C, D \in M_{n\times n}$ son diagonales, entonces CD es diagonal y $CD = DC$, es decir, el producto entre matrices diagonales es conmutativo. Esto es simple de verificar ya que*

$$(CD)_{ij} = (C_{ii}\vec{e}_i) \cdot (D_{jj}\vec{e}_j) = (C_{ii}D_{jj})(\vec{e}_i \cdot \vec{e}_j) = \begin{cases} C_{ii}D_{jj} \text{ si } i = j, \\ 0 \text{ si } i \neq j. \end{cases}$$

Observación 6.38 *Evidentemente, si $C, D \in M_{n\times n}$ son diagonales, $C + D$ es diagonal. Adicionalmente, si C es diagonal y todas sus entradas en la diagonal son no nulas, la matriz D cuyas entradas diagonales son $D_{ii} = \frac{1}{C_{ii}}$ será un inverso multiplicativo de C.*

Ejemplo 6.39 *Si*

$$C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7 \end{bmatrix},$$

entonces la matriz

$$D = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{7} \end{bmatrix}$$

es tal que

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{7} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7 \end{bmatrix}.$$

La matriz D es, de facto, inversa multiplicativa de C. Este ejemplo nos da la pauta para construir inversos multiplicativos de una matriz diagonal para la cual $C_{ii} \neq 0$ para toda $i = 1, \ldots, n$. Esta última condición es, por supuesto, equivalente a que sus columnas formen un conjunto LI.

Las observaciones anteriores implican que el conjunto de matrices diagonales en $M_{n\times n}$ es cerrado bajo la suma, el producto (que en este caso es conmutativo), contiene neutros aditivos y multiplicativos e inversos aditivos. Además, si una matriz diagonal es tal que sus columnas forman un conjunto LI, ésta también tiene un inverso multiplicativo. Ahora bien, dada una matriz $A \in M_{n\times n}$, no necesariamente diagonal, ¿cuándo tendrá ésta un inverso multiplicativo? En el siguiente capítulo abordaremos precisamente esta cuestión.

Matrices elementales

Dentro de las matrices de $n \times n$ existe una clase destacada cuyos elementos se conocen como matrices elementales. Estas matrices son útiles para representar las operaciones elementales entre renglones en términos de productos de matrices. Su definición es sumamente simple y es la siguiente.

Definición 6.40 *Una matriz $E \in M_{n\times n}$ es una **matriz elemental** si E puede obtenerse a partir de la matriz identidad I realizando una (y sólo una) operación elemental entre renglones (de los tres tipos descritos en la Definición 2.16). Si esta operación elemental es de tipo i, (en donde $i = 1, 2, 3$) diremos que E es una **matriz elemental de tipo** i.*

Ejemplo 6.41 *La matriz*

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

es una matriz elemental de tipo 1 puesto que se obtiene a partir de la identidad intercambiando los renglones.

Ejemplo 6.42 *La matriz*

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

es una matriz elemental de tipo 2 ya que se obtiene de la identidad multiplicando el segundo renglón por π.

Ejemplo 6.43 *La matriz*

$$\begin{bmatrix} 1 & 0 & 666 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

es una matriz elemental de tipo 3 pues se obtiene de la identidad sumando al primer renglón 666 *veces el tercero.*

Ejemplo 6.44 *La matriz*

$$\begin{bmatrix} 0 & -7 \\ 1 & 0 \end{bmatrix}$$

no es una matriz elemental puesto que se requieren dos operaciones elementales, sobre los renglones de la matriz identidad, para obtenerla.

Ejemplo 6.45 *Consideremos la matriz*

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

y a las matrices elementales

$$E(1) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; E(2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 666 \end{bmatrix},$$
$$E(3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 666 \\ 0 & 0 & 1 \end{bmatrix},$$

en donde $E(1)$ *se obtiene de la identidad intercambiando los renglones 1 y 2,* $E(2)$ *se obtiene de la identidad multiplicando el tercer renglón por* 666 *y* $E(3)$ *se obtiene de la identidad sumando* 666 *veces el tercer renglón al segundo. A continuación se evaluan los productos* $E(1)A, E(2)A$

y $E(3)A$.

$$
\begin{aligned}
E(1)A &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 1 & 1 \end{bmatrix}, \\
|E(2)A &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 666 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & -1 \\ 1 & 2 \\ 666 & 666 \end{bmatrix}, \\
E(3)A &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 666 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}. \\
&= \begin{bmatrix} 0 & -1 \\ 667 & 668 \\ 1 & 1 \end{bmatrix}
\end{aligned}
$$

El Ejemplo 6.45 muestra que multiplicar, por la izquierda, la matriz A por la matriz elemental $E(i)$, $i = 1, 2, 3$, es lo mismo que realizar sobre A la misma operación elemental que se utilizó para obtener a $E(i)$ a partir de I. Esta propiedad de matrices elementales puede enunciarse como sigue.

Proposición 6.46 *Sean* $A \in M_{m\times n}$ *y* $E \in M_{m\times m}$ *tales que* E *es una matriz elemental. Entonces* EA *es una matriz que se obtiene de* A *realizando la misma operación elemental sobre los renglones de* A *que se utilizó para obtener* E *a partir de* I.

Demostración. Observemos que la matriz identidad de $m \times m$ tiene como renglones a los vectores[5] $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m$ de la base estándar de $\mathbb{R}^m$. Adicionalmente, si E es una matriz elemental de $m \times m$, se tiene que, por la Proposición 6.25,

$$E_i A = (EA)_i \,. \qquad (\spadesuit)$$

Para el primer tipo de operación elemental, supongamos que $E(1)$ se obtiene de I intercambiando los renglones $\vec{e}_k$ y $\vec{e}_l$. Entonces, utilizando

[5] Aquí estamos pensando a los vectores de la base estándar como vectores renglón y no como vectores columna.

(♠), tenemos que los renglones del producto $E(1)A$ están dados por

$$(E(1)A)_i = \begin{cases} \vec{e}_i A = A_i \text{ si } i \neq k, l, \\ \vec{e}_l A = A_l \text{ si } i = k, \\ \vec{e}_k A = A_k \text{ si } i = l. \end{cases}$$

Así, la matriz $E(1)A$ se obtiene de A intercambiando los renglones k y l. Para el segundo tipo de operación elemental, sea $E(2)$ la matriz elemental que se obtiene de I multiplicando el renglón k por el escalar $c \neq 0$. Como antes, los renglones del producto $E(2)A$ son,

$$(E(2)A)_i = \begin{cases} \vec{e}_i A = A_i \text{ si } i \neq k, \\ c\vec{e}_k A = cA_k \text{ si } i = k, \end{cases}$$

y $E(2)A$ se obtiene de A multiplicando el $k-\textit{ésimo}$ renglón por el escalar c. Finalmente, para el tercer tipo de operación elemental, sea $E(3)$ la matriz elemental que se obtiene de I sumando c veces el renglón l al renglón k. Los renglones del producto $E(3)A$ están dados por

$$(E(3)A)_i = \begin{cases} \vec{e}_i A = A_i \text{ si } i \neq k, \\ (c\vec{e}_l + \vec{e}_k)\, A = cA_l + A_k \text{ si } i = k, \end{cases}$$

es decir, $E(3)A$ se obtiene de A sumando c veces el renglón l al renglón k. ■

Corolario 6.47 *Sea $A \in M_{m\times n}$ y $\tilde{A}$ su FER, entonces existen matrices elementales $E(1), E(2), \ldots, E(k) \in M_{m\times m}$ tales que*

$$\tilde{A} = E(k)E(k-1)\ldots E(1)A.$$

Ejemplo 6.48 *La FER de la matriz*

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

puede obtenerse haciendo las siguientes operaciones elementales:

$$\begin{bmatrix} 1 & -1 & 2 \\ 1 & -1 & 3 \\ 2 & -1 & 1 \end{bmatrix} \underset{\sim}{-R_1+R_2} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 2 & -1 & 1 \end{bmatrix} \underset{\sim}{-2R_1+R_3} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & -3 \end{bmatrix} \underset{\sim}{3R_2+R_3}$$
$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \underset{\sim}{-2R_2+R_1} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \underset{\sim}{R_3+R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \underset{\sim}{R_2 \leftrightarrow R_3}$$
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Alternativamente, cada una de estas operaciones elementales entre renglones puede sustituirse por la multiplicación por la matriz elemental correspondiente. Esto se muestra a continuación.

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{-R_1+R_2} \begin{bmatrix} 1 & -1 & 2 \\ 1 & -1 & 3 \\ 2 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 2 & -1 & 1 \end{bmatrix},$$

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}}_{-2R_1+R_3} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 2 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & -3 \end{bmatrix},$$

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}}_{3R_2+R_3} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & -3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

$$\underbrace{\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{-2R_2+R_1} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

$$\underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{R_3+R_1} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}_{R_2\leftrightarrow R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Concluimos esta sección analizando las transpuestas de matrices elementales. Supongamos que $E(1), E(2)$ y $E(3)$ son tales que

$$\begin{aligned} I &\overset{R_k\leftrightarrow R_l}{\sim} E(1), \\ I &\overset{aR_k}{\sim} E(2), \\ I &\overset{bR_k+R_l}{\sim} E(3), \end{aligned}$$

en donde $a, b \in \mathbb{R}$, $a \neq 0$. Es sencillo notar que

$$\begin{aligned} E(1)^T &= E(1), \\ E(2)^T &= E(2). \end{aligned}$$

En el caso de $E(3)$, esta matriz coincide con la identidad en todas sus entradas con excepción de $E(3)_{lk} = b$. Así, se tiene que su transpuesta es igual a la identidad excepto por la entrada kl que es igual a b. Podemos expresar esto como

$$I \overset{bR_l+R_k}{\sim} E(3)^T,$$

es decir, $E(3)^T$ es una matriz elemental del mismo tipo que $E(3)$.

Ejemplo 6.49 *Las transpuestas de las matrices elementales*

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

están dadas por,

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$
$$\begin{bmatrix} 1 & 5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Unicidad de la FER(opcional)

En la demostración del Teorema 2.28 vimos que si A es una matriz de $m \times n$, entonces existe una matriz en forma escalonada reducida que se obtiene a partir de A aplicando operaciones elementales. Tal matriz es conocida como "la FER de A" y ahora justificaremos que es apropiado utilizar "la" (en lugar del indeterminado "una") en este nombre[6].

Proposición 6.50 *Sea $A \in M_{m\times n}$, entonces su FER es única.*

Demostración. Procederemos por inducción sobre n. Si $n = 1$, entonces la FER de A es el vector $\vec{0}$ en $\mathbb{R}^m$ (en caso de que A sea igual a dicho vector), o bien el vector $\vec{e}_1$ (si A tiene al menos una componente no nula). Supongamos ahora que la unicidad es cierta para cualquier

[6] El ingenioso argumento para demostrar la unicidad de la FER aparece en: Yuster, T., "The Reduced Row Echelon Form of a Matrix is Unique: A Simple Proof", *Mathematics Magazine*, Vol. 57, No. 2, pp. 93-94, March 1984.

matriz con $n-1$ columnas $(n \geq 2)$ y mostremos que lo mismo sucede para n. Sea pues, $A \in M_{m\times n}$ y supongamos - en búsqueda de una contradicción - que U y V son ambas FER de A con $U \neq V$. Sean $\hat{A}, \hat{U}$ y $\hat{V}$ las matrices de $m \times (n-1)$ que se obtienen a partir de A, U y V eliminando su respectiva $n-\acute{e}sima$ columna. Así, $\hat{U}$ y $\hat{V}$ son ambas FER de $\hat{A}$ y, gracias a nuestra hipótesis inductiva, podemos asegurar que $\hat{U} = \hat{V}$. Como U y V pueden diferir únicamente en su última columna, debe existir algún i, $1 \leq i \leq m$, tal que

$$u_{in} - v_{in} \neq 0. \tag{$\spadesuit$}$$

Sea

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

una solución cualquiera del sistema homogéneo $A\vec{x} = \vec{0}$. Entonces,

$$U\vec{x} = V\vec{x} = \vec{0}$$

y en consecuencia,

$$(U-V)\vec{x} = \vec{0}.$$

Notamos ahora que $U - V$ es de la forma

$$U - V = \begin{bmatrix} 0 & \cdots & 0 & u_{1n} - v_{1n} \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & u_{mn} - v_{mn} \end{bmatrix},$$

de manera que

$$\begin{aligned} (U-V)\vec{x} &= \begin{bmatrix} 0 & \cdots & 0 & u_{1n} - v_{1n} \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & u_{mn} - v_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} (u_{1n} - v_{1n})x_n \\ \vdots \\ (u_{mn} - v_{mn})x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}. \end{aligned}$$

Por ($\spadesuit$) tenemos que $u_{in} - v_{in} \neq 0$ de manera que x_n está forzada a ser cero. En particular, x_n es una variable restringida y las columnas U^n y V^n contienen unos principales. Sea k el menor entero entre 1 y m tal que $\hat{U}_k$ es nulo. Como las primeras $n-1$ columnas de U y V son idénticas, k también es el entero más pequeño para el cual $\hat{V}_k$ es nulo. Por definición de "estar en forma escalonada reducida", $u_{kn} = 1 = v_{kn}$ y $u_{jn} = 0 = v_{jn}$ si $j \neq k$. Dado que U y V también coinciden en su última columna, llegamos al absurdo de que $U = V$. ■

Ejercicios

Ejercicio 6.1 *Calcular (en caso de que sea posible) el resultado de las operaciones indicadas entre las matrices que se presentan a continuación.*

$$A = \begin{bmatrix} 1 & 3 & -4 \\ 0 & 3 & 8 \\ 0 & 0 & 1 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 2 & -4 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix}, \ C = \begin{bmatrix} 0 & 1 \\ 5 & 0 \\ 3 & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} 7 & 1 & 0 \\ 1 & -4 & 9 \end{bmatrix}, \ E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 5 & 8 \\ -3 & 1 \end{bmatrix},$$

$$F = \begin{bmatrix} 6 & 6 \\ -5 & 8 \end{bmatrix}, \ G = \begin{bmatrix} 8 & 6 \\ -5 & 10 \end{bmatrix}, \ H = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},$$

$$L = \begin{bmatrix} -2 & 7 \end{bmatrix}, \ J = \begin{bmatrix} -10 & 3 \\ 7 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \ K = \begin{bmatrix} \frac{1}{11} & -\frac{3}{55} \\ \frac{1}{22} & \frac{4}{55} \end{bmatrix}.$$

a. DC, $DC(F-G)$, $DC-4I$, *donde* I *es la matriz identidad de* 2×2.

b. CD, $CD+A$, C^TD.

c. BE, $BE+2F^T$.

d. DE, $C+E$.

e. $(DA)^T$, AC.

f. HL, LH, L^TH^T.

g. BJ, JB.

h. GK, KG.

Ejercicio 6.2 *Sean*

$$S:\mathbb{R}^2\to\mathbb{R}^2,\ T:\mathbb{R}^2\to\mathbb{R}^4,\ U:\mathbb{R}^4\to\mathbb{R}^3$$

las transformaciones lineales dadas por

$$S\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)=\begin{bmatrix}-4x_1+x_2\\x_1-x_2\end{bmatrix},\ T\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)=\begin{bmatrix}x_1+x_2\\x_2\\3x_2\\0\end{bmatrix},$$
$$U\left(\begin{bmatrix}w_1\\w_2\\w_3\\w_4\end{bmatrix}\right)=\begin{bmatrix}w_3\\w_2\\w_1\end{bmatrix}.$$

Utilizando directamente la definición de composición, calcular

$$(S\circ S)(\vec{x}),(T\circ S)(\vec{x}),(U\circ T)(\vec{x}).$$

Ejercicio 6.3 *Sean*

$$T_1:\mathbb{R}^2\to\mathbb{R}^2,\ T_2:\mathbb{R}^2\to\mathbb{R}^2,\ T_3:\mathbb{R}^3\to\mathbb{R}^2$$

las transformaciones lineales dadas por

$$T_1\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)=\begin{bmatrix}6x_1-7x_2\\7x_2\end{bmatrix},\ T_2\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)=\begin{bmatrix}-x_1-\pi x_2\\x_1+9x_2\end{bmatrix},$$
$$T_3\left(\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}\right)=\begin{bmatrix}y_2\\y_1\end{bmatrix}.$$

a. *Encontrar dos matrices A y B, distintas de la identidad, tales que el producto AB es igual a la matriz estándar que representa a*

$$T_1\circ T_2:\mathbb{R}^2\to\mathbb{R}^2.$$

b. *Encontrar dos matrices C y D, distintas de la identidad, tales que el producto CD es igual a la matriz estándar que representa a*

$$T_2\circ T_1:\mathbb{R}^2\to\mathbb{R}^2.$$

c. *Encontrar dos matrices E y F, distintas de la identidad, tales que el producto EF es igual a la matriz estándar que representa a*

$$T_1\circ T_1:\mathbb{R}^2\to\mathbb{R}^2.$$

d. *Encontrar dos matrices G y H, distintas de la identidad, tales que el producto GH es igual a la matriz estándar que representa a*

$$T_2\circ T_3:\mathbb{R}^3\to\mathbb{R}^2.$$

Ejercicio 6.4 *Sean T y S transformaciones de $\mathbb{R}^n$ en $\mathbb{R}^m$ y sea c un número real. Demostrar que*

a. $T + S : \mathbb{R}^n \to \mathbb{R}^m$ *es una transformación lineal.*

b. $cT : \mathbb{R}^n \to \mathbb{R}^m$ *es una transformación lineal.*

Ejercicio 6.5 *Decimos que una matriz cuadrada A es* ***antisimétrica*** *syss $A^T = -A$. Pruebe que si A es una matriz cuadrada, entonces $A - A^T$ es una matriz antisimétrica. Utilizar lo anterior y el resultado del Ejemplo 6.12 para probar lo siguiente: si A es una matriz cuadrada, entonces existen una matriz simétrica B y una matriz antisimétrica C tales que $A = B + C$.*

Ejercicio 6.6 *Demostrar las propiedades del producto punto dadas en la Proposición 6.17.*

Ejercicio 6.7 *Probar que si $T : \mathbb{R}^n \to \mathbb{R}^p$ y $S : \mathbb{R}^p \to \mathbb{R}^m$ son transformaciones lineales, entonces la composición $S \circ T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal.*

Ejercicio 6.8 *Probar los siguientes enunciados utilizando el hecho de que dos matrices son iguales syss coinciden entrada a entrada.*

a. Si $A, B \in M_{m\times n}$ y $C \in M_{n\times p}$, entonces $(A + B)C = AC + BC$.

b. Si $A \in M_{m\times n}$ y $B, C \in M_{n\times p}$, entonces $A(B + C) = AB + AC$.

c. Si $A \in M_{m\times n}$ e I_m denota a la matriz identidad de $m \times m$, entonces $I_m A = A$.

d. Si $A \in M_{m\times n}$ e I_n denota a la matriz identidad de $n\times n$, entonces $AI_n = A$.

e. Si $A \in M_{m\times n}$, $B \in M_{n\times p}$ y α y β son números reales, entonces $(\alpha A)(\beta B) = (\alpha\beta)(AB)$.

Ejercicio 6.9 *Utilizando que el producto de matrices corresponde a la composición de las transformaciones lineales asociadas demostrar que si $A \in M_{m\times n}$, entonces $AI_n = A$.*

Ejercicio 6.10 *Demostrar (en general) o refutar (con un contraejemplo) los siguientes enunciados (en cada caso se debe indicar claramente "verdadero" o "falso").*

a. Sean $A \in M_{m\times n}$, $B \in M_{n\times p}$. Si el i-ésimo renglón de A es un renglón de ceros , entonces el i-ésimo renglón del producto AB es un renglón de ceros *b.* *Sean A y B matrices de 2×2. Si el primer renglón de B es un renglón de ceros , entonces el primer renglón del producto AB es un renglón de ceros.*

c. *Sean A y B matrices de 2×2. Si todas las componentes de A y todas las componentes de B son distintas de cero, entonces el producto AB tiene al menos una componente que es distinta de cero.*

d. *Si A y B son matrices cuadradas de $n \times n$, entonces $(AB)^T = A^T B^T$.*

e. *Si A es una matriz cuadrada, entonces $A^T A$ es simétrica.*

f. *Si A es una matriz cuadrada, entonces $(A-I)(A+I) = A^{(2)} - I$, donde $A^{(2)}$ denota el producto AA.*

g. *Si A y B son matrices de $n \times n$, entonces $(A-B)(A+B) = A^{(2)} + AB - BA - B^{(2)}$.*

h. *Si A y B son matrices de $n \times n$, entonces $(A-B)(A+B) = A^{(2)} - B^{(2)}$.*

Nota. La notación $A^{(2)}$ puede parecer un tanto peculiar, pero es simplemente para distinguir la segunda potencia de A de la segunda columna de A (en este texto hemos reservado el símbolo A^2 para referirnos a dicha columna).

Ejercicio 6.11 *Determinar cuáles de las siguientes matrices son matrices elementales:*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \ B = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \ C = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \pi & 1 \end{bmatrix}, \ E = \begin{bmatrix} 666 & 0 \\ 0 & 1 \end{bmatrix}, \ F = \begin{bmatrix} 1 & 7 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Ejercicio 6.12 *Sean*

$$A = \begin{bmatrix} 1 & 1 & 8 \\ 0 & 2 & -2 \\ 4 & 4 & 0 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 3 & 3 & 6 & 9 \end{bmatrix}$$

las matrices del Ejercicio 2.3 y sean $\tilde{A}$ y $\tilde{B}$, respectivamente, sus formas escalonadas. Encontrar matrices elementales $E(i_1), E(i_2), \ldots, E(i_k)$ y $E(j_1), E(j_2), \ldots, E(j_l)$, de manera que

$$\tilde{A} = E(i_1)E(i_2)\cdots E(i_k)A,$$
$$\tilde{B} = E(j_1)E(j_2)\cdots E(j_l)B.$$

Ejercicio 6.13 *Encontrar la matriz estándar de las transformaciones lineales de* $\mathbb{R}^2$ *en* $\mathbb{R}^2$ *que representen las siguientes operaciones geométricas e ilustra la imagen del cuadrado con vértices* $(0,0),(1,0),(1,1)$ *y* $(0,1)$ *bajo las mismas.*

a. Primero una rotación por un ángulo $\frac{\pi}{4}$ *y después una reflexión sobre la recta* $y=x$.

b. Una contracción del 50 % a lo largo del eje horizontal, seguida por una rotación por un ángulo de $\frac{\pi}{4}$.

Capítulo 7

Inversas

Introducción

Como vimos en el capítulo anterior, las operaciones entre matrices corresponden a operaciones análogas entre sus transformaciones lineales asociadas. El producto de matrices, por ejemplo, tiene una traducción natural en términos de composición de transformaciones lineales.

Recordemos que una función $f : \mathcal{A} \longrightarrow \mathcal{B}$ es biyectiva syss existe su función inversa $f^{-1} : \mathcal{B} \longrightarrow \mathcal{A}$. Si ese es el caso, tenemos que para todo $x \in \mathcal{A}$ y $y \in \mathcal{B}$ se cumple que

$$\left(f^{-1} \circ f\right)(x) = x \quad \text{y} \quad \left(f \circ f^{-1}\right)(y) = y,$$

de manera que

$$f^{-1} \circ f : \mathcal{A} \longrightarrow \mathcal{A} \;\; \text{y} \;\; f \circ f^{-1} : \mathcal{B} \longrightarrow \mathcal{B}$$

son, respectivamente, las funciones identidad

$$I_{\mathcal{A}} : \mathcal{A} \longrightarrow \mathcal{A} \;\; \text{e} \;\; I_{\mathcal{B}} : \mathcal{B} \longrightarrow \mathcal{B}.$$

Una pregunta natural es si las transformaciones lineales, o equivalentemente sus matrices asociadas, pueden tener inversas de la misma naturaleza (transformaciones lineales o matrices, respectivamente) y, de ser así, cómo podemos encontrarlas. A continuación procedemos a responder esta pregunta.

Matrices inversas

Comenzamos examinanando la existencia de inversos multiplicativos de matrices para posteriormente relacionarla con la descripción de las inversas de sus transformaciones lineales correspondientes. Es importante recordar que si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal biyectiva, entonces $n = m$ (ver inciso (e) del Ejercicio 5.18) y por lo tanto, la matriz estándar que representa a T es una matriz cuadrada. En consecuencia, debe ocurrir que si una matriz es invertible, ésta satisface - entre otras cosas - el ser cuadrada.

Definición 7.1 *Sea $A \in M_{n\times n}$ una matriz cuadrada. Decimos que A es **invertible**(o **no singular**) syss existe una matriz $B \in M_{n\times n}$ tal que*

$$AB = I = BA.$$

Observación 7.2 *Cuando no exista confusión, denotaremos a la matriz identidad simplemente por I al igual que su transformación lineal asociada.*

Ejemplo 7.3 *Como se vio en el capítulo anterior, dada la matriz diagonal*

$$A = \begin{bmatrix} \pi & 0 \\ 0 & 666 \end{bmatrix},$$

se tiene que

$$B = \begin{bmatrix} \pi^{-1} & 0 \\ 0 & 666^{-1} \end{bmatrix}$$

satisface $AB = I = BA$, con lo cual A es invertible.

Ejemplo 7.4 *Consideremos la matriz*

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

Esta matriz es invertible pues

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

Ejemplo 7.5 *Consideremos la matriz*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Entonces A es invertible pues si tomamos

$$B = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix},$$

es sencillo verificar que $AB = I = BA$.

Ejemplo 7.6 *La matriz*

$$\begin{bmatrix} 1 & 0 & -1 \\ \pi & 666 & 7 \end{bmatrix}$$

no es invertible pues ni siquiera es una matriz cuadrada.

Ejemplo 7.7 *Sea*

$$A = \begin{bmatrix} 4 & 9 \\ 0 & 0 \end{bmatrix}.$$

A no es invertible, pues si existiese $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ tal que

$$\begin{bmatrix} 4 & 9 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

se tendría que

$$\begin{bmatrix} 4b_{11} + 9b_{21} & 4b_{12} + 9b_{22} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

lo cual es imposible.

Proposición 7.8 *Si $A \in M_{n\times n}$ es una matriz invertible, entonces existe una única matriz B tal que $AB = I = BA$. Tal matriz B es conocida como la (matriz)* ***inversa*** *de A y es denotada por A^{-1}.*

Demostración. Probamos algo más general; de hecho vemos que si A tiene una inversa por la derecha C (esto es, $AC = I$) y una inversa por la izquierda B (esto es, $BA = I$), entonces $C = B$. Esto es inmediato de la siguientes igualdades:

$$C = IC = (BA)C = B(AC) = BI = B.$$

Por lo tanto, si existe alguna matriz que es simultáneamente inversa derecha e izquierda de A, esa matriz debe ser única. En efecto, si hubiese dos matrices con esta propiedad, en particular una de ellas es inversa derecha de A y la otra es inversa izquierda de A, por lo que ambas deben ser la misma matriz. ■

Observación 7.9 *Observemos que $I^{-1} = I$ ya que $II = I$ y la matriz nula Θ de $n \times n$ no puede ser invertible pues $\Theta A = \Theta$ para cualquier $A \in M_{n\times n}$.*

A continuación veremos que si $A \in M_{n\times n}$ tiene una inversa por la izquierda, entonces tendrá una inversa por la derecha y, como acabamos de ver, ambas coinciden y son iguales a A^{-1}. En otras palabras, A es invertible syss A tiene un inverso izquierdo (más adelante veremos que podemos sustituir "izquierdo" por "derecho").

Lema 7.10 *Si $A \in M_{n\times n}$ tiene una inversa por la izquierda, es decir, existe $B \in M_{n\times n}$ tal que $BA = I$, entonces las columnas de A forman un conjunto LI (o lo que es lo mismo, la FER de A es la matriz identidad).*

Demostración. Consideremos el sistema lineal

$$A\vec{x} = \vec{0}.$$

Ahora bien, multiplicando por la izquierda ambos lados de $A\vec{x} = \vec{0}$ por la matriz B obtenemos que

$$\vec{x} = I\vec{x} = (BA)\vec{x} = B(A\vec{x}) = B\vec{0} = \vec{0}.$$

Por lo tanto, las columnas de A forman un conjunto LI. ∎

Ejemplo 7.11 *Sea*

$$A = \begin{bmatrix} 1 & -3 \\ -1 & 3 \end{bmatrix},$$

Ésta es una matriz cuadrada de 2×2; *sin embargo, como sus columnas son LD, no puede tener inversa por la izquierda y por lo tanto, no es invertible.*

Utilizando el mismo argumento que el empleado en el Lema 7.10 se obtiene un método conceptualmente sencillo para resolver sistemas de ecuaciones de $n \times n$ (para el caso en el cual la correspondiente matriz de coeficientes tiene una inversa por la izquierda).

Proposición 7.12 *Sean $A, B \in M_{n\times n}$ tales que $BA = I$ (B es inversa izquierda de A). Sea $\vec{b} \in \mathbb{R}^n$ y consideremos el sistema lineal $A\vec{x} = \vec{b}$. Entonces este sistema tiene solución única y más aún,*

$$\vec{x} = B\vec{b}.$$

Demostración. Multiplicamos por la izquierda ambos lados de la expresión $A\vec{x} = \vec{b}$ por la matriz B, obteniendo

$$\vec{x} = I\vec{x} = (BA)\,\vec{x} = B\,(A\vec{x}) = B\vec{b}$$

y concluyendo así con la demostración. ∎

Ejemplo 7.13 *Sea*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},$$

y supongamos que queremos resolver los sistemas

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad A \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$$

utilizando que (por el Ejemplo 7.5)

$$A^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

es un inverso izquierdo de A.

En el primer caso multiplicamos (por la izquierda) ambos lados de la ecuación matricial en cuestión por A^{-1}, obteniendo que

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = 1 \begin{bmatrix} -2 \\ \frac{3}{2} \end{bmatrix} + 4 \begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 2 \\ -\frac{1}{2} \end{bmatrix}.$$

Análogamente, en el segundo caso multiplicamos (por la izquierda) ambos lados de la ecuación matricial en cuestión por A^{-1} obteniendo que

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 \\ 5 \end{bmatrix} = 0 \begin{bmatrix} -2 \\ \frac{3}{2} \end{bmatrix} + 5 \begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 5 \\ -\frac{5}{2} \end{bmatrix}.$$

El ejemplo anterior muestra que el método de multiplicar por la inversa izquierda es una forma rápida de resolver sistemas de ecuaciones de $n \times n$ que comparten la misma matriz de coeficientes. Sin embargo, el problema está en que tenemos que conocer de antemano esta matriz inversa. El siguiente resultado nos dice que la existencia de un inverso izquierdo es suficiente para garantizar la invertibilidad de una matriz cuadrada. Adicionalmente, su demostración nos proporciona un método relativamente simple de encontrar, explícitamente, la matriz inversa.

Proposición 7.14 *Sea $A \in M_{n\times n}$. Si las columnas de A forman un conjunto LI, entonces A tiene una inversa por la derecha. En particular, si A tiene inversa por la izquierda, entonces tendrá inversa por la derecha, ambas coinciden y A es invertible.*

Demostración. Como las columnas de A forman un conjunto LI, dado cualquier $\vec{b} \in \mathbb{R}^n$, el sistema $A\vec{x} = \vec{b}$ tiene solución. En particular, si tomamos los vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ de la base estándar en $\mathbb{R}^n$, tendremos que existen $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ en $\mathbb{R}^n$ de tal suerte que

$$A\vec{x}_1 = \vec{e}_1, \; A\vec{x}_2 = \vec{e}_2, \ldots, A\vec{x}_n = \vec{e}_n.$$

Sea C la matriz cuyas columnas son $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$; entonces, utilizando la Proposición 6.21 tenemos que,

$$AC = A\,[\vec{x}_1 \; \vec{x}_2 \; \cdots \; \vec{x}_n] = [A\vec{x}_1 \; A\vec{x}_2 \; \cdots \; A\vec{x}_n] = [\vec{e}_1 \; \vec{e}_2 \; \cdots \; \vec{e}_n] = I,$$

de manera que C es inversa derecha de A. Finalmente, la demostración de la Proposición 7.8 nos dice que si B es inversa izquierda de A, entonces $B = C = A^{-1}$. ■

Corolario 7.15 *Si $A \in M_{n\times n}$ es invertible, entonces A^{-1} también lo es y $(A^{-1})^{-1} = A$.*

Demostración. Como $AA^{-1} = I$, entonces A es una inversa por la izquierda para A^{-1} y por la Proposición 7.14 tenemos que $A = (A^{-1})^{-1}$ y concluimos la demostración. ∎

Si $A \in M_{n\times n}$, entonces la matriz transpuesta A^T también es una matriz de $n \times n$. Ahora bien, si A es invertible, ¿lo será también A^T? El siguiente corolario responde a esta pregunta.

Corolario 7.16 *Si $A \in M_{n\times n}$ es invertible, entonces A^T también lo es y $(A^T)^{-1} = (A^{-1})^T$.*

Demostración. Como $AA^{-1} = I$, tomando la transpuesta de ambos lados y utilizando el punto 4 de la Proposición 6.10 tenemos

$$(A^{-1})^T A^T = \left(AA^{-1}\right)^T = I^T = I.$$

Gracias a la Proposición 7.14 concluimos que $(A^T)^{-1} = (A^{-1})^T$. ∎

Ejemplo 7.17 *La matriz*

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

del Ejemplo 7.4 tiene por inversa a

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

Así pues, gracias al Corolario 7.16 tenemos que

$$(A^T)^{-1} = \left(A^{-1}\right)^T = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}.$$

Proposición 7.18 *Si $A \in M_{n\times n}$ y C es una inversa derecha de A, entonces C^T es una inversa izquierda de A^T. Si $A \in M_{n\times n}$ y B es un inversa izquierda de A, entonces B^T es una inversa derecha de A^T.*

Demostración. La primera afirmación se prueba notando que

$$C^T A^T = (AC)^T = I^T = I$$

y la segunda se verifica de forma similar. ∎

La Proposición 7.18 nos permite establecer las versiones análogas del Lema 7.10 y de la Proposición 7.14, pero esta vez partiendo de la existencia de inversos por la derecha.

Corolario 7.19 *Si $A \in M_{n\times n}$ tiene una inversa por la derecha, es decir, existe $C \in M_{n\times n}$ tal que $AC = I$, entonces los renglones de A forman un conjunto LI.*

Demostración. Supongamos que C es la inversa derecha de A. Por la Proposición 7.18, C^T es inversa izquierda de A^T. Utilizando el Lema 7.10 para A^T se tiene que sus columnas forman un conjunto LI. Dado que las columnas de A^T coinciden con los renglones de A, se concluye que los renglones de A forman un conjunto LI. ■

Corolario 7.20 *Si $A \in M_{n\times n}$ tiene una inversa por la derecha, entonces A tendrá una inversa por la izquierda y ambas coinciden.*

Demostración. Supongamos que C es la inversa derecha de A. Por la Proposición 7.18, C^T es inversa izquierda de A^T y aplicando la Proposición 7.14 a la matriz A^T, concluimos que ésta tiene una inversa por la derecha, digamos D. Una vez más, por la Proposición 7.18, D^T es inversa izquierda de $\left(A^T\right)^T = A$. Finalmente, por la demostración de la Proposición 7.8, $D^T = C = A^{-1}$. ■

El siguiente resultado se sigue del Lema 7.10, de la Proposición 7.14 y de los Corolarios 7.19 y 7.20.

Proposición 7.21 *Dada $A \in M_{n\times n}$, las siguientes son equivalentes:*

1. *A tiene una inversa por la izquierda.*
2. *Las columnas de A forman un conjunto LI.*
3. *La FER de A es la matriz identidad.*
4. *A tiene una inversa por la derecha.*
5. *Los renglones de A forman un conjunto LI.*
6. *A es invertible.*

Dada una matriz invertible A, las columnas de su matriz inversa pueden encontrarse explícitamente utilizando la demostración de la Proposición 7.14. Concretamente, si $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ son las columnas de la matriz inversa, éstas satisfacen los siguientes sistemas de ecuaciones:

$$A\vec{x}_1 = \vec{e}_1,\ A\vec{x}_2 = \vec{e}_2, \ldots, A\vec{x}_n = \vec{e}_n.$$

Es decir, se tienen n sistemas de ecuaciones (uno por cada columna). Afortunadamente, como comparten la misma matriz de coeficientes, todos estos sistemas pueden condensarse en uno sólo para resolverlos mediante el método de Gauss - Jordan estudiado en el Capítulo 2. Este procedimiento se ilustra en los siguientes ejemplos.

Ejemplo 7.22 *Sea*

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix}$$

y supongamos que queremos encontrar $C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = A^{-1}$*. Para este efecto resolvemos los siguientes sistemas de ecuaciones:*

$$\begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} c_{11} \\ c_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} c_{12} \\ c_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Como los sistemas poseen la misma matriz de coeficientes, podemos considerarlos como un sólo sistema con dos columnas de términos independientes obteniendo la FER como sigue:

$$\left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & 4 & 0 & 1 \end{array}\right] \overset{\frac{1}{2}R_1}{\sim} \left[\begin{array}{cc|cc} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 4 & 0 & 1 \end{array}\right] \overset{\frac{1}{4}R_2}{\sim}$$
$$\left[\begin{array}{cc|cc} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 & \frac{1}{4} \end{array}\right] \overset{-\frac{1}{2}R_2+R_1}{\sim} \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & -\frac{1}{8} \\ 0 & 1 & 0 & \frac{1}{4} \end{array}\right].$$

Por lo tanto, la matriz

$$C = \begin{bmatrix} \frac{1}{2} & -\frac{1}{8} \\ 0 & \frac{1}{4} \end{bmatrix}$$

es una inversa por la derecha para A (AC = I). Por la Proposición 7.21, la matriz A será invertible, de manera que $C = A^{-1}$.

Ejemplo 7.23 *Sea*

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

y supongamos que deseamos encontrar A^{-1}*. El lector paciente puede comprobar que*

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & 1 & -\frac{1}{2} & 0 & \frac{1}{2} \end{array}\right],$$

por lo que

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}.$$

Ejemplo 7.24 *Si tomamos la matriz*

$$A = \begin{bmatrix} 1 & -3 \\ -1 & 3 \end{bmatrix}$$

del Ejemplo 7.11, podemos obtener su FER con el procedimiento anterior como sigue,

$$\left[\begin{array}{rr|rr} 1 & -3 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{array}\right] \overset{R_1+R_2}{\sim} \left[\begin{array}{rr|rr} 1 & -3 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

Claramente el sistema (o más bien los dos sistemas compactados en uno) es inconsistente y no existe la matriz inversa de A.

Observación 7.25 *El procedimiento anterior queda resumido como sigue: si A es invertible y queremos encontrar su inversa utilizando el método de Gauss Jordan, simplemente se comienza con la matriz aumentada*

$$[A \mid I]$$

y se aplican las operaciones usuales para los renglones para llegar a la FER dada por

$$\left[I \mid A^{-1}\right].$$

Inversas de productos

Dadas dos matrices invertibles, una pregunta natural es si su producto también es invertible, el siguiente resultado nos contesta esta pregunta en sentido afirmativo.

Proposición 7.26 *Si $A, B \in M_{n\times n}$ son invertibles, entonces el producto AB también lo es y $(AB)^{-1} = B^{-1}A^{-1}$.*

Demostración. Por la Proposición 7.21 basta probar que

$$(B^{-1}A^{-1})(AB) = I.$$

Pero,

$$(B^{-1}A^{-1})(AB) = B^{-1}\left(A^{-1}A\right)B = B^{-1}IB = B^{-1}B = I,$$

lo cual concluye la demostración. ■

Ejemplo 7.27 *Si*

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

entonces

$$AB = \begin{bmatrix} 1 & 2 \\ 5 & 8 \end{bmatrix}$$

y

$$(AB)^{-1} = B^{-1}A^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} -4 & 1 \\ \frac{5}{2} & -\frac{1}{2} \end{bmatrix}.$$

Corolario 7.28 *Si* $A(1), A(2), \ldots, A(k) \in M_{n\times n}$ *son invertibles, entonces el producto* $A(1)A(2)\ldots A(k)$ *también lo es y, de hecho,*

$$(A(1)A(2)\ldots A(k))^{-1} = A(k)^{-1}A(k-1)^{-1}\ldots A(1)^{-1}.$$

Proposición 7.29 *Sean* $A, B \in M_{n\times n}$*, entonces* AB *es invertible syss* A *y* B *son invertibles.*

Demostración. Si A y B son invertibles, por la Proposición 7.26 el producto AB también lo es. Ahora bien, si AB es invertible, se tiene que existe $(AB)^{-1}$ tal que

$$(AB)(AB)^{-1} = I.$$

Por asociatividad del producto podemos reescribir esta igualdad como

$$A\left(B(AB)^{-1}\right) = I.$$

Por lo tanto, A tiene una inversa derecha y por la Proposición 7.21 A es invertible y

$$A^{-1} = B(AB)^{-1}.$$

En forma análoga, como

$$(AB)^{-1}(AB) = I,$$

se tiene que

$$\left((AB)^{-1}A\right)B = I,$$

por lo que B es invertible y

$$B^{-1} = (AB)^{-1}A,$$

concluyendo así la demostración. ■

Inversas de matrices elementales

Si interpretamos las operaciones entre renglones en términos de multiplicación por matrices elementales, la Proposición 7.21 y el Corolario 6.47 nos llevan al siguiente resultado.

Proposición 7.30 $A \in M_{n\times n}$ *es invertible syss existen matrices elementales* $E(1), E(2), \ldots, E(k)$ *tales que*

$$E(k)E(k-1)\ldots E(1)A = I.$$

En particular dado que, por definición, las matrices elementales se obtienen de la identidad mediante una operación elemental entre renglones, cualquier matriz elemental es invertible. Más aún, el inverso de una matriz elemental es otra matriz elemental ya que el inverso de cualquier operación elemental entre renglones es otra operación elemental del mismo tipo como veremos a continuación.

Proposición 7.31 *Si $E \in M_{n\times n}$ es una matriz elemental, entonces E es invertible y E^{-1} es elemental del mismo tipo que E. Más aún,*

1. *Si E es de tipo 1, entonces $E^{-1} = E$.*
2. *Si E es e tipo 2 e $I \overset{cR_i}{\sim} E$ con $c \neq 0$, entonces $I \overset{\frac{1}{c}R_i}{\sim} E^{-1}$.*
3. *Si E es de tipo 3 e $I \overset{cR_i+R_j}{\sim} E$, entonces $I \overset{-cR_i+R_j}{\sim} E^{-1}$.*

Demostración. A manera de ejemplo probamos 1 y 3. Si E es de tipo 1 e $I \overset{R_i\leftrightarrow R_j}{\sim} E$, entonces $E \overset{R_i\leftrightarrow R_j}{\sim} I$. La Proposición 6.46 asegura que $EE = I$, por lo que $E = E^{-1}$. Si E es de tipo 3 e $I \overset{cR_i+R_j}{\sim} E$, entonces $E \overset{-cR_i+R_j}{\sim} I$. Una vez más, por la Proposición 6.46, se tiene que la matriz elemental F obtenida como $I \overset{-cR_i+R_j}{\sim} F$ satisface $FE = I$, por lo tanto $E^{-1} = F$. ■

Ejemplo 7.32 *La matriz elemental*

$$E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

que se obtiene a partir de la identidad intercambiando los renglones, es su propia inversa (simplemente se vuelven a intercambiar los renglones). En efecto,

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

Ejemplo 7.33 *La matriz*

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

que se obtiene de la identidad multiplicando el segundo renglón por π, tiene como inversa la matriz que se obtiene de la identidad multiplicando el segundo renglón por $\frac{1}{\pi}$:

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

Ejemplo 7.34 *La matriz*

$$E = \begin{bmatrix} 1 & 0 & 666 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

tiene por inversa a

$$E^{-1} = \begin{bmatrix} 1 & 0 & -666 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Ejemplo 7.35 *En el ejemplo 7.22 podemos representar la sucesión de operaciones elementales sobre los renglones como productos por matrices elementales como sigue:*

$$\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

de manera que

$$\left(\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \right) = \begin{bmatrix} \frac{1}{2} & -\frac{1}{8} \\ 0 & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix}^{-1}.$$

Corolario 7.36 *Si A es una matriz de $n \times n$, entonces A es invertible syss A es el producto de un número finito*[1] *de matrices elementales.*

Demostración. Supongamos primero que A es invertible. Por la proposición 7.30, sabemos que existen matrices elementales $E(1), E(2), \cdots E(k)$ tales que

$$E(k)E(k-1)\dots E(1)A = I.$$

Así pues,

$$A^{-1} = E(k)E(k-1)\dots E(1)$$

y consecuentemente A es igual al siguiente producto de matrices elementales,

$$A = E(1)^{-1}E(2)^{-1}\dots E(k)^{-1}.$$

Por otro lado, como las matrices elementales son invertibles, por el Corolario 7.28 A es invertible. ■

Ejemplo 7.37 *En el Ejemplo 7.35 hemos visto que si*

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix},$$

entonces

$$A^{-1} = \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}.$$

Por lo tanto,

$$A = (A^{-1})^{-1} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix}.$$

[1] Tal número puede ser 1,2,3,..., etcétera.

Transformaciones inversas

Gracias al inciso (e) del Ejercicio 5.18, sabemos que si tenemos una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$, con $n \neq m$, entonces dicha transformación no es biyectiva. Ahora bien, dados $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una transformación lineal biyectiva y cualquier vector $\vec{y} \in \mathbb{R}^n$, por suprayectividad tenemos que existe $\vec{x} \in \mathbb{R}^n$ tal que $T(\vec{x}) = \vec{y}$. Asimismo, como T es inyectiva, solamente hay un vector $\vec{x}$ con esta propiedad. Esta transformación biyectiva tiene una función inversa[2]

$$T^{-1} : \mathbb{R}^n \longrightarrow \mathbb{R}^n,$$

definida como sigue: si $\vec{y} \in \mathbb{R}^n$, entonces $T^{-1}(\vec{y}) = \vec{x}$, en donde $\vec{x}$ es el único elemento tal que $T(\vec{x}) = \vec{y}$. Si I es la transformación identidad, es claro que se cumplen

$$\begin{aligned}(T^{-1} \circ T)(\vec{x}) &= T^{-1}(T(\vec{x})) = \vec{x} = I(\vec{x}),\\ (T \circ T^{-1})(\vec{y}) &= T(T^{-1}(\vec{y})) = \vec{y} = I(\vec{y}).\end{aligned}$$

Lo anterior nos dice que $T^{-1} \circ T = I = T \circ T^{-1}$.

El Teorema 5.46 afirma que si $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es una transformación lineal biyectiva y $A \in M_{n\times n}$ es su matriz estándar asociada, las columnas de A forman un conjunto LI, y por la Proposición 7.21, A es invertible. Sea A^{-1} la inversa de A y sea

$$S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

la transformación lineal asociada, es decir, $S(\vec{y}) = A^{-1}\vec{y}$ para $\vec{y} \in \mathbb{R}^n$. Si I_n representa la matriz identidad de $n \times n$, se tiene que para todo $\vec{x}, \vec{y} \in \mathbb{R}^n$

$$\begin{aligned}(S \circ T)(\vec{x}) &= S(T(\vec{x})) = A^{-1}A\vec{x} = I_n\vec{x} = I(\vec{x}) = \vec{x},\\ (T \circ S)(\vec{y}) &= T(S(\vec{y})) = AA^{-1}\vec{y} = I_n\vec{y} = I(\vec{y}) = \vec{y}.\end{aligned}$$

Notemos que dado cualquier vector $\vec{y} \in \mathbb{R}^n$, la transformación S queda definida por $S(\vec{y}) = \vec{x}$, en donde $\vec{x} \in \mathbb{R}^n$ es el único vector tal que $T(\vec{x}) = \vec{y}$. Por lo tanto, S coincide con la definición de T^{-1} (como función entre conjuntos) y $S = T^{-1}$. Concluimos así que T^{-1} es la **transformación inversa** de T. La discusión anterior puede resumirse en el siguiente corolario.

Corolario 7.38 *Si $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es una transformación lineal biyectiva y $A \in M_{n\times n}$ es su matriz estándar asociada, entonces*

$$T^{-1} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

[2] Aún no sabemos si resultará ser o no una transformación lineal.

es una transformación lineal biyectiva y A^{-1} es la matriz estándar que la representa.

Ejemplo 7.39 *Sea $T:\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por*

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} y \\ x \end{bmatrix}.$$

Esta transformación lineal es claramente biyectiva y por lo tanto invertible. Supongamos que deseamos calcular la matriz estándar que representa a T^{-1} y $T^{-1}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$. La matriz estándar, A, que representa a T está dada por

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

Gracias al Corolario 7.38, sabemos que A^{-1} es la matriz que representa a T^{-1}. Para calcular esta matriz inversa notamos que

$$\left[\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right] \overset{R_1 \leftrightarrow R_2}{\sim} \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right],$$

por lo que

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

Finalmente,

$$T^{-1}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = A^{-1}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}.$$

Esto es, $T^{-1} = T$ y $A^{-1} = A$, lo cual tiene mucho sentido pues es obvio que la composición $T \circ T$ es igual a la transformación identidad y que AA es igual a la matriz identidad.

Ejemplo 7.40 *Dada la transformación lineal*

$$S\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y+z \\ x-z \\ y+z \end{bmatrix},$$

queremos averiguar si es invertible y en tal caso, encontrar su transformación inversa. Para este efecto, notamos que la matriz estándar asociada a S está dada por

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

El lector puede comprobar fácilmente que

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & -1 & 1 & 2 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{array}\right],$$

por lo que A es invertible y

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

De aquí se concluye que S es invertible y S^{-1} está dada como

$$S^{-1}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 2 \\ 1 & -1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - z \\ -x + y + 2z \\ x - y - z \end{bmatrix}.$$

Para concluir este capítulo enunciamos los resultados correspondientes a las Proposiciones 7.26, 7.29 y a los Corolarios 7.15 y 7.28 en el contexto de transformaciones lineales. Las demostraciones son inmediatas y se dejan como ejercicio al lector.

Proposición 7.41 *Si $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es una transformación lineal biyectiva, entonces $(T^{-1})^{-1} = T$.*

Proposición 7.42 *Si $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ y $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ son transformaciones lineales biyectivas, entonces $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$.*

Proposición 7.43 *Si $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ y $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ son transformaciones lineales, entonces S y T son biyectivas syss $S \circ T$ es biyectiva.*

Corolario 7.44 *Si $T_i : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $i = 1, 2, \ldots, k$ son transformaciones lineales biyectivas, se cumple*

$$(T_1 \circ T_2 \circ \cdots \circ T_k)^{-1} = T_k^{-1} \circ T_{k-1}^{-1} \circ \cdots \circ T_1^{-1}.$$

Ejercicios

Ejercicio 7.1 *Calcular (en caso de que sea posible) la inversas de las matrices que se presentan a continuación.*

$$A = \begin{bmatrix} 5 & 2 \\ -6 & -3 \end{bmatrix},\ B = \begin{bmatrix} 0.4 & 2 \\ 2 & 10 \end{bmatrix},\ C = \begin{bmatrix} 2 & 4 & 0 \\ 8 & 6 & 5 \\ 0 & -10 & 1 \end{bmatrix},$$

$$D = \begin{bmatrix} 0 & 0 & 8 \\ 1 & 0 & 7 \\ -3 & 6 & 3 \end{bmatrix},\ E = \begin{bmatrix} 4 & 6 & 0 \\ 1 & 0 & -1 \\ 5 & 10 & 5 \end{bmatrix},\ F = \begin{bmatrix} 1 & 1 & 6 \\ 1 & 2 & 7 \\ 1 & 3 & 8 \end{bmatrix},$$

$$G = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},\ H = \begin{bmatrix} \pi & 0 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ \sqrt{2} & 2 & 3 & 5 \\ -6 & 0 & 0 & 0 \end{bmatrix},$$

$$J = \begin{bmatrix} 1 & 0 \\ -1 & 2 \\ 5 & 5 \end{bmatrix},\ K = \begin{bmatrix} -2 & -2 \\ 1 & 1 \end{bmatrix}.$$

Ejercicio 7.2 *Considerar las matrices A, C y D del ejercicio anterior y sean*

$$\vec{a} = \begin{bmatrix} 6 \\ 3 \end{bmatrix},\ \vec{c} = \begin{bmatrix} 10 \\ 40 \\ 0 \end{bmatrix},\ \vec{d} = \begin{bmatrix} 8 \\ 2 \\ 6 \end{bmatrix}.$$

Resolver los sistemas $A\vec{x} = \vec{a}$, $C\vec{y} = \vec{c}$ y $D\vec{z} = \vec{d}$ (por supuesto, $\vec{x}$ es un vector con dos variables mientras que $\vec{y}$, $\vec{z}$ son vectores con tres variables) utilizando las matrices inversas encontradas.

Ejercicio 7.3 *Sea A una matriz simétrica. Probar que si A es invertible, entonces A^{-1} es simétrica.*

Ejercicio 7.4 *Demostrar que si A es una matriz invertible y α es un escalar distinto de cero, entonces αA es invertible y $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$.*

Ejercicio 7.5 *Sean A y B matrices de $n \times n$. Probar que si $AB = 7I$, entonces $BA = 7I$.*

Ejercicio 7.6 *Probar que si A es una matriz invertible y antisimétrica (es decir, $A^T = -A$), entonces A^{-1} es antisimétrica.*

Ejercicio 7.7 *Demostrar (en general) o refutar (con un contraejemplo) los siguientes enunciados (en cada caso se debe indicar claramente "verdadero" o "falso").*

a. *Sean A y B matrices de $n \times n$. Si A y B son matrices invertibles, entonces $(A+B)^{-1} = A^{-1} + B^{-1}$.*

b. *Sean A y B matrices de $n \times n$. Si A y B son matrices invertibles, entonces $(AB)^{-1} = A^{-1}B^{-1}$.*

c. *Sean A y B matrices de $n \times n$. Si A y B son distintas de la matriz cero, entonces el producto AB es distinto de la matriz cero.*

d. *Sean A y B matrices de $n \times n$. Si A es invertible y B es distinta de la matriz cero, entonces el producto AB es distinto de la matriz cero.*

e. *Sean A, B, $C \in M_{2\times 2}$ tres matrices cada una de ellas distinta de la matriz cero. Si $AB = AC$, entonces $B = C$.*

f. *Sean A, B y C tres matrices de $n \times n$. Si A es invertible y $AB = AC$, entonces $B = C$.*

g. *Sea A una matriz de 3×3. Si A es distinta de la identidad y de la matriz cero, entonces $AA \neq A$.*

h. *Sea A una matriz de 7×7. Si A es invertible y A es distinta de la matriz identidad, entonces $AA \neq A$.*

Ejercicio 7.8 *Sean $A, B, C, D, X \in M_{n\times n}$ matrices invertibles. Despejar la matriz X en las siguientes ecuaciones.*

a. $AX = B$.

b. $AX = B - CX$.

c. $XA = B + XD$.

d. $B(A + X^{-1})^{-1} = C$.

e. $2X - D = AX$.

f. $B(A + X)^T = X^T$.

Ejercicio 7.9 *Encontrar $S^{-1}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$ y $T^{-1}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$, donde S y T son la transformaciones lineales de $\mathbb{R}^2$ en $\mathbb{R}^2$ dadas por*

$$S\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x-y \end{bmatrix},\ T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 4x+3y \\ 3x+2y \end{bmatrix}.$$

Ejercicio 7.10 *Encontrar las inversas de las siguientes matrices elementales*

$$E(1)=\begin{bmatrix}1&0&0\\0&1&666\\0&0&1\end{bmatrix},\ E(2)=\begin{bmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{bmatrix},\ E(3)=\begin{bmatrix}\pi&0&0\\0&1&0\\0&0&1\end{bmatrix}.$$

Ejercicio 7.11 *Expresar a las siguientes matrices invertibles como productos de matrices elementales.*

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

Ejercicio 7.12 *Demostrar las Proposiciones 7.41, 7.42, 7.43 y el Corolario 7.44.*

Capítulo 8

Determinantes

Introducción

En este capítulo introduciremos un criterio numérico para decidir si una matriz cuadrada es o no invertible. Para ello, comenzamos notando que si la matriz

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

satisface $ad - bc \neq 0$, entonces la matriz A es invertible. En efecto, si $ad - bc \neq 0$, necesariamente $ad \neq 0$ o bien $bc \neq 0$. Sin pérdida de generalidad, supongamos que $ad \neq 0$ (el caso $bc \neq 0$ es análogo) de manera que $a \neq 0 \neq d$. La FER de la matriz aumentada

$$\left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right]$$

se encuentra como sigue

$$\left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right] \overset{-\frac{c}{a}R_1+R_2}{\sim} \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & d-\frac{bc}{a} & -\frac{c}{a} & 1 \end{array}\right] \overset{\frac{1}{a}R_1,\ \frac{a}{ad-bc}R_2}{\sim}$$
$$\left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right] \overset{-\frac{b}{a}R_2+R_1}{\sim} \left[\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right].$$

Por lo tanto,

$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

lo cual se verifica fácilmente ya que,

$$\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Ahora veremos que si

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

es invertible entonces $ad - bc \neq 0$. Si A es invertible, entonces sus columnas forman un conjunto LI y por lo tanto, $a \neq 0$ o $b \neq 0$ (Si $a = b = 0$, las columnas forman un conjunto LD). Sin pérdida de generalidad supondremos que $a \neq 0$ (el lector puede verificar que el caso cuando $b \neq 0$ es completamente análogo). Al multiplicar el primer renglón de A por $-\frac{c}{a}$ y sumárselo al segundo obtenemos la matriz

$$\begin{bmatrix} a & b \\ 0 & d - \frac{bc}{a} \end{bmatrix}.$$

Notamos ahora que $d - \frac{bc}{a} \neq 0$, porque en caso contrario la matriz no sería invertible. Por lo tanto,

$$ad - bc \neq 0,$$

tal y como queríamos demostrar.

La función determinante

El razonamiento de la sección anterior nos conduce directamente a la siguiente definición.

Definición 8.1 *Si A es una matriz de 2×2 de la forma*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

entonces se define el ***determinante*** *de A, denotado por* $\det(A)$, *como*

$$\det(A) = ad - bc.$$

Utilizando esta terminología tenemos el siguiente resultado.

Proposición 8.2 *Si A es una matriz de 2×2 de la forma*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

entonces A es invertible syss $\det(A) \neq 0$. *En cuyo caso,*

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Ejemplo 8.3 *Si*

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

entonces A no es invertible (pues $\det(A) = 3(-4) - (-2)6 = 0$*) y B sí es invertible (pues* $\det(B) = 1(4) - (2)(-3) = 10$*). Más aún,*

$$B^{-1} = \frac{1}{10} \begin{bmatrix} 4 & -2 \\ 3 & 1 \end{bmatrix}.$$

Ejemplo 8.4 *La matriz identidad* $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ *es invertible pues* $\det(I) = 1 \neq 0$.

Observación 8.5 *Dada*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

la matriz

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

*se denomina la **matriz adjunta** de A y se denota por $adj(A)$. Esta matriz tiene la propiedad de que si A es invertible, entonces*

$$A^{-1} = \frac{1}{\det(A)} adj(A).$$

Como puede observarse en el caso en el cual $A \in M_{2\times 2}$, la matriz $adj(A)$ es sumamente simple de obtener. El caso general es más elaborado y se abordará mas adelante en este capítulo.

A continuación proporcionamos una definición del determinante basada en la dada por E. Artin[1]. Más adelante veremos que ésta concide con la que hemos proporcionado anteriormente para matrices de 2×2.

[1] Galois Theory, Dover Publications, Inc, NY, 1998.

Definición 8.6 *Decimos que una función*

$$F : M_{n\times n} \longrightarrow \mathbb{R}$$

*es **multilineal en los renglones** si se cumplen las siguientes dos propiedades para toda $i = 1, \ldots, n$.*

1. *Dadas $A, A^{'} \in M_{n\times n}$ y $c \in \mathbb{R}$ tales que,*

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{bmatrix} \quad y \quad A^{'} = \begin{bmatrix} A_1 \\ \vdots \\ cA_i \\ \vdots \\ A_n \end{bmatrix};$$

   entonces,

$$F(A^{'}) = cF(A).$$

2. *Dadas $A, A^{'}, B \in M_{n\times n}$ tales que,*

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{bmatrix}, \ A' = \begin{bmatrix} A_1 \\ \vdots \\ A_i^{'} \\ \vdots \\ A_n \end{bmatrix} \quad y \quad B = \begin{bmatrix} A_1 \\ \vdots \\ A_i + A_i^{'} \\ \vdots \\ A_n \end{bmatrix};$$

   entonces se tiene que

$$F(B) = F(A) + F(A').$$

Observación 8.7 *Lo anterior no implica que F sea una función lineal puesto que, dados $A, B \in M_{n\times n}$ y $c \in \mathbb{R}$, no se exige ni*

$$F(cA) = cF(A)$$

ni

$$F(A + B) = F(A) + F(B).$$

La multilinealidad implica que F es lineal en cada renglón. Sin embargo, si variamos única y exclusivamente el renglón i de A y dejamos a los demás fijos, la multininealidad da lugar a una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}$.

Ejemplo 8.8 *La función determinante de la Definición 8.1 es multilineal en los renglones. En efecto, observamos que*

$$\begin{aligned}
\det\left(\begin{bmatrix} a_1+a_2 & b_1+b_2 \\ c & d \end{bmatrix}\right) &= (a_1+a_2)\,d-(b_1+b_2)\,c \\
&= (a_1d-b_1c)+(a_2d-b_2c) \\
&= \det\left(\begin{bmatrix} a_1 & b_1 \\ c & d \end{bmatrix}\right)+\det\left(\begin{bmatrix} a_2 & b_2 \\ c & d \end{bmatrix}\right)
\end{aligned}$$

y dada $r \in \mathbb{R}$,

$$\begin{aligned}
\det\left(\begin{bmatrix} ra & rb \\ c & d \end{bmatrix}\right) &= rad-rbc \\
&= r(ad-bc) = r\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right).
\end{aligned}$$

Evidentemente, las mismas propiedades se cumplen para el segundo renglón y concluimos que la función $\det : M_{2\times 2} \longrightarrow \mathbb{R}$ *es multilineal.*

Ejemplo 8.9 *Sean*

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},\ \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix},\ \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \in M_{2\times 2}$$

y $\det : M_{2\times 2} \longrightarrow \mathbb{R}$ *la función determinante de la Definición 8.1. Entonces,*

$$\begin{aligned}
\det\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right) &= -2 = -1+(-1) \\
&= \det\left(\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}\right)+\det\left(\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\right)
\end{aligned}$$

y

$$\det\left(\begin{bmatrix} 1 & 2 \\ -6 & -8 \end{bmatrix}\right) = -2\det\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right) = 4.$$

Definición 8.10 *Una función determinante es una función*

$$\det : M_{n\times n} \longrightarrow \mathbb{R}$$

que satisface las siguientes propiedades:

1. $\det$ *es multilineal en los renglones.*
2. $\det(A) = 0$ *cuando A tiene dos renglones consecutivos iguales.*
3. $\det(I) = 1$.

Gracias a la Proposición 8.2 y a los Ejemplos 8.4 y 8.8, tenemos que la función det de la Definición 8.1 es, efectivamente una función determinante de acuerdo a las condiciones de la Definición 8.10. En general, dada $n \geq 2$, adoptaremos la notación det para referirnos a cualquier función que satisfaga dichas condiciones. Esto no constituye un abuso de notación porque, como veremos más adelante, para cada $n \geq 2$, existe una única función de este tipo. Consecuentemente, dada $A \in M_{n\times n}$, nos referiremos a $\det(A)$ como el **determinante** de A (mismo que también suele ser denotado por $|A|$ en otros textos).

Proposición 8.11 *Si $A \in M_{n\times n}$ tiene un renglón de ceros,*

$$\det(A) = 0.$$

Demostración. Supongamos que A_i es un renglón de ceros. Como A se obtiene de A multiplicando A_i por cero, la multilinealidad nos dice que

$$\det(A) = 0 \times \det(A) = 0$$

y concluimos con la demostración. ■

Proposición 8.12 *Supongamos que A' se obtiene a partir de A intercambiando dos renglones consecutivos, digamos el i y el $i+1$, es decir,*

$$A \overset{R_i \leftrightarrow R_{i+1}}{\sim} A',$$

entonces

$$\det(A') = -\det(A).$$

Demostración. Utilizando las propiedades (1) y (2) del determinante se tiene que,

$$\begin{aligned}
0 &= \det \begin{bmatrix} \vdots \\ A_i + A_{i+1} \\ A_i + A_{i+1} \\ \vdots \end{bmatrix} \begin{array}{l} \\ \longleftarrow i \\ \longleftarrow i+1 \\ \\ \end{array} \\
&= \det \begin{bmatrix} \vdots \\ A_i \\ A_i + A_{i+1} \\ \vdots \end{bmatrix} + \det \begin{bmatrix} \vdots \\ A_{i+1} \\ A_i + A_{i+1} \\ \vdots \end{bmatrix} \\
&= \det \begin{bmatrix} \vdots \\ A_i \\ A_i \\ \vdots \end{bmatrix} + \det \begin{bmatrix} \vdots \\ A_i \\ A_{i+1} \\ \vdots \end{bmatrix} + \det \begin{bmatrix} \vdots \\ A_{i+1} \\ A_i \\ \vdots \end{bmatrix} + \det \begin{bmatrix} \vdots \\ A_{i+1} \\ A_{i+1} \\ \vdots \end{bmatrix} \\
&= 0 + \det(A) + \det(A') + 0,
\end{aligned}$$

por lo tanto $\det(A') = -\det(A)$. ■

Corolario 8.13 *La propiedad (2) de la Definición 8.10 puede generalizarse como sigue.*

(2)´ $\det(A) = 0$ *cuando A tiene dos renglones iguales (no necesariamente consecutivos).*

Demostración. Supongamos que en la matriz A los renglones i y k son iguales con $i < k$, es decir $A_i = A_k$ y A_k está por debajo de A_i. Construimos la matriz A' intercambiando A_k en forma sucesiva con los renglones que lo preceden hasta llegar a la posición $i + 1$. De acuerdo a la Proposición 8.12 se tendrá que

$$\det(A') = \pm \det(A),$$

en donde el signo depende del número de veces que intercambiamos renglones. Dado que A' tiene dos renglones consecutivos iguales (el i y el $i + 1$), la propiedad (2) garantiza que

$$\det(A') = 0 = \det(A).$$

Por lo tanto, las propiedades (2) y (2)´ son equivalentes. ■

Utilizando la propiedad $(2)'$, el siguiente corolario es inmediato sustituyendo $i + 1$ por k en la demostración de la Proposición 8.12.

Corolario 8.14 *Si A' se obtiene a partir de A intercambiando dos renglones, digamos el $i -$ ésimo y el $k -$ ésimo; es decir,*

$$A \overset{R_i \leftrightarrow R_k}{\sim} A',$$

entonces

$$\det(A') = -\det(A).$$

Las propiedades que hemos demostrado hasta ahora son suficientes para garantizar que las Definiciones 8.1 y 8.10 son coherentes. En efecto, si

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

y calculamos el determinante de A *à l'Artin* (de acuerdo a la Definición 8.10), tenemos que

$$\begin{aligned} \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} &= \det \begin{bmatrix} a & 0 \\ c & d \end{bmatrix} + \det \begin{bmatrix} 0 & b \\ c & d \end{bmatrix} \\ &= a \det \begin{bmatrix} 1 & 0 \\ c & d \end{bmatrix} + b \det \begin{bmatrix} 0 & 1 \\ c & d \end{bmatrix} \\ &= ac \det \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + ad \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &+ bc \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + bd \det \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \\ &= 0 + ad(1) + bc(-1) + 0 = ad - bc. \end{aligned}$$

Es decir, la función $\det(A)$ no solo existe sino que es única por lo que coincide con la expresión dada en la Definición 8.1.

Con el fin de encontrar una caracterización similar para $\det(A)$ cuando $A \in M_{n\times n}$ con $n \geq 3$, es necesario seguir explorando propiedades adicionales de la función determinante. El Corolario 8.14 describe su interacción con operaciones elementales de tipo 1 y a continuación analizaremos lo que sucede en los dos casos restantes.

Proposición 8.15 *Si A' se obtiene a partir de A multiplicando el renglón i por un real $c \neq 0$, es decir*

$$A \overset{cR_i}{\sim} A',$$

entonces

$$\det(A') = c\det(A).$$

Demostración. *La demostración es inmediata por la multilinealidad (para renglones) de la función determinante.* ∎

Proposición 8.16 *Si $c \in \mathbb{R}$ y A' se obtiene a partir de A sumando c veces el renglón i al renglón j, es decir*

$$A \overset{cR_i+R_j}{\sim} A',$$

entonces

$$\det(A') = \det(A).$$

Demostración. Por la multilinealidad del determinante se tiene que

$$\det(A') = \det\begin{bmatrix} \vdots \\ A_i \\ \vdots \\ cA_i + A_j \\ \vdots \end{bmatrix} \begin{array}{l} \\ \longleftarrow i \\ \\ \longleftarrow j \\ \\ \end{array} = \det\begin{bmatrix} \vdots \\ A_i \\ \vdots \\ cA_i \\ \vdots \end{bmatrix} + \det\begin{bmatrix} \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \end{bmatrix}$$

$$= c\det\begin{bmatrix} \vdots \\ A_i \\ \vdots \\ A_i \\ \vdots \end{bmatrix} + \det(A) = 0 + \det(A) = \det(A).$$

En el último paso se utilizó que el determinante de una matriz con dos renglones iguales es nulo. Concluimos que $\det(A) = \det(A')$. ∎

Corolario 8.17 *La función*

$$\det : M_{n\times n} \longrightarrow \mathbb{R}$$

cumple las siguientes propiedades en el contexto de matrices elementales:

1. *Si* $E(1) \in M_{n\times n}$ *es tal que* $I \overset{R_i \leftrightarrow R_j}{\sim} E(1)$, *entonces*
$$\det(E(1)) = -1.$$

2. *Si* $E(2) \in M_{n\times n}$ *es tal que* $I \overset{cR_i}{\sim} E(2)$, *donde* $c \in \mathbb{R}$, $c \neq 0$, *entonces*
$$\det(E(2)) = c.$$

3. *Sea* $c \in \mathbb{R}$, *si* $E(3) \in M_{n\times n}$ *es tal que* $I \overset{cR_i+R_j}{\sim} E(3)$, *donde* $c \in \mathbb{R}$, *entonces*
$$\det(E(3)) = 1.$$

Demostración. Las demostraciones son inmediatas tomando $A = I$ en el Corolario 8.14 y las Proposiciones 8.15 y 8.16. ■

Proposición 8.18 *Sea* $A \in M_{n\times n}$ *y sean* $E(1), E(2), E(3) \in M_{n\times n}$ *las matrices elementales del Corolario 8.17. Entonces*

$$\begin{aligned}
\det(E(1)A) &= -\det(A) = \det(E(1))\det(A),\\
\det(E(2)A) &= c\det(A) = \det(E(2))\det(A),\\
\det(E(3)A) &= \det(A) = \det(E(3))\det(A),
\end{aligned}$$

es decir, la función determinante preserva productos de matrices elementales.

Demostración. Por lo visto en la Proposición 6.46, la multiplicación por matrices elementales equivale a las operaciones elementales sobre los renglones de A. Por lo tanto, aplicando los Corolarios 8.14 y 8.17 se obtiene que

$$\det(E(1)A) = -\det(A) = \det(E(1))\det(A).$$

Similarmente,

$$\det(E(2)A) = c\det(A) = \det(E(2))\det(A)$$

se obtiene de la Proposición 8.15 y el Corolario 8.17 y

$$\det(E(3)A) = \det(A) = \det(E(3))\det(A)$$

se sigue de la Proposición 8.16 y el Corolario 8.17. ■

Corolario 8.19 *Sean* $E(1)$, $E(2), \ldots$, $E(k) \in M_{n\times n}$ *cualesquiera matrices elementales y* $A \in M_{n\times n}$. *Entonces*

$$\det\left(E(k)E(k-1)\cdots E(1)A\right) = \det\left(E(k)\right)\det\left(E(k-1)\right)\cdots\det\left(E(1)\right)\det(A).$$

Demostración. La prueba es inmediata aplicando la Proposición 8.18 y el principio de inducción. ■

Corolario 8.20 *Sean* $E(1)$, $E(2), \ldots$, $E(k) \in M_{n\times n}$ *cualesquiera matrices elementales, entonces*

$$\det\left(E(k)E(k-1)\cdots E(1)\right) = \det\left(E(k)\right)\det\left(E(k-1)\right)\cdots\det\left(E(1)\right),$$

es decir, la función det *preserva productos de matrices elementales.*

Demostración. La demostración es inmediata tomando $A = I$ en el Corolario 8.19. ■

Ahora estamos listos para extender el resultado que se tenía para matrices de 2×2 que proporciona la relación entre el determinante y la invertibilidad de una matriz.

Proposición 8.21 *Si* $A \in M_{n\times n}$, *entonces* A *es invertible syss*

$$\det(A) \neq 0.$$

Demostración. Sean $\tilde{A}$ la FER de A y $E(1), E(2), \ldots E(k)$ matrices elementales tales que

$$A = E(1)E(2)\cdots E(k)\tilde{A}.$$

Ahora bien, si A es invertible $\tilde{A} = I$. En tal caso, los Corolarios 8.17 y 8.19, garantizan que

$$\begin{aligned}\det(A) &= \det\left(E(1)E(2)\cdots E(k)I\right)\\ &= \det\left(E(1)\right)\det\left(E(2)\right)\cdots\det\left(E(k)\right) \neq 0.\end{aligned}$$

Si A no es invertible, entonces $\tilde{A}$ tiene un renglón de ceros y la Proposición 8.11 nos dice que $\det(\tilde{A}) = 0$. Por lo tanto,

$$\begin{aligned}\det(A) &= \det\left(E(1)E(2)\cdots E(k)\tilde{A}\right)\\ &= \det\left(E(1)\right)\det\left(E(2)\right)\cdots\det\left(E(k)\right)\det(\tilde{A}) = 0,\end{aligned}$$

con lo cual se concluye la demostración. ■

Ejemplo 8.22 *La matriz invertible*

$$A = \begin{bmatrix} 0 & 1 \\ 5 & -1 \end{bmatrix},$$

puede expresarse como producto de matrices elementales:

$$\begin{bmatrix} 0 & 1 \\ 5 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}.$$

Entonces, dado que

$$\det\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) = -1, \ \det\left(\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\right) = 1, \det\left(\begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}\right) = 5,$$

se tiene que

$$\det(A) = (-1)(1)(5) = -5.$$

El siguiente resultado es uno de los más importantes en relación a la función determinante.

Proposición 8.23 *Si A y B son matrices de $n \times n$, entonces*

$$\det(AB) = \det(A)\det(B).$$

Demostración. Si A o B no es invertible, entonces AB tampoco lo es y por la Proposición 8.21 se cumple

$$0 = \det(AB) = \det(A)\det(B) = 0.$$

Si A y B son invertibles, AB también lo es y la prueba se sigue del Corolario 8.20, recordando que las matrices invertibles son productos de matrices elementales. ■

Corolario 8.24 *Si A es invertible,*

$$\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(I) = 1,$$

por lo que

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Ejemplo 8.25 *Sea*

$$A = \begin{bmatrix} -1 & 1 \\ 3 & 2 \end{bmatrix},$$

entonces $\det(A) = -5$*. La matriz inversa está dada por*

$$A^{-1} = \begin{bmatrix} -\frac{2}{5} & \frac{1}{5} \\ \frac{3}{5} & \frac{1}{5} \end{bmatrix}$$

y

$$\det(A^{-1}) = -\frac{1}{5} = \frac{1}{\det(A)}.$$

De acuerdo a la discusión final del Capítulo 6 tenemos que si $E(i) \in M_{n\times n}$ es una matriz elemental de tipo $i = 1, 2, 3$, entonces

$$\begin{aligned} E(1)^T &= E(1), \\ E(2)^T &= E(2) \end{aligned}$$

y $E(3)^T$ es una matriz elemental de tipo 3. De esta forma,

$$\det(E(i)^T) = \det(E(i)),\ i = 1, 2, 3.$$

Estas consideraciones nos llevan al siguiente resultado.

Proposición 8.26 *Si $A \in M_{n\times n}$, entonces*

$$\det(A) = \det(A^T).$$

Demostración. Si A no es invertible A^T tampoco lo es y se tiene

$$0 = \det(A) = \det(A^T) = 0.$$

Si A es invertible, entonces A^T también lo es. Adicionalmente, A puede expresarse como producto de matrices elementales, digamos

$$A = E(1)E(2)\cdots E(k).$$

Por lo tanto,

$$A^T = (E(1)E(2)\cdots E(k))^T = E(k)^T E(k-1)^T \cdots E(1)^T$$

y

$$\begin{aligned} \det(A^T) &= \det\left(E(k)^T E(k-1)^T \cdots E(1)^T\right) \\ &= \det\left(E(k)^T\right)\det\left(E(k-1)^T\right)\cdots\det\left(E(1)^T\right) \\ &= \det\left(E(k)\right)\det\left(E(k-1)\right)\cdots\det\left(E(1)\right) \\ &= \det(A), \end{aligned}$$

lo cual concluye la demostración. ■

Observación 8.27 *Como consecuencia inmediata de este resultado se tiene que las condiciones (1) y (2) de la Definición 8.10 pueden expresarse, alternativamente, en términos de las columnas de la matriz como sigue.*

1c. La función det *es* ***multilineal en las columnas****.*

2c. $\det(A) = 0$ *cuando A tiene dos columnas consecutivas iguales.*

Antes de mostrar una función que cumpla todas las propiedades de la función determinante para $n > 2$, demostraremos que dicha función es única. En realidad, esto es relativamente simple dadas todas las propiedades que ya conocemos de la función determinante como se ve a continuación.

Teorema 8.28 (Unicidad de la función determinante) *Si existe la función* det *de la Definición 8.10, ésta es única.*

Demostración. Supongamos que existen dos funciones, det y det′, que cumplen con las tres condiciones de la Definición 8.10. Dada $A \in M_{n\times n}$, si A no es invertible, por la Proposición 8.21 se tiene que

$$\det(A) = 0 = \det'(A) \ .$$

Si A es invertible, existen $E(1), E(2), \ldots, E(k)$ matrices elementales de $n \times n$ tales que

$$A = E(1)E(2)\cdots E(k).$$

Por el Corolario 8.17 y la Proposición 8.23 se tiene que las funciones det y det´coinciden en las matrices elementales y preservan productos, por lo tanto

$$\begin{aligned}\det(A) &= \det\left(E(1)\right)\det\left(E(2)\right)\cdots\det\left(E(k)\right)\\ &= \det'(E(1))\det'(E(2))\cdots\det'(E(k))\\ &= \det'(A)\end{aligned}$$

y ambas funciones coinciden. ■

Observación 8.29 *La descomposición de A como producto de matrices elementales no es única ya que es posible llegar a la FER de A de diversas maneras. Esto podría ocasionar problemas para definir la función determinante; sin embargo, el teorema de unicidad que acabamos de demostrar no tiene que ver con la existencia y definición del determinante. Simplemente nos dice que si existen las funciones* det *y* det′, *entonces tienen que coincidir.*

Hasta el momento, todo el desarrollo acerca de las propiedades de la función determinante ha sido un simple ejercicio teórico. Después de todo, es posible que esta función no exista para $n > 2$. Afortunadamente, éste no es el caso y en la siguiente sección construiremos una función explícita para calcular el deteminante de una matriz de $n \times n$ con $n > 2$.

Existencia del determinante

Antes de definir el determinante en el caso general se requiere de la siguiente definición.

Definición 8.30 *Dados $A \in M_{n\times n}$ y dos números naturales i y j en $\{1, 2, \ldots, n\}$, definimos $\widetilde{A}_{ij} \in M_{(n-1)\times(n-1)}$ como la matriz que se obtiene a partir de A eliminando el renglón i y la columna j.*

Ejemplo 8.31 *Consideremos las matrices*

$$A = \begin{bmatrix} 1 & -7 & 0 \\ -2 & 6 & 9 \\ 8 & -1 & -2 \end{bmatrix} \text{ y } B = \begin{bmatrix} 2 & 5 & 7 & 9 \\ 3 & 2 & 0 & \pi \\ 1 & 8 & 6 & 4 \\ 0 & 0 & 1 & 9 \end{bmatrix}.$$

Entonces

$$\widetilde{A}_{13} = \begin{bmatrix} -2 & 6 \\ 8 & -1 \end{bmatrix}, \ \tilde{A}_{31} = \begin{bmatrix} -7 & 0 \\ 6 & 9 \end{bmatrix}, \text{ y } \widetilde{B}_{43} = \begin{bmatrix} 2 & 5 & 9 \\ 3 & 2 & \pi \\ 1 & 8 & 4 \end{bmatrix}.$$

Recordemos que en el caso de matrices

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}.$$

la función determinante está dada por

$$\det(A) = ad - bc.$$

Definición 8.32 *Dados $A \in M_{n\times n}$ y j cualquier número natural entre 1 y n, la función*

$$D : M_{n\times n} \longrightarrow \mathbb{R}$$

se define como

$$D(A) = \det(A)$$

para el caso $n = 2$ y recursivamente para $n > 2$ como

$$D(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} D(\widetilde{A}_{ij}). \qquad (\bigstar)$$

Observemos que la función que acabamos de definir parece depender de la j elegida. Lo sorprendente (y el primero en notarlo fue P.S. Laplace) es que esta elección es irrelevante para el desarrollo de $D(A)$.

La razón es muy simple: independientemente de cual sea j, la expresión ($\bigstar$) satisface las condiciones de la Definición 8.10 y por lo tanto, la unicidad del determinante garantiza que

$$D(A) = \det(A).$$

Procedamos a verificar que, en efecto, esto sucede.

Proposición 8.33 *Para todo número natural $n \geq 2$, la función*

$$D : M_{n\times n} \longrightarrow \mathbb{R},$$

dada en la Definición 8.32, satisface las condiciones de la Definición 8.10, independientemente de la elección de j.

Demostración. Como ya se mencionó anteriormente, en el texto que sigue a la Definición 8.10, el resultado es válido para $n = 2$. Sea ahora $n \geq 3$ y supongamos, de manera inductiva, que el resultado es válido para $n - 1 \geq 2$. Verificaremos que la función $D : M_{n\times n} \longrightarrow \mathbb{R}$ cumple la primera propiedad de multilinealidad y la segunda condición de la Definición 8.10. Las propiedades restantes se demuestran en forma semejante y se dejan como ejercicio al lector. Sea pues, $A \in M_{n\times n}$ y sea A' dada por

$$A \overset{cR_k}{\sim} A',$$

en donde c es un escalar diferente del cero. Observemos que si $i \neq k$ las matrices $\widetilde{A'}_{ij}$ son de tamaño $(n-1) \times (n-1)$ y se obtienen de $\widetilde{A}_{ij}$ multiplicando uno de sus renglones por la constante c. Entonces, por hipótesis de inducción,

$$D(\widetilde{A'}_{ij}) = cD(\widetilde{A}_{ij})$$

para toda $i \neq k$. En el caso para el cual $i = k$, se tiene que $\widetilde{A'}_{ij} = \widetilde{A}_{ij}$ pero $a'_{kj} = ca_{kj}$. De esta forma se cumple

$$\begin{aligned} D(A') &= \sum_{i=1}^{n} (-1)^{i+j} a'_{ij} D(\widetilde{A'}_{ij}) \\ &= (-1)^{k+j} ca_{kj} D(\widetilde{A}_{kj}) + \sum_{\substack{i=1 \\ i\neq k}}^{n} (-1)^{i+j} a_{ij} cD(\widetilde{A}_{ij}) \\ &= cD(A). \end{aligned}$$

Supongamos ahora que la matriz A tiene dos renglones consecutivos iguales, digamos $A_k = A_{k+1}$. Si i es distinto de k y de $k+1$, las matrices

$\widetilde{A}_{ij}$ son de tamaño $(n-1) \times (n-1)$ y también tienen dos renglones consecutivos iguales. Por hipótesis de inducción,

$$D(\widetilde{A}_{ij}) = 0$$

para toda $i \neq k$, $i \neq k+1$. Por lo tanto, en el desarrollo de $D(A)$ sólo aparecen los términos correspondientes a $i = k$ e $i = k+1$. Adicionalmente, $a_{kj} = a_{(k+1)j}$ y $\widetilde{A}_{kj} = \widetilde{A}_{(k+1)j}$, de manera que

$$D(A) = (-1)^{k+j} a_{kj} D(\widetilde{A}_{kj}) + (-1)^{(k+1)+j} a_{kj} D(\widetilde{A}_{kj}) = 0,$$

concluyendo así con la demostración. ■

Como se mencionó anteriormente, la unicidad de la función determinante implica que, $D(A) = \det(A)$, independientemente de la columna j que haya sido elegida para calcular $D(A)$. Se tiene como corolario el resultado conocido como **expansión de Laplace** para el determinante. Éste nos dice que el determinante puede calcularse no sólo por medio de cualquier columna, sino también de cualquier renglón.

Teorema 8.34 *Sea* $n \geq 3$ *y sea* $A \in M_{n \times n}$. *Para cualesquiera* i, $j = 1, 2, \ldots, n$ *se tiene que*

$$\begin{aligned}\det(A) &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(\widetilde{A}_{ij}) \\ &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\widetilde{A}_{ij}).\end{aligned}$$

Demostración. La demostración se sigue de la Proposición 8.26 que nos dice que $\det(A) = \det(A^T)$. ■

En este contexto, el determinante $\det(\widetilde{A}_{ij})$ se conoce como el **menor** ij de A. Definiendo

$$C_{ij} = (-1)^{i+j} \det(\widetilde{A}_{ij}),$$

el número C_{ij} es conocido como el **cofactor** ij de A. Asimismo, la suma

$$\sum_{i=1}^{n} C_{ij} a_{ij}$$

es llamada la expansión por menores o por cofactores a lo largo de la $j-\textit{ésima}$ columna de A. Similarmente, la suma

$$\sum_{j=1}^{n} C_{ij} a_{ij}$$

es **la expansión por menores o cofactores** a lo largo del $i-\textit{ésimo}$ renglón de A. El Teorema de Laplace nos dice que

$$\det(A) = \sum_{i=1}^{n} C_{ij} a_{ij} = \sum_{j=1}^{n} C_{ij} a_{ij}.$$

Cálculo de determinantes

En esta sección se utiliza el método de Laplace para el cálculo de determinantes.

Ejemplo 8.35 *Consideremos la matriz*

$$A = \begin{bmatrix} 1 & 3 & -3 \\ -3 & -5 & 2 \\ -4 & 4 & -6 \end{bmatrix}$$

y supongamos que deseamos calcular el determinante de A *y fijamos* $i = 1$*, es decir, el desarrollo por cofactores del determinante se hace utilizando el primer renglón. Para ello, notamos que*

$$\begin{aligned} \det(A) &= (-1)^{1+1} \times 1 \times \det \begin{bmatrix} -5 & 2 \\ 4 & -6 \end{bmatrix} \\ &+ (-1)^{1+2} \times 3 \times \det \begin{bmatrix} -3 & 2 \\ -4 & -6 \end{bmatrix} \\ &+(-1)^{1+3} \times (-3) \times \det \begin{bmatrix} -3 & -5 \\ -4 & 4 \end{bmatrix} \\ &= (30-8) + (-3)(18+8) + (-3)(-12-20) \\ &= 22 - 78 + 96 = 40. \end{aligned}$$

Si alternativamente el desarrollo se realiza tomando la segunda columna de A*, entonces* $j = 2$ *y tenemos que*

$$\begin{aligned} \det(A) &= (-1)^{1+2} \times 3 \times \det \begin{bmatrix} -3 & 2 \\ -4 & -6 \end{bmatrix} \\ &+ (-1)^{2+2} \times (-5) \times \det \begin{bmatrix} 1 & -3 \\ -4 & -6 \end{bmatrix} \\ &+(-1)^{3+2} \times 4 \times \det \begin{bmatrix} 1 & -3 \\ -3 & 2 \end{bmatrix} \\ &= (-3)(18+8) + (-5)(-6-12) - 4(2-9) \\ &= -78 + 90 + 28 = 40, \end{aligned}$$

obteniendo lo mismo que antes.

Ejemplo 8.36 *Consideremos la matriz*

$$B = \begin{bmatrix} 0 & 1 & 3 & -3 \\ 2 & 0 & 0 & 0 \\ 1 & -3 & -5 & 2 \\ 0 & -4 & 4 & -6 \end{bmatrix}$$

y supongamos que deseamos calcular el determinante de B. Como el segundo renglón contiene el mayor número de entradas iguales a cero, tomamos $i = 2$ y notamos que,

$$\begin{aligned}\det(B) &= \sum_{j=1}^{4}(-1)^{2+j} b_{2j} \det(\widetilde{B}_{2j}) \\ &=(-1)^{2+1} \times 2 \times \det \begin{bmatrix} 1 & 3 & -3 \\ -3 & -5 & 2 \\ -4 & 4 & -6 \end{bmatrix} + 0 + 0 + 0 \\ &= (-1)^{2+1} \times 2 \times 40 = -80.\end{aligned}$$

La penúltima igualdad se debe a que en el ejemplo anterior habíamos mostrado que $\det(A) = 40$ *(notar que $\widetilde{B}_{21} = A$).*

Ejemplo 8.37 *Consideremos la matriz,*

$$A = \begin{bmatrix} 8 & 4 & 0 \\ 8 & 4 & 1 \\ 1 & 3 & 0 \end{bmatrix}$$

y supongamos que deseamos calcular el determinante de A utilizando la expansión por cofactores a lo largo de la tercera columna de A (debido a que en ella aparecen el mayor número de ceros). Entonces, tomamos $j = 3$ para obtener

$$\begin{aligned}\det(A) &= (-1)^{1+3} \times 0 \times \det \begin{bmatrix} 8 & 4 \\ 1 & 3 \end{bmatrix} \\ &+ (-1)^{2+3} \times 1 \times \det \begin{bmatrix} 8 & 4 \\ 1 & 3 \end{bmatrix} \\ &+ (-1)^{3+3} \times 0 \times \det \begin{bmatrix} 8 & 4 \\ 8 & 4 \end{bmatrix} \\ &= -(24 - 4) = -20.\end{aligned}$$

Ejemplo 8.38 *Consideremos la matriz*

$$B = \begin{bmatrix} 8 & 0 & 8 & 1 \\ 4 & 0 & 4 & 3 \\ 0 & 7 & 1 & 0 \\ -1 & 2 & 3 & 0 \end{bmatrix}$$

y supongamos que deseamos calcular el determinante de B utilizando la expansión por cofactores a lo largo de la segunda columna de B (simplemente porque en esa columna aparecen “muchos ceros”). Para *ello notamos que*

$$
\begin{aligned}
\det(B) &= \sum_{k=1}^{4} (-1)^{k+2} b_{k2} \det(\widetilde{B}_{k2}) \\
&= (-1)^{1+2} \times 0 \times \det \begin{bmatrix} 4 & 4 & 3 \\ 0 & 1 & 0 \\ -1 & 3 & 0 \end{bmatrix} + (-1)^{2+2} \times 0 \times \det \begin{bmatrix} 8 & 8 & 1 \\ 0 & 1 & 0 \\ -1 & 3 & 0 \end{bmatrix} \\
&+ (-1)^{3+2} \times 7 \times \det \begin{bmatrix} 8 & 8 & 1 \\ 4 & 4 & 3 \\ -1 & 3 & 0 \end{bmatrix} + (-1)^{4+2} \times 2 \times \det \begin{bmatrix} 8 & 8 & 1 \\ 4 & 4 & 3 \\ 0 & 1 & 0 \end{bmatrix} \\
&= -7(-80) + 2(-20) = 520.
\end{aligned}
$$

La penúltima igualdad se debe a que en el ejemplo anterior habíamos mostrado que $\det(A) = -20$ *(notar que* $\widetilde{B}_{42} = A^T$ *y utilizar la Proposición 8.26) y a que*

$$
\det \begin{bmatrix} 8 & 8 & 1 \\ 4 & 4 & 3 \\ -1 & 3 & 0 \end{bmatrix} =
$$

$$
(-1)^{3+1} \times (-1)(24-4) + (-1)^{3+2} \times 3(24-4) + 0 = -80.
$$

Ejemplo 8.39 *Si*

$$
A = \begin{bmatrix} 0 & 0 & \pi & 0 \\ 1 & 2 & 7 & 3 \\ 1 & 1 & -5 & 1 \\ 2 & 3 & 6 & 4 \end{bmatrix},
$$

desarrollando con respecto al primer renglón se tiene que

$$
\det(A) = (-1)^{1+3} \times \pi \times \det \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 2 & 3 & 4 \end{bmatrix} = 0,
$$

ya que los renglones de la submatriz $\tilde{A}_{13}$ *forman, claramente, un conjunto LD. En particular,* A *no es invertible y sus renglones también forman un conjunto LD, lo cual no es evidente a primera vista*[2]*.*

Recordemos que una matriz cuadrada D es diagonal syss todas sus componentes fuera de la diagonal principal son cero (i.e., syss $d_{ij} = 0$

---

[2] El lector paciente puede verificar que, por ejemplo

$$
A_1 = \frac{\pi}{4} (A_4 - A_2 - A_3).
$$

para toda $i \neq j$). Entonces, si

$$D = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix}$$

es una matriz diagonal, por la multilinealidad del determinante se tiene

$$\begin{aligned} \det(A) &= d_{11}d_{22}\cdots d_{nn} \det \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \\ &= d_{11}d_{22}\cdots d_{nn}. \end{aligned}$$

De esta forma, el determinante de una matriz diagonal es, simplemente el producto de los elementos de la diagonal. Este resultado puede extenderse a matrices un poco más generales como se ve a continuación.

Definición 8.40 *Sea $A \in M_{n\times n}$. Decimos que*

*(a) A es **triangular superior** syss todas sus componentes por debajo de la diagonal principal son cero (i.e., syss $a_{ij} = 0$ para toda $i > j$).*

*(b) A es **triangular inferior** syss todas sus componentes por arriba de la diagonal principal son cero (i.e., syss $a_{ij} = 0$ para toda $i < j$).*

Observación 8.41 *Si A es una matriz de $n \times n$, entonces*

1. *A es triangular superior syss A^T es triangular inferior.*

2. *A es diagonal syss A es triangular superior y triangular inferior.*

Ejemplo 8.42 *Las matrices*

$$\begin{bmatrix} 8 & 7 \\ 0 & \pi \end{bmatrix} \ y \ \begin{bmatrix} 0 & 7 & 6 \\ 0 & -1 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

son triangulares superiores. De igual manera la matriz identidad y la matriz cero de $n \times n$ son triangulares superiores.

Ejemplo 8.43 *Las matrices*

$$\begin{bmatrix} 8 & 0 \\ 9 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 2 & 666 & 0 \\ 7 & 0 & 9 \end{bmatrix} \ y \ \begin{bmatrix} -1 & 0 & 0 \\ 0 & \pi & 0 \\ 8 & 0 & 0 \end{bmatrix}$$

son triangulares inferiores. De igual manera la matriz identidad y la matriz cero de $n \times n$ son triangulares inferiores.

Proposición 8.44 *Sea $A \in M_{n\times n}$. Si A es triangular superior o triangular inferior, entonces*

$$\det(A) = a_{11}a_{22}\ldots a_{nn}.$$

Esto es, el determinante de A es igual al producto de todas las componentes de A que están en su diagonal principal.

Demostración. Se utiliza inducción sobre n. Si $n = 2$, el resultado es inmediato. Supongamos que el resultado es válido para matrices triangulares de tamaño $(n-1)\times(n-1)$. Dada $A \in M_{n\times n}$, una matriz triangular inferior, se tiene que A es de la forma

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ a_{(n-1)1} & \cdots & a_{(n-1)(n-1)} & 0 \\ a_{n1} & \cdots & a_{n(n-1)} & a_{nn} \end{bmatrix}.$$

Realizando el desarrollo de Laplace con respecto al primer renglón se tiene que,

$$\det(A) = a_{11}\det\left(\begin{bmatrix} a_{22} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ a_{(n-1)2} & \cdots & a_{(n-1)(n-1)} & 0 \\ a_{n2} & \cdots & a_{n(n-1)} & a_{nn} \end{bmatrix}\right).$$

Por hipótesis de inducción, el determinante de la matriz del lado derecho está dado como,

$$\det\left(\begin{bmatrix} a_{22} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ a_{(n-1)2} & \cdots & a_{(n-1)(n-1)} & 0 \\ a_{n2} & \cdots & a_{n(n-1)} & a_{nn} \end{bmatrix}\right) = a_{22}\cdots a_{nn},$$

de manera que $\det(A) = a_{11}a_{22}\ldots a_{nn}$. Si la matriz A fuese triangular superior, el único cambio en la demostración de arriba es que se realiza el desarrollo de Laplace con respecto a la primera columna de A. ■

Ejemplo 8.45 *Como la matriz*

$$A = \begin{bmatrix} 7 & 8 & 1 \\ 0 & \pi & -666 \\ 0 & 0 & 9 \end{bmatrix}$$

es triangular superior, $\det(A) = 63\pi$.

Ejemplo 8.46 *Sea*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 1 & 4 & 4 \end{bmatrix}.$$

Utilizando operaciones elementales entre renglones podemos transformar la matriz A en una triangular como sigue.

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 1 & 4 & 4 \end{bmatrix} \overset{-2R_1+R_2}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 1 & 4 & 4 \end{bmatrix} \overset{-R_1+R_3}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}$$
$$\overset{\frac{1}{2}R_3}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 1 & \frac{1}{2} \end{bmatrix} \overset{R_2 \leftrightarrow R_3}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}.$$

Las primeras dos operaciones no alteran el determinante. Utilizando la multilinealidad y el Corolario 8.14 se tiene,

$$\det(A) = 2(-1)\det\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}\right) = -2.$$

Una generalización del cálculo de determinantes para matrices triangulares se refiere a matrices de $n \times n$ cuya estructura "por bloques" es de la forma

$$\begin{bmatrix} A & B \\ \Phi & C \end{bmatrix}, \tag{$\spadesuit$}$$

en donde $A \in M_{p\times p}$, $B \in M_{p\times(n-p)}$, $C \in M_{(n-p)\times(n-p)}$ y $\Phi \in M_{(n-p)\times p}$ es la matriz cero. Por ejemplo, las matrices de 4×4 dadas como

$$M = \begin{bmatrix} 1 & -1 & \pi & 2 \\ -3 & 5 & -7 & 666 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}, \; N = \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 5 & -7 & 666 \\ 1 & 0 & -2 & \pi \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

tienen esta estructura; en efecto, en el caso de la matriz M se tiene que

$$A = \begin{bmatrix} 1 & -1 \\ -3 & 5 \end{bmatrix}, \; B = \begin{bmatrix} \pi & 2 \\ -7 & 666 \end{bmatrix}, \; C = \begin{bmatrix} -2 & 3 \\ 1 & 2 \end{bmatrix}, \; \Phi = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

y para la matriz N,

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 5 & -7 \\ 1 & 0 & -2 \end{bmatrix}, \; B = \begin{bmatrix} -2 \\ 666 \\ \pi \end{bmatrix}, \; C = [2], \; \Phi = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}.$$

Los determinantes de este tipo de matrices son relativamente simples de calcular como se muestra a continuación.

Proposición 8.47 *Dada una matriz M de $n \times n$ con la estructura dada por* ($\spadesuit$)*, se tiene que*

$$\det(M) = \det(A)\det(C).$$

Demostración. La demostración se hace por inducción sobre n. Si $n = 2$, el resultado es inmediato. Sea $n > 2$ y supongamos que el resultado es válido para matrices de $k \times k$ con $k < n$. Desarrollando el determinante con respecto a la primera columna de M se obtiene:

$$\det(M) = \sum_{i=1}^{p} (-1)^{i+1} a_{i1} \det(\tilde{M}_{i1}),$$

en donde utilizamos que

$$m_{i1} = \begin{cases} a_{i1} \text{ si } i = 1, \ldots, p, \\ 0 \text{ si } i = p+1, \ldots, n. \end{cases}$$

Las matrices $\tilde{M}_{i1}$ tienen la misma estructura por bloques dada por ($\spadesuit$) y son de tamaño $(n-1) \times (n-1)$; entonces, por hipótesis de inducción se tiene que para toda $i = 1, \ldots, p$,

$$\det(\tilde{M}_{i1}) = \det(\tilde{A}_{i1})\det(C).$$

Por lo tanto,

$$\begin{aligned} \det(M) &= \sum_{i=1}^{p} (-1)^{i+1} a_{i1} \det(\tilde{A}_{i1})\det(C) = \left(\sum_{i=1}^{p} (-1)^{i+1} a_{i1} \det(\tilde{A}_{i1}) \right) \det(C) \\ &= \det(A)\det(C), \end{aligned}$$

quedando demostrado el resultado. ∎

Ejemplo 8.48 *Para la matriz*

$$M = \begin{bmatrix} 1 & -1 & \pi & 2 \\ -3 & 5 & -7 & 666 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

se tiene que

$$\det(M) = \det\left(\begin{bmatrix} 1 & -1 \\ -3 & 5 \end{bmatrix} \right) \det\left(\begin{bmatrix} -2 & 3 \\ 1 & 2 \end{bmatrix} \right) = 2 \times (-7) = -14.$$

Para la matriz

$$N = \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 5 & -7 & 666 \\ 1 & 0 & -2 & \pi \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

tenemos que

$$\det(N) = \det\left(\begin{bmatrix} 1 & 0 & -1 \\ 0 & 5 & -7 \\ 1 & 0 & -2 \end{bmatrix}\right)\det([2]) = -5 \times 2 = -10.$$

La siguiente serie de equivalencias (todas han sido demostradas previamente) recapitula gran parte del trabajo hecho hasta ahora.

Proposición 8.49 *Sea $A \in M_{n\times n}$. Las siguientes afirmaciones son equivalentes.*

1. A es invertible.
2. Las columnas de A forman un conjunto LI.
3. La FER de A es la matriz identidad.
4. La ecuación $A\vec{x} = \vec{0}$ tiene una única solución (a saber, $\vec{x} = \vec{0}$).
5. La ecuación $A\vec{x} = \vec{b}$ tiene solución única para toda $\vec{b} \in \mathbb{R}^n$.
6. A^T es invertible.
7. Los renglones de A forman un conjunto LI.
8. $\det(A) \neq 0$

Ejemplo 8.50 *Consideremos la matriz*

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 \\ \pi & 10 & 12 & 14 \end{bmatrix}$$

y supongamos que deseamos calcular el determinante de B utilizando la expansión por cofactores a lo largo del primer renglón de B. Para ello, notamos que

$$\det(B) = (1) \times \det\begin{bmatrix} 3 & 4 & 5 \\ 7 & 8 & 9 \\ 10 & 12 & 14 \end{bmatrix} = 1 \times 0 = 0,$$

pues

$$\det\begin{bmatrix} 3 & 4 & 5 \\ 7 & 8 & 9 \\ 10 & 12 & 14 \end{bmatrix} = 0$$

ya que el tercer renglón de esta matriz es combinación lineal de los otros dos renglones.

Ejemplo 8.51 *Supongamos que deseamos encontrar todos los valores de a tales que el conjunto*

$$\left\{ \begin{bmatrix} 0 \\ a-3 \\ a^4+9 \end{bmatrix}, \begin{bmatrix} a-2 \\ a \\ \pi \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix} \right\}$$

es LI. Para ello construimos la matriz cuyas columnas son estos vectores, ordenados de tal suerte que el determinante sea lo más simple posible de calcular; por ejemplo sea

$$A = \begin{bmatrix} a-2 & 0 & 0 \\ a & a-3 & 0 \\ \pi & a^4+9 & 6 \end{bmatrix}.$$

Lo que estámos buscando son todos los valores de a tales que $det(A) \neq 0$. Como A es triangular inferior,

$$\det(A) = (a-2)(a-3)6,$$

por lo tanto nuestro conjunto es LI syss $a \neq 2$ y $a \neq 3$.

Matriz adjunta

Si $a \in \mathbb{R}$, definimos $\det(a) = a$, en donde el escalar a se interpreta como una matriz de 1×1. De esta forma, la definición del cofactor ij de una matriz A proporcionada en la discusión posterior al Teorema 8.34, tiene sentido si $A \in M_{n\times n}$, con $n \geq 2$. Así, el cofactor ij de A está dado como,

$$C_{ij} = (-1)^{i+j} \det(\widetilde{A}_{ij}).$$

Observamos que si $A \in M_{2\times 2}$, entonces $\det(\widetilde{A}_{ij}) = \widetilde{A}_{ij}$ ya que $\widetilde{A}_{ij} \in \mathbb{R}$.

Ejemplo 8.52 *Si* $A = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$, *sus cofactores están dados por*

$$\begin{aligned} C_{11} &= 1 \times 3 = 3, \quad C_{12} = -1 \times 2 = -2 \\ C_{21} &= -1 \times (-1) = 1, \quad C_{22} = 1 \times 1 = 1. \end{aligned}$$

Ejemplo 8.53 *Los cofactores de* $A = \begin{bmatrix} 1 & 2 & -2 \\ -1 & 0 & 1 \\ 0 & 3 & 1 \end{bmatrix}$ *están dados por*

$$\begin{aligned}
C_{11} &= 1 \times \det\left(\begin{bmatrix} 0 & 1 \\ 3 & 1 \end{bmatrix}\right), \ C_{12} = -1 \times \det\left(\begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}\right), \\
C_{13} &= 1 \times \det\left(\begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}\right), \ C_{21} = -1 \times \det\left(\begin{bmatrix} 2 & -2 \\ 3 & 1 \end{bmatrix}\right), \\
C_{22} &= 1 \times \det\left(\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\right), \ C_{23} = -1 \times \det\left(\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}\right), \\
C_{31} &= 1 \times \det\left(\begin{bmatrix} 2 & -2 \\ 0 & 1 \end{bmatrix}\right), \ C_{32} = -1 \times \det\left(\begin{bmatrix} 1 & -2 \\ -1 & 1 \end{bmatrix}\right), \\
C_{33} &= 1 \times \det\left(\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}\right).
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
&C_{11} = -3, \ C_{12} = 1, \ C_{13} = -3, \\
&C_{21} = -8, \ C_{22} = 1, \ C_{23} = -3, \\
&C_{31} = 2, \ C_{32} = 1, \ C_{33} = 2.
\end{aligned}$$

Dada una matriz A de 2×2, en la Observación 8.5 definimos la matriz adjunta de A, denotada por $adj(A)$, y demostramos que si A es invertible, entonces se cumple la relación $A^{-1} = \left(\frac{1}{\det(A)}\right) adj(A)$. A continuación generalizamos este resultado para matrices de $n \times n$.

Definición 8.54 *Dada $A \in M_{n\times n}$, se define la **matriz adjunta**[3] de A, denotada por $adj(A)$, como la transpuesta de la matriz cuyas entradas son los cofactores de A. Es decir,*

$$(adj(A))_{ij} = C_{ji}.$$

La siguiente proposición, nos dice que la matriz adjunta de A está íntimamente relacionada con $\det(A)$.

Proposición 8.55 *Sea $A \in M_{n\times n}$, entonces se cumple*

$$adj(A)A = Aadj(A) = \det(A)I.$$

[3]Cabe notar que en inglés esta matriz se conoce como "adjugate matrix". También existe el concepto de "adjoint matrix", sin embargo se refiere a una matriz distinta.

Demostración. Sean $i, j \in \{1, 2, \ldots, n\}$ y consideremos la expresión

$$\sum_{k=1}^{n} a_{ik} C_{jk}.$$

Si $i = j$, tenemos que $\sum_{k=1}^{n} a_{ik} C_{ik} = \det(A)$. Ahora bien, sea $i \neq j$ y sin pérdida de generalidad supongamos que $j > i$. Se define la matriz

$$A^{'} = \begin{array}{cl} \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_i \\ \vdots \\ A_n \end{bmatrix} & \begin{array}{l} \\ \\ \longleftarrow i \\ \\ \longleftarrow j \\ \\ \\ \end{array} \end{array}$$

como aquella que se obtiene de A sustituyendo el renglón j por el renglón i. Las entradas y los cofactores asociados al $j - \acute{e}simo$ renglón de $A^{'}$están dados como

$$a_{jk}^{'} = a_{ik} \ \text{ y } C_{jk}^{'} = C_{jk}.$$

Asimismo, como $A^{'}$ tiene dos renglones iguales su detrminante es cero. Se tiene así que si desarrollamos este determinante a lo largo del renglón j se tiene que

$$0 = \det(A^{'}) = \sum_{k=1}^{n} a_{jk}^{'} C_{jk}^{'} = \sum_{k=1}^{n} a_{ik} C_{jk}.$$

Se tiene así que

$$\sum_{k=1}^{n} a_{ik} C_{jk} = \begin{cases} \det(A) \text{ si } i = j, \\ \quad 0 \quad \text{ si } i \neq j. \end{cases}$$

Ahora bien, las entradas del producto $Aadj(A)$ están dadas como

$$(Aadj(A))_{ij} = A_i adj(A)^j = \sum_{k=1}^{n} a_{ik} C_{jk} = \begin{cases} \det(A) \text{ si } i = j, \\ \quad 0 \quad \text{ si } i \neq j, \end{cases}$$

de manera que se cumple la igualdad

$$Aadj(A) = \det(A) I.$$

La igualdad $adj(A)A = \det(A)I$ se deja como ejercicio al lector. ∎

Si la matriz A es invertible, se satisface $\det(A) \neq 0$ y se tiene el siguiente corolario que ya conocíamos para el caso de matrices de 2×2.

Corolario 8.56 *Sea $A \in M_{n\times n}$ una matriz invertible, entonces se cumple*

$$A^{-1} = \left(\frac{1}{\det(A)}\right) adj(A).$$

Ejemplo 8.57 *Para la matriz $A = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$ del Ejemplo 8.52 el lector puede fácilmente verificar que*

$$adj(A) = \begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}$$

y

$$A^{-1} = \frac{1}{5}\begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}.$$

Ejemplo 8.58 *Para la matriz $A = \begin{bmatrix} 1 & 2 & -2 \\ -1 & 0 & 1 \\ 0 & 3 & 1 \end{bmatrix}$ del Ejemplo 8.53 puede verificarse que*

$$A^{-1} = \frac{1}{5}\begin{bmatrix} -3 & -8 & 2 \\ -1 & 1 & 1 \\ -3 & -3 & 2 \end{bmatrix}.$$

Regla de Cramer

En esta sección se introduce la llamada **Regla de Cramer**, un método sumamente popular (aunque no muy eficiente) para resolver sistemas de ecuaciones lineales con el mismo número de ecuaciones que de variables. La utilidad de este resultado es más bien en el cálculo diferencial de varias variables, dado que se utiliza para una demostración de un teorema importante: el Teorema de la Función Implícita.

Teorema 8.59 (Regla de Cramer) *Sea $A \in M_{n\times n}$ una matriz invertible, sea $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ un vector de $\mathbb{R}^n$ y consideremos el sistema lineal $A\vec{x} = \vec{b}$ en donde*

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

es la única solución del sistema. Entonces, para cada $k \in \{1, \ldots, n\}$ se tiene que

$$x_k = \frac{\det(A^{k \leftrightarrow \vec{b}})}{\det(A)},$$

en donde $A^{k \leftrightarrow \vec{b}}$ es la matriz

$$\left[A^1 \cdots A^{k-1} \quad \vec{b} \quad A^{k+1} \cdots A^n \right]$$

que resulta de reemplazar la k-ésima columna de A por el vector $\vec{b}$ de términos independientes.

Demostración. Sea $\{\vec{e}_1, \ldots, \vec{e}_n\}$ la base estándar de $\mathbb{R}^n$. Realizamos el producto de la matriz A con la matriz que se obtiene de la identidad reemplazando su $k - \acute{e}sima$ columna por $\vec{x}$ como sigue,

$$\begin{aligned} A\left[\vec{e}_1 \cdots \vec{e}_{k-1} \ \vec{x} \ \vec{e}_{k+1} \cdots \vec{e}_n\right] &= \left[A\vec{e}_1 \cdots A\vec{e}_{k-1} \ A\vec{x} \ A\vec{e}_{k+1} \cdots A\vec{e}_n\right] \\ &= \left[A^1 \cdots A^{k-1} \ \vec{b} \ A^{k+1} \cdots A^n\right] = A^{k \leftrightarrow \vec{b}}. \end{aligned}$$

Tomando el determinante de ambos lados de esta expresión obtenemos

$$\begin{gathered} \det\left(A\left[\vec{e}_1 \cdots \vec{e}_{k-1} \ \vec{x} \ \vec{e}_{k+1} \cdots \vec{e}_n\right]\right) = \\ \det(A) \det\left(\left[\vec{e}_1 \cdots \vec{e}_{k-1} \ \vec{x} \ \vec{e}_{k+1} \cdots \vec{e}_n\right]\right) = \det\left(A^{k \leftrightarrow \vec{b}}\right). \end{gathered}$$

Es simple verificar que

$$\det\left(\left[\vec{e}_1 \cdots \vec{e}_{k-1} \ \vec{x} \ \vec{e}_{k+1} \cdots \vec{e}_n\right]\right) = x_k \det(I_{n-1}) = x_k$$

al evaluar este determinante a lo largo del $k - \acute{e}simo$ renglón. De esta forma, se obtiene que

$$x_x \det(A) = \det\left(A^{k \leftrightarrow \vec{b}}\right)$$

o bien, dado que $\det(A) \neq 0$ tenemos que

$$x_k = \frac{\det(A^{k \leftrightarrow \vec{b}})}{\det(A)},$$

concluyéndose así la demostración. ■

Ejemplo 8.60 *Supongamos que deseamos encontrar la solución del sistema*

$$\begin{aligned} 3x_1 + x_2 + 0 &= 5, \\ -x_1 + 2x_2 + x_3 &= -2, \\ 0 - x_2 + 2x_3 &= -1. \end{aligned}$$

Para ello, tomamos

$$A = \begin{bmatrix} 3 & 1 & 0 \\ -1 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix} \text{ y } \vec{b} = \begin{bmatrix} 5 \\ -2 \\ -1 \end{bmatrix},$$

por lo que nuestro sistema puede reescribirse como $A\vec{x} = \vec{b}$. *Notamos ahora que como*

$$\det(A) = 17 \neq 0,$$

puede utilizarse la Regla de Cramer para encontrar a x_1, x_2 *y* x_3. *Tenemos así que*

$$x_1 = \frac{\det(A^{1\leftrightarrow\vec{b}})}{\det(A)} = \frac{1}{17} \det \begin{bmatrix} 5 & 1 & 0 \\ -2 & 2 & 1 \\ -1 & -1 & 2 \end{bmatrix} = \frac{28}{17},$$

$$x_2 = \frac{\det(A^{2\leftrightarrow\vec{b}})}{\det(A)} = \frac{1}{17} \det \begin{bmatrix} 3 & 5 & 0 \\ -1 & -2 & 1 \\ 0 & -1 & 2 \end{bmatrix} = \frac{1}{17},$$

y

$$x_3 = \frac{\det(A^{3\leftrightarrow\vec{b}})}{\det(A)} = \frac{1}{17} \det \begin{bmatrix} 3 & 1 & 5 \\ -1 & 2 & -2 \\ 0 & -1 & -1 \end{bmatrix} = \frac{-8}{17}.$$

Ejercicios

Ejercicio 8.1 *Demostrar que la función*

$$F : M_{2\times 2} \longrightarrow \mathbb{R},$$

definida como

$$F\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = a_{11}a_{12},$$

no es multilineal (en los renglones).

Ejercicio 8.2 *Demostrar que la función*

$$F : M_{2\times 2} \longrightarrow \mathbb{R},$$

definida como

$$F\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = a_{11}a_{21},$$

es multilineal (en los renglones).

Ejercicio 8.3 *Si*

$$\det\left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}\right) = 5,$$

con $n \geq 2$, *calcular*

$$\det\left(\begin{bmatrix} a_{n1} & a_{n2} & \cdots & a_{nn} \\ a_{(n-1)1} & a_{(n-1)2} & \cdots & a_{(n-1)n} \\ \vdots & \vdots & & \vdots \\ a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix}\right),$$

es decir, el determinante de la matriz obtenida invirtiendo el orden de los renglones.

Ejercicio 8.4 *Sean* $A, B, C \in M_{n\times n}$, *tales que* $\det(A) = -2, \det(B) = 3$ *y* $\det(C) = 1$. *Encontrar,*

a. $\det(AB)$.

b. $\det(BC^T)$.

c. $\det(ABA^{-1})$.

d. $\det(2C)$.

e. $\det(3A^{(2)}B)$, *en donde* $A^{(2)} = AA$.

Ejercicio 8.5 *Sea* $A \in M_{n\times n}$ *una matriz antisimétrica, es decir,* $A^T = -A$. *Demostrar que si* n *es impar, entonces* $\det(A) = 0$.

Ejercicio 8.6 *Calcular los siguientes determinantes utilizando únicamente las propiedades de la función determinante (sin utilizar el método de Laplace).*

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \ B = \begin{bmatrix} 0 & 1 & -2 \\ 1 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}, \ C = \begin{bmatrix} -7 & 1 & 1 \\ 2 & 0 & -2 \\ -5 & 1 & -1 \end{bmatrix}.$$

Ejercicio 8.7 *Utilizar el método de cofactores de Laplace para calcular los determinantes de las siguientes matrices:*

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 3 & 5 \\ 0 & 1 & 3 \end{bmatrix}, \ B = \begin{bmatrix} 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & -5 & -6 \\ \pi & 666 & 7 & 11 & 111 \\ 1 & 1 & -1 & 2 & 2 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 2 & -1 & 5 \\ 1 & 0 & -2 & 2 \\ -1 & 1 & 2 & 0 \\ 3 & -1 & 0 & 1 \end{bmatrix}, \ D = \begin{bmatrix} 1 & 432 & \pi \\ 0 & 666 & 0 \\ 0 & 19 & 1 \end{bmatrix}.$$

Ejercicio 8.8 *Evaluar los determinates de las siguientes matrices con cualquier método (válido).*

$$A = \begin{bmatrix} -1 & 8 & 9 & 10 \\ 0 & 7 & -13 & -11 \\ 0 & 0 & 4 & -12 \\ 0 & 0 & 0 & -\frac{1}{7} \end{bmatrix}, \ B = \begin{bmatrix} 0 & -2 & 7 \\ 0 & 1 & 0 \\ 1 & 13 & 15 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 4 & 8 \\ 1 & 5 & 3 \\ 0 & 1 & 7 \end{bmatrix}, \ D = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 2 & 0 & -1 & 2 \\ 1 & 1 & 0 & 3 \\ 1 & -1 & 0 & 0 \end{bmatrix},$$

$$F = \begin{bmatrix} 1 & 0 & -2 \\ -1 & 3 & 7 \\ 1 & 6 & 8 \end{bmatrix}, \ G = \begin{bmatrix} 1 & 2 & 2 & -1 \\ 2 & 7 & -1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 10 & 0 \end{bmatrix},$$

$$H = \begin{bmatrix} 7 & 2 & -6 \\ 7 & 2 & -7 \\ 7 & 3 & -8 \end{bmatrix}, \ J = \begin{bmatrix} 0 & x & y & z \\ x & 1 & 0 & 0 \\ y & 0 & 1 & 0 \\ z & 0 & 0 & 1 \end{bmatrix}.$$

Ejercicio 8.9 *Encontrar todos los valores de λ tales que satisfacen las siguientes ecuaciones.*

a. $\det\left(\begin{bmatrix} 1-\lambda & 3 \\ -1 & 1+\lambda \end{bmatrix}\right) = 0.$

b. $\det\left(\begin{bmatrix} -\lambda & 0 & 1 \\ 1 & -\lambda & -3 \\ 0 & 1 & 3-\lambda \end{bmatrix}\right) = 0.$

Ejercicio 8.10 *Utilizar la descomposición de la matriz* $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 2 & 4 & 1 \end{bmatrix}$ *como producto de matrices elementales, encontrada en el Ejercicio 7.11, para calcular su determinante.*

Ejercicio 8.11 *Dada las matrices*

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

calcular

a. $adj(A)$ *y* $adj(B)$.

b. $\det(A)$ *y* $\det(B)$.

c. Utilizando los incisos anteriores calcular A^{-1} *y* B^{-1}.

Ejercicio 8.12 *Intentar resolver los siguientes sistemas mediante la regla de Cramer. Si no es posible, explicar cuál es el problema y resolverlos utilizando algún método alternativo.*

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

b. $\begin{cases} x+2y=1, \\ 2x+4y=2. \end{cases}$

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

Ejercicio 8.13 *Utilizar la regla de Cramer para calcular el valor de y en el siguiente sistema.*

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

Ejercicio 8.14 *Sean $A, B \in M_{n\times n}$. Probar (en general) o refutar (con un contraejemplo) los siguientes enunciados (en cada caso se debe indicar claramente "verdadero" o "falso").*

a. *La función*
$$F : M_{n\times n} \longrightarrow \mathbb{R},$$
definida como $F(A) = A_{11}A_{22}\cdots A_{nn}$, satisface las tres propiedades de la Definición 8.10 (definición de la función determinante).

b. *Si B se obtiene de A intercambiando la columna i con la columna j, entonces* $\det(B) = -\det(A)$.

c. $\det(A + B) = \det(A) + \det(B)$.

d. $\det(AA^T) \geq 0$.

e. *Si* $c \in \mathbb{R}$, $\det(cA) = c^n \det(A)$.

f. *Si* $c \in \mathbb{R}$, $\det(cA) = c \det(A)$.

g. *Si A es invertible, entonces* $\det(adj(A)) = \det^{n-1}(A)$.

h. *Si* $\det(A) = 0$, *entonces todos los cofactores de A son iguales a cero.*

i. *Si E es una matriz elemental, entonces* $\det(E) = \pm 1$.

j. *Sea $\vec{b} \in \mathbb{R}^n$. El sistema de ecuaciones $A\vec{x} = \vec{b}$ siempre puede resolverse por el método de Cramer.*

Ejercicio 8.15 *Completar la demostración de la Proposición 8.33.*

Ejercicio 8.16 *Probar que dada $A \in M_{n\times n}$ se satisface* $adj(A)A = \det(A)I$.

Capítulo 9

Factorización PLU(opcional)

Introducción

Supongamos que $A \in M_{n\times n}$ es una matriz triangular superior con $a_{ii} \neq 0$, $i = 1, 2, \ldots, n$, de manera que

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}.$$

Si consideramos el sistema lineal $A\vec{x} = \vec{b}$ con

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in \mathbb{R}^n,$$

observamos que las ecuaciones correspondientes están dadas por

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{nn}x_n &= b_n. \end{aligned}$$

Este sistema es consistente y puede resolverse recursivamente comenzando con la última ecuación.

Análogamente, si $A \in M_{n\times n}$ es una matriz triangular inferior, con $a_{ii} \neq 0$, $i = 1, 2, \ldots, n$, y de la forma

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

las ecuaciones correspondientes al sistema lineal $A\vec{x} = \vec{b}$ están dadas por

$$\begin{aligned} a_{11}x_1 &= b_{1,} \\ a_{21}x_1 + a_{22}x_2 &= b_{2,} \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n. \end{aligned}$$

En este caso, el sistema es consistente y puede resolverse recursivamente comenzando con la primera ecuación.

Ejemplo 9.1 *Sean*

$$A = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{bmatrix}, \ \vec{b} = \begin{bmatrix} -2 \\ 15 \\ 10 \end{bmatrix}.$$

El sistema de ecuaciones correspondiente está dado por

$$\begin{aligned} -x_1 + 2x_2 + x_3 &= -2, \\ 2x_2 - x_3 &= 15, \\ 2x_3 &= 10. \end{aligned}$$

De la tercera ecuación se tiene que $x_3 = 5$; *sustituyendo este resultado en la segunda ecuación se tiene que*

$$2x_2 - 5 = 15,$$

por lo tanto, $x_2 = 10$. *Finalmente, sustituyendo los valores de* x_2 *y* x_3 *en la primer ecuación tenemos que*

$$-x_1 + 2 \times 10 + 5 = -2,$$

por lo que $x_1 = 27$. *En este caso hemos resuelto el sistema de ecuaciones de "abajo hacia arriba" (lo que se denomina sustitución hacia atrás) y, si la matriz de coeficientes hubiese sido triangular inferior, se resolvería de "arriba hacia abajo" (o mediante sustitución hacia adelante).*

Factorización LU

Las consideraciones anteriores resaltan la simplicidad de aquellos sistemas lineales cuyas matrices de coeficientes son triangulares. Además, el número de operaciones elementales requerido para convertir a una matriz cuadrada en una matriz triangular (superior o inferior) es menor que el empleado para encontrar su FER. Por ésta y otras razones, es deseable -computacionalmente hablando - optar por la descomposición de una matriz en matrices triangulares en las situaciones donde efectivamente existe tal posibilidad. Una de ellas es cuando la matriz A de $n \times n$ satisface ciertas hipótesis (mismas que haremos explícitas más adelante) que permiten encontrar una matriz triangular inferior L y una superior U, tales que

$$A = LU,$$

en donde L es un producto de matrices elementales E de la forma

$$I \overset{cR_i+R_j}{\sim} E$$

con $i < j$.

El proceso para encontrar estas matrices triangulares es siempre el mismo de manera que, por simplicidad, supondremos que $n = 4$. Así pues, sea

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

Si $a_{11} \neq 0$, entonces podemos realizar sucesivamente las operaciones elementales

$$\begin{gathered} -\frac{a_{21}}{a_{11}}R_1 + R_2, \\ -\frac{a_{31}}{a_{11}}R_1 + R_3, \\ -\frac{a_{41}}{a_{11}}R_1 + R_4 \end{gathered}$$

sobre los renglones de A obteniendo como resultado a la matriz

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & b_{22} & b_{23} & b_{24} \\ 0 & b_{32} & b_{33} & b_{34} \\ 0 & b_{42} & b_{43} & b_{44} \end{bmatrix},$$

la cual coincide con A en el primer renglón. Gracias a la Proposición 6.46 tenemos que si E, F y G son las matrices elementales que se obtienen a partir de la identidad $I \in M_{4\times 4}$ realizando, respectivamente, las tres operaciones elementales que se acaban de mencionar, entonces

$$(GFE)A = B,$$

en donde

$$GFE = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & 0 & 0 \\ -\frac{a_{31}}{a_{11}} & 0 & 1 & 0 \\ -\frac{a_{41}}{a_{11}} & 0 & 0 & 1 \end{bmatrix}.$$

Si $b_{22} \neq 0$, entonces también existe una matriz

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ 0 & c_{22} & c_{23} & c_{24} \\ 0 & 0 & c_{33} & c_{34} \\ 0 & 0 & c_{43} & c_{44} \end{bmatrix}$$

que coincide con A y B en los renglones uno y dos, respectivamente, así como dos matrices elementales H y J con

$$I \overset{-\frac{b_{32}}{b_{22}}R_2+R_3}{\sim} H \quad \text{e} \quad I \overset{-\frac{b_{42}}{b_{22}}R_2+R_4}{\sim} J,$$

tales que

$$JHB = C.$$

Esta igualdad puede escribirse como

$$(JHGFE)A = C,$$

donde

$$JHGFE = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & 0 & 0 \\ -\frac{a_{31}}{a_{11}} & -\frac{b_{32}}{b_{22}} & 1 & 0 \\ -\frac{a_{41}}{a_{11}} & -\frac{b_{42}}{b_{22}} & 0 & 1 \end{bmatrix}.$$

Finalmente, si $c_{33} \neq 0$, entonces existe una matriz triangular superior

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{bmatrix}$$

(que bien podía haber sido llamada D) que coincide con los renglones uno dos y tres de A, B y C, respectivamente, así como una matriz elemental K con

$$I \overset{-\frac{c_{43}}{c_{33}}R_3+R_4}{\sim} K$$

tales que

$$KC = U.$$

Esta igualdad puede escribirse como

$$(KJHGFE)A = U,$$

donde

$$KJHGFE = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & 0 & 0 \\ -\frac{a_{31}}{a_{11}} & -\frac{b_{32}}{b_{22}} & 1 & 0 \\ -\frac{a_{41}}{a_{11}} & -\frac{b_{42}}{b_{22}} & -\frac{c_{43}}{c_{33}} & 1 \end{bmatrix}.$$

y su inversa es

$$(KJHGFE)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{a_{21}}{a_{11}} & 1 & 0 & 0 \\ \frac{a_{31}}{a_{11}} & \frac{b_{32}}{b_{22}} & 1 & 0 \\ \frac{a_{41}}{a_{11}} & \frac{b_{42}}{b_{22}} & \frac{c_{43}}{c_{33}} & 1 \end{bmatrix}.$$

Si denotamos por L a la matriz $(KJHGFE)^{-1}$, entonces es claro que

$$A = LU$$

con L triangular inferior y U triangular superior[1].

Ejemplo 9.2 *Sea*

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 0 & 3 \\ -3 & 1 & -4 \end{bmatrix};$$

entonces utilizando únicamente operaciones elementales del tipo 3 tenemos,

$$\begin{bmatrix} 1 & -1 & 2 \\ 2 & 0 & 3 \\ -3 & 1 & -4 \end{bmatrix} \overset{-2R_1+R_2}{\sim} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -1 \\ -3 & 1 & -4 \end{bmatrix} \overset{3R_1+R_3}{\sim} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -1 \\ 0 & -2 & 2 \end{bmatrix}$$
$$\overset{R_2+R_3}{\sim} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{bmatrix}.$$

Se tiene así que

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

[1] En el caso particular para el cual $L = U^T$, esta factorización se denomina **factorización de Cholesky**.

y como $A = LU$, L puede calcularse como

$$L = AU^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & -1 & 1 \end{bmatrix}.$$

Si $A = LU$, donde L y U son, respectivamente, triangular inferior y superior, entonces el sistema

$$A\vec{x} = \vec{b}$$

puede resolverse haciendo

$$U\vec{x} = \vec{y}$$

y notando que

$$L\vec{y} = L(U\vec{x}) = (LU)\vec{x} = A\vec{x} = \vec{b}.$$

De esta forma, primero tenemos que resolver $L\vec{y} = \vec{b}$ y posteriormente $U\vec{x} = \vec{y}$, con la ventaja de que en ambos casos podemos utilizar sustituciones regresivas tal y como explicamos al inicio de este capítulo.

Ejemplo 9.3 *Supongamos que deseamos resolver el sistema lineal*

$$\begin{bmatrix} 1 & -1 & 2 \\ -2 & 5 & -4 \\ 2 & -5 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 6 \end{bmatrix}$$

y que se tiene la siguiente descomposición LU :

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

Entonces, procedemos como sigue. Primero reescribimos el sistema original como

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 6 \end{bmatrix},$$

definimos

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

y observamos que del sistema

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 6 \end{bmatrix},$$

se obtiene directamente que

$$\begin{aligned} x' &= 3, \\ -2x' + y' &= 0, \ \textit{por lo que } y' = 6, \\ 2x' - y' + z' &= 6, \ \textit{por lo que } z' = 6. \end{aligned}$$

Después, utilizamos estos valores para encontrar x, y y z a partir del sistema

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 6 \end{bmatrix},$$

del cual también se obtiene en forma inmediata que

$$\begin{aligned} z &= 6, \\ 3y &= 6, \ \textit{por lo que } y = 2, \\ x - y + 2z &= 3, \ \textit{por lo que } x = -7, \end{aligned}$$

quedando así resuelto el sistema original.

Factorización PLU

Supongamos que tenemos una matriz $A \in M_{n\times n}$ que no admite una factorización LU. Por ejemplo, la matriz

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

no puede llevarse a una matriz triangular superior utilizando únicamente operaciones elementales del tipo 3 (el problema es que $a_{11} = 0$). En este caso se requiere intercambiar los renglones 1 y 2 para llegar a la matriz triangular superior. Ahora bien, mediante el intercambio adecuado de renglones, una matriz invertible siempre puede transformarse en una matriz que admite una factorización LU.

Ejemplo 9.4 *Es inmediato verificar que la matriz*

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

no admite una factorización LU. Sin embargo, intercambiando los renglones 2 y 3 obtenemos la matriz

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

Ahora bien, como

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \overset{-R_1+R_3}{\sim} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

A' tiene la siguiente factorización LU :

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

Definición 9.5 *Decimos que $P \in M_{n\times n}$ es una* ***matriz de permutación*** *syss P es un producto de matrices elementales de tipo 1.*

Una matriz de permutación es la matriz

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

la cual tiene la propiedad de que si A es una matriz con cuatro renglones, entonces PA se obtiene a partir de A intercambiando primero los renglones 2 y 4 y luego los renglones 1 y 3. Asimismo, la matriz

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

es la matriz de permutación tal que en el Ejemplo 9.4 $A' = QA$ tiene factorización LU. Esto último es un caso particular de un resultado más general que enunciamos a continuación.

Proposición 9.6 *Sea $A \in M_{n\times n}$, entonces existen una matriz triangular inferior (con unos en su diagonal principal) L, una matriz triangular superior U y una matriz de permutación Q tales que se cumple*

$$QA = LU.$$

Es decir, existe una matriz de permutación Q tal que QA admite una factorización LU.

Observación 9.7 *En el Ejercicio 9.6 se pide demostrar que si Q es una matriz de permutación, entonces $Q^{-1} = Q^T$. Así pues, si en la proposición anterior definimos a P como*

$$P = Q^{-1} = Q^T,$$

entonces P también es una matriz de permutación y además,

$$A = PLU.$$

A esta factorización de la matriz A se le conoce como ***factorización PLU****.*

Ejemplo 9.8 *La matriz*

$$A = \begin{bmatrix} 0 & -1 & 2 \\ -2 & 0 & -4 \\ 2 & -1 & 1 \end{bmatrix}$$

no tiene una descomposición LU. Sin embargo la matriz dada por

$$\underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{Q} \begin{bmatrix} 0 & -1 & 2 \\ -2 & 0 & -4 \\ 2 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 & -4 \\ 0 & -1 & 2 \\ 2 & -1 & 1 \end{bmatrix},$$

obtenida intercambiando los renglones 1 y 2 puede descomponerse como

$$\begin{bmatrix} -2 & 0 & -4 \\ 0 & -1 & 2 \\ 2 & -1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} -2 & 0 & -4 \\ 0 & -1 & 2 \\ 0 & 0 & -5 \end{bmatrix}}_{U}.$$

Si $P = Q^{-1}$, tenemos que $P = Q^T$. En este caso muy particular también es cierto que

$$A = PLU = QLU.$$

Esto es,

$$\begin{bmatrix} 0 & -1 & 2 \\ -2 & 0 & -4 \\ 2 & -1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} -2 & 0 & -4 \\ 0 & -1 & 2 \\ 0 & 0 & -5 \end{bmatrix}}_{U}.$$

Aquí es natural hacer notar que si Q es una matriz de permutación con un único factor, entonces $Q^T = Q$ (lo cual no es necesariamente cierto si Q tiene al menos dos factores). Finalizamos este capítulo con un resultado que funciona independientemente del número de factores que tenga Q.

Lema 9.9 *Sea $Q \in M_{n\times n}$ una matriz de permutación con q factores (es decir, es producto de q matrices elementales de tipo 1), entonces,* $\det(Q) = (-1)^q$.

Demostración. La demostración es inmediata pues el determinante de cada uno de los factores es igual a -1. ■

Proposición 9.10 *Si A, Q, L y U son como en el enunciado de la Proposición 9.6, q es el número de factores de Q y $P = Q^{-1}$ (es decir $P = Q^T$), entonces*

$$\det(A) = (-1)^q(u_{11}u_{22}\cdots u_{nn}).$$

Demostración. La demostración es inmediata recordando que el determinante de las matrices triangulares es simplemente el producto de los elementos de su diagonal y que preserva productos. En efecto,

$$\begin{aligned}\det(A) &= \det(PLU)\\ &= \det(P)\det(L)\det(U)\\ &= \det(Q^T)\det(L)\det(U)\\ &= \det(Q)\det(L)\det(U)\\ &= (-1)^q(u_{11}u_{22}\cdots u_{nn})\end{aligned}$$

con lo cual concluimos la demostración. ■

Ejemplo 9.11 *Utilizando la descomposición del Ejemplo 9.8 se tiene que*

$$\det\left(\begin{bmatrix} 0 & -1 & 2\\ -2 & 0 & -4\\ 2 & -1 & 1\end{bmatrix}\right) =$$

$$\det\left(\begin{bmatrix} 0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ -1 & 1 & 1\end{bmatrix}\begin{bmatrix} -2 & 0 & -4\\ 0 & -1 & 2\\ 0 & 0 & -5\end{bmatrix}\right)$$

$$= (-1)^1\left((2)(-1)(-5)\right) = 10.$$

Ejercicios

Ejercicio 9.1 *Encontrar, de ser posible, la factorización LU de las siguientes matrices:*

$$A = \begin{bmatrix} \pi & \pi \\ \pi & 0 \end{bmatrix}, \; B = \begin{bmatrix} 0 & 666 \\ 666 & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 7 & -1 & 8 \end{bmatrix}, \; D = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 2 & 1 \\ -2 & 1 & 1 & -2 \\ 0 & 0 & 0 & 10 \end{bmatrix}.$$

Ejercicio 9.2 *Utilizando la descomposición LU para la matriz C del ejercicio anterior, resolver el sistema dado por,*

$$\begin{aligned} x + 2y + 3z &= 0, \\ 5y + 6z &= 2, \\ 7x - y + z &= 4. \end{aligned}$$

Ejercicio 9.3 *Encontrar la descomposición PLU de las siguientes matrices:*

$$A = \begin{bmatrix} 0 & \pi \\ 666 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 2 & 1 \end{bmatrix}, \; C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 1 \\ -2 & 1 & 1 & -2 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Ejercicio 9.4 *Calcular* $\det(B)$ *y* $\det(C)$, *utilizando la descomposición PLU obtenida para las matrices B y C del ejercicio anterior y las propiedades del determinante.*

Ejercicio 9.5 *Probar (en general) o refutar (con un contraejemplo) los siguientes enunciados (en cada caso se debe indicar claramente "verdadero" o "falso").*

a. La matriz $\begin{bmatrix} \pi & 0 & 0 \\ 1 & 2 & 0 \\ -1 & 2 & 1 \end{bmatrix}$ *tiene factorización LU.*

b. Si A es invertible, entonces tiene factorización LU.

c. Si A es triangular superior, la factorización LU es $A = IA$.

d. Si A tiene factorización PLU, entonces A es invertible.

Ejercicio 9.6 *Demostrar que si Q es una matriz de permutación se tiene que $Q^{-1} = Q^T$.*

Capítulo 10

Subespacios

Introducción

Cuando se tienen conjuntos con alguna estructura algebraica es común que éstos posean subconjuntos con el mismo tipo de estructura. Tenemos así, que los racionales ($\mathbb{Q}$) son un subconjunto de los reales ($\mathbb{R}$) y ambos conjuntos de números, junto con las operaciones usuales de suma y multiplicación, constituyen lo que se conoce como un **campo.** De esta forma, los racionales no solo forman un subconjunto de los reales, sino que constituyen un "subcampo" de éstos. En el caso de los espacios vectoriales $\mathbb{R}^n$, éstos también tienen subconjuntos no vacíos que son cerrados bajo las operaciones de suma y multiplicación por escalares.

Definición 10.1 *Sea S un subconjunto no vacío de $\mathbb{R}^n$. Decimos que S es un* ***subespacio*** *de $\mathbb{R}^n$ syss se cumplen las siguientes tres condiciones:*

1. *$\vec{0}_n \in S$, donde $\vec{0}_n$ denota al vector cero de $\mathbb{R}^n$.*
2. *S es cerrado bajo sumas. Esto es, si $\vec{x}$ y $\vec{y}$ son elementos de S, entonces $\vec{x} + \vec{y} \in S$.*
3. *S es cerrado bajo multiplicación escalar. Esto es, si $\vec{x}$ es un elemento de S y c es un escalar, entonces $c\vec{x} \in S$.*

Observación 10.2 *El lector debe notar que si S es un subespacio de $\mathbb{R}^n$, entonces los elementos del conjunto S satisfacen las propiedades V1 a V8 dadas en 3.6. De esta forma, S es en sí mismo un espacio vectorial contenido en $\mathbb{R}^n$.*

Un resultado inmediato a partir de las propiedades de la definición de subespacio es el siguiente.

Proposición 10.3 *Si $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ son elementos de un subespacio S, entonces span$\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\} \subset S$.*

Demostración. Supongamos que $\vec{v} \in span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$. Esto es, existen escalares $c_1, c_2, \ldots, c_m$ tales que

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_m\vec{v}_m.$$

Debemos probar que $\vec{v} \in S$. Sin embargo, como $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m \in S$ y S es cerrado bajo multiplicación escalar se tiene que

$$c_1\vec{v}_1, c_2\vec{v}_2, \ldots, c_m\vec{v}_m \in S.$$

Esto último y el hecho de que S es cerrado bajo sumas implican que

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_m\vec{v}_m \in S,$$

por lo que $\vec{v} \in span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$. ■

Ahora determinaremos si algunos subconjuntos de $\mathbb{R}^n$ son o no subespacios. En vista de la proposición anterior, el primer candidato a examinar es el conjunto generado por un cierto número de vectores.

Proposición 10.4 *Si $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ son vectores de $\mathbb{R}^n$, entonces*

$$span\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p\}$$

es un subespacio de $\mathbb{R}^n$.

Demostración. Como

$$\vec{0}_n = 0\vec{v}_1 + 0\vec{v}_2 + \cdots + 0\vec{v}_p$$

se tiene que $\vec{0}_n \in span\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p\}$. Si $\vec{x}, \vec{y} \in span\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p\}$, entonces existen a_i, b_i, $i = 1, \ldots, p$, escalares, tales que

$$\begin{aligned} \vec{x} &= a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_p\vec{v}_p, \\ \vec{y} &= b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_p\vec{v}_p. \end{aligned}$$

Por lo tanto,

$$\vec{x} + \vec{y} = (a_1 + b_1)\,\vec{v}_1 + (a_2 + b_2)\,\vec{v}_2 + \cdots + (a_p + b_p)\,\vec{v}_p$$

y $\vec{x}+\vec{y} \in span\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p\}$. Finalmente, si $c \in \mathbb{R}$ y $\vec{x} \in span\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p\}$ es como antes, se tiene que

$$c\vec{x} = (ca_1)\,\vec{v}_1 + (ca_2)\,\vec{v}_2 + \cdots + (ca_p)\,\vec{v}_p$$

y $c\vec{x} \in span\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p\}$. Por lo tanto, $span\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p\}$ es un subespacio de $\mathbb{R}^n$. ■

Corolario 10.5 *$\{\vec{0}_n\}$ y $\mathbb{R}^n$ son subespacios de $\mathbb{R}^n$.*

Demostración. Como

$$\{\vec{0}_n\} = span\{\vec{0}_n\}$$

y

$$\mathbb{R}^n = span\{\vec{e}_1, \vec{e}_2, \ldots \vec{e}_n\},$$

se sigue el resultado. ■

Definición 10.6 *$\{\vec{0}_n\}$ y $\mathbb{R}^n$ son conocidos como los **subespacios triviales** de $\mathbb{R}^n$.*

Es claro que todos los espacios $\mathbb{R}^n$ contienen al menos a los subespacios triviales. ¿Qué otros subespacios contienen? En el caso de $n = 1$, es decir, $\mathbb{R}^1 = \mathbb{R}$, no existen subespacios no triviales; en efecto, si $S \subset \mathbb{R}$ y $S \neq \{\vec{0}_n\}$ es un subespacio distinto del cero, entonces existe un vector $\vec{v} \in S$, $\vec{v} \neq \vec{0}$ y por la Proposición 10.3

$$span\{\vec{v}\} \subset S.$$

Como $\{\vec{v}\}$ es un conjunto LI en $\mathbb{R}$, el Corolario 4.30 nos dice que

$$span\{\vec{v}\} = \mathbb{R},$$

de manera que $S = \mathbb{R}$.

Si $n = 2$ y S es un subespacio no trivial de $\mathbb{R}^2$, existe un vector $\vec{v} \neq \vec{0}$ tal que $\vec{v} \in S$. Más aún, cualquier otro elemento de S debe ser múltiplo de $\vec{v}$ porque en caso contrario, existiría $\vec{w} \in S$ tal que $\vec{w} \notin span\{\vec{v}\}$, con lo cual $\{\vec{v}, \vec{w}\}$ sería un conjunto LI y, por el Corolario 4.30 y la Proposición 10.3, tendríamos que

$$\mathbb{R}^2 = span\{\vec{v}, \vec{w}\} \subset S,$$

lo cual es imposible pues $S \subsetneq \mathbb{R}^2$. De aquí que si S es un subespacio no trivial de $\mathbb{R}^2$, S puede visualizarse en el plano como una recta que pasa por el origen.

De manera análoga, puede justificarse que si S es un subespacio no trivial de $\mathbb{R}^3$, entonces, o bien S es de la forma $span\{\vec{v}\}$ con $\vec{v} \neq \vec{0}$

(S se visualiza en el espacio como una recta que pasa por el origen) o $S = span\{\vec{v}, \vec{w}\}$ con $\{\vec{v}, \vec{w}\}$ un conjunto LI (S representa un plano que contiene al origen). S no puede poseer más de dos vectores linealmente independientes ya que si $\{\vec{u}, \vec{v}, \vec{w}\}$ fuese tal conjunto de vectores, entonces, una vez más por el Corolario 4.30 y la Proposición 10.3 tendríamos

$$\mathbb{R}^3 = span\{\vec{u}, \vec{v}, \vec{w}\} \subset S,$$

lo cual es imposible ya que $S \subsetneq \mathbb{R}^3$.

Generalizando los ejemplos anteriores, más adelante veremos que los subespacios no triviales de $\mathbb{R}^n$ pueden expresarse como

$$S = span\{v_1, v_2, \ldots, v_m\},$$

con $m < n$ y donde el conjunto $\{v_1, v_2, \ldots, v_m\}$ es LI.

Retomando el caso de $\mathbb{R}^3$ tenemos que la intersección de dos planos (distintos) que contienen al origen es una recta que también pasa por el origen. Así, dicha intersección también es un subespacio de $\mathbb{R}^3$. De forma más general, tenemos el siguiente resultado para la intersección de subespacios.

Proposición 10.7 *Si S_1 y S_2 son subespacios de $\mathbb{R}^n$, entonces $S_1 \cap S_2$ es subespacio de $\mathbb{R}^n$.*

Demostración. Verificamos, primero que $\vec{0}_n \in S_1 \cap S_2$, lo cual es obvio pues al ser S_1 y S_2 subespacios, $\vec{0}_n \in S_1$ y $\vec{0}_n \in S_2$. Veamos ahora que $S_1 \cap S_2$ es cerrado bajo sumas. Si $\vec{x}$ y $\vec{y}$ son elementos de $S_1 \cap S_2$, debemos verificar que

$$\vec{x} + \vec{y} \in S_1 \cap S_2.$$

Ahora bien, como

$$\vec{x}, \vec{y} \in S_i, \ i = 1, 2$$

y S_1, S_2 son cerrados bajo sumas se tiene que

$$\vec{x} + \vec{y} \in S_1, \ \vec{x} + \vec{y} \in S_2,$$

por lo que

$$\vec{x} + \vec{y} \in S_1 \cap S_2.$$

Finalmente, la demostración de que $S_1 \cap S_2$ es cerrado bajo multiplicación escalar se deja como ejercicio para el lector. ■

Concluimos esta sección con algunos ejemplos de subconjuntos que satisfacen dos de las propiedades que definen a un subespacio pero que no son subespacios.

Ejemplo 10.8 *Sean*

$$S_1 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x, y \geq 0 \right\} \text{ y } S_2 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x, y \leq 0 \right\}$$

y supongamos que deseamos determinar si los siguientes conjuntos son subespacios de $\mathbb{R}^2$: S_1, S_2 y $S_1 \cup S_2$. En el caso de S_1 es claro que satisface la primera y la segunda propiedad de subespacio. Sin embargo,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \in S_1 \text{ y } -2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \notin S_1,$$

por lo que S_1 no es subespacio de $\mathbb{R}^2$ (pues S_1 no es cerrado bajo multiplicación escalar). Análogamente puede verificarse que S_2 tampoco es subespacio de $\mathbb{R}^2$ (pues S_2 no es cerrado bajo multiplicación escalar). Finalmente, en el caso de $S_1 \cup S_2$ observamos que se satisfacen la primera y la tercera propiedad, pero

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \in S_1 \cup S_2, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \in S_1 \cup S_2, \text{ y } \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \notin S_1 \cup S_2,$$

por lo que $S_1 \cup S_2$ no es cerrado bajo sumas y, en consecuencia, no es subespacio de $\mathbb{R}^2$.

Ejemplo 10.9 *Consideremos los subespacios (es fácil verificar que lo son)*

$$S_1 = \left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} \mid x \in \mathbb{R} \right\} \text{ y } S_2 = \left\{ \begin{bmatrix} 0 \\ y \end{bmatrix} \mid y \in \mathbb{R} \right\}.$$

y supongamos que deseamos determinar si $S_1 \cup S_2$ también es subespacio de $\mathbb{R}^2$. Para ello, notamos que

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \in S_1 \cup S_2, \begin{bmatrix} 0 \\ \pi \end{bmatrix} \in S_1 \cup S_2, \text{ pero } \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \pi \end{bmatrix} = \begin{bmatrix} 1 \\ \pi \end{bmatrix} \notin S_1 \cup S_2,$$

por lo que $S_1 \cup S_2$ no es subespacio de $\mathbb{R}^2$ (pues $S_1 \cup S_2$ no es cerrado bajo sumas).

Espacio nulo y rango

Dada una transformación lineal,

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m,$$

se tienen dos subespacios, uno de $\mathbb{R}^n$ y otro de $\mathbb{R}^m$, naturalmente asociados a ella. Uno de ellos es

$$Rango(T) = \{T(\vec{x}) \mid \vec{x} \in \mathbb{R}^n\},$$

el cual puede reescribirse como

$$Rango(T) = span\{A^1, \ldots, A^n\}$$

en donde $A \in M_{m\times n}$ es la matriz estándar que representa a T. Así pues, la Proposición 10.4 nos dice que $Rango(T)$ es un subespacio de $\mathbb{R}^m$.

Ejemplo 10.10 *Sea* $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ *la transformación lineal dada por*

$$T\left(\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}\right) = \begin{bmatrix} x - w \\ w \end{bmatrix}$$

y supongamos que queremos encontrar un conjunto LI de $\mathbb{R}^2$ *cuyo span sea igual al rango de* T*. La matriz* A *que representa a* T *está dada por*

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

y como $Rango(T) = span\{A^1, A^2, A^3, A^4\}$*, es claro que el conjunto*

$$\mathcal{B}_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

es LI y genera a $Rango(T)$*.*

EL segundo subespacio que tenemos en mente también ya ha sido introducido previamente en el Capítulo 5. Recordemos que si

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m,$$

es una transformación lineal, entonces el **espacio nulo** se define como:

$$Null(T) = \{\vec{x} \in \mathbb{R}^n \mid T(\vec{x}) = \vec{0}_m\}.$$

Proposición 10.11 *Si* $T : \mathbb{R}^n \longrightarrow R^m$ *una transformación lineal, entonces* $Null(T)$ *es un subespacio de* $\mathbb{R}^n$*.*

Demostración. Como $T(\vec{0}_n) = \vec{0}_m$ se tiene que $\vec{0}_n \in Null(T)$. Ahora bien, si $\vec{x}, \vec{y} \in Null(T)$, entonces

$$T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y}) = \vec{0}_m + \vec{0}_m = \vec{0}_m,$$

por lo que

$$\vec{x} + \vec{y} \in Null(T).$$

Finalmente, dados $\vec{x} \in Null(T)$ y c un escalar, se tiene que

$$T(c\vec{x}) = cT(\vec{x}) = c\vec{0}_m = \vec{0}_m,$$

de forma que $c\vec{x} \in Null(T)$. Concluimos que $Null(T)$ cumple las tres propiedades requeridas de un subespacio. ■

Observación 10.12 *En algunos textos, al espacio nulo de T se le conoce como el **kernel*** [1] *o **núcleo** de T.*

Definición 10.13 *Sea $A \in M_{m\times n}$. Definimos el **espacio nulo** de A, denotado por $Null(A)$, como el conjunto*

$$Null(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}_m\},$$

es decir, $Null(A)$ es el conjunto de soluciones del sistema homogéneo.

Sean $A \in M_{m\times n}$ y $T : \mathbb{R}^n \to \mathbb{R}^m$ la transformación lineal dada por $T(\vec{x}) = A\vec{x}$. Entonces $Null(A) = Null(T)$ y se tiene una consecuencia inmediata de la Proposición 10.11.

Corolario 10.14 *Dada $A \in M_{m\times n}$, $Null(A)$ es un subespacio de $\mathbb{R}^n$.*

Ejemplo 10.15 *Sea*

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 5 \end{bmatrix}$$

y supongamos que deseamos expresar a $Null(A)$ como un cierto “span”. Resolvemos el sistema homogéneo

$$\begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

notando que la FER de la matriz aumentada del sistema está dada por

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right].$$

[1] La palabra “kernel” del idioma inglés se refiere al centro o corazón de una fruta o vegetal.

Por lo tanto, si $\vec{x} \in Null(A)$, sus componentes satisfacen

$$x_1 + 2x_3 = 0,\ x_2 + x_3 = 0,\ x_3 \in \mathbb{R}$$

y concluimos que

$$Null(A) = \left\{ x_3 \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \mid x_3 \in \mathbb{R} \right\} = span \left\{ \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

Ejemplo 10.16 *Sea $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ la transformación lineal del Ejemplo 10.10 dada por*

$$T\left(\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}\right) = \begin{bmatrix} x - w \\ w \end{bmatrix},$$

representada por la matriz estándar

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

Supongamos que queremos encontrar un conjunto LI de $\mathbb{R}^4$ cuyo span sea igual al espacio nulo de T.

Como $Null(T) = Null(A)$, resolvemos el sistema homogéneo

$$\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

La FER de la matriz aumentada $[A \mid \vec{0}]$ está dada por

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right],$$

de manera que $\vec{x} \in Null(A) = Null(T)$ syss sus componentes satisfacen

$$x_1 = 0,\ x_4 = 0,\ x_2 \in \mathbb{R},\ x_3 \in \mathbb{R}.$$

Concluimos que

$$\begin{aligned} Null(T) &= \left\{ x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \mid x_2, x_3 \in \mathbb{R} \right\} \\ &= span \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}. \end{aligned}$$

En los ejemplos anteriores se encontraron conjuntos LI que generaban a ciertos subespacios de $\mathbb{R}^2$ y de $\mathbb{R}^4$. Estos conjuntos generadores son de suma importancia en la teoría de espacios vectoriales y serán el motivo de estudio del siguiente capítulo.

Ejercicios

Ejercicio 10.1 *Determinar si* $\mathbb{Z}^2 = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 \mid a, b \in \mathbb{Z} \right\}$ *es un subespacio de* $\mathbb{R}^2$.

Ejercicio 10.2 *Determinar si cada uno de lo subconjuntos que se presentan a continuación son subespacios de* $\mathbb{R}^2$.

$$S_1 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x = 2y \right\}, \; S_2 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x^2 - y^2 = 0 \right\},$$
$$S_3 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid xy \leq 0 \right\}.$$

Ejercicio 10.3 *Determinar si cada uno de lo subconjuntos que se presentan a continuación son subespacios de* $\mathbb{R}^3$.

$$S_4 = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid 2x + y - z = 4 \right\}, \; S_5 = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid \pi x + 666y = 7z \right\}.$$

Ejercicio 10.4 *Determinar si cada uno de lo subconjuntos que se presentan a continuación son subespacios de* $\mathbb{R}^4$.

$$S_6 = \left\{ k \begin{bmatrix} 1 \\ -1 \\ -2 \\ 1 \end{bmatrix} \mid k \in \mathbb{R} \right\},$$
$$S_7 = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \mid a + 2b + 3c = 0 \;\&\; c + d = 0 \right\}.$$

Ejercicio 10.5 *Considerar los conjuntos de los Ejercicios 10.2,10.3 y 10.4 y encuentrar matrices* $A \in M_{1\times 2}$, $B \in M_{1\times 3}$, $C \in M_{3\times 4}$ *y* $D \in M_{2\times 4}$ *tales que* $Null(A) = S_1$, $Null(B) = S_5$, $Null(C) = S_6$ *y* $Null(D) = S_7$. *¿Existe alguna matriz cuyo espacio nulo sea igual al conjunto* S_2*?.*

Ejercicio 10.6 *Encontrar el espacio nulo de las matrices que se presentan a continuación.*

$$E = \begin{bmatrix} 3 & 2 \\ \pi & 1 \end{bmatrix}, \; F = \begin{bmatrix} 2 & 8 \\ \frac{3}{2} & 6 \end{bmatrix}, \; G = \begin{bmatrix} 2 & 8 & 6 \\ \frac{3}{2} & 6 & \frac{9}{2} \end{bmatrix}.$$

Ejercicio 10.7 *Sean T y U las transformaciones lineales de $\mathbb{R}^3$ en $\mathbb{R}^3$ dadas por*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 3x - y + z \\ 2y + 6z \\ x \end{bmatrix} \text{ y } U\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + 2y \\ -2x + 3y + 7z \\ x + 5y + 3z \end{bmatrix}.$$

Encontrar $Null(T)$ y $Null(U)$.

Ejercicio 10.8 *Encontrar mil subespacios de $\mathbb{R}^2$ distintos entre sí.*

Ejercicio 10.9 *Determinar todos los subespacios de $\mathbb{R}$ (donde $\mathbb{R}$ es el espacio vectorial $\mathbb{R}^1$).*

Ejercicio 10.10 *Pruebe que si S es un subespacio de $\mathbb{R}^n$ y $S \neq \{\vec{0}_n\}$, entonces S tiene un número infinito de elementos.*

Ejercicio 10.11 *Sean W_1 y W_2 subespacios de $\mathbb{R}^n$ y considerar al conjunto $W_1 + W_2$ definido como*

$$W_1 + W_2 = \{\vec{u} + \vec{z} \mid \vec{u} \in W_1, \vec{z} \in W_2\}.$$

Probar que $W_1 \subset W_1 + W_2$, $W_2 \subset W_1 + W_2$ y que $W_1 + W_2$ es un subespacio de $\mathbb{R}^n$

Ejercicio 10.12 *Calcular $W_1 + W_2$, donde W_1 y W_2 son los subespacios de $\mathbb{R}^3$ que se presentan a continuación.*

$$W_1 = \left\{ \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \mid x \in \mathbb{R} \right\} \text{ y } W_2 = \left\{ \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} \mid y \in \mathbb{R} \right\}.$$

Ejercicio 10.13 *Calcular $W_1 + W_2$, donde W_1 y W_2 son los subespacios de $\mathbb{R}^3$ que se presentan a continuación.*

$$W_1 = \left\{ \begin{bmatrix} k \\ k \\ 0 \end{bmatrix} \mid k \in \mathbb{R} \right\} \text{ y } W_2 = \left\{ \begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix} \mid t \in \mathbb{R} \right\}.$$

Ejercicio 10.14 *Probar (en general) o refutar (con un contraejemplo) los siguientes enunciados (en cada caso debe indicarse claramente "verdadero" o "falso").*

a. *Sean S_1 y S_2 subconjuntos no vacíos de $\mathbb{R}^2$. Si $S_1 \cup S_2$ es subespacio de $\mathbb{R}^2$, entonces S_1 es subespacio de $\mathbb{R}^2$ o S_2 es subespacio de $\mathbb{R}^2$.*

b. *Si* $n \geq 2$, *entonces* $\mathbb{R}^n$ *tiene un número infinito de subespacios distintos entre sí.*

c. *Si* S *es un subespacio no trivial de* $\mathbb{R}^n$ *y* $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ *son* n *elementos de* S *distintos entre sí, entonces* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ *es un conjunto LD.*

d. *Sean* $\vec{v}_1$ *y* $\vec{v}_2$ *vectores de* $\mathbb{R}^n$ *y sea* S *un subespacio de* $\mathbb{R}^n$. *Si* $\vec{v}_1 + \vec{v}_2 \in S$, *entonces* $\vec{v}_1 \in S$ *o* $\vec{v}_2 \in S$.

e. *Sean* $\vec{v}_1$ *y* $\vec{v}_2$ *vectores de* $\mathbb{R}^n$ *y sea* S *un subespacio de* $\mathbb{R}^n$. *Si* $\vec{v}_1 + \vec{v}_2 \in S$ *y* $\vec{v}_2 \in S$, *entonces* $\vec{v}_1 \in S$.

f. *Sea* A *una matriz de* $m \times n$, *sea* $\vec{b}$ *un vector de* $\mathbb{R}^m$ *y sea* S *el conjunto solución de la ecuación matricial* $A\vec{x} = \vec{b}$. *Si* $\vec{b}$ *es distinto del vector cero de* $\mathbb{R}^m$, *entonces* S *no es subespacio de de* $\mathbb{R}^n$.

Ejercicio 10.15 *Dados*

$$M = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + y - z = 0 \right\} \subset \mathbb{R}^3,$$
$$N = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + y + z = 0 \right\} \subset \mathbb{R}^3,$$

demostrar que M *y* N *son subespacios y encontrar un vector* $\vec{v} \in \mathbb{R}^3$ *tal que* $M \cap N = span\{\vec{v}\}$.

Ejercicio 10.16 *Sean* V, W *subespacios de* $\mathbb{R}^n$. *Probar que* $V \cap W = V$ *syss* $V + W = W$.

Capítulo 11

Bases y dimensión

Introducción

Dado cualquier subespacio $S \subset \mathbb{R}^n$, decimos que el conjunto

$$\{\vec{v}_1, \ldots, \vec{v}_p\}$$

genera a S si

$$S = span\{\vec{v}_1, \ldots, \vec{v}_p\}.^1$$

Ahora bien, si $\{\vec{v}_1, \ldots, \vec{v}_p\}$ fuese LD, la intuición nos dice que deberían poder extraerse algunos vectores de manera que el conjunto resultante, $\{\vec{v}_{i_1}, \ldots, \vec{v}_{i_m}\}$, fuese LI y

$$S = span\{\vec{v}_{i_1}, \ldots, \vec{v}_{i_m}\}, \quad m < p.$$

En otras palabras, se eliminan los vectores "redundantes" para generar a S de manera más eficiente. Estos conjuntos linealmente independientes de generadores son de suma importancia en el estudio de los espacios vectoriales. Anteriormente ya hemos introducido conjuntos como éstos, concretamente, el conjunto

$$\{\vec{e}_1, \ldots, \vec{e}_n\},$$

al que denominamos la base estándar de $\mathbb{R}^n$ es un conjunto LI que genera a $\mathbb{R}^n$, es decir,

$$span\{\vec{e}_1, \ldots, \vec{e}_n\} = \mathbb{R}^n.$$

[1] Posteriormente, veremos que todos los subespacios son de esta forma.

Este tipo de conjuntos serán el objeto de estudio de este capítulo.

Bases

Definición 11.1 *Sea S un subespacio de $\mathbb{R}^n$ y sean $\vec{v}_1, \ldots, \vec{v}_p$ vectores de S. Decimos que el conjunto $\{\vec{v}_1, \ldots, \vec{v}_p\}$ es una* ***base*** *para S syss se cumplen las siguientes dos condiciones.*

1. *$span\{\vec{v}_1, \ldots, \vec{v}_p\} = S$.*

2. *$\{\vec{v}_1, \ldots, \vec{v}_p\}$ es LI.*

Sabemos (por el Corolario 4.30) que si un conjunto de vectores de $\mathbb{R}^n$ es LI y tiene exactamente n elementos, entonces ese conjunto genera a todo $\mathbb{R}^n$. Por lo tanto, el siguiente resultado se sigue en forma inmediata.

Proposición 11.2 *Sean $\vec{v}_1, \ldots, \vec{v}_n$ vectores de $\mathbb{R}^n$. Si $\{\vec{v}_1, \ldots, \vec{v}_n\}$ es LI, entonces $\{\vec{v}_1, \ldots, \vec{v}_n\}$ es una base para $\mathbb{R}^n$. En particular, el conjunto $\{\vec{e}_1, \ldots, \vec{e}_n\} \subset \mathbb{R}^n$, conocido como la base estándar de $\mathbb{R}^n$, es efectivamente una base para $\mathbb{R}^n$ en el sentido de la Definición 11.1.*

Ejemplo 11.3 *Los conjuntos*

$$\mathcal{B}_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \text{ y } \mathcal{B}_2 = \left\{ \begin{bmatrix} 1 \\ 9 \end{bmatrix}, \begin{bmatrix} \pi \\ -2 \end{bmatrix} \right\}$$

son bases para $\mathbb{R}^2$. Los conjuntos

$$\mathcal{C}_1 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right\} \text{ y } \mathcal{C}_2 = \left\{ \begin{bmatrix} 5 \\ 0 \end{bmatrix} \right\}$$

no son bases para $\mathbb{R}^2$ ya que $\mathcal{C}_1$ es LD y

$$span\ \mathcal{C}_2 = \left\{ \begin{bmatrix} a \\ 0 \end{bmatrix} \mid a \in \mathbb{R} \right\} \subsetneq \mathbb{R}^2.$$

Sin embargo, $\mathcal{C}_2$ es una base para el subespacio

$$\left\{ \begin{bmatrix} a \\ 0 \end{bmatrix} \mid a \in \mathbb{R} \right\}.$$

Ejemplo 11.4 *Los conjuntos*

$$\mathcal{B}_3 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \quad y \quad \mathcal{B}_4 = \left\{ \begin{bmatrix} 1 \\ 9 \\ 0 \end{bmatrix}, \begin{bmatrix} \pi \\ -2 \\ 0 \end{bmatrix} \right\}$$

son bases para el subespacio

$$\left\{ \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} \mid a, b \in \mathbb{R} \right\}.$$

Ejemplo 11.5 *Sea*

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

y supongamos que deseamos encontrar una base para

$$Null(A) = \{\vec{x} \mid A\vec{x} = \vec{0}\}.$$

Para ello, basta notar que la FER de la matriz aumentada $\left[A \mid \vec{0}\right]$ *es igual a la matriz*

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Por lo tanto, si $\vec{x} \in Null(A)$*, entonces sus componentes satisfacen*

$$x_1 = -2x_2 - x_4,\ x_3 = x_4,\ x_2 \in \mathbb{R},\ x_4 \in \mathbb{R}$$

y concluimos que

$$Null(A) = \left\{ \begin{bmatrix} -2x_2 - x_4 \\ x_2 \\ x_4 \\ x_4 \end{bmatrix} \mid x_2, x_4 \in \mathbb{R} \right\}$$

$$= \left\{ x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \mid x_3, x_4 \in \mathbb{R} \right\} = span \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Además, el conjunto

$$\mathcal{B} = \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

es LI pues cada uno de los vectores tiene un uno en la posición de la variable libre que lo ha inducido, mientras que el otro tiene un cero en esa misma posición. Por lo tanto, $\mathcal{B}$ es una base para $Null(A)$.

Por supuesto que si S es un subespacio, $\{\vec{v}_1, \ldots, \vec{v}_p\}$ es una base para S y $\vec{v} \in S$, entonces $\vec{v}$ puede escribirse como combinación lineal de $\vec{v}_1, \ldots, \vec{v}_p$. El siguiente resultado dice que, salvo por el orden de los términos, existe una **única** manera de expresar a este vector $\vec{v}$ como combinación lineal de $\vec{v}_1, \ldots, \vec{v}_p$.

Proposición 11.6 *Sean S un subespacio de $\mathbb{R}^n$, $\{\vec{v}_1, \ldots, \vec{v}_p\}$ una base para S y $\vec{v} \in S$. Si $c_1, \ldots, c_p$ y $d_1, \ldots, d_p$ son escalares tales que*

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_p\vec{v}_p \quad y \quad \vec{v} = d_1\vec{v}_1 + \cdots + d_p\vec{v}_p,$$

entonces $c_j = d_j$, para toda j.

Demostración. Por hipótesis,

$$\begin{aligned}\vec{0} &= \vec{v} - \vec{v} = c_1\vec{v}_1 + \cdots + c_p\vec{v}_p - (d_1\vec{v}_1 + \cdots + d_p\vec{v}_p)\\ &= (c_1 - d_1)\vec{v}_1 + \cdots + (c_p - d_p)\vec{v}_p\end{aligned}$$

Ahora bien, como $\{\vec{v}_1, \ldots, \vec{v}_p\}$ es LI,

$$c_1 - d_1 = \cdots = c_p - d_p = 0.$$

Por lo tanto, $c_1 = d_1, \ldots, c_p = d_p$. ■

Dimensión

A continuación, probaremos que cualesquiera dos bases de un subespacio tienen exactamente el mismo número de elementos. Este resultado es esencial para definir el concepto de dimensión de un espacio vectorial.

Teorema 11.7 *Si S es un subespacio de $\mathbb{R}^n$ tal que $span\{\vec{v}_1, \ldots, \vec{v}_k\} = S$ y $\{\vec{w}_1, \ldots, \vec{w}_k, \vec{w}_{k+1}\}$ es un subconjunto de S con $k+1$ elementos, entonces $\{\vec{w}_1, \ldots, \vec{w}_k, \vec{w}_{k+1}\}$ es LD.*

Demostración. Sea $\{\vec{w}_1, \ldots, \vec{w}_k, \vec{w}_{k+1}\}$ cualquier subconjunto de S con $k+1$ elementos. Como $S = span\{\vec{v}_1, \ldots, \vec{v}_k\}$ se tiene que para toda $i = 1, 2, \ldots, k+1$,

$$\vec{w}_i \in span\{\vec{v}_1, \ldots, \vec{v}_k\}.$$

En particular, si definimos a la matriz $A \in M_{n\times k}$ como la matriz cuyas columnas son los vectores $\vec{v}_j$, es decir

$$A = [\vec{v}_1\ \vec{v}_2 \cdots \vec{v}_k],$$

se tiene que existen $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{k+1} \in \mathbb{R}^k$ tales que

$$A\vec{x}_i = \vec{w}_i,\ i = 1, 2, \ldots, k+1. \tag{$\bigstar$}$$

Ahora bien, definiendo las matrices B y C como

$$\begin{aligned} B &= [\vec{x}_1\ \vec{x}_2 \cdots \vec{x}_{k+1}] \in M_{k\times(k+1)}, \\ C &= [\vec{w}_1\ \vec{w}_2 \cdots \vec{w}_{k+1}] \in M_{n\times(k+1)}, \end{aligned}$$

se tiene que, por la Proposición 6.21 y las igualdades dadas en ($\bigstar$), se cumple, para toda $j = 1, 2, \ldots, k+1$

$$(AB)^j = AB^j = A\vec{x}_j = \vec{w}_j = C^j,$$

por lo tanto

$$AB = C. \tag{$\spadesuit$}$$

Observemos que las columnas de B forman un conjunto LD ya que $\{\vec{x}_1,\ \vec{x}_2, \cdots, \vec{x}_{k+1}\}$ es un conjunto de $k+1$ vectores en $\mathbb{R}^k$. Entonces, existe $\vec{x} \in \mathbb{R}^{k+1}$, $\vec{x} \neq \vec{0}_{k+1}$, tal que se cumple

$$B\vec{x} = \vec{0}_k.$$

En virtud de ($\spadesuit$) se tiene que

$$C\vec{x} = AB\vec{x} = \vec{0}_n$$

y como $\vec{x} \neq \vec{0}_{k+1}$ concluimos que el conjunto de columnas de la matriz C, es decir $\{\vec{w}_1, \ldots, \vec{w}_k, \vec{w}_{k+1}\}$, es LD. ■

Corolario 11.8 *Si S es un subespacio de $\mathbb{R}^n$ y $\{\vec{v}_1, \ldots, \vec{v}_k\}$, $\{\vec{w}_1, \ldots, \vec{w}_l\}$ son bases para S, entonces $k = l$. Este número, denotado por $dim(S)$, será llamado la* ***dimensión*** *de S.*

Demostración. Como $span\{\vec{v}_1, \ldots, \vec{v}_k\} = S$ y $\{\vec{w}_1, \ldots, \vec{w}_l\}$ es LI tenemos, por el Teorema 11.7, que $l \leq k$ (de otra forma $\{\vec{w}_1, \ldots, \vec{w}_l\}$ sería LD). De igual manera, como $span\{\vec{w}_1, \ldots, \vec{w}_l\} = S$ y $\{\vec{v}_1, \ldots, \vec{v}_k\}$ es LI tenemos, por el Teorema 11.7, que $k \leq l$, por lo que podemos concluir que $k = l$. ■

Observación 11.9 *La dimensión nos dice cuántos elementos son necesarios y suficientes para generar un espacio vectorial. Esto puede interpretarse como una medida de qué tan grande es este espacio.*

Ejemplo 11.10 *Como la base estándar de* $\mathbb{R}^n$ *(es decir, el conjunto* $\{\vec{e}_1, \ldots \vec{e}_n\}$*) es una base para* $\mathbb{R}^n$*, es claro que* $dim(\mathbb{R}^n) = n$.

Ejemplo 11.11 *Como el conjunto*

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

es una base para el subespacio

$$S = \left\{ \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} \mid a, b \in \mathbb{R} \right\},$$

se tiene que $dim(S) = 2$.

Ejemplo 11.12 *Como el conjunto*

$$\left\{ \begin{bmatrix} 5 \\ 0 \end{bmatrix} \right\}$$

es una base para el subespacio

$$S = \left\{ \begin{bmatrix} a \\ 0 \end{bmatrix} \mid a \in \mathbb{R} \right\},$$

entonces $dim(S) = 1$.

A la luz de los tres ejemplos anteriores y por motivos de conveniencia, resulta natural asignarle una dimensión nula al espacio vectorial cuyo único elemento es el vector cero.

Definición 11.13 *Si* $\vec{0}_n$ *denota al vector cero de* $\mathbb{R}^n$*, entonces*

$$\dim\left(\left\{\vec{0}_n\right\}\right) = 0.$$

Observación 11.14 *Al asignarle dimensión nula al subespacio* $S = \left\{\vec{0}_n\right\}$ *estamos diciendo, literalmente, que cualquier base de* S *tiene cero elementos, por lo tanto, la única base posible es el conjunto vacío:* $\varnothing$.

Observación 11.15 *Consideremos ahora el conjunto*

$$\mathcal{D} = \left\{ \begin{bmatrix} 7 \\ 8 \\ 9 \\ 6 \end{bmatrix}, \begin{bmatrix} -9 \\ 7 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 23 \\ 22 \\ 13 \end{bmatrix} \right\}$$

y supongamos que deseamos probar que $span\mathcal{D} \subsetneq \mathbb{R}^4$*. Para ello, basta notar que* $dim(span\mathcal{D}) \leq 3$ *(de hecho,* $dim(span\mathcal{D}) = 2$ *porque el tercer elemento de* $\mathcal{D}$ *es combinación lineal de los otros dos) mientras que* $dim(\mathbb{R}^4) = 4$*. Por lo tanto,* $span\mathcal{D} \subsetneq \mathbb{R}^4$.

La siguiente proposición es una generalización de la observación anterior.

Proposición 11.16 (El famoso pecado por defecto) *Sea S un subespacio de $\mathbb{R}^n$. Si $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ son vectores de S con $p < \dim(S)$, entonces*

$$span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} \subsetneq S.$$

Demostración. Como $dim(span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}) \leq p < n = dim(\mathbb{R}^n)$, es claro que $span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} \subsetneq \mathbb{R}^n$. ■

Ahora justificaremos que todo subespacio efectivamente tiene una base.

Proposición 11.17 *Si $S \subset \mathbb{R}^n$ es un subespacio, $S \neq \{\vec{0}\}$ y p es el máximo número natural tal que existen $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p \in S$ de forma que el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es LI; entonces $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es una base para S.*

Demostración. La demostración se llevará a cabo por contradicción. Primero observamos que $1 \leq p$ (pues S tiene al menos un vector distinto del vector cero) y que $p \leq n$ (pues cualquier subconjunto LI de $\mathbb{R}^n$ tiene a lo más n elementos y porque $S \subset \mathbb{R}^n$). Ahora bien, si $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} \subset S$ no fuese una base para S, entonces

$$span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} \subsetneq S$$

y existiría un vector $\vec{v} \in S$ tal que

$$\vec{v} \notin span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}.$$

Por lo tanto, $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p, \vec{v}\}$ sería un subconjunto linealmente independiende de S con $p + 1$ elementos, lo cual contradice la supuesta maximalidad de p. ■

Corolario 11.18 *Si $S \subset \mathbb{R}^n$ es un subespacio con $dim(S) = p > 0$ y $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} \subset S$ es un conjunto linealmente independiente con p elementos, entonces $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es una base para S.*

Demostración. Gracias al Teorema 11.7 sabemos cualquier subconjunto de S de tamaño $p + 1$ es LD (pues S es generado por cualquiera de sus bases y todas ellas son de tamaño p). De esta forma, p y $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ son como en las hipótesis de la Proposición 11.17 y por lo tanto, $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es una base para S. ■

Ejemplo 11.19 *Sea S el subespacio de $\mathbb{R}^4$ que consiste de todos aquellos vectores cuya cuarta componente es igual a cero y supongamos que deseamos mostrar que el conjunto $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ es una base para S, donde*

$$\vec{w}_1 = \begin{bmatrix} 1 \\ \pi \\ 7 \\ 0 \end{bmatrix}, \vec{w}_2 = \begin{bmatrix} -1 \\ 2 \\ -1 \\ 0 \end{bmatrix}, \vec{w}_3 = \begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix}.$$

Para ello, y en vista del corolario anterior, basta notar que $dim(S) = 3$ (pues $\{\vec{e}_1, \vec{e}_2, \vec{e}_3)$ es una base para S), que w_1, w_2 y w_3 son elementos de S y que el conjunto $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ es un conjunto LI.

Finalizamos esta sección con un resultado bastante intuitivo.

Proposición 11.20 *Si S_1 y S_2 son subespacios de $\mathbb{R}^n$ tales que $S_1 \subset S_2$, entonces $dim(S_1) \leq dim(S_2)$. En caso de que los subespacios S_1 y S_2 sean tales que $S_1 \subset S_2$ y $dim(S_1) = dim(S_2)$, entonces $S_1 = S_2$.*

Demostración. Sean $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ y $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_q\}$ bases para S_1 y S_2, respectivamente, mismas que existen por la Proposición 11.17. Como $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es un conjunto LI cuyos elementos están en S_1 (y por tanto, en S_2) y como $S_2 = span\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_q\}$, el Teorema 11.7 asegura que $p \leq q$, es decir, $dim(S_1) \leq dim(S_2)$.

Supongamos ahora que $S_1 \subset S_2$ y que $dim(S_1) = dim(S_2)$. Una vez más tenemos que $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es un conjunto LI cuyos elementos están en S_1 (y por tanto, en S_2). Como el tamaño de este conjunto es igual a la dimensión de S_2, el Corolario 11.18 asegura que $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ también es base de S_2. Por lo tanto,

$$S_1 = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} = S_2,$$

quedando demostrado el resultado. ■

Minería (extracción) de bases

En el caso de la Observación 11.15, el lector minucioso seguramente habrá observado que, a pesar de que $\mathcal{D}$ tiene 3 elementos y $dim(span\mathcal{D}) = 2$, es posible encontrar una base para $span\mathcal{D}$ cuyos vectores son elementos de $\mathcal{D}$. De hecho, gracias a la advertencia de que "el tercer elemento de $\mathcal{D}$ es combinación lineal de los otros dos", es claro que una base para dicho subespacio es el conjunto

$$\left\{ \begin{bmatrix} 7 \\ 8 \\ 9 \\ 6 \end{bmatrix}, \begin{bmatrix} -9 \\ 7 \\ 4 \\ 1 \end{bmatrix} \right\}.$$

De forma más general, supongamos que $S = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$. Si $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es LI, por definición se tiene que $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es una base para S. En caso contrario, alguno de los elementos de este conjunto es combinación lineal de los otros $p-1$ elementos. Sin pérdida de generalidad, podemos suponer que $v_1 \in span\{\vec{v}_2, \ldots, \vec{v}_p\}$ (de no ser así, simplemente se reordenan los elementos). Entonces, $S = span\{\vec{v}_2, \ldots, \vec{v}_p\}$ y tenemos dos posibilidades:

1. $\{\vec{v}_2, \ldots, \vec{v}_p\}$ es LI y por lo tanto, este conjunto con $p-1$ elementos constituye una base para S, o bien

2. $\{\vec{v}_2, \ldots, \vec{v}_p\}$ es LD y por lo tanto, es posible seguir extrayendo un número finito elementos de este conjunto hasta obtener un subconjunto LI que genere a todo S (i.e., una base para S).

El siguiente teorema es una consecuencia inmediata del razonamiento anterior.

Teorema 11.21 *Si $S = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$, entonces existe un subconjunto de $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ que es base para S y por lo tanto, $dim(S) \leq p$.*

El teorema anterior nos dice que dado un conjunto de generadores para un subespacio S, es posible extraer un subconjunto LI que también genera a S. Sin embargo, es conveniente ir un paso más allá y disponer de métodos efectivos que nos permitan encontrar estas bases. Para este propósito se requiere de los siguientes resultados preliminares.

Proposición 11.22 *Sea*

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} \in M_{m \times n}$$

y sea

$$\tilde{A} = \begin{bmatrix} \tilde{A}_1 \\ \tilde{A}_2 \\ \vdots \\ \tilde{A}_m \end{bmatrix}$$

su FER. Entonces $span\{A_1, A_2, \ldots, A_m\} = span\{\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_m\}$, *es decir, los renglones de ambas matrices generan el mismo subespacio en* $\mathbb{R}^n$.

Demostración. Todos los renglones de $\tilde{A}$ se obtienen de los renglones de A mediante operaciones elementales sobre los mismos. Concretamente, dado cualquier renglón $\tilde{A}_i$ de $\tilde{A}$, necesariamente este renglón será una combinación lineal de renglones de A y por lo tanto,

$$span\{\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_m\} \subset span\{A_1, A_2, \ldots, A_m\}.$$

Ahora bien, cada paso utilizado para obtener la FER de A es reversible, simplemente utilizando la operación elemental inversa (que es del mismo tipo que la operación original). De aquí que cualquier renglón A_i de A será una combinación lineal de renglones de $\tilde{A}$, de manera que

$$span\{A_1, A_2, \ldots, A_m\} \subset span\{\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_m\}$$

y se tiene la igualdad

$$span\{A_1, A_2, \ldots, A_m\} = span\{\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_m\}$$

con lo cual se concluye la demostración. ■

El conjunto de los renglones de la matriz $\tilde{A}$ puede dividirse en dos subconjuntos: el subconjunto de aquellos renglones cuyas entradas son todas cero y el subconjunto de aquellos renglones que tienen un uno principal. Sin embargo, es claro que el segundo subconjunto es LI y genera a todos los elementos del primero. Por ello, es más preciso decir que el conjunto generado por los renglones de A es igual al conjunto generado por los renglones de $\tilde{A}$ que tienen un uno principal y que tal conjunto es LI. En resumen, el conjunto generado por los renglones de A tiene como base al conjunto formado por los renglones de $\tilde{A}$ que tienen un uno principal. El siguiente corolario es inmediato a partir de esta observación.

Corolario 11.23 *Consideremos la matriz de* $m \times n$ *dada por*

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}$$

y sea

$$\tilde{A} = \begin{bmatrix} \tilde{A}_1 \\ \tilde{A}_2 \\ \vdots \\ \tilde{A}_m \end{bmatrix}$$

su FER. Si $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_l$ *con* $l \leq m$*, son los renglones de* $\tilde{A}$ *que tienen unos principales, entonces* $\{\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_l\}$ *es LI y*

$$span\{A_1, A_2, \ldots, A_m\} = span\{\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_l\}.$$

Ejemplo 11.24 *Si*

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & -2 & 5 \\ -4 & 4 & -9 \end{bmatrix},$$

entonces su FER se obtiene haciendo los siguientes pasos (cada uno de ellos concentra dos operaciones):

$$\begin{bmatrix} 1 & -1 & 2 \\ 2 & -2 & 5 \\ -4 & 4 & -9 \end{bmatrix} \overset{\substack{-2R_1+R_2 \\ 4R_1+R_3}}{\sim} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix} \overset{\substack{R_2+R_3 \\ -2R_2+R_1}}{\sim} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \tilde{A},$$

Dado que

$$\begin{aligned} \tilde{A}_1 &= A_1 - 2\,(-2A_1 + A_2) = 5A_1 - 2A_2, \\ \tilde{A}_2 &= -2A_1 + A_2, \\ \tilde{A}_3 &= (-2A_1 + A_2) + (4A_1 + A_3) = 2A_1 + A_2 + A_3, \end{aligned}$$

entonces es claro que $span\{\tilde{A}_1, \tilde{A}_2, \tilde{A}_3\} \subset span\{A_1, A_2, A_3\}$ *. Por otro lado, utilizando las operaciones elementales inversas se obtiene*

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \overset{\substack{-R_2+R_3 \\ 2R_2+R_1}}{\sim} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix} \overset{\substack{2R_1+R_2 \\ -4R_1+R_3}}{\sim} \begin{bmatrix} 1 & -1 & 2 \\ 2 & -2 & 5 \\ -4 & 4 & -9 \end{bmatrix} = A,$$

por lo que también es cierto que

$$span\{A_1, A_2, A_3\} \subset span\{\tilde{A}_1, \tilde{A}_2, \tilde{A}_3\}$$

y por lo tanto,

$$span\{A_1, A_2, A_3\} = span\{\tilde{A}_1, \tilde{A}_2, \tilde{A}_3\}.$$

Más aún, los renglones de $\tilde{A}$ *que tienen unos principales son* $\tilde{A}_1$ *y* $\tilde{A}_2$ *de manera que* $\{\tilde{A}_1, \tilde{A}_2\}$ *es LI y*

$$span\{A_1, A_2, A_3\} = span\{\tilde{A}_1, \tilde{A}_2\}.$$

A continuación presentaremos dos algoritmos muy efectivos para encontrar bases de subespacios S de la forma $S = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$. El primero de ellos se limita a encontrar una base (aunque la base no tenga ninguna relación con el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$) mientras que el segundo, efectivamente extrae una base del conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ en el sentido del Teorema 11.21.

ALGORITMO 1. El primer algoritmo está basado en el Corolario 11.23. Así, si S es un subespacio de $\mathbb{R}^n$ de la forma

$$S = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$$

y queremos encontrar una base para S, consideramos primero la matriz $A \in M_{m\times n}$ tal que,[2]

$$A_1 = \vec{v}_1,\ A_2 = \vec{v}_2, \ldots, A_m = \vec{v}_m.$$

Después, encontramos la FER de A; supongamos que $\tilde{A}$ es dicha FER. Finalmente, notamos que

$$S = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\} = span\{\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_l\}$$

en donde $\{\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_l\}$ es el conjunto formado por los renglones de $\tilde{A}$ que tienen un uno principal. Dado que este conjunto es LI, concluimos que $\{\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_l\}$ es una base para S.

Ejemplo 11.25 *Sea*

$$S = Span\left\{\begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 5 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -10 \end{bmatrix}\right\}$$

y supongamos que deseamos encontrar una base para S. Para utilizar el Algoritmo 1, basta notar que la FER de la matriz

$$\begin{bmatrix} -1 & -1 & 0 \\ -1 & -1 & 5 \\ -1 & -1 & -10 \end{bmatrix},$$

cuyos renglones son estos vectores, es la matriz

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

por lo que el conjunto (de renglones con unos principales)

$$\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$$

es una base para S.

[2] Al igual que en la demostración de la Proposición 6.46, aquí estamos pensando a los vectores $\vec{v} \in S$ como vectores renglón y no como vectores columna.

El algoritmo que acabamos de utilizar no es útil cuando queremos obtener una base a partir de un conjunto de generadores, constituida por elementos de dicho conjunto. Para lograr este objetivo analizamos un nuevo enfoque basado en el hecho de que las columnas de una matriz y las columnas de su FER tienen, estructuralmente, las mismas relaciones de dependencia/independencia lineal.

Recordemos que si $A = \begin{bmatrix} A^1 & A^2 & \cdots & A^n \end{bmatrix} \in M_{m\times n}$ y $\tilde{A} = \begin{bmatrix} \tilde{A}^1 & \tilde{A}^2 & \cdots & \tilde{A}^n \end{bmatrix}$ es su FER, entonces $Null(A) = Null(\tilde{A})$, es decir

$$A\vec{x} = \vec{0} \ \text{ syss } \ \tilde{A}\vec{x} = \vec{0}.$$

Equivalentemente, tomando $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$,

$$x_1A^1+x_2A^2+\cdots+x_nA^n = \vec{0} \quad \text{syss} \quad x_1\tilde{A}^1+x_2\tilde{A}^2+\cdots+x_n\tilde{A}^n = \vec{0}. \ (\bigstar)$$

La equivalencia dada por $(\bigstar)$ nos dice que la relación de dependencia lineal entre los vectores columna de A y los de $\tilde{A}$ es exactamente la misma. Por lo tanto, un subconjunto de columnas de A es LI syss el subconjunto de columnas correspondientes de $\tilde{A}$ es LI.

Ejemplo 11.26 *Sea*

$$A = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 6 & 2 \\ 1 & 2 & 4 \end{bmatrix}$$

cuya FER está dada por

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

Por lo tanto,

$$Null(A) = Null(\tilde{A}) = \left\{ \begin{bmatrix} -2x_3 \\ -x_3 \\ x_3 \end{bmatrix} \mid x_3 \in \mathbb{R} \right\} \subset \mathbb{R}^3.$$

Como

$$\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} \in Null(A) = Null(\tilde{A}),$$

se cumple que

$$2\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \begin{bmatrix} -2 \\ 6 \\ 2 \end{bmatrix} - \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$
$$2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

de manera que

$$2\underbrace{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}}_{A^1} + \underbrace{\begin{bmatrix} -2 \\ 6 \\ 2 \end{bmatrix}}_{A^2} = \underbrace{\begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}}_{A^3},$$
$$2\underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{\tilde{A}^1} + \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\tilde{A}^2} = \underbrace{\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}}_{\tilde{A}^3}.$$

Asimismo, se observa que los conjuntos $\{A^1, A^2\}$ *y* $\left\{\tilde{A}^1, \tilde{A}^2\right\}$ *son LI.*

Supongamos que $A \in M_{m\times n}$ es una matriz que no tiene columnas nulas y tal que su FER $\tilde{A}$ tiene exactamente k unos principales, $k \leq n$. Sean $\tilde{A}^{j_1}, \tilde{A}^{j_2}, \ldots, \tilde{A}^{j_k}$ las columnas que contienen los unos principales. Por construcción, para $l = 1, 2, \ldots k$, cada una de estas columnas es de la forma

$$\tilde{A}^{j_l} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \vec{e}_l,$$

en donde el "1" se encuentra en la posición l. Las $n-k$ columnas que no poseen unos principales se denotan por $\tilde{A}^{j_{(k+1)}}, \tilde{A}^{j_{(k+2)}}, \ldots, \tilde{A}^{j_{(n)}}$. Dada cualquiera de estas columnas, digamos $\tilde{A}^{j_{(k+i)}}$ con $i = 1, \ldots, n-k$, tomamos el conjunto $\{\tilde{A}^{j_1}, \tilde{A}^{j_2}, \ldots, \tilde{A}^{j_{p_i}}\}$ consistente de las columnas con unos principales que se encuentran a la izquierda de $\tilde{A}^{j_{(k+i)}}$, es decir,

$$j_1 < j_2 < \cdots < j_{p_i} < j_{(k+i)}.$$

Entonces, tenemos que

$$\tilde{A}^{j_{(k+i)}} \in span\{\tilde{A}^{j_1}, \tilde{A}^{j_2}, \ldots, \tilde{A}^{j_{p_i}}\} = span\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_{p_i}\},$$

ya que de no ser así existiría un renglón A_q, en donde $q > p_i$, de la forma
$$A_q = [0 \;\; 0 \; \cdots \; 0 \; 1 \; * \cdots \;\; *],$$
en donde el uno principal está en la columna $\tilde{A}^{j_{(k+i)}}$. Esto es imposible ya que, por construcción, esta columna no contiene unos principales. Por lo tanto,
$$\tilde{A}^{j_{(k+i)}} \in span\{\tilde{A}^{j_1}, \tilde{A}^{j_2}, \ldots, \tilde{A}^{j_{p_i}}\}$$
para todo $i = 1, \ldots, n-k$. Como el conjunto $\left\{\tilde{A}^{j_1}, \tilde{A}^{j_2}, \ldots, \tilde{A}^{j_k}\right\}$ es LI, éste constituye una base para el conjunto generado por las columnas de $\tilde{A}$. Como vimos anteriormente, la relación de dependencia lineal entre los vectores columna de A y de $\tilde{A}$ es la misma (ver equivalencia ($\bigstar$)), por lo que el conjunto $\{A^{j_1}, A^{j_2}, \ldots, A^{j_k}\}$ de columnas de A que corresponden a las columnas con unos principales de $\tilde{A}$ constituye una base para el subespacio $span\{A^1, A^2, \ldots, A^n\}$.

Proposición 11.27 *Consideremos una matriz de $m \times n$ dada como*
$$A = \left[\; A^1 \quad A^2 \quad \cdots \quad A^n \;\right]$$
y
$$\tilde{A} = \left[\; \tilde{A}^1 \quad \tilde{A}^2 \quad \cdots \quad \tilde{A}^n \;\right]$$
su FER. Entonces si $\tilde{A}^{j_1}, \tilde{A}^{j_2}, \ldots, \tilde{A}^{j_k}$ son las columnas de $\tilde{A}$ que poseen unos principales, éstas forman un conjunto LI. Asimismo, el conjunto correspondiente $\{A^{j_1}, A^{j_2}, \ldots, A^{j_k}\}$ de columnas de A también es LI. Adicionalmente,
$$span\left\{\tilde{A}^{j_1}, \tilde{A}^{j_2}, \ldots, \tilde{A}^{j_k}\right\} = span\left\{\tilde{A}^1, \tilde{A}^2, \ldots, \tilde{A}^n\right\}$$
y
$$span\left\{A^{j_1}, A^{j_2}, \ldots, A^{j_k}\right\} = span\left\{A^1, A^2, \ldots, A^n\right\}.$$

Este resultado nos lleva al siguiente algoritmo.

ALGORITMO 2. Sea S un subespacio de $\mathbb{R}^m$ de la forma
$$S = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}.$$

Supongamos que queremos encontrar una base $\mathcal{B}$ para S con la restricción adicional de que ésta debe ser un subconjunto del conjunto original $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. La filosofía del algoritmo que resuelve este problema se basa en la Proposición 11.27. Sea A la matriz cuyas columnas son $A^1 = \vec{v}_1, A^2 = \vec{v}_2, \ldots, A^n = \vec{v}_n$ y sea $\tilde{A}$ la FER de A. Entonces, la columna A^j (i.e., el vector $\vec{v}_j$) estará en nuestra base $\mathcal{B}$ syss la columna $\tilde{A}^j$ tiene uno principal.

Ejemplo 11.28 *Sean*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -1 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \vec{v}_4 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ -1 \end{bmatrix}, \vec{v}_5 = \begin{bmatrix} 0 \\ 4 \\ 0 \\ 4 \end{bmatrix}$$

vectores en $\mathbb{R}^4$ *y supongamos que deseamos encontrar una base para*

$$span\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4, \vec{v}_5\},$$

con la restricción de que dicha base debe de ser un subconjunto de

$$\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4, \vec{v}_5\}.$$

Para hacer uso del Algoritmo 2 calculamos la FER de la matriz cuyas columnas son los vectores $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4, \vec{v}_5$. *Sea*

$$A = \begin{bmatrix} 1 & -1 & 0 & 1 & 0 \\ -2 & 2 & 0 & 2 & 4 \\ 1 & -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & -1 & 4 \end{bmatrix}$$

dicha matriz, cuya FER está dada por

$$\tilde{A} = \begin{bmatrix} 1 & 0 & 1 & 0 & 4 \\ 0 & 1 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Esta matriz tiene unos principales en la primera, segunda y cuarta columnas, por lo que el Algoritmo 2 asegura que el conjunto

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \\ -1 \end{bmatrix} \right\}$$

es una base para S. *Si el lector quisiera una verificación de que* $\mathcal{B}$ *es LI, sería suficiente hacerle notar que la FER de la matriz aumentada*

$$[A^1 \ A^2 \ A^4 \mid \ \vec{0}]$$

es la matriz

$$[\tilde{A}^1 \ \tilde{A}^2 \ \tilde{A}^4 \mid \ \vec{0}],$$

De igual manera, si el lector quisiera verificar, por ejemplo, que $\vec{v}_5 \in span\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$, *sería suficiente hacerle notar que la FER de la matriz aumentada* $[A^1 \ A^2 \ A^4 \mid \ A^5]$ *es la matriz*

$$[\tilde{A}^1 \ \tilde{A}^2 \ \tilde{A}^4 \mid \ \tilde{A}^5],$$

la cual representa un sistema consistente. Similarmente, puede verificarse que $\vec{v}_3 \in span\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$ *y* $\vec{v}_4 \in span\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$.

Ejemplo 11.29 *Consideremos el subconjunto de* $\mathbb{R}^3$ *dado por*

$$\mathcal{A} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Es claro que $\mathcal{A}$ *es LD y span*$\mathcal{A} = \mathbb{R}^3$. *Si deseamos extraer de* $\mathcal{A}$ *una base para* $\mathbb{R}^3$, *tomamos la matriz cuyas columnas son los elementos de* $\mathcal{A}$, *es decir*

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

cuya FER está dada por

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

Los unos principales están en las columnas uno dos y cinco y, por lo tanto, por el Algoritmo 2

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

es una base para $\mathbb{R}^3$ *formada por elementos de* $\mathcal{A}$.

Observemos que si reordenamos los elementos de $\mathcal{A}$ *como*

$$\mathcal{A} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\},$$

la matriz asociada será ahora

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

cuya FER se obtiene fácilmente como

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

En consecuencia

$$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

es otro subconjunto LI de $\mathcal{A}$ *tal que span*$\mathcal{C} = \mathbb{R}^3$.

Observación 11.30 *En el ejemplo anterior vemos que dado un conjunto $\mathcal{A}$ pueden extraerse múltiples bases para span$\mathcal{A}$. Asimismo, si $\mathcal{B}$ es un subconjunto LI de $\mathcal{A}$ y queremos garantizar que, mediante el Algoritmo 2, se obtiene una base para span$\mathcal{A}$ que incluya a todos los elementos de $\mathcal{B}$, entonces reordenamos los elementos de $\mathcal{A}$ de manera que los vectores de $\mathcal{B}$ sean las primeras columnas de la matriz en cuestión.*

Extensiones de conjuntos LI

Los Algoritmos 1 y 2 nos permiten encontrar una base a partir de cualquier conjunto finito de generadores de un subespacio $S \subset \mathbb{R}^n$. Esencialmente, dado uno de estos conjuntos, se eliminan vectores hasta llegar a un conjunto minimal de generadores que constituye una base para S. Ahora supongamos que, sin tener un conjunto finito de generadores, queremos encontrar una base para el subespacio S. En este caso, se construye una base para S agregando uno a uno sus elementos hasta obtener un conjunto maximal LI ¿Cómo procederíamos?

Sea $S \subset \mathbb{R}^n$ cualquier subespacio. Si $S = \{\vec{0}\}$, entonces el conjunto vacío, $\emptyset$, es la base buscada. Si $S \neq \{\vec{0}\}$, existe $\vec{v}_1 \in S$ con $\vec{v}_1 \neq \vec{0}$ y tenemos dos posibilidades:

$$span\{\vec{v}_1\} = S \text{ o } span\{\vec{v}_1\} \subsetneq S.$$

En el primer caso $\{\vec{v}_1\}$ es una base para S y el proceso termina. En el segundo, existe $\vec{v}_2 \in S$ con $\vec{v}_2 \notin span\{\vec{v}_1\}$ de manera que $\{\vec{v}_1, \vec{v}_2\}$ es LI. Una vez más hay dos posibilidades:

$$span\{\vec{v}_1, \vec{v}_2\} = S \text{ o } span\{\vec{v}_1, \vec{v}_2\} \subsetneq S.$$

Al igual que antes, en el primer caso $\{\vec{v}_1, \vec{v}_2\}$ es una base para S y en el segundo, existe $\vec{v}_3 \in S$ con $\vec{v}_2 \notin span\{\vec{v}_1, \vec{v}_2\}$ de manera que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es LI. Continuamos agregando vectores hasta que

$$span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} = S.$$

El proceso termina, necesariamente, ya que $S \subset \mathbb{R}^n$ y, por lo tanto, el número de vectores que se agregan, debe cumplir $p \leq n$. El siguiente teorema y su corolario se siguen del razonamiento anterior.

Teorema 11.31 *Sea $S \subset \mathbb{R}^n$ cualquier subespacio diferente del $\{\vec{0}\}$, entonces existen $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p \in S$, tales que $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es una base de S.*

Corolario 11.32 *Sea $S \subset \mathbb{R}^n$ cualquier subespacio diferente del $\{\vec{0}\}$ y sea $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} \subset S$ cualquier subconjunto LI. Si $\dim(S) = r > p$, entonces existen vectores $\vec{v}_{p+1}, \vec{v}_{p+2}, \ldots, \vec{v}_r \in S$ tales que*

$$\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} \cup \{\vec{v}_{p+1}, \vec{v}_{p+2}, \ldots, \vec{v}_r\}$$

es una base para S. En otras palabras, cualquier subconjunto LI de S puede completarse a una base.

Ejemplo 11.33 *Supongamos que se desea encontrar una base para $\mathbb{R}^3$ a partir del conjunto LI dado por*

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Tenemos que

$$\begin{aligned} & span \left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \\ &= \left\{ r \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \middle| r, s \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} r \\ -r+s \\ 2r \end{bmatrix} \middle| r, s \in \mathbb{R} \right\}. \end{aligned}$$

Observando que la tercera coordenada de cualquier elemento de este span es dos veces su primera coordenada se tiene que

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \notin span \left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

y el conjunto

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

es LI independiente. En este caso, el proceso termina aquí pues este conjunto LI tiene tres elementos y forma una base para $\mathbb{R}^3$.

Ejemplo 11.34 *Otra forma de obtener una base para $\mathbb{R}^3$ a partir del conjunto LI*

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

se basa en el Algoritmo 2, notando que estos dos vectores junto con $\vec{e}_1$ *y* $\vec{e}_3$ *forman un conjunto de generadores para* $\mathbb{R}^3$*. Construimos pues la matriz* $A \in M_{3\times 4}$ *cuyas columnas son estos cuatro vectores.*

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

La FER de A *está dada por*

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

y por el Algoritmo 2 tenemos que el conjunto de vectores

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

forma una base para $\mathbb{R}^3$.

Ejercicios

Ejercicio 11.1 *En los siguientes casos, determinar si el conjunto de vectores $\mathcal{B}$ es una base del subespacio S. Justificar la respuesta.*

a. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$, $S = \mathbb{R}^2$.

b. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$, $S = \mathbb{R}^2$.

c. $\mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$, $S = \mathbb{R}^3$.

d. $\mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$, $S = \mathbb{R}^3$.

e. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} \pi \\ \pi \\ \pi \end{bmatrix}, \begin{bmatrix} 666 \\ 777 \\ 888 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$, $S = \mathbb{R}^3$.

f. $\mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$, $S = Null\left(\begin{bmatrix} \pi & 0 & \pi \\ 2 & 0 & 2 \\ -7 & 0 & -7 \end{bmatrix} \right)$.

g. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$, $S = span\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \end{bmatrix} \right\}$.

h. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}$, $S = \mathbb{R}^4$.

i. $\mathcal{B} = \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}$,

$$S = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \mid a + 2b + 3c = 0 \ \& \ c + d = 0 \right\}.$$

Ejercicio 11.2 *j.* $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\}, S = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + y - z = 0 \right\}.$

Ejercicio 11.3 *Proporcionar cien bases distintas para* $\mathbb{R}^2$.

Ejercicio 11.4 *Proporcionar una base para los siguientes subespacios.*

a. $S = span \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} \pi \\ \pi \end{bmatrix} \right\} \subset \mathbb{R}^2.$

b. $S = span \left\{ \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\} \subset \mathbb{R}^3.$

c. $S = span \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\} \subset \mathbb{R}^4.$

Ejercicio 11.5 *Encontrar una base para el espacio nulo de las siguientes matrices.*

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

b. $B = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$

c. $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$

Ejercicio 11.6 *Encontrar la dimensión de los siguientes subespacios (observar que los primeros cuatro aparecen en el ejercicio 11.1).*

a. $S_1 = Null \left(\begin{bmatrix} \pi & 0 & \pi \\ 2 & 0 & 2 \\ -7 & 0 & -7 \end{bmatrix} \right) \subset \mathbb{R}^3.$

b. $S_2 = span \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \end{bmatrix} \right\} \subset \mathbb{R}^2.$

c. $S_3 = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \mid a + 2b + 3c = 0 \ \& \ c + d = 0 \right\} \subset \mathbb{R}^4.$

d. $S_4 = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + y - z = 0 \right\} \subset \mathbb{R}^3.$

e. $S_5 = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} \mid a - c - e = 0 \right\} \subset \mathbb{R}^5.$

Ejercicio 11.7 *Encontrar una base para $\mathbb{R}^3$ que contenga al conjunto dado.*

a. $\left\{ \begin{bmatrix} \pi \\ \pi \\ \pi \end{bmatrix} \right\}.$

b. $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \right\}.$

Ejercicio 11.8 *Proporcionar ejemplos de subespacios que cumplan con las propiedades dadas.*

a. *Un subespacio S de $\mathbb{R}^n$ (con $n > 2$) tal que* $\dim(S) = n - 2$.

b. *Dos subespacios S_1 y S_2 de $\mathbb{R}^4$ tales que* $\dim(S_1) = \dim(S_2) = 2$ *y* $S_1 \cap S_2 = \{\vec{0}\}$.

Ejercicio 11.9 *Sea $S \subset \mathbb{R}^n$ un subespacio y $\{\vec{u}, \vec{v}, \vec{w}\}$ una base de S.*

a. *Probar que si a, b, c son escalares no nulos, entonces $\{a\vec{u}, b\vec{v}, c\vec{w}\}$ es una base de S.*

b. *Probar que $\{\vec{u}, \vec{u} + \vec{v}, \vec{u} + \vec{v} + \vec{w}\}$ es una base de S.*

Ejercicio 11.10 *Utilizar el Algoritmo 1 para encontrar una base para los siguientes subespacios.*

a. $S_1 = span \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -4 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \\ 2 \end{bmatrix} \right\} \subset \mathbb{R}^3.$

b. $S_2 = span \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 2 \\ 2 \end{bmatrix} \right\} \subset \mathbb{R}^4.$

Ejercicio 11.11 *Utilizar el Algoritmo 2 para encontrar las bases que se piden a continuación.*

a. *Una base para el subespacio S_1 del ejercicio anterior que sea un subconjunto de*

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -4 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \\ 2 \end{bmatrix} \right\}.$$

b. *Una base para el subespacio S_2 del ejercicio anterior que contenga a los vectores* $\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ *y* $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$.

c. *Una base para el subespacio S_2 del ejercicio anterior que contenga al vector* $\begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$.

d. *Una base para Rango(T) en donde $T : \mathbb{R}^3 \to \mathbb{R}^3$ está dada por*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = \begin{bmatrix} x + y + 3z \\ x - y + z \\ x + 2z \end{bmatrix}.$$

Ejercicio 11.12 *Probar (en general) o refutar (con un contraejemplo) los siguientes enunciados (en cada caso se debe indicar claramente "verdadero" o "falso").*

a. *Si S es un subespacio de $\mathbb{R}^2$ con $dim(S) = 1$ y $\vec{u}$, $\vec{v}$ son vectores de $\mathbb{R}^2$ tales que $span\{\vec{u}, \vec{v}\} = \mathbb{R}^2$, entonces $\{\vec{u}\}$ es base para S o $\{\vec{v}\}$ es base para S.*

b. *Si S_1 y S_2 son subespacios de $\mathbb{R}^n$, entonces $dim(S_1 \cap S_2) \leq dim(S_1)$ y $dim(S_1 \cap S_2) \leq dim(S_2)$.*

c. *Si S_1 y S_2 son subespacios de $\mathbb{R}^3$, entonces $dim(S_1 \cap S_2) = dim(S_1)$ o $dim(S_1 \cap S_2) = dim(S_2)$.*

d. *Si S_1 y S_2 son subespacios de $\mathbb{R}^n$ y $dim(S_1 \cap S_2) = dim(S_1)$, entonces $S_1 \subseteq S_2$.*

e. *Si S es un subespacio de $\mathbb{R}^n$ tal que $S = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$, entonces* $\dim(S) = k$.

f. *Sea S es un subespacio de $\mathbb{R}^n$ y $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\} \subset S$ un conjunto LI. Entonces $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ es una base de S o bien existen $\vec{v}_{k+1}, \vec{v}_{k+2}, \ldots, \vec{v}_{k+l} \in S$ tales que*

$$\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k, \vec{v}_{k+1}, \vec{v}_{k+2}, \ldots, \vec{v}_{k+l}\}$$

es una base de S.

g. *Sea S es un subespacio de $\mathbb{R}^n$, si $\dim(S) = n$, entonces $S = \mathbb{R}^n$.*

h. *Sean S_1, S_2 subespacios de $\mathbb{R}^n$, si $\dim(S_1) = \dim(S_2)$, entonces $S_1 = S_2$.*

Ejercicio 11.13 *Demostrar que si $S \subset \mathbb{R}^n$ es un subespacio de dimensión $p > 0$ y span$\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} = S$, entonces $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es una base para S.*

Capítulo 12

Subespacios asociados a una matriz (opcional)

Espacio nulo y nulidad

En este breve capítulo mostraremos la relación que existe, en términos de dimensión, entre los tres espacios fundamentales asociados a una matriz. Recordemos, una vez más, que si A es una matriz de $m \times n$, entonces el espacio nulo de A es el conjunto de soluciones a la ecuación homogénea, es decir,

$$Null(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}_m\}.$$

Definición 12.1 *Si $A \in M_{m\times n}$, entonces la **nulidad** de A, denotada por $nulidad(A)$, se define como la dimensión del espacio nulo de A.*

La siguiente proposición generaliza el último argumento utilizado en el Ejemplo 11.5, en donde se justifica que el conjunto cuyos elementos son los dos vectores inducidos por las variables libres es LI.

Proposición 12.2 *Si $A \in M_{m\times n}$, entonces la nulidad de A es igual al número de columnas de la FER de A que no tienen un uno principal; es decir, coincide con el número de variables libres del sistema $A\vec{x} = \vec{0}$.*

Demostración. Sea $\tilde{A}$ es la FER de A y supongamos que ésta tiene k unos principales. Sean

$$\tilde{A}^{l_1}, \tilde{A}^{l_2}, \cdots, \tilde{A}^{l_{n-k}}$$

con $l_1 < l_2 < \cdots < l_{n-k}$ las $n-k$ columnas correspondientes a las variables libres de $[A \mid \vec{0}]$. Despejando las k variables restringidas en $[\tilde{A} \mid \vec{0}]$ es posible expresar a $Null(A)$ como

$$Null(A) = \left\{ x_{l_1} v_{l_1} + x_{l_2} v_{l_2} + \cdots + x_{l_{n-k}} v_{l_{n-k}} \mid x_{l_i} \in \mathbb{R} \right\},$$

en donde $x_{l_1}, x_{l_2}, \ldots, x_{l_{n-k}}$ corresponden a las variables libres y $\vec{v}_{l_1}, \vec{v}_{l_2}, \ldots, \vec{v}_{l_{n-k}} \in \mathbb{R}^n$. Por lo tanto,

$$Null(A) = span \left\{ \vec{v}_{l_1}, \vec{v}_{l_2}, \ldots, \vec{v}_{l_{n-k}} \right\}.$$

Además, el conjunto $\left\{ \vec{v}_{l_1}, \vec{v}_{l_2}, \ldots, \vec{v}_{l_{n-k}} \right\}$ es LI puesto que $\vec{v}_{l_j}$ tiene un uno en la entrada l_j mientras que los otros vectores de este conjunto tienen un cero en dicha posición. Así,

$$\left\{ \vec{v}_{l_1}, \vec{v}_{l_2}, \ldots, \vec{v}_{l_{n-k}} \right\}$$

es una base para $Null(A)$ con $n-k$ elementos y se cumple que

$$nulidad(A) = \# \text{ de variables libres} = n - k,$$

quedando así demostrada la proposición. ■

Espacio columna y rank

Ahora veremos que el número de variables restringidas de la matriz $A \in M_{m \times n}$ es igual a la dimensión del espacio generado por las columnas de A y, sorprendentemente, también coincide con la dimensión del espacio generado por los renglones de A. Observemos que estos dos subespacios no son necesariamente iguales incluso si $n = m$ (evidentemente si $n \neq m$, ambos subespacios son distintos pues están contenidos en distintos espacios).

Definición 12.3 *Sea $A \in M_{m \times n}$. Definimos el* ***espacio columna*** *de A, denotado por $col(A)$, como el conjunto*

$$\begin{aligned} col(A) &= span\{A^1, \ldots, A^n\} \\ &= \{\vec{b} \in \mathbb{R}^m \mid \textit{el sistema } A\vec{x} = \vec{b} \textit{ es consistente}\} \end{aligned}$$

y a su dimensión como

$$rank(A) = \dim(\mathrm{col}(A)).$$

De acuerdo al Algoritmo 2 del capítulo anterior, si queremos hallar una base para $col(A)$, con la restricción de que dicha base sea un subconjunto de $\{A^1, \ldots, A^n\}$, primero encontramos la FER de A; supongamos que $\tilde{A}$ es dicha FER. Segundo, notamos que la columna A^j estará en nuestra base $\mathcal{B}$ syss la columna $\tilde{A}^j$ contiene un uno principal. Esto nos lleva al siguiente resultado.

Proposición 12.4 *Sea $A \in M_{m\times n}$, entonces $rank(A)$ es igual al número de unos principales de la FER de A.*

Ejemplo 12.5 *Consideremos a la matriz*

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

y supongamos que deseamos encontrar bases para $Null(A)$ y $col(A)$, así como la nulidad de A y $rank(A)$. Para ello, notamos que la FER de A es la matriz

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

por lo que el Algoritmo 2 asegura que el conjunto

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

es una base para $col(A)$ y $rank(A) = 2$. La matriz $\tilde{A}$ también nos dice que la nulidad de A es igual a 3 (pues tiene tres variables libres). Más aún, como la FER de la matriz aumentada del sistema homogéneo $A\vec{x} = \vec{0}$ es igual a $(\tilde{A} \mid \vec{0})$ tenemos que $\vec{x} \in Null(A)$ syss sus componentes satisfacen

$$\begin{aligned} x_1 - x_2 &= 0, \\ x_3 - x_4 + 2x_5 &= 0, \\ x_2, x_4, x_5 &\in \mathbb{R}. \end{aligned}$$

Por lo tanto,

$$null(A) = \left\{ \begin{bmatrix} x_2 \\ x_2 \\ x_4 - 2x_5 \\ x_4 \\ x_5 \end{bmatrix} \mid x_2, x_4, x_5 \in \mathbb{R} \right\}$$

$$= \left\{ x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 \\ 0 \\ -2 \\ 0 \\ 1 \end{bmatrix} \mid x_2, x_4, x_5 \in \mathbb{R} \right\}$$

$$= span \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Concluimos pues, que el conjunto

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

es una base para $Null(A)$ por lo que la nulidad de A es igual a 3.

Teorema de la dimensión y espacio renglón

El Ejemplo 12.5 ilustra como el número de variables libres (nulidad de A) más el número de variables restringidas ($rank(A)$) coincide, evidentemente, con el número total de columnas (o variables). Esto puede enunciarse en forma general como sigue.

Teorema 12.6 (Teorema de la dimensión.) *Si A es una matriz de $m \times n$, entonces*

$$n = rank(A) + nulidad(A).$$

Demostración. Supogamos que $\tilde{A}$ es la FER de A y que k es el número de unos principales que allí aparecen. Entonces, de acuerdo a la Proposición 12.4, $rank(A) = k$. Por otro lado, la matriz aumentada $(\tilde{A} \mid \vec{0})$ correspondiente al sistema $\tilde{A}\vec{x} = \vec{0}$, tiene exactamente $n - k$ variables libres. Por lo tanto,

$$nulidad(A) = n - k = n - rank(A)$$

o lo que es lo mismo,

$$n = rank(A) + nulidad(A),$$

con lo cual se concluye la demostración. ■

Definición 12.7 *Sea $A \in M_{m\times n}$. Definimos el* ***espacio renglón*** *de A, denotado por $row(A)$, como el conjunto*

$$row(A) = span\{A_1, \ldots, A_m\}.$$

Proposición 12.8 *Si $A \in M_{m\times n}$, entonces*

$$\dim(\mathrm{row}(A)) = \dim(\mathrm{col}(A)) = rank(A),$$

es decir, los espacios renglón y columna de una matriz tienen la misma dimensión: $rank(A)$.

Demostración. Sea $\tilde{A}$ la FER de A. De acuerdo al Algoritmo 1, los renglones de $\tilde{A}$ que contienen un uno principal forman una base para $\mathrm{row}(A)$. Por lo tanto, la dimensión de $\mathrm{row}(A)$ es igual al número de unos principales de $\tilde{A}$. Finalmente, la Proposición 12.4 asegura que este número coincide con $\dim(\mathrm{col}(A)) = rank(A)$. ■

El siguiente corolario se sigue de la Proposición 12.8, dado que las columnas de la matriz A corresponden a los renglones de la matriz A^T.

Corolario 12.9 *Dada $A \in M_{m\times n}$, se tiene que $rank(A) = rank(A^T)$.*

Ejemplo 12.10 *Consideremos la matriz*

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 5 & 6 & 7 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

cuya FER es la matriz

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

Dado que hay dos unos principales se tiene que $rank(A) = 2$. Más aún, el Algoritmo 2 asegura que el conjunto

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix} \right\} \subset \mathbb{R}^3$$

es una base para $col(A)$. Por otro lado, el Algoritmo 1 asegura que el conjunto

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\} \subset \mathbb{R}^4$$

es una base para $row(A)$. *Como la FER de* A *tiene dos columnas que corresponden a las variables libres del sistema homogéneo* $A\vec{x} = \vec{0}$, *se tiene que* $nulidad(A) = 2$. *Asimismo, el espacio nulo está dado por*

$$Null(A) = \left\{ x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \end{bmatrix} \mid x_3, x_4 \in \mathbb{R} \right\}.$$

Claramente se satisface el teorema de la dimensión pues

$$rank(A) + nulidad(A) = 2 + 2 = 4$$

Tomemos ahora

$$A^T = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 5 & 2 \\ 3 & 6 & 3 \\ 4 & 7 & 4 \end{bmatrix},$$

cuya FER es la matriz

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

Entonces $rank(A^T) = 2$ *y una base para* $\mathrm{col}(A^T)$ *está dada por*

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ 6 \\ 7 \end{bmatrix} \right\}.$$

Observemos que esta base es distinta a la que se obtuvo anteriormente para los renglones de A *ya que en ese caso se utilizó el Algoritmo 1. El espacio* $\mathrm{row}(A^T)$ *tiene como base a*

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

que también difiere de la primera base obtenida para las columnas de A. *Finalmente, es claro que* $nulidad(A^T) = 1$ *y que*

$$Null(A^T) = \left\{ x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \mid x_3 \in \mathbb{R} \right\}.$$

En este caso el Teorema de la dimensión simplemente afirma que

$$rank(A^T) + nulidad(A^T) = 2 + 1 = 3.$$

Ejemplo 12.11 *Consideremos la matriz*

$$B = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

cuya FER está dada por

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

Utilizando los Algoritmos 1 y 2 se tiene que

$$\mathrm{row}(B) = span \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \; y$$

$$\mathrm{col}(B) = span \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}.$$

Ahora presentamos un resultado importante cuya demostración está autocontenida en el enunciado del mismo.

Corolario 12.12 *Sea $A \in M_{n\times n}$. Las siguientes afirmaciones son equivalentes.*

1. *A es invertible.*
2. *Las columnas de A forman un conjunto LI.*
3. *La ecuación $A\vec{x} = \vec{0}$ tiene una única solución (a saber, $\vec{x} = \vec{0}$).*
4. *$Null(A) = \{\vec{0}\}$.*
5. *$nulidad(A) = 0$.*
6. *$rank(A) = n$.*
7. *Los renglones de A forman un conjunto LI.*

Observación 12.13 *Si $A \in M_{m\times n}$ y $T : \mathbb{R}^n \rightharpoonup \mathbb{R}^m$ es la transformación lineal dada por $T(\vec{x}) = A\vec{x}$, claramente*

$$nulidad(A) = \dim(Null(T))$$

y

$$rank(A) = \dim(Rango(T)).$$

De esta forma, el Teorema de la Dimensión puede expresarse en términos de la transformación lineal T como

$$\dim(Rango(T)) + \dim(Null(T)) = n.$$

Finalizamos este capítulo con algunos ejemplos que muestran cómo los conceptos de rank y nulidad delimitan las posibilidades de que existan cierto tipo de matrices y transformaciones.

Ejemplo 12.14 *Supongamos que* $A \in M_{7\times 16}$. *¿ Es posible que la nulidad de* A *sea igual a 8? Para responder a esta preguntar notamos que si existiese tal matriz* A*, entonces*

$$rank(A) = 16 - nulidad(A) = 16 - 8 = 8,$$

por lo que la FER de A *tendría al menos 8 renglones (distintos de cero).* A *partir de esta contradicción concluimos que si* A *es una matriz de* 7×16*, no es posible que su nulidad sea igual a* 8.

Ejemplo 12.15 *Supongamos que* $A \in M_{5\times 13}$. *¿Cuál es el posible valor máximo para* $rank(A)$ *y cuál es el posible valor mínimo para* $nulidad(A)$*? Para responder la primera pregunta recordamos que* $rank(A)$ *es igual a la dimensión del espacio renglón de* A*, por lo que dicha dimensión puede ser como mucho 5. Por otro lado, como*

$$13 = rank(A) + nulidad(A),$$

tenemos que si $rank(A)$ *toma su valor máximo, entonces* $nulidad(A)$ *toma su valor mínimo y viceversa. Por lo tanto, el valor mínimo para la nulidad de* A *es 8.*

Ejemplo 12.16 *Sea* $T : \mathbb{R}^5 \longrightarrow \mathbb{R}^4$ *y supongamos que la FER de la matriz estándar asociada a* T *tiene exactamente 3 renglones diferentes de cero. ¿Cuáles son las dimensiones de* $Null(T)$ *y de* $Rango(T)$*? Si* $\tilde{A}$ *es la FER mencionada, entonces* $rank(\tilde{A}) = 3 = \dim(Rango(T))$. *Por el Teorema de la Dimensión se tiene que*

$$\dim(Null(T)) = 5 - 3 = 2.$$

Observemos que T *no es suprayectiva (pues* $\dim(Rango(T)) < 4$*) ni tampoco inyectiva (pues* $\dim(Null(T)) > 0$*).*

Ejercicios

Ejercicio 12.1 *Para cada una de las matrices que se presentan a continuación encontrar:1) una base para su espacio columna, 2) una base para su espacio renglón, 3) una base para su espacio nulo, 4) su rank y 5) su nulidad.*

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 2 \end{bmatrix},\ B = \begin{bmatrix} 1 & \pi \\ 0 & 0 \\ 1 & \pi \end{bmatrix},\ C = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \\ -1 & 1 & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},\ E = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},\ F = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$G = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \\ -1 & 1 & 0 \\ 1 & 3 & 2 \end{bmatrix},\ H = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 1 & 0 \\ -1 & -2 & 0 & 1 \end{bmatrix},\ J = \begin{bmatrix} 0 & 1 & -2 \\ 0 & 3 & 3 \\ 0 & 1 & 7 \\ 0 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Ejercicio 12.2 *Encontrar los valores de a y b de tal suerte que $rank(A) = 2 = rank(B)$, en donde*

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

Ejercicio 12.3 *Probar que si $A \in M_{m\times p}$ y $B \in M_{p\times n}$, entonces $nulidad(B) \leq nulidad(AB)$ y $rank(AB) \leq rank(B)$.*

Ejercicio 12.4 *Probar que si $A \in M_{m\times p}$ y $B \in M_{p\times n}$, entonces $rank(AB) \leq rank(A)$.*

Ejercicio 12.5 *Probar que si $A \in M_{m\times p}$, $B \in M_{p\times n}$ y $rank(A) = p$, entonces $nulidad(B) = nulidad(AB)$ y $rank(AB) = rank(B)$.*

Ejercicio 12.6 *Probar que si $A \in M_{m\times p}$, $B \in M_{p\times n}$ y $rank(B) = p$, entonces $rank(AB) = rank(A)$.*

Ejercicio 12.7 *Probar que si $A \in M_{m\times n}$ y $m \neq n$, entonces $nulidad(A) \neq nulidad(A^T)$.*

Ejercicio 12.8 *Probar (en general) o refutar (con un contraejemplo) los siguientes enunciados (en cada caso se debe indicar claramente "verdadero" o "falso").*

a. *Si A y B son matrices de 2×2 y $\vec{x}$ está en el espacio nulo del producto AB, entonces $\vec{x}$ está en el espacio nulo de A o $\vec{x}$ está en el espacio nulo de B.*

b. *Sean A y B matrices de $n \times n$. Si A es invertible y $\vec{x}$ está en el espacio nulo de AB, entonces $\vec{x}$ está en el espacio nulo de B.*

c. *Si $A \in M_{m\times n}$ y $B \in M_{m\times n}$ son tales que B se obtiene a partir de A aplicando operaciones elementales, entonces el espacio columna de A es igual al espacio columna de B.*

d. *Si $A \in M_{m\times n}$ y $B \in M_{m\times n}$ son tales que B se obtiene a partir de A aplicando operaciones elementales, entonces el espacio renglón de A es igual al espacio renglón de B.*

e. *No existe una matriz $A \in M_{9\times 18}$ tal que $nulidad(A) = 8$.*

f. *Si $A \in M_{m\times p}$ y $B \in M_{p\times n}$, entonces $rank(B) = rank(AB)$.*

g. *Si $A \in M_{m\times p}$ y $B \in M_{p\times n}$, entonces $rank(A) = rank(AB)$.*

h. *Si $A \in M_{n\times n}$ y $B \in M_{n\times n}$ son matrices invertibles, entonces $rank(A + B) = n$.*

i. *Si $A \in M_{n\times n}$ y $B \in M_{n\times n}$ son matrices invertibles, entonces $rank(AB) = n$.*

j. *Si $A \in M_{m\times n}$ es tal que $rank(A) < n$, entonces las columnas de A forman un conjunto LD.*

k. *Si $A \in M_{m\times n}$ es tal que el sistema $A\vec{x} = \vec{b}$ siempre tiene solución independientemente de la elección del vector $\vec{b} \in \mathbb{R}^m$, entonces $rank(A) = m$.*

Capítulo 13

Eigenvalores y eigenvectores

Introducción

Comenzamos este capítulo con una observación bastante natural. Supongamos que

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

es una transformación lineal y que $\vec{x} \in \mathbb{R}^n$. ¿En qué sentido podemos decir que la transformación T no afecta "mucho" al vector $\vec{x} \in \mathbb{R}^n$? Queda claro que si

$$T(\vec{x}) = \vec{x} = 1\vec{x},$$

entonces T no cambia en nada al vector $\vec{x}$. Análogamente, podríamos decir que T "perturba poco" a $\vec{x}$ si **existe** un escalar λ tal que[1]

$$T(\vec{x}) = \lambda\vec{x}.$$

Si A es la matriz estándar asociada a T, lo anterior equivale a

$$A\vec{x} = \lambda\vec{x}.$$

[1]En este caso T simplemente reescala y/o cambia la dirección del vector $\vec{x}$.

Esta situación se ilustra en la Figura 13.1.

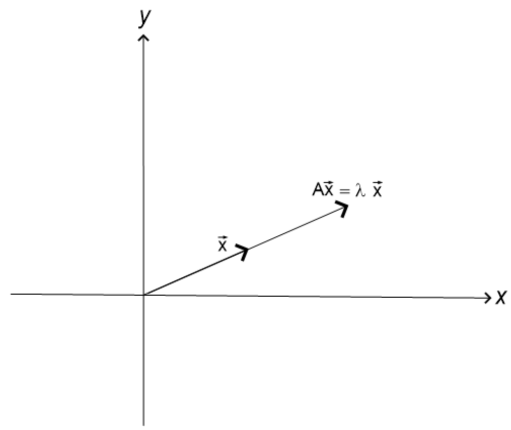


Figura 13.1: $A\vec{x} = \lambda\vec{x}$.

En el caso del vector cero siempre se cumple

$$A\vec{0} = \vec{0} = \lambda\vec{0}$$

para cualquier escalar λ, por lo que no es un caso de interés para este análisis y únicamente se consideran vectores $\vec{x} \neq \vec{0}$. Nuestro objetivo es determinar si, dada una transformación lineal, existen o no vectores (no nulos) que se perturban poco, en el sentido que acabamos de explicar. En las definiciones que siguen, la matriz A puede ser sustituida por su transformación lineal asociada T como se hará explícito más adelante.

Obtención de eigenvalores y eigenvectores

Definición 13.1 *Sean A es una matriz de $n \times n$ y λ un escalar. Decimos que λ es un **eigenvalor (**también llamado **valor propio o valor característico)** de A, si existe un vector $\vec{x}$ de $\mathbb{R}^n$, con $\vec{x} \neq \vec{0}$, tal que $A\vec{x} = \lambda\vec{x}$. A tal vector $\vec{x}$ se le denomina **eigenvector (o vector propio o característico)** de A con respecto a λ.*

Observación 13.2 *Los eigenvectores nunca son nulos, sin embargo el real 0 sí puede ser un eigenvalor. Notemos que $\vec{x} \neq \vec{0}$ es un eigenvector con eigenvalor cero syss $\vec{x} \in Null(A)$.*

Definición 13.3 *Sea A una matriz de $n \times n$ y sea λ un escalar. El* ***eigenespacio*** *de A asociado a λ, denotado por $Eig(\lambda)$, se define como el conjunto*

$$Eig(\lambda) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \lambda\vec{x}\},$$

es decir, como la colección de todos los eigenvectores con eigenvalor λ junto con el vector $\vec{0}$.

Observación 13.4 *En la Definición 13.3, λ es un eigenvalor de A syss $Eig(\lambda) \neq \{\vec{0}\}$.*

Dada una matriz $A \in M_{n\times n}$, ahora proporcionaremos un método para encontrar sus eigenvalores y eigenvectores, si es que éstos existen.

Proposición 13.5 *Si A es una matriz de $n \times n$ y λ es un escalar, entonces*

$$Eig(\lambda) = Null(A - \lambda I),$$

y por lo tanto, $Eig(\lambda)$ es un subespacio de $\mathbb{R}^n$.

Demostración. Basta notar las siguientes equivalencias: $\vec{x} \in Eig(\lambda)$ syss

$$A\vec{x} = \lambda\vec{x} = \lambda I\vec{x}$$

syss

$$A\vec{x} - \lambda I\vec{x} = \vec{0},$$

syss

$$(A - \lambda I)\vec{x} = \vec{0},$$

syss

$$\vec{x} \in Null(A - \lambda I),$$

quedando demostrado el resultado. ■

Corolario 13.6 *Sea A una matriz de $n \times n$. Entonces las siguientes afirmaciones son equivalentes.*

(a) λ es un eigenvalor de A.

(b) Existe un vector distinto cero que pertence a

$$Eig(\lambda) = Null(A - \lambda I).$$

(c) $dim\,(Eig(\lambda)) = \dim(Null(A - \lambda I) \geq 1$.

(d) Las columnas de $A - \lambda I$ forman un conjunto LD.

(e) La matriz $A - \lambda I$ no es invertible.

(f) $det(A - \lambda I) = 0$.

(g) *El sistema homogéneo* $(A - \lambda I)\vec{x} = \vec{0}$ *tiene una infinidad de soluciones.*

Ejemplo 13.7 *Si*

$$A = \begin{bmatrix} 1 & 9 \\ 1 & 1 \end{bmatrix},$$

entonces $\lambda \in \mathbb{R}$ *es un eigenvalor de* A *syss*

$$0 = \det(A - \lambda I) = \det\left(\begin{bmatrix} 1 & 9 \\ 1 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)$$

$$\det\left(\begin{bmatrix} 1-\lambda & 9 \\ 1 & 1-\lambda \end{bmatrix} \right) = (1-\lambda)(1-\lambda) - 9 = (\lambda - 4)(\lambda + 2).$$

Así pues, los únicos eigenvalores de A *son* $\lambda_1 = 4$ *y* $\lambda_2 = -2$*. Además*

$$\begin{aligned} Eig(\lambda_1) &= Null\left(\begin{bmatrix} -3 & 9 \\ 1 & -3 \end{bmatrix} \right), \\ Eig(\lambda_2) &= Null\left(\begin{bmatrix} 3 & 9 \\ 1 & 3 \end{bmatrix} \right). \end{aligned}$$

Es fácil ver que

$$\begin{aligned} Eig(\lambda_1) &= span\left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}, \\ Eig(\lambda_2) &= span\left\{ \begin{bmatrix} -3 \\ 1 \end{bmatrix} \right\}. \end{aligned}$$

En particular, puede verificarse que si $\vec{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ *y* $\vec{v}_2 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ *se tiene que*

$$\begin{aligned} A\vec{v}_1 &= \begin{bmatrix} 1 & 9 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 12 \\ 4 \end{bmatrix} \\ &= 4 \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \lambda_1 \vec{v}_1, \\ A\vec{v}_2 &= \begin{bmatrix} 1 & 9 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ -2 \end{bmatrix} \\ &= -2 \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \lambda_2 \vec{v}_2. \end{aligned}$$

Definición 13.8 *Sea* $A \in M_{n\times n}$ *y* λ *una variable. Definimos al* ***polinomio característico*** *de* A *como*

$$p_A(\lambda) = det(A - \lambda I).$$

Utilizando un argumento inductivo es posible demostrar que, efectivamente, $p_A(\lambda)$ es un polinomio y que su grado es igual a n. Recordemos que si

$$p(x) = a_n x^n + \cdots + a_1 x + a_0 = a_n x^n + \cdots + a_1 x + a_0 x^0$$

es un polinomio con coeficientes reales distinto del polinomio constante cero, entonces el **grado** de $p(x)$, denotado por $\delta(p(x))$, se define como el máximo $m \leq n$ tal que $a_m \neq 0$. Por ejemplo,

$$\delta(0x^8 + 6x^2 + 7) = 2 \text{ y } \delta(6) = \delta(6x^0) = 0.$$

En lo anterior se hace énfasis en que

$$a_0 = a_0 \times 1 = a_0 x^0.$$

También recordamos que si

$$p(x) = a_n x^n + \cdots + a_1 x + a_0$$

es un polinomio y r es un número real, decimos que r es una **raíz (**o un **"**cero**")** de $p(x)$ syss $p(r) = 0$. Utilizando esta terminología, el resultado anterior (o más bien parte de él) puede reescribirse como sigue.

Corolario 13.9 *Si A es una matriz de $n \times n$ y λ^* es un número real, entonces λ^* es un eigenvalor de A syss λ^* es una raíz del polinomio característico $p_A(\lambda)$.*

Ejemplo 13.10 *Consideremos la matriz*

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

y supongamos que deseamos determinar los eigenvalores de A así como una base para cada uno de sus eigenespacios. Para ello, notamos que el polinomio característico de A es igual a

$$\begin{aligned} p_A(\lambda) = det(A - \lambda I) &= det\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \\ &= det \begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 - 1. \end{aligned}$$

Así, $p_A(\lambda) = 0$ syss $\lambda = 1$ o $\lambda = -1$ y los dos eigenvalores de A son $\lambda_1 = 1$ y $\lambda_2 = -1$. Además,

$$Eig(\lambda_1) = Null(A - \lambda_1 I) = Null \begin{bmatrix} -\lambda_1 & 1 \\ 1 & -\lambda_1 \end{bmatrix} = Null \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Por lo tanto,

$$Eig(\lambda_1) = span \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

y

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

es una base para $Eig(\lambda_1)$. *Como*

$$Eig(\lambda_2) = Null(A - \lambda_2 I) = Null \begin{bmatrix} -\lambda_2 & 1 \\ 1 & -\lambda_2 \end{bmatrix} = Null \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

se tiene que

$$Eig(\lambda_2) = span \left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

y

$$\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

es una base para $Eig(\lambda_2)$.

Debido a la correspondencia entre matrices y transformaciones lineales las siguientes dos definiciones resultan bastante naturales.

Definición 13.11 *Sea* $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *una transformación lineal y sea* $A \in M_{n\times n}$ *la matriz estándar que representa a* T *(esto es,* $T(\vec{x}) = A\vec{x}$*). Dado un escalar* λ*, decimos que* λ *es un* ***eigenvalor*** *de* T *syss* λ *es un eigenvalor de* A*. Dado* $\vec{x} \in \mathbb{R}^n$ *también decimos que* $\vec{x}$ *es un* ***eigenvector*** *de* T *con respecto a* λ *syss* $\vec{x}$ *es un eigenvector de* A *con respecto a* λ*.*

Definición 13.12 *Sean* $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *una transformación lineal,* $A \in M_{n\times n}$ *la matriz estándar que representa a* T *y* λ *un escalar. El* ***eigenespacio*** *de* T *asociado a* λ *se define como el eigenespacio de* A *asociado a* λ*.*

Ejemplo 13.13 *Supongamos que* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *es la transformación lineal que satisface*

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = \begin{bmatrix} x + y \\ -2x + 4y \end{bmatrix}$$

y supongamos que deseamos determinar los eigenvalores de T *así como una base para cada uno de sus eigenespacios. Para ello, notamos primero que la matriz estándar que representa a* T *es igual a*

$$A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$$

y su polinomio característico es igual a

$$\begin{aligned} p_A(\lambda) &= det(A - \lambda I) \\ &= det\left(\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \\ &= det\begin{bmatrix} 1-\lambda & 1 \\ -2 & 4-\lambda \end{bmatrix} = (1-\lambda)(4-\lambda) + 2 \\ &= 4 - \lambda - 4\lambda + \lambda^2 + 2 = \lambda^2 - 5\lambda + 6. \end{aligned}$$

Por lo tanto, $p_A(\lambda) = 0$ syss $\lambda = 3$ o $\lambda = 2$ y los dos eigenvalores de T (que son los mismos que los de A) son $\lambda_1 = 3$ y $\lambda_2 = 2$. Como

$$\begin{aligned} Eig(\lambda_1) &= Null(A - \lambda_1 I) \\ &= Null\begin{bmatrix} 1-\lambda_1 & 1 \\ -2 & 4-\lambda_1 \end{bmatrix} = Null\begin{bmatrix} -2 & 1 \\ -2 & 1 \end{bmatrix} \\ &= \left\{ y \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix} \in \mathbb{R}^2 \mid y \in \mathbb{R} \right\} = \left\{ s \begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathbb{R}^2 \mid s \in \mathbb{R} \right\} \\ &= span\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}, \end{aligned}$$

se tiene que

$$\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$$

es una base para $Eig(\lambda_1)$. Análogamente,

$$\begin{aligned} Eig(\lambda_2) &= Null(A - \lambda_2 I) = Null\begin{bmatrix} 1-\lambda_2 & 1 \\ -2 & 4-\lambda_2 \end{bmatrix} = Null\begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix} \\ &= span\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \end{aligned}$$

y

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

es una base para $Eig(\lambda_2)$.

Hasta ahora sólo hemos considerado ejemplos donde los eigenespacios en cuestión tienen dimensión uno (la menor posible), pero el siguiente ejemplo involucra a un eigenespacio de dimensión mayor.

Ejemplo 13.14 *Consideremos la matriz*

$$B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

y supongamos que deseamos determinar los eigenvalores de B así como una base para cada uno de sus eigenespacios. Para ello, notamos que el polinomio característico de B es igual a

$$\begin{aligned}
p_B(\lambda) &= det(B - \lambda I) \\
&= det\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \\
&= det \begin{bmatrix} -\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{bmatrix} = -\lambda(1-\lambda)^2.
\end{aligned}$$

Por lo tanto, $p_B(\lambda) = 0$ syss $\lambda = 0$ o $\lambda = 1$ y los dos eigenvalores de B son $\lambda_1 = 0$ y $\lambda_2 = 1$. Además,

$$\begin{aligned}
Eig(\lambda_1) &= Null(B - \lambda_1 I) \\
&= Null \begin{bmatrix} 0-\lambda_1 & 0 & 0 \\ 0 & 1-\lambda_1 & 0 \\ 1 & 0 & 1-\lambda_1 \end{bmatrix} = Null \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \\
&= \left\{ z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^3 \mid z \in \mathbb{R} \right\} = span \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}
\end{aligned}$$

y

$$\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

es una base para $Eig(\lambda_1)$. Análogamente,

$$\begin{aligned}
Eig(\lambda_2) &= Null(B - \lambda_2 I) \\
&= Null \begin{bmatrix} 0-\lambda_2 & 0 & 0 \\ 0 & 1-\lambda_2 & 0 \\ 1 & 0 & 1-\lambda_2 \end{bmatrix} = Null \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\
&= span \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}
\end{aligned}$$

y

$$\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

es una base para $Eig(\lambda_2)$.

Como consecuencia directa del siguiente resultado podemos afirmar que la matriz B del ejemplo anterior no es invertible.

Proposición 13.15 *Si A es una matriz de $n \times n$, entonces A no es invertible syss 0 es un eigenvalor de A.*

Demostración. Basta notar que A no es invertible syss

$$0 = det(A) = det(A - 0I) = p_A(0).$$

Por lo tanto, A no es invertible syss 0 es un eigenvalor de A. ∎

Ejemplo 13.16 *Consideremos la matriz*

$$C = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

y supongamos que deseamos determinar los eigenvalores de C así como una base para cada uno de sus eigenespacios. Para ello, notamos que el polinomio característico de C es igual a

$$\begin{aligned} p_C(\lambda) &= det(C - \lambda I) \\ &= det\left(\begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \\ &= det\begin{bmatrix} 4-\lambda & 0 & 0 \\ 0 & 1-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{bmatrix} = (4-\lambda)(1-\lambda)^2. \end{aligned}$$

Por lo tanto, $p_C(\lambda) = 0$ syss $\lambda = 4$ o $\lambda = 1$ y los dos eigenvalores de B son $\lambda_1 = 4$ y $\lambda_2 = 1$. Además,

$$\begin{aligned} Eig(\lambda_1) &= Null(C - \lambda_1 I) \\ &= Null\begin{bmatrix} 4-\lambda_1 & 0 & 0 \\ 0 & 1-\lambda_1 & 1 \\ 0 & 0 & 1-\lambda_1 \end{bmatrix} = Null\begin{bmatrix} 0 & 0 & 0 \\ 0 & -3 & 1 \\ 0 & 0 & -3 \end{bmatrix} \\ &= \left\{ x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^3 \mid x \in \mathbb{R} \right\} = span\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\} \end{aligned}$$

y

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

es una base para $Eig(\lambda_1)$. *Por otro lado,*

$$\begin{aligned} Eig(\lambda_2) &= Null(C-\lambda_2 I) \\ &= Null \begin{bmatrix} 4-\lambda_2 & 0 & 0 \\ 0 & 1-\lambda_2 & 1 \\ 0 & 0 & 1-\lambda_2 \end{bmatrix} = Null \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \left\{ y \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \in \mathbb{R}^3 \mid y \in \mathbb{R} \right\} = span \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \end{aligned}$$

y

$$\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

es una base para $Eig(\lambda_2)$.

Notamos que el polinomio característico de C puede reescribirse como

$$p_C(\lambda) = -(\lambda-4)(\lambda-1)^2,$$

por lo que podemos decir que $\lambda_1 = 4$ tiene multiplicidad 1 mientras que $\lambda_2 = 1$ tiene multiplicidad 2 en este polinomio. En general, si $p(x)$ es un polinomio y r es una raíz de $p(x)$, definimos a la **multiplicidad** de r en el polinomio $p(x)$ como el máximo número natural m tal que $(x-r)^m$ divide al polinomio $p(x)$. Por ejemplo, si

$$p(x) = (x-5)^3(x+2)(x-666)^7,$$

entonces las raíces $5, -2$ y 666 tienen multiplicidades $3, 1$ y 7, respectivamente. Dada r una raíz de $p(x)$, denotamos a su multiplicidad por $\mu(r)$.

Los Ejemplos 13.14 y 13.16 sugieren el siguiente resultado cuya demostración omitimos.

Proposición 13.17 *Sea* $A \in M_{n\times n}$ *y supongamos que* λ^* *es un eigenvalor de* A *(i.e.,* λ^* *es una raíz de* $p_A(\lambda)$*). Si* λ^* *tiene multiplicidad* m *en* $p_A(\lambda)$*, entonces* $1 \leq dim(Eig(\lambda^*)) \leq m$.

Ejemplo 13.18 *En el Ejemplo 13.14 encontramos que el polinomio característico de* B *es igual a*

$$p_B(\lambda) = -\lambda(1-\lambda)^2 = -(\lambda-0)(\lambda-1)^2$$

con raices $\lambda_1 = 0$ *y* $\lambda_2 = 1$. *Además,*

$$\begin{aligned} dim(Eig(\lambda_1)) &= 1 = \mu(\lambda_1), \\ dim(Eig(\lambda_2)) &= 2 = \mu(\lambda_2). \end{aligned}$$

Ejemplo 13.19 *En el Ejemplo 13.16 encontramos que el polinomio característico de C es igual a*

$$p_C(\lambda) = (4-\lambda)(1-\lambda)^2 = -(\lambda-4)(\lambda-1)^2$$

con raices $\lambda_1 = 4$ y $\lambda_2 = 1$. Además,

$$\begin{aligned} dim(Eig(\lambda_1)) &= 1 = \mu(\lambda_1), \\ dim(Eig(\lambda_2)) &= 1 \leq 2 = \mu(\lambda_2). \end{aligned}$$

Corolario 13.20 *Sea $A \in M_{n\times n}$ y supongamos que λ^* es un eigenvalor de A. Si λ^* tiene multiplicidad 1 en $p_A(\lambda)$, entonces $dim(Eig(\lambda^*)) = 1$.*

Demostración. Como λ^* es un eigenvalor de A, es claro que

$$dim(Eig(\lambda^*)) \geq 1.$$

Por otro lado, gracias a la Proposición 13.17, se tiene que,

$$dim(Eig(\lambda^*)) \leq 1$$

por lo que $dim(Eig(\lambda^*)) = 1$. ■

Ejemplo 13.21 *Si deseamos determinar los eigenvalores de la matriz*

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

comenzamos notando que

$$p_A(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 + 1.$$

Ahora bien, como la ecuación $\lambda^2+1 = 0$ no tiene solución en $\mathbb{R}$, diremos que A no posee eignvalores (reales).

Existe una extensión de $\mathbb{R}$, llamada el campo de los **números complejos** (denotado por $\mathbb{C}$), donde la ecuación anterior sí tiene solución: i y $-i$, en donde i es tal que $i^2 = -1$. El famoso Teorema Fundamental del Álgebra garantiza que cualquier polinomio de grado $n \geq 1$ con coeficientes complejos (en particular con coeficientes reales) tiene al menos una raíz compleja. Dado que en este texto únicamente consideramos escalares reales, no abordaremos el caso de eigenvalores y eigenvectores complejos. En consecuencia, hay matrices, como la dada en el Ejemplo 13.21, para las cuales simplemente diremos que no poseen eigenvalores o eigenvectores (reales).

Geométricamente, la matriz A representa una rotación por un ángulo de $\frac{\pi}{2}$, lo cual es un caso particular del Ejemplo 5.55 del Capítulo 5. Es claro que es imposible tener eigenvectores reales ya que todos los vectores no nulos serán rotados (por $\frac{\pi}{2}$ en este ejemplo) al ser multiplicados por la matriz A.

Diagonalización

Desde una perspectiva meramente operativa no hay duda de que las matrices cuadradas más simples son las matrices diagonales. Ya hemos visto, por ejemplo, que es muy sencillo multiplicarlas entre sí, encontrar su inversa (si ésta existe) o calcular su determinante. En esta sección veremos cuándo una matriz $A \in M_{n\times n}$, es "similar", en un sentido que precisaremos más adelante, a una matriz diagonal. Concretamente, probaremos que tal equivalencia es sinónimo de la existencia de una base de $\mathbb{R}^n$ consistente de eigenvectores de A. El siguiente resultado es el primer paso en esta dirección.

Proposición 13.22 *Sea A una matriz de $n \times n$ y sean $\lambda_1, \lambda_2, \ldots, \lambda_k$ eigenvalores distintos de A. Supongamos que $\vec{v}_i$ es un eigenvector para λ_i, $i = 1, 2, \ldots, k$. Entonces el conjunto*

$$\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$$

es LI.

Demostración. La demostración se hará por inducción sobre k, el número de eigenvectores. Si $k = 1$, el conjunto $\{\vec{v}_1\}$ es LI, puesto que $\vec{v}_1 \neq \vec{0}$. Supongamos que el resultado es válido para $k-1$, es decir, asumimos que el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{k-1}\}$ es LI. Para probar que $\{\vec{v}_1, \vec{v}_2 \ldots, \vec{v}_k\}$ es LI, consideramos la expresión

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \ldots + a_k\vec{v}_k = \vec{0}. \qquad (\bigstar)$$

Multiplicando ambos lados de esta expresión por la matriz A y utilizando el hecho de que $A\vec{v}_i = \lambda_i\vec{v}_i$ para $i = 1, 2, \ldots, k$, se tiene que

$$a_1\lambda_1\vec{v}_1 + a_2\lambda_2\vec{v}_2 + \ldots + a_k\lambda_k\vec{v}_k = A\vec{0} = \vec{0}. \qquad (\spadesuit)$$

Multiplicando $(\bigstar)$ por λ_k y restando la expresión resultante a $(\spadesuit)$ se obtiene

$$a_1(\lambda_1 - \lambda_k)\vec{v}_1 + a_2(\lambda_2 - \lambda_k)\vec{v}_2 + \ldots + a_{k-1}(\lambda_{k-1} - \lambda_k)\vec{v}_{k-1} = \vec{0}.$$

Por hipótesis de inducción estos $k-1$ vectores forman un conjunto LI y, por ende, $a_i(\lambda_i - \lambda_k) = 0$ con $\lambda_i \neq \lambda_k$ para $i = 1, 2, \ldots, k-1$. Por lo tanto, $a_i = 0$ para $i = 1, 2, \ldots, k-1$. Esto último y $(\bigstar)$ implican que $a_k\vec{v}_k = \vec{0}$ y como $\vec{v}_k \neq \vec{0}$, necesariamente $a_k = 0$. Concluimos pues que el conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ es LI. ∎

Corolario 13.23 *Sea A una matriz de $n \times n$. Si A tiene exactamente n eigenvalores distintos, entonces existe una base para $\mathbb{R}^n$ que consiste de eigenvectores de A.*

Demostración. Supongamos que $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los n eigenvalores distintos de A y supongamos que para, $i \in \{1, \ldots, n\}$, $\vec{v}_i$ es un eigenvector de A con respecto a λ_i. Por la proposición anterior tenemos ahora que el conjunto

$$\{\vec{v}_1, \vec{v}_2 \ldots, \vec{v}_n\}$$

es LI y, por lo tanto, una base para $\mathbb{R}^n$ que consiste de eigenvectores de A. ■

Ejemplo 13.24 *En el Ejemplo 13.10 notamos que la matriz*

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

tiene dos eigenvalores: $\lambda_1 = 1$ *y* $\lambda_2 = -1$*. También notamos que los conjuntos*

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \; y \; \left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

son, respectivamente, una base para Eig(λ_1) *y Eig*(λ_2)*. Por lo tanto, el Corolario 13.23 afirma que el conjunto*

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

es una base para $\mathbb{R}^2$ *que consiste de eigenvectores de* A*.*

Corolario 13.25 *Sea* A *una matriz de* $n \times n$ *y sean* $\lambda_1, \lambda_2, \ldots, \lambda_k$ *eigenvalores distintos de* A*. Supongamos que* $\vec{v}_i \in$*Eig*(λ_i), $i = 1, 2, \ldots, k$*. Entonces*

$$\vec{v}_1 + \vec{v}_2 + \cdots + \vec{v}_k = \vec{0} \quad syss \quad \vec{v}_i = \vec{0} \text{ para } i = 1, 2, \ldots, k.$$

Demostración. Si algunos de estos vectores fuesen no nulos, sin pérdida de generalidad podemos suponer que $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_j$ con $j \leq k$ son precisamente esos vectores. Entonces

$$\vec{v}_1 + \vec{v}_2 + \cdots + \vec{v}_j = \vec{0}$$

y el vector $\vec{0}$ puede expresarse como una combinación lineal no trivial de eigenvectores con eigenvalores distintos, lo cual contradice la Proposición 13.22. ■

Dada una matriz $A \in M_{n \times n}$, probaremos que dos eigenespacios provenientes de eigenvalores distintos de A no tienen elementos comunes diferentes del vector nulo.

Lema 13.26 *Si* λ_1 *y* λ_2 *son eigenvalores distintos de* A*, entonces*

$$Eig(\lambda_1) \cap Eig(\lambda_2) = \{\vec{0}\}.$$

Demostración. Sea $\vec{v} \in Eig(\lambda_1) \cap Eig(\lambda_2)$, entonces se cumple

$$A\vec{v} = \lambda_1 \vec{v} = \lambda_j \vec{v},$$

por lo tanto

$$(\lambda_1 - \lambda_2)\vec{v} = \vec{0}$$

y como $\lambda_1 - \lambda_2 \neq 0$, necesariamente $\vec{v} = \vec{0}$. ■

Proposición 13.27 *Sea A una matriz de $n \times n$ y supongamos que $\lambda_1, \lambda_2, \ldots, \lambda_k$ son todos los k eigenvalores distintos de A ($k \leq n$). Si n_i es la dimensión de $Eig(\lambda_i)$, entonces $\mathbb{R}^n$ tiene una base que consiste de eigenvectores de A syss*

$$n_1 + n_2 + \cdots n_k = n.$$

En tal caso, si $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$ son, respectivamente, bases para los eigenespacios $Eig(\lambda_1)$, $Eig(\lambda_2), \ldots$, $Eig(\lambda_k)$, entonces

$$\mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots \cup \mathcal{B}_k$$

es una base para $\mathbb{R}^n$ que consiste de eigenvectores de A.

Demostración. Es claro que si

$$n_1 + n_2 + \cdots n_k < n,$$

entonces no puede existir una base para $\mathbb{R}^n$ consistente de eigenvectores (pues esta suma acota el tamaño de cualquier conjunto LI cuyos elementos son eigenvectores de A). Supongamos ahora que

$$n_1 + n_2 + \cdots n_k = n$$

y sean $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$, respectivamente, bases para

$$Eig(\lambda_1), Eig(\lambda_2), \ldots, Eig(\lambda_k).$$

Por el Lema 13.26, estas bases no tienen elementos en común y, por lo tanto, el conjunto

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots \cup \mathcal{B}_k$$

posee exactamente $n_1 + n_2 + \cdots + n_k = n$ elementos. Así pues, la demostración queda completada si verificamos que $\mathcal{B}$ es LI. Supongamos que

$$\mathcal{B}_i = \{v_{i1}, v_{i2}, \ldots, v_{in_i}\}$$

y que

$$\sum_{j=1}^{n_1} a_{1j} v_{1j} + \sum_{j=1}^{n_2} a_{2j} v_{2j} + \cdots + \sum_{j=1}^{n_k} a_{kj} v_{kj} = \vec{0}$$

es una combinación lineal de elementos de $\mathcal{B}$ igual al vector nulo (por conmutatividad y asociatividad siempre podemos agrupar de esta forma). Como

$$\sum_{j=1}^{n_i} a_{ij}v_{ij} \in Eig(\lambda_i),$$

el Corolario 13.25 nos dice que

$$\sum_{j=1}^{n_i} a_{ij}v_{ij} = \vec{0},\ i = 1, 2, \ldots, k.$$

Como $\mathcal{B}_i$ es LI, debe tenerse que $a_{ij} = 0$ para todo $i = 1, 2, \ldots, k$, y $j = 1, 2, \ldots, n_i$, por lo que $\mathcal{B}$ es un conjunto LI. De esta forma, $\mathcal{B}$ es una base para $\mathbb{R}^n$ que consiste de eigenvectores de A. ■

Si analizamos detenidamente el argumento anterior, nos damos cuenta de que la conclusión de que $\mathcal{B}$ es LI depende exclusivamente del hecho de que cada $\mathcal{B}_i$ es un conjunto LI. Por lo tanto, también es cierto que la unión de conjuntos LI contenidos en eigenespacios distintos vuelve a ser un conjunto LI.

Ejemplo 13.28 *En el Ejemplo 13.14 vimos que la matriz*

$$B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

tiene dos eigenvalores: $\lambda_1 = 0$ *y* $\lambda_2 = 1$. *También notamos que los conjuntos*

$$\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\} \ y \ \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

son, respectivamente, bases para $Eig(\lambda_1)$ *y* $Eig(\lambda_2)$. *Por lo tanto, la Proposición 13.27 asegura que el conjunto*

$$\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

es una base para $\mathbb{R}^3$ *que consiste de eigenvectores de* B.

En este punto es importante recordar que una matriz cuadrada es **diagonal** syss todas sus componentes fuera de la diagonal principal son cero. Dada la simplicidad inherente de las matrices diagonales es importante caracterizar cuando una matriz puede "convertirse" en una matriz diagonal en el siguiente sentido.

Definición 13.29 *Sea $A \in M_{n\times n}$. Decimos que A es* ***diagonalizable*** *syss existe una matriz diagonal $D \in M_{n\times n}$ y una matriz invertible P tal que*

$$D = P^{-1}AP.$$

Ejemplo 13.30 *Si B es una matriz diagonal, entonces B es diagonalizable pues $B = I^{-1}BI$.*

En general, dadas dos matrices A, $B \in M_{n\times n}$, decimos que A es **similar**[2] a B syss existe una matriz invertible P tal que

$$B = P^{-1}AP.$$

Por lo tanto, A es diagonalizable syss A es similar a una matriz diagonal. Antes de enunciar el resultado principal de esta sección es conveniente recordar el siguiente hecho observado en la demostración del Lema 5.11.

Lema 13.31 *Sean $B \in M_{m\times n}$ y $\{\vec{e}_1, \ldots, \vec{e}_n\}$ la base estándar de $\mathbb{R}^n$, entonces*

$$B\vec{e}_1 = 1B^1 + 0B^2 + \ldots + 0B^n = B^1$$

$$\vdots$$

$$B\vec{e}_n = 0B^1 + 0B^2 + \ldots + 1B^n = B^n.$$

Teorema 13.32 *Sea $A \in M_{n\times n}$. Entonces A es diagonalizable syss existe una base para $\mathbb{R}^n$ formada por eigenvectores de A.*

Demostración. Por definición, A es diagonalizable syss existen una matriz invertible $P \in M_{n\times n}$ y una matriz diagonal $D \in M_{n\times n}$ tales que $D = P^{-1}AP$ o equivalentemente, si $AP = PD$. Sea

$$D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix},$$

en donde las λ_i no son necesariamente distintas, y notemos que las columnas de D pueden expresarse como

$$D^j = \lambda_j \vec{e}_j, \quad j = 1, 2, \ldots, n.$$

[2]Esta terminología está motivada por el hecho de que si A es similar a B, entonces A y B tienen muchas propiedades en común (ver, por ejemplo, la Proposición 13.38 y el Ejercicio 13.9).

Se tiene entonces que A es diagonalizable syss existe una matriz invertible P y escalares $\lambda_1, \ldots, \lambda_n$ tales que

$$\begin{aligned} AP &= \left[AP^1\ AP^2 \cdots AP^n\right] = PD \\ &= \left[PD^1\ PD^2 \cdots PD^n\right] = \left[P\lambda_1\vec{e}_1 \ \ P\lambda_2\vec{e}_2 \ \cdots \ P\lambda_n\vec{e}_n\right] \\ &= \left[\lambda_1 P\vec{e}_1 \ \ \lambda_2 P\vec{e}_2 \ \cdots \ \lambda_n P\vec{e}_n\right] = \left[\lambda_1 P^1 \ \ \lambda_2 P^2 \ \cdots \ \lambda_n P^n\right], \end{aligned}$$

en donde la última igualdad se sigue del Lema 13.31. De lo anterior se tiene que, para toda $j = 1, 2, \ldots, n$ se cumple

$$AP^j = \lambda_j P^j,$$

o lo que es lo mismo, P^j es un eigenvector con eigenvalor λ_j. Como la matriz P es invertible, sus columnas forman un conjunto LI y por lo tanto, una base para $\mathbb{R}^n$. Concluimos pues que A es diagonalizable syss existe una base

$$\{P^1, \ldots, P^n\}$$

para $\mathbb{R}^n$ que consiste de eigenvectores de A. Más aún, dada esta base, si P es la matriz cuya j-ésima columna es P^j y D es la matriz diagonal tal que la $j - ésima$ entrada de su diagonal principal es el eigenvalor correspondiente a P^j, entonces $D = P^{-1}AP$. ■

Ejemplo 13.33 *En el Ejemplo 13.24 notamos que la matriz*

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

es tal que el conjunto

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

es una base para $\mathbb{R}^2$ que consiste de eigenvectores de A. El Teorema 13.32 afirma que A es diagonalizable. Más aún, gracias al enunciado final de la demostración de dicho teorema, para encontrar P y D tales que D es diagonal y $D = P^{-1}AP$ se procede como sigue. En primer lugar, definimos a P como la matriz cuyas columnas son los elementos de la base de eigenvectores de A. Concretamente,

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Luego, como los eigenvalores de $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ son 1 y -1, respectivamente, definimos a la matriz diagonal como

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

Es fácil ver que la matriz inversa P^{-1} está dada por

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix};$$

por lo tanto,

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

Otra posibilidad es invertir el orden de los eigenvectores definiendo

$$\hat{P} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$$

y, consecuentemente (cuidando de cambiar también el orden de los eigenvalores),

$$\hat{D} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Por supuesto, también se satisface que $\hat{D} = \hat{P}^{-1} A \hat{P}$.

Observación 13.34 *En el ejemplo anterior podríamos haber tomado otra base para los eigenespacios, por ejemplo*

$$\left\{ \begin{bmatrix} 7 \\ 7 \end{bmatrix} \right\}$$

para $Eig(1)$ y

$$\left\{ \begin{bmatrix} -666 \\ 666 \end{bmatrix} \right\}$$

para $Eig(-1)$. La base de $\mathbb{R}^2$ estaría dada por

$$\left\{ \begin{bmatrix} 7 \\ 7 \end{bmatrix}, \begin{bmatrix} -666 \\ 666 \end{bmatrix} \right\}$$

de manera que la matriz invertible estaría dada por

$$\tilde{P} = \begin{bmatrix} 7 & -666 \\ 7 & 666 \end{bmatrix}$$

y la matriz diagonal sería, una vez más

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

Ejemplo 13.35 *En el Ejemplo 13.28 notamos que la matriz*

$$B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

es tal que el conjunto

$$\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

es una base para $\mathbb{R}^3$ *que consiste de eigenvectores de* B. *El Teorema 13.32 afirma que* B *es diagonalizable. Así,*

$$P = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad y \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Al igual que en el caso anterior, existen varias posibilidades para diagonalizar a la matriz en cuestión. Por ejemplo, si

$$\hat{P} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad y \quad \hat{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

se tiene que $\hat{D} = \hat{P}^{-1} B \hat{P}$.

Ejemplo 13.36 *Observamos que la matriz*

$$C = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

cuyo polinomio característico es

$$p(\lambda) = (4 - \lambda)(1 - \lambda)^2,$$

tiene dos eigenvalores: $\lambda_0 = 4$ *y* $\lambda_1 = 1$. *También notamos que los conjuntos*

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\} \; y \; \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

son respectivamente bases para $Eig(\lambda_0)$ *y* $Eig(\lambda_1)$. *Gracias al Corolario 13.27 y dado que*

$$dim(Eig(\lambda_0)) + dim(Eig(\lambda_1)) = 1 + 1 = 2 < 3,$$

tenemos que no existe una base para $\mathbb{R}^3$ *consistente de eigenvectores de* C. *Este hecho, junto con el Teorema 13.32 implican que* C *no es diagonalizable.*

Observación 13.37 *La matriz*

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

dada en el Ejemplo 13.21 no tiene eigenvalores (reales)[3] *y por lo tanto, no es diagonalizable de acuerdo a nustras convenciones, aunque sí lo es dentro de un contexto más amplio que admite la posibilidad de escalares complejos.*

Finalizamos esta sección con la siguiente proposición.

Proposición 13.38 *Si A, B y P son matrices de $n \times n$ tales que $B = P^{-1}AP$, entonces $det(B) = det(A)$.*

Demostración. Basta notar que

$$\begin{aligned} \det(B) &= \det(P^{-1}AP) = \det(P^{-1}) \det(A) \det(P) \\ &= \frac{1}{\det(P)} \det(P) \det(A) = \det(A), \end{aligned}$$

con lo cual queda demostrado el resultado. ■

Ejemplo 13.39 *Si*

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

entonces el polinomio característico de A está dado por

$$p_A(\lambda) = -(\lambda - 1)(\lambda - 3)(\lambda + 2)$$

de manera que los eigenvalores de A son $\lambda_1 = 1$, $\lambda_2 = 3$ y $\lambda_3 = -2$. Sin necesidad de encontrar los eigenespacios correspondientes, sabemos que, $\det(A) = \det(D) = -6$, *donde D es la matriz diagonal*

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

[3] El polinomio característico está dado por $p_A(\lambda) = \lambda^2 + 1$.

Cadenas de Markov (opcional)

En ocasiones es necesario estudiar la evolución en el tiempo de las características de un conjunto o población. Por ejemplo, supongamos que una población puede dividirse en n grupos disjuntos y x_{ti} representa el porcentaje de la población en el tiempo t que pertenece al grupo i. El estado de esta población en t puede representarse por el vector de $\mathbb{R}^n$ dado por

$$\vec{x}_t = \begin{bmatrix} x_{t1} \\ x_{t2} \\ \vdots \\ x_{tn} \end{bmatrix}.$$

Observemos que, por construcción, todas las coordenadas de $\vec{x}_t$ son no negativas y adicionalmente se cumple la igualdad

$$x_{t1} + x_{t2} + \cdots + x_{tn} = 1.$$

Una forma simple en la cual $\vec{x}_t$ puede evolucionar en el tiempo es que el estado futuro $\vec{x}_{t+1}$ dependa únicamente del estado en el pasado inmediato $\vec{x}_t$ y esta dependencia sea lineal. En este caso existirá una matriz $A \in M_{n\times n}$, de tal suerte que

$$\vec{x}_{t+1} = A\vec{x}_t.$$

Observemos que si $t = 0$ la población se encuentra en el estado inicial $\vec{x}_0$, entonces los estados subsecuentes están dados como

$$\begin{aligned} \vec{x}_1 &= A\vec{x}_0, \\ \vec{x}_2 &= A\vec{x}_1 = A(A\vec{x}_0), \\ \vec{x}_3 &= A\vec{x}_2 = A(A(A\vec{x}_0)), \\ &\vdots \end{aligned}$$

y requerimos del producto repetido o potencias de la matriz A para representar la evolución de $\vec{x}_t$.

Notación: denotamos por $A^{(p)}$ al producto de A consigo misma p veces, es decir,

$$A^{(p)} = \underbrace{A \cdots A}_{p\ veces}.$$

Por ejemplo, $A^{(1)}$ es igual a A y $A^{(2)}$ es igual al producto AA. La notación $A^{(p)}$ puede parecer peculiar, pero es simplemente para distinguir a la potencia p de la p-ésima columna de A, denotada en este texto por A^p.

Dado que una matriz de $n \times n$ da lugar a una transformación lineal del espacio $\mathbb{R}^n$ en sí mismo, la potencia de una matriz no es más que la aplicación repetida de esta transformación lineal. La siguiente proposición nos da una relación simple entre los eigenvalores de A y de $A^{(p)}$.

Proposición 13.40 *Sean A una matriz de $n \times n$ y p un entero positivo. Si $\vec{x}$ es un eigenvector de A con eigenvalor λ, entonces $\vec{x}$ es un eigenvector de $A^{(p)}$ con eigenvalor λ^p.*

Demostración. Utilizamos inducción sobre p. Si $p = 1$ el resultado es válido por nuestra hipótesis sobre $\vec{x}$. Sea $p > 1$ y asumimos que el resultado válido para $p - 1$. Entonces,

$$A^{(p)}\vec{x} = (\underbrace{A \cdots A}_{p \text{ veces}})\vec{x} = A(\underbrace{A \cdots A}_{(p-1) \text{ veces}})\vec{x}$$

y por hipótesis de inducción esto puede reescribirse como

$$A(\lambda^{p-1}\vec{x}) = \lambda^{p-1}(A\vec{x}) = \lambda^{p-1}(\lambda\vec{x}) = \lambda^p\vec{x},$$

con lo cual concluimos la demostración. ■

En el siguiente ejemplo ilustramos la evolución en el tiempo del estado de una población.

Ejemplo 13.41 *Tomemos como nuestra población a la clase política mundial (misma que asumimos uniforme en número y comportamiento). El comportamiento diario de cada miembro de esta clase puede ser coherente (C) o errático (E). Se ha observado que únicamente en el 20 % de los casos una conducta coherente es seguida por otra del mismo tipo, en contraste, en el 80 % restante es seguida por un proceder totalmente errático. Asimismo, un comportamiento errático es seguido de otro errático en el 70 % de los casos y en el otro 30 %, por fortuna, es seguido por una conducta coherente. Podemos representar esta situación con la tabla 13.1 en la cual C y E se refieren al tipo de comportamiento: coherente o errático, respectivamente.*

Tabla 13.1

	C	***E***
C	*0.2*	*0.3*
E	*0.8*	*0.7*

Supongamos que en un día inicial dado, el 60 % de los políticos se comporta en forma coherente y el 40 % lo hace en forma errática. Al día siguiente se tiene que como

$$(0.2 \times 0.6) + (0.3 \times 0.4) = 0.24 \text{ } y$$
$$(0.8 \times 0.6) + (0.7 \times 0.4) = 0.76,$$

el 24 % serán coherentes y el 76 % serán erráticos. Observemos que si representamos la distribución o estado inicial de la clase política por el vector

$$\vec{x}_0 = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix},$$

entonces el estado en el periodo siguiente, digamos $\vec{x}_1$, puede obtenerse como

$$\vec{x}_1 = \begin{bmatrix} 0.24 \\ 0.76 \end{bmatrix} = \begin{bmatrix} 0.2 & 0.3 \\ 0.8 & 0.7 \end{bmatrix} \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} = P\vec{x}_0,$$

en donde

$$P = \begin{bmatrix} 0.2 & 0.3 \\ 0.8 & 0.7 \end{bmatrix}.$$

Siguiendo este mismo procedimiento tenemos que

$$\begin{aligned} \vec{x}_2 &= P\vec{x}_1 = P^{(2)}\vec{x}_0, \\ \vec{x}_3 &= P\vec{x}_2 = P^{(3)}\vec{x}_0, \\ &\vdots \\ \vec{x}_k &= P\vec{x}_{k-1} = P^{(k)}\vec{x}_0. \end{aligned}$$

Como

$$\begin{aligned} P^{(2)} &= \begin{bmatrix} 0.28 & 0.27 \\ 0.72 & 0.73 \end{bmatrix}, \\ P^{(3)} &= \begin{bmatrix} 0.272 & 0.273 \\ 0.728 & 0.727 \end{bmatrix}, \\ P^{(4)} &= \begin{bmatrix} 0.2728 & 0.2727 \\ 0.7272 & 0.7273 \end{bmatrix}, \end{aligned}$$

pueden calcularse fácilmente $\vec{x}_2, \vec{x}_3$ y $\vec{x}_4$ como sigue,

$$\begin{aligned} \vec{x}_2 &= P^{(2)}\vec{x}_0 = \begin{bmatrix} 0.276 \\ 0.724 \end{bmatrix}, \\ \vec{x}_3 &= P^{(3)}\vec{x}_0 = \begin{bmatrix} 0.272\,4 \\ 0.727\,6 \end{bmatrix}, \\ \vec{x}_4 &= P^{(4)}\vec{x}_0 = \begin{bmatrix} 0.272\,76 \\ 0.727\,24 \end{bmatrix}. \end{aligned}$$

Intuitivamente, la sucesión $\vec{x}_0, \vec{x}_1, \vec{x}_2, \ldots$ nos proporciona la evolución del comportamiento de la clase política a lo largo del tiempo.

Formalizaremos ahora la situación descrita en este ejemplo.

Definición 13.42 *Un vector $\vec{x} \in \mathbb{R}^n$ se denomina un* ***vector de probabilidad*** *si todas sus entradas son no negativas y suman 1. Asimismo, decimos que una matriz $P \in M_{n\times n}$ es una* ***matriz de transición*** *(también llamada* ***matriz estocástica o matriz de Markov)*** *si sus columnas son vectores de probabilidad.*

Observación 13.43 *En particular, si $\vec{x}_t$ representa el estado de una población en el tiempo t, tal y como lo hemos descrito anteriormente, $\vec{x}_t$ es un vector de probabilidad.*

Ejemplo 13.44 *los vectores*

$$\begin{bmatrix} \frac{1}{7} \\ \frac{6}{7} \end{bmatrix}, \begin{bmatrix} 0.1 \\ 0.2 \\ 0.3 \\ 0.4 \end{bmatrix}$$

son vectores de probabilidad. La matriz

$$\begin{bmatrix} 0 & 0.7 & 0.4 \\ 0.5 & 0.1 & 0.3 \\ 0.5 & 0.2 & 0.3 \end{bmatrix}$$

es una matriz de transición, pero la matriz

$$\begin{bmatrix} 0.6 & 0.75 \\ 0.5 & 0.25 \end{bmatrix}$$

no lo es pues su primera columna no es un vector de probabilidad.

Lema 13.45 *Sea $P \in M_{n\times n}$ una matriz tal que todas sus entradas son no negativas y sea $\vec{1} \in \mathbb{R}^n$, el vector cuyas entradas son todas igual a uno. Entonces P es una matriz de transición syss*

$$P^T\vec{1} = \vec{1}.$$

Demostración. Si P es una matriz de transición, por definición se tiene que se cumplen las n igualdades siguientes para $j = 1, 2, \ldots, n$:

$$p_{1j} + \ p_{2j} + \cdots \ p_{nj} = \sum_{i=1}^{n} p_{ij} = 1,$$

mismas que pueden expresarse en forma vectorial como

$$P^T\vec{1} = \begin{bmatrix} \sum_{i=1}^{n} p_{i1} \\ \vdots \\ \sum_{i=1}^{n} p_{in} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \vec{1},$$

concluyendo así la demostración. ■

La siguiente proposición nos proporciona algunos resultados en relación a los vectores de probabilidad y las matrices de transición.

Proposición 13.46 *Dadas las matrices de transición P y $Q \in M_{n\times n}$ y un vector de probabilidad $\vec{x}_0 \in \mathbb{R}^n$, se cumplen las siguientes.*

1. *PQ es una matriz de transición.*
2. *$P^{(k)}$ es una matriz de transición para todo $k \in \mathbb{N}$.*
3. *Dado $\vec{x}_0$, si definimos $\vec{x}_{i+1} = P\vec{x}_i$, entonces los vectores $\vec{x}_1, \vec{x}_2, \ldots$ son vectores de probabilidad.*
4. *P tiene un eigenvector con eigenvalor $\lambda = 1$.*
5. *Si λ es eigenvalor de P, entonces $|\lambda| \leq 1$.*

Demostración. Para demostrar (1), primero hacemos notar que, por la definición del producto de matrices, dado que todas las entradas de P y Q son no negativas, entonces las entradas de PQ también lo son. Adicionalmente, utilizando el Lema 13.45 se tiene que

$$(PQ)^T\vec{1} = Q^TP^T\vec{1} = Q^T\vec{1} = \vec{1},$$

con lo cual PQ es una matriz de transición.

El punto (2) es inmediato tomando $Q = P$ en (1) y utilizando inducción sobre k.

Para probar (3) recordamos que $\vec{x}_0$ es un vector de probabilidad y para $k = 1, 2, \ldots$ se tiene que

$$\vec{x}_k = P^{(k)}\vec{x}_0 = \begin{bmatrix} P_1^{(k)}\vec{x}_0 \\ P_2^{(k)}\vec{x}_0 \\ \vdots \\ P_n^{(k)}\vec{x}_0 \end{bmatrix}.$$

Por (2), las matrices $P^{(k)}$ son de probabilidad, por lo que cada una de las entradas de los vectores $\vec{x}_k$ son no negativas. Adicionalmente, se cumplen las siguientes igualdades:

$$P_1^{(k)}\vec{x}_0+P_2^{(k)}\vec{x}_0+\cdots+P_n^{(k)}\vec{x}_0 = \left(P_1^{(k)} + P_2^{(k)} + \cdots + P_n^{(k)}\right)\vec{x}_0 = \vec{1}^T\vec{x}_0 = 1,$$

concluyéndose que para $k = 1, 2, \ldots$ los vectores $\vec{x}_k$ son de probabilidad ya que la suma de sus entradas es igual a 1.

Para demostrar (4) observemos que por el Lema 13.45

$$P^T\vec{1} = \vec{1},$$

de manera que $\vec{1}$ es un eigenvector de P^T con eigenvalor uno. Por el Ejercicio 13.2 se tiene que P y P^T comparten los mismos eigenvalores por lo que uno es eigenvalor de P.

Finalmente, sea $\lambda \in \mathbb{R}$ cualquier eigenvalor de la matriz P^T y sea

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in Eig(\lambda).$$

Por definición, se cumple la igualdad

$$\lambda \vec{x} = P^T \vec{x},$$

misma que puede reescribirse como

$$\lambda \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} p_{i1} x_i \\ \vdots \\ \sum_{i=1}^{n} p_{in} x_i \end{bmatrix}.$$

Igualando coordenada a coordenada se tienen las siguientes n igualdades:

$$\begin{aligned} \lambda x_1 &= p_{11} x_1 + p_{21} x_2 + \cdots + p_{n1} x_n, \\ &\vdots \\ \lambda x_n &= p_{1n} x_1 + p_{2n} x_2 + \cdots + p_{nn} x_n. \end{aligned}$$

Sea x_k tal que $|x_k| = \max\limits_{i=1,\ldots,n} \{|x_i|\}$, es decir, x_k es la coordenada de $\vec{x}$ de "tamaño" máximo, misma que es distinta del cero ya que $\vec{x} \neq \vec{0}$ por tratarse de un eigenvector. Tomando valores absolutos en la $k - \acute{e}sima$ igualdad de arriba tenemos que[4]

$$\begin{aligned} |\lambda| \, |x_k| = |\lambda x_k| &= |p_{1k} x_1 + p_{2k} x_2 + \cdots + p_{nk} x_n| \\ &\leq |p_{1k} x_1| + |p_{2k} x_2| + \cdots + |p_{nk} x_n| \\ &= |p_{1k}| \, |x_1| + |p_{2k}| \, |x_2| + \cdots + |p_{nk}| \, |x_n| \\ &\leq |p_{1k}| \, |x_k| + |p_{2k}| \, |x_k| + \cdots + |p_{nk}| \, |x_k| \\ &= (p_{1k} + p_{2k} + \cdots + p_{nk}) \, |x_k| \\ &= |x_k| \end{aligned}$$

y podemos concluir que $|\lambda| \leq 1$. ■

[4] Aquí se utiliza que dados $a, b \in \mathbb{R}$ se cumple la desigualdad

$$|a + b| \leq |a| + |b|,$$

lo cual es el caso particular, para $n = 1$, de la Desigualdad del triángulo dada en la Proposición 14.15 del siguiente capítulo.

Definición 13.47 *Dados $P \in M_{n\times n}$ una matriz de transición y $\vec{x}_0 \in \mathbb{R}^n$ un vector de probabilidad o estado inicial, se define la siguiente sucesión de vectores de probabilidad:*

$$\vec{x}_{t+1} = P\vec{x}_t, \ t = 0, 1, 2, \ldots$$

A esta sucesión se le conoce como la ***cadena de Markov*** *generada por P y $\vec{x}_0$.*

Ejemplo 13.48 *Si tomamos P y $\vec{x}_0$ como en el Ejemplo 13.41, la sucesión $\{\vec{x}_0, \vec{x}_1, \vec{x}_2, \ldots\}$ es una cadena de Markov.*

Ejemplo 13.49 *Si la matriz de transición P y $\vec{x}_0$ están dados por*

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ \vec{x}_0 = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix},$$

entonces la cadena de Markov asociada es la sucesión alternante

$$\left\{ \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}, \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix} \cdots \right\}.$$

En virtud de los dos ejemplos anteriores, una pregunta natural es si dada una matriz de transición P y un estado inicial $\vec{x}_0$ sucede que los vectores de la cadena de Markov se estabilizan, es decir, que la sucesión $\{\vec{x}_0, \vec{x}_1, \vec{x}_2, \ldots\}$ converge[5] a algún vector $\vec{x}$.

El futuro es estable (a veces)

Definición 13.50 *Dada una matriz de transición $P \in M_{n\times n}$, decimos que $\vec{x} \in \mathbb{R}^n$ es un* ***vector estacionario de*** *P syss $\vec{x}$ es un vector de probabilidad con la propiedad adicional de que es eigenvector de P con respecto a uno, es decir, tal que*

$$P\vec{x} = \vec{x}.$$

[5] Formalmente se tiene que

$$\lim_{k\to\infty} \vec{x}_k = \lim_{k\to\infty} P^{(k)}\vec{x}_0 = \vec{x},$$

sin embargo para definir el límite de una sucesión de vectores es necesario tener el concepto de distancia entre vectores, mismo que se define en el capítulo siguiente. Aquí nos conformaremos con la noción intuitiva de límite.

Observación 13.51 *La Proposición 13.46 nos dice que toda matriz de transición posee un eigenvector con eigenvalor uno por lo que* $\dim Eig(1) \geq 1$. *Además, una variante del famoso Teorema de Perron-Frobenius garantiza que para cualquier matriz de transición, el eigenespacio correspondiente posee una base de vectores cuyas entradas son no negativas. Si* $\vec{y}$ *es algún elemento de esta base, entonces el eigenvector*

$$\vec{x} = \left(\frac{1}{x_1 + x_2 + \cdots + x_n} \right) \vec{y}$$

es un vector estacionario de P. *De aquí se sigue que* P *tiene un único vector estacionario syss la dimensión de su correspondiente eigenespacio* $Eig(1)$ *es igual a uno.*

Dada una matriz de transición $P \in M_{n\times n}$, nos interesa saber cuando ésta posee un vector estacionario $\vec{x}$ y si éste se comporta como un "atractor" en el sentido de que, para cualquier estado inicial $\vec{x}_0$, la cadena de Markov

$$\{\vec{x}_0, \vec{x}_1, \vec{x}_2, \ldots\}$$

siempre converge a $\vec{x}$. Debido a que cualesquiera dos vectores estacionarios $\vec{x}$ y $\vec{y}$ de P dan lugar a dos cadenas de Markov constantes, una condición necesaria para la existencia de un vector estacionario atractor es que, tal y como se mencionó en la Observación 13.51, la dimensión de $Eig(1)$ sea igual a uno.

Ejemplo 13.52 *Es sencillo verificar que la matriz de transición*

$$Q = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix}$$

es tal que

$$Eig(1) = span \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Como $\dim Eig(1) > 1$, Q *no posee ningún atractor. Más aún, todos los vectores de la forma*

$$\begin{bmatrix} 0 \\ p \\ 1-p \end{bmatrix},$$

con $p \in [0, 1]$, *son vectores estacionarios.*

La condición $\dim Eig(1) = 1$ es necesaria para la existencia de un atractor pero no es suficiente, como se ilustra en el siguiente ejemplo.

Ejemplo 13.53 *La matriz*

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

es tal que

$$\dim Eig(1) = span \left\{ \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \right\}$$

y

$$\vec{x} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

es el único vector estacionario. Sin embargo, si tomamos $p \in [0,1]$, *con* $p \neq \frac{1}{2}$ *y consideramos el estado inicial dado por*

$$\vec{x}_0 = \begin{bmatrix} p \\ 1-p \end{bmatrix},$$

entonces la cadena de Markov asociada es la sucesión alternante

$$\left\{ \begin{bmatrix} p \\ 1-p \end{bmatrix}, \begin{bmatrix} 1-p \\ p \end{bmatrix}, \begin{bmatrix} p \\ 1-p \end{bmatrix}, \ldots \right\},$$

misma que no converge a $\vec{x}$ *o a ningún otro vector, de manera que* $\vec{x}$ *no es un atractor.*

En lo que sigue introduciremos una clase de matrices de transición para las cuales sí es posible garantizar la existencia de un atractor

Ejemplo 13.54 *Consideremos la matriz*

$$\begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix},$$

la cual tiene a

$$\begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$$

como su único vector estacionario. En este caso, dado cualquier estado inicial $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$, *la cadena de Markov asociada siempre converge al atractor*

$$\begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}.$$

Por ejemplo, para el estado inicial $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$ *se tiene que*

$$\begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix}^{(10)} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0.338\,04 \\ 0.661\,96 \end{bmatrix}$$

y para el $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *también puede comprobarse que*

$$\begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix}^{(14)} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.337\,85 \\ 0.662\,15 \end{bmatrix}.$$

Definición 13.55 *Sea $P \in M_{n\times n}$ una matriz de transición. Decimos que P es **regular** syss para algún número natural $k \geq 1$ la matriz $P^{(k)}$ es tal que todas sus entradas son estrictamente positivas.*

Nótese que las matrices de los Ejemplos 13.41 y 13.54 son regulares pues podemos tomar $k = 1$.

Ejemplo 13.56 *La matriz*

$$\begin{bmatrix} 0 & 0.5 \\ 1 & 0.5 \end{bmatrix}$$

es regular ya que tomando $k = 2$ se tiene que

$$\begin{bmatrix} 0 & 0.5 \\ 1 & 0.5 \end{bmatrix} \begin{bmatrix} 0 & 0.5 \\ 1 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.25 \\ 0.5 & 0.75 \end{bmatrix}.$$

Ejemplo 13.57 *La matriz*

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

no es regular puesto que para cualquier $k \geq 1$,

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{(k)} = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \textit{si } k \textit{ es par,} \\ \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & \textit{si } k \textit{ es impar.} \end{cases}$$

Ejemplo 13.58 *La matriz*

$$\begin{bmatrix} \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix}$$

no es regular pues sus potencias siempre tendrán entradas nulas.

El siguiente teorema, cuya demostración omitimos, dice que las matrices de transición regulares poseen un único vector estacionario atractor.

Teorema 13.59 *Si $P \in M_{n\times n}$ es una matriz de transición regular, entonces P tiene un único vector estacionario atractor y se cumplen los siguientes:*

(a) Existe un único vector de probabilidad $\vec{x}$ con la propiedad de que $P\vec{x} = \vec{x}$.

(b) Si $\vec{x}_0 \in \mathbb{R}^n$ es cualquier vector de probabilidad inicial, la cadena de Markov generada por P y $\vec{x}_0$ converge a $\vec{x}$, independientemente de la elección de $\vec{x}_0$.

Ejemplo 13.60 *Para encontrar el vector estacionario atractor en el caso de la matriz regular $P \in M_{2\times 2}$ del Ejemplo 13.41, encontramos al único vector de probabilidad en el eigenespacio*

$$\begin{aligned} Eig(1) &= Null\begin{bmatrix} 0.2-1 & 0.3 \\ 0.8 & 0.7-1 \end{bmatrix} \\ &= Null\begin{bmatrix} -0.8 & 0.3 \\ 0.8 & -0.3 \end{bmatrix}. \end{aligned}$$

Resolviendo el correspondiente sistema homogéneo tenemos que

$$\begin{aligned} Eig(1) &= \left\{ b\begin{bmatrix} \frac{3}{8} \\ 1 \end{bmatrix} \in \mathbb{R}^2 \mid b \in \mathbb{R} \right\} \\ &= \left\{ s\begin{bmatrix} 3 \\ 8 \end{bmatrix} \in \mathbb{R}^2 \mid s \in \mathbb{R} \right\}. \end{aligned}$$

Entonces

$$\vec{x} = \frac{1}{8+3}\begin{bmatrix} 3 \\ 8 \end{bmatrix} = \begin{bmatrix} \frac{3}{11} \\ \frac{8}{11} \end{bmatrix}$$

es el vector estacionario atractor. Esto quiere decir que, sin importar cuál es la distribución inicial de políticos erráticos y coherentes, conforme avanza el tiempo, el porcentaje de políticos erráticos se acerca a $\frac{8}{11} = 0.\overline{72}$ y el de coherentes será únicamente $\frac{3}{11} = 0.\overline{27}$.

PageRank

En los inicios del internet las herramientas de búsqueda eran bastante deficientes. Lo usual era que cualquier consulta realizada arrojara una lista de páginas *web*, la mayoría de las cuales poco o nada tenían que ver con nuestra indagatoria original. Esta situación cambió radicalmente en 1998 cuando Sergey Brin, Larry Page, Rajeev Motwani y Terry Winograd utilizaron su conocimiento del álgebra lineal para crear el algoritmo conocido como *PageRank*[6]. El propósito del algoritmo es

[6] El algoritmo se dio a conocer en el artículo “The Pagerank Citation Ranking: Bringing Order to the Web”, Technical Report, Stanford infoLab, enero 1998.

asignar a cada página un grado de "importancia" para ordenarlas de acuerdo a su relevancia. A continuación se describe una versión simplificada de este proceso.

Lo que se desea es que el grado de importancia o popularidad de una página *web* sea proporcional a la cantidad de páginas que hacen referencia a ella. Así, si la página i contiene ligas a otras k páginas, el algoritmo prescribe que la página i transfiere un $\frac{1}{k} - \acute{e}simo$ de su importancia a cada una de estas páginas. Por ejemplo, si pensamos en las páginas que hablan de tenis, las páginas de Rafael Nadal y Serena Williams tendrán un alto grado de importancia (casi cualquier página que hable de tenis hará referencia a ellas); en cambio el grado de importancia de las páginas de aquellos tenistas que no están entre los 100 primeros del mundo será bajo, la razón es que habrá muy pocas páginas que tengan ligas a ellas.

La medida de este grado de importancia es un número entre cero y uno que puede interpretarse de la siguiente forma: si un internauta explora aleatoriamente todas las páginas *web* en existencia, la probabilidad de que este explorador del mundo digital visite una página dada es, precisamente, el grado de importancia de la misma. Podemos pensar al conjunto de páginas *web* como los nodos de una gráfica unidos por flechas o aristas dirigidas. Una flecha de la página i a la j significa que la página i contiene una liga o referencia a la página j. Para ilustrar lo anterior, imaginemos un universo de 4 páginas web como se muestra en la Figura 13.2.

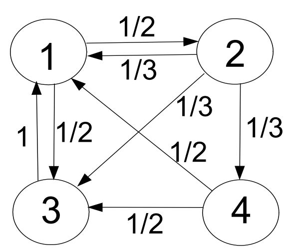


Figura 13.2: Referencias entre 4 páginas *web*.

En este caso, la página 1 contiene ligas a las páginas 2 y 3, la 2 hace referencia a todas las demás, la 3 hace referencia a la 1 y la página 4 contiene ligas a la 1 y a la 3. El número $p \leq 1$, asignado a cada arista de esta gráfica, representa la importancia que transfiere una página a otra. Esta situación puede caracterizarse mediante la siguiente matriz de transición, a la que llamamos **matriz de importancia**, en la cual la entrada a_{ij} representa la importancia que recibe la página i de la página j.

$$A = \begin{bmatrix} 0 & \frac{1}{3} & 1 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 & 0 \end{bmatrix}.$$

Supongamos que inicialmente se la asigna la misma importancia a cada una de las páginas; entonces, el internauta que las visita aleatoriamente realiza su primer clic del ratón en alguna de ellas con la misma probabilidad. Así, el estado inicial de este sistema de búsqueda está dado por el vector de probabilidad

$$\vec{x}_0 = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix}.$$

Para el segundo clic aleatorio del internauta, las probabilidades de visitar las páginas son

$$\vec{x}_1 = \begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix} = \begin{bmatrix} 0.458\,3 \\ 0.125 \\ 0.333\,3 \\ 0.0833 \end{bmatrix}.$$

Para el tercer y cuarto clic se tiene, respectivamente,

$$\vec{x}_2 = \begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 & 0 \end{bmatrix}^{(2)} \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix} = \begin{bmatrix} 0.416\,7 \\ 0.229\,2 \\ 0.312\,5 \\ 0.0442 \end{bmatrix},$$

$$\vec{x}_3 = \begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 & 0 \end{bmatrix}^{(3)} \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix} = \begin{bmatrix} 0.409\,7 \\ 0.208\,3 \\ 0.305\,6 \\ 0.0763 \end{bmatrix}$$

y así sucesivamente.

Puede verificarse que

$$A^{(5)} = \begin{bmatrix} \frac{1}{3} & \frac{89}{216} & \frac{19}{36} & \frac{7}{18} \\ \frac{19}{72} & \frac{5}{24} & \frac{1}{8} & \frac{11}{48} \\ \frac{13}{36} & \frac{17}{54} & \frac{17}{72} & \frac{47}{144} \\ \frac{1}{24} & \frac{7}{108} & \frac{1}{9} & \frac{1}{18} \end{bmatrix},$$

de manera que la matriz A es regular y la cadena de Markov $\{\vec{x}_0, \vec{x}_1, \vec{x}_2, \ldots\}$ converge al vector estacionario atractor dado por

$$\vec{x}^* = \begin{bmatrix} 0.413\,8 \\ 0.206\,9 \\ 0.310\,3 \\ 0.069 \end{bmatrix}.$$

Al vector $\vec{x}^*$ se le conoce como el **vector PageRank** de nuestro pequeño internet conformado por 4 páginas *web*. Así, el internauta que vive dando clics aleatorios para accesar estas 4 páginas, visita las páginas 1,2,3 y 4 con probabilidades de 0.3478, 0.2609, 0.3043 y 0.0869, respectivamente. La probabilidad de cada página refleja su importancia o popularidad dentro de la red. En este ejemplo la página 1 es la más popular, seguida por la 3, la 2 y la 4.

En nuestro internet simplificado, la matriz de importancia es una matriz regular, pero evidentemente esto podría no suceder. Por ejemplo, este sería el caso si existe un renglón o una columna de ceros. Esto sucede si una página no es referida por ninguna otra o bien si ésta no contiene referencias a las demás páginas. El algoritmo puede modificarse de manera que siempre se tenga una matriz regular que permita la obtención del vector PageRank. Asimismo, en la realidad existen miles de millones de páginas *web* y los cálculos para obtener el vector PageRank están lejos de ser triviales; no obstante, la idea detrás de este algoritmo puede entenderse con álgebra lineal elemental.

Ejercicios

Ejercicio 13.1 *Encontrar el polinomio característico, los eigenvalores y una base para cada uno de los eigenespacios de las matrices que se presentan a continuación.*

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

$$F = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix},\ G = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix},$$
$$H = \begin{bmatrix} 1 & 2 & -3 & 1 \\ 2 & 1 & 7 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

Ejercicio 13.2 *Sea $A \in M_{n\times n}$. Probar que el polinomio característico de A es igual al polinomio característico de A^T (y que por tanto, A y A^T tienen los mismos eigenvalores).*

Ejercicio 13.3 *Sea $A \in M_{n\times n}$ una matriz invertible. Probar que si λ es un eigenvalor de A, entonces $\frac{1}{\lambda}$ es un eigenvalor de A^{-1}.*

Ejercicio 13.4 *Sean A y B matrices de $n \times n$. Probar que si $\vec{v}$ es un eigenvector de A con eigenvalor λ y $\vec{v}$ es un eigenvector de B con eigenvalor μ, entonces $\vec{v}$ es un eigenvector de AB con eigenvalor $\mu\lambda$.*

Ejercicio 13.5 *Probar (en general) o refutar (con un contraejemplo) los siguientes enunciados (en cada caso se debe indicar claramente "verdadero" o "falso"). Recordar que en este texto, "eigenvalor" se refiere siempre a eigenvalores reales.*

a. Sean A y B matrices de $n \times n$. Si B es la FER de A, entonces A y B tienen los mismos eigenvalores.

b. Si A no tiene eigenvalores, entonces A es invertible.

c. Si A es invertible, entonces A no tiene eigenvalores.

d. Si λ es un eigenvalor de A, entonces 0 es un eigenvalor de $A-\lambda I$.

e. *Si A es una matriz triangular, entonces λ es un eigenvalor de A syss λ es una componente de la diagonal principal de A.*

f. *Si $A \in M_{2\times 2}$ es una matriz triangular inferior invertible, entonces A es diagonalizable.*

g. *Si A es una matriz de $n \times n$ y A tiene exactamente m eigenvalores con $m < n$, entonces A no es diagonalizable.*

Ejercicio 13.6 *Determinar si cada una de las matrices del Ejercicio 13.1 es diagonalizable. En caso afirmativo, encontrar una matriz diagonal y una matriz invertible que lo testifiquen.*

Ejercicio 13.7 *Sea $A \in M_{n\times n}$. Probar que si A es diagonalizable, entonces A^T es diagonalizable.*

Ejercicio 13.8 *Sea $A \in M_{n\times n}$ una matriz invertible. Probar que si A es diagonalizable, entonces A^{-1} es diagonalizable.*

Ejercicio 13.9 *Mostrar que si A, B y $P \in M_{n\times n}$ son matrices tales que $B = P^{-1}AP$, entonces $p_B(\lambda) = p_A(\lambda)$.*

Ejercicio 13.10 *a.* *Mostrar que si $A, D, P \in M_{n\times n}$ son matrices tales que $A = PDP^{-1}$, entonces $A^{(m)} = PD^{(m)}P^{-1}$ para toda $m \geq 1$.*

b. *Encontrar $A^{(11)}$ y $B^{(8)}$, en donde*

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \; B = \begin{bmatrix} 2 & 1 & 3 \\ 0 & 1 & 6 \\ 0 & 0 & -1 \end{bmatrix}.$$

Ejercicios de cadenas de Markov (opcional)

Ejercicio 13.11 *Determinar si los siguientes son vectores de probabilidad.*

$$a) \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}, \; b) \begin{bmatrix} 0.1 \\ 0.2 \\ 0.6 \end{bmatrix}, \; c) \begin{bmatrix} \frac{2}{7} \\ -\frac{3}{7} \\ \frac{8}{7} \end{bmatrix}, \; d) \begin{bmatrix} \frac{1}{n} \\ \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{bmatrix} \in \mathbb{R}^n.$$

Ejercicio 13.12 *Determinar si las siguientes son matrices de transición.*

$$a) \begin{bmatrix} 0.2 & 0 \\ 0.8 & 1 \end{bmatrix}, \; b) \begin{bmatrix} \frac{3}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{11} & \frac{4}{11} & \frac{5}{11} \end{bmatrix}, \; c) \begin{bmatrix} 0.2 & 0.1 & 0 \\ 0.3 & 0.1 & 1 \\ 0.5 & 0.8 & 0 \end{bmatrix}.$$

Ejercicio 13.13 *Calcular* $P^{(2)}$ *y* $P^{(3)}$ *si*

$$P = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Ejercicio 13.14 *Determinar si las siguientes matrices de transición son regulares.*

$$a) \begin{bmatrix} 1 & \frac{1}{4} \\ 0 & \frac{3}{4} \end{bmatrix}, \; b) \begin{bmatrix} \frac{1}{4} & 1 \\ \frac{3}{4} & 0 \end{bmatrix},$$
$$c) \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.9 & 0.8 & 0.7 \\ 0 & 0 & 0.2 \end{bmatrix}, \; d) \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{2}{3} & 0 \end{bmatrix}.$$

Ejercicio 13.15 *Encontrar una matriz de transición* $P \in M_{2\times 2}$ *para la cual* $\begin{bmatrix} \frac{1}{4} \\ \frac{3}{4} \end{bmatrix}$ *sea su único vector estacionario atractor.*

Ejercicio 13.16 *Encontrar una matriz de transición* $P \in M_{3\times 3}$ *y un vector de probabilidad inicial* $\vec{x}_0$*, tales que la cadena de Markov asociada no sea convergente.*

Ejercicio 13.17 *Probar que la matriz*

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

es regular y encontrar su único vector estacionario atractor.

Ejercicio 13.18 *Demuestra que si* $P \in M_{n\times n}$ *es una matriz de transición regular, entonces para* $k \to \infty$ *las columnas de la matriz* $P^{(k)}$ *convergen al único vector estacionario. (Sugerencia: si* $\{\vec{e}_1, \ldots, \vec{e}_n\}$ *es la base estándar de* $\mathbb{R}^n$*, analizar el comportamiento de la sucesión* $P\vec{e}_j, P^{(2)}\vec{e}_j, P^{(3)}\vec{e}_j, \ldots$ *).*

Ejercicio 13.19 *Supongamos que simplificamos los estados del tiempo en días soleados (o despejados) y días nublados. En la Ciudad de México, si hay un día soleado, la probabilidad de que el día siguiente sea soleado es del 80 % y si hay un día nublado el día siguiente será nublado con un probabilidad del 25 %. Asimismo, en la ciudad de Edimburgo estas mismas probabilidades son del 30 % y del 90 %, respectivamente.*

a. *Encontrar las matrices de transición (regulares) que representan la evolución de los días soleados y nublados para ambas ciudades.*

b. Si hoy es un día soleado en la Ciudad de México, ¿cuál es la probabilidad de que mañana sea un día nublado?

c. Si hoy es un día nublado en la ciudad de Edinburgo, ¿cuál es la probabilidad de que en dos días siga estando nublado?

d. Encontrar los vectores estacionarios atractores para ambas matrices.

Ejercicio 13.20 *Durante cualquier día del año escolar, los estudiantes pueden no asistir a la escuela por alguna causa. Se ha visto que si un estudiante asiste a la escuela en un día dado, la probabilidad de que asista el siguiente día (hábil) es del 90 %. Asimismo, si un estudiante falta a la escuela un día, la probabilidad de que falte al día siguiente es del 60 %.*

a. Encontrar la matriz de transición (regular) que representa la evolución de las poblaciones de estudiantes que asisten o faltan a clase a lo largo del tiempo.

b. Encontrar el vector estacionario atractor de dicha matriz.

c. Si al inicio del año escolar el 10 % de los alumnos faltaron a la escuela, que porcentaje de alumnos se esperaría que faltaran al final del año escolar.

Capítulo 14

Longitud y dirección

Introducción

Dado un vector $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$, podemos visualizarlo en el plano como en la Figura 14.1.

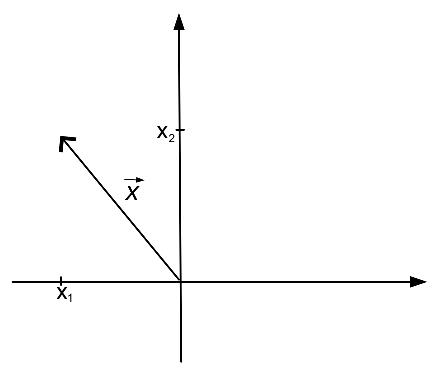


Figura 14.1: Vector $\vec{x}$ en $\mathbb{R}^2$.

Utilizando el Teorema de Pitágoras, resulta natural definir la **norma** (o longitud) de $\vec{x}$, denotada por $||\vec{x}||$, como

$$||\vec{x}|| = \sqrt{x_1^2 + x_2^2}.$$

Por lo tanto,

$$||x||^2 = x_1^2 + x_2^2.$$

De esta forma, $||\vec{x}||$ coincide con el "tamaño" del vector $\vec{x}$. Dados dos puntos A y B en el plano, correpondientes a los puntos finales de los vectores $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ y $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ en $\mathbb{R}^2$, puede definirse fácilmente la distancia entre los puntos A y B como la longitud del segmento que los conecta. Por lo tanto, la distancia entre A y B es igual a

$$\begin{aligned} ||\vec{x} - \vec{y}|| &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \\ &= \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2} = ||\vec{y} - \vec{x}||. \end{aligned}$$

En particular, la **distancia** entre A y B es igual a la distancia entre B y A. Las consideraciones anteriores se ilustran en la Figura 14.2.

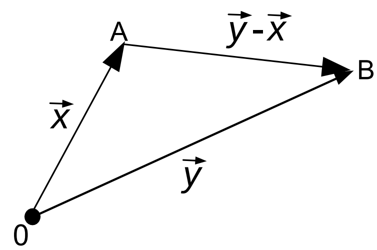


Figura 14.2: Distancia entre dos puntos.

Ahora generalizaremos estas nociones geométricas en el contexto de $\mathbb{R}^n$. Recordemos que el producto punto entre los vectores

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \ \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n$$

está dado por

$$\vec{x} \cdot \vec{y} = \sum_{i=1}^{n} x_i y_i,$$

o si pensamos a los vectores como matrices, $\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y}$. Es inmediato comprobar que si $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$, entonces

$$\vec{x} \cdot \vec{x} = x_1 x_1 + x_2 x_2 = x_1^2 + x_2^2 = ||\vec{x}||^2.$$

Esto nos sugiere una forma de abordar las nociones de norma y distancia en $\mathbb{R}^n$ como sigue.

Definición 14.1 *Sean $\vec{u}$ y $\vec{v}$ vectores de $\mathbb{R}^n$.*

- *La* ***norma***[1] *de* $\vec{u}$ *se define como*

$$||\vec{u}|| = \sqrt{\vec{u} \cdot \vec{u}}.$$

- *La* ***distancia*** *entre los vectores* $\vec{u}$ *y* $\vec{v}$

$$d(\vec{u}, \vec{v}) = ||\vec{v} - \vec{u}||.$$

Ejemplo 14.2 *Sean*

$$\vec{u} = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix} \; y \; \vec{v} = \begin{bmatrix} -1 \\ -5 \\ -3 \\ 0 \end{bmatrix},$$

entonces

$$\vec{u} \cdot \vec{u} = 1^2 + 2^2 + (-3)^2 + 4^2 = 30$$

y se tiene

$$||\vec{u}|| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{30}.$$

Asimismo, la distancia entre los puntos determinados por $\vec{u}$ *y* $\vec{v}$ *es igual a*

$$\begin{aligned} ||\vec{v} - \vec{u}|| &= \left\| \begin{bmatrix} -1 \\ -5 \\ -3 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -2 \\ -7 \\ 0 \\ -4 \end{bmatrix} \right\| \\ &= \sqrt{2^2 + 7^2 + 4^2} = \sqrt{69}. \end{aligned}$$

Observación 14.3 *Si* $\vec{v} \in \mathbb{R}$, *entonces* $\vec{v}$ *puede identificarse con un número real y su norma coincide con el valor absoluto. Por ejemplo, si* $\vec{v} = -5$, *entonces*

$$||\vec{v}|| = \sqrt{(-5)^2} = |-5| = 5.$$

Recordemos ahora las siguientes propiedades del producto punto mencionadas previamente en el Capítulo 6.

Proposición 14.4 *Si* $\vec{u}$, $\vec{v}$ *y* $\vec{w}$ *son vectores de* $\mathbb{R}^n$ *y* $\beta \in \mathbb{R}$, *entonces*

1. $\vec{u} \cdot \vec{u} = u_1^2 + \ldots u_n^2$ *y por lo tanto,* $\vec{u} \cdot \vec{u} = 0$ *syss* $\vec{u} = \vec{0}$.

2. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$.

3. $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$.

[1] También conocida como **magnitud**, **longitud** o **módulo** de $\vec{u}$.

4. $(\beta\vec{u})\cdot\vec{v} = \beta(\vec{u}\cdot\vec{v}) = \vec{u}\cdot(\beta\vec{v})$.

Corolario 14.5 *Si $\vec{u}$ es un vector de $\mathbb{R}^n$ y $\beta\in\mathbb{R}$, entonces*

- $||\vec{u}|| = \sqrt{\vec{u}\cdot\vec{u}} \geq 0$ *y* $||\vec{u}|| = 0$ *syss* $\vec{u} = \vec{0}$.
- $||\beta\vec{u}|| = \sqrt{(\beta\vec{u})\cdot(\beta\vec{u})} = \sqrt{\beta^2(\vec{u}\cdot\vec{u})} = \sqrt{\beta^2}\sqrt{\vec{u}\cdot\vec{u}} = (|\beta|)\,||\vec{u}||$. *En particular,*

  $$||\vec{v}-\vec{u}|| = (|-1|)\,||\vec{v}-\vec{u}|| = ||(-1)(\vec{v}-\vec{u})|| = ||\vec{u}-\vec{v}||,$$

  de manera que

  $$d(\vec{u},\vec{v}) = ||\vec{v}-\vec{u}|| = ||\vec{u}-\vec{v}|| = d(\vec{v},\vec{u}).$$

Ejemplo 14.6

$$\left\|\begin{bmatrix}-1000\\-2000\\0\end{bmatrix}\right\| = \left\|(-1000)\begin{bmatrix}1\\2\\0\end{bmatrix}\right\| = (|-1000|)\left\|\begin{bmatrix}1\\2\\0\end{bmatrix}\right\| = 1000\sqrt{5}.$$

En lo que sigue, extenderemos la noción de ortogonalidad (perpendicularidad) entre vectores y de hecho, el concepto global de medición de ángulos en $\mathbb{R}^n$.

Ortogonalidad

Supongamos que los vectores $\vec{x},\vec{y}$ y $\vec{x}\pm\vec{y}$ en $\mathbb{R}^2$, forman un triángulo rectángulo de acuerdo a alguna de las configuraciones geométricas de la Figura 14.3.

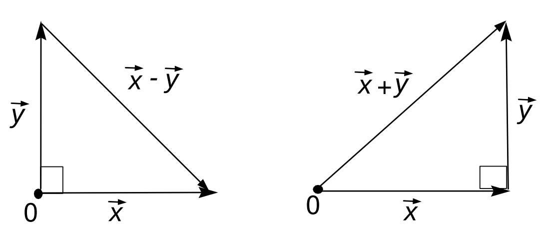


Figura 14.3: Triángulo rectángulo.

Por un lado, el Teorema de Pitágoras nos dice que

$$\|\vec{x}\pm\vec{y}\|^2 = \|x\|^2 + \|y\|^2\,.$$

Por otro lado, gracias a las propiedades del producto punto también se tiene que

$$\begin{aligned}\|\vec{x} \pm \vec{y}\|^2 &= (\vec{x} \pm \vec{y}) \cdot (\vec{x} \pm \vec{y}) \\ &= (\vec{x} \cdot \vec{x}) \pm (\vec{x} \cdot \vec{y}) \pm (\vec{y} \cdot \vec{x}) + (\vec{y} \cdot \vec{y}) \\ &= \|\vec{x}\|^2 \pm 2\,(\vec{x} \cdot \vec{y}) + \|\vec{y}\|^2 ,\end{aligned}$$

por lo que necesariamente se cumple que $\vec{x} \cdot \vec{y} = 0$.

Definición 14.7 *Decimos que los vectores $\vec{x}, \vec{y} \in \mathbb{R}^n$ son* ***ortogonales*** *o* ***perpendiculares,*** *denotado por $\vec{x} \perp \vec{y}$, syss*

$$\vec{x} \cdot \vec{y} = 0.$$

Utilizando nuevamente las propiedades del producto punto (ahora en el contexto de $\mathbb{R}^n$) junto con esta definición puramente algebraica, hemos llegado a la siguiente generalización del Teorema de Pitágoras.

Proposición 14.8 *Si $\vec{x}, \vec{y} \in \mathbb{R}^n$ son ortogonales, entonces*

$$\|\vec{x} \pm \vec{y}\|^2 = \|x\|^2 + \|y\|^2 .$$

Ejemplo 14.9 *Los vectores*

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

son ortogonales ya que

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0.$$

Análogamente, en el contexto de $\mathbb{R}^n$, los vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ son ortogonales dos a dos ya que el producto punto de cualesquiera dos vectores distintos es igual a cero.

Ejemplo 14.10 *Los vectores*

$$\begin{bmatrix} -1 \\ 2 \\ -3 \\ \pi \end{bmatrix} \ y \ \begin{bmatrix} \pi \\ 3 \\ 2 \\ 1 \end{bmatrix}$$

son ortogonales pues su producto punto es $-\pi + 6 - 6 + \pi = 0$.

Ejemplo 14.11 *Los vectores*

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \ y \ \begin{bmatrix} -\pi \\ 7 \\ 666 \end{bmatrix}$$

son ortogonales pues $0 + 0 + 0 = 0$.

Ahora veremos que si $\vec{u}$ y $\vec{v}$ son vectores en $\mathbb{R}^n$ con $\vec{u} \neq \vec{0}$, entonces $\vec{v}$ puede escribirse como una suma de vectores en la cual uno de los sumandos es un múltiplo escalar de $\vec{u}$ y el otro es ortogonal a $\vec{u}$, tal y como sugiere la siguiente figura.

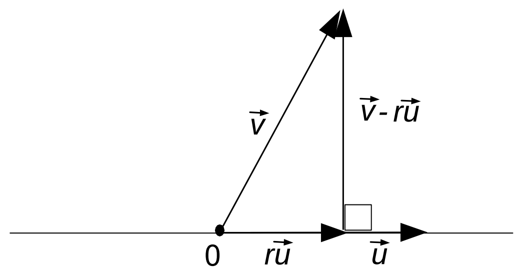


Figura 14.4: Descomposición del vector $\vec{v}$.

A continuación veremos como encontrar explícitamente esta descomposición del vector $\vec{v}$.

Proposición 14.12 *Si $\vec{u}$ y $\vec{v}$ son vectores en $\mathbb{R}^n$ con $\vec{u} \neq \vec{0}$, entonces el único r tal que*[2]

$$\vec{v} = r\vec{u} + (\vec{v} - r\vec{u})$$

y con la propiedad adicional de que

$$(\vec{v} - r\vec{u}) \perp \vec{u}$$

es

$$r = \frac{\vec{v} \cdot \vec{u}}{\|\vec{u}\|^2}.$$

Demostración. Buscamos un número real r tal que

$$0 = (\vec{v} - r\vec{u}) \cdot \vec{u} = \vec{v} \cdot \vec{u} - r\vec{u} \cdot \vec{u} = \vec{v} \cdot \vec{u} - r\|\vec{u}\|^2 .$$

Dado que $\|\vec{u}\|^2 > 0$ se tiene que

$$r = \frac{\vec{v} \cdot \vec{u}}{\|\vec{u}\|^2}$$

es el único real con las características deseadas. ■

Ejemplo 14.13 *Sean*

$$\vec{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

[2]Evidentemente $\vec{v} = r\vec{u} + (\vec{v} - r\vec{u})$ se cumple para cualquier real r.

y supongamos que deseamos expresar a $\vec{v}$ como una suma de dos vectores: el primero perteneciente al conjunto generado por $\vec{u}$ y el segundo ortogonal a $\vec{u}$. Para ello encontramos

$$\begin{aligned} r &= \left(\frac{\vec{v}\cdot\vec{u}}{\|\vec{u}\|^2}\right) = \frac{1}{2}, \\ r\vec{u} &= \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \\ \vec{v} - r\vec{u} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}. \end{aligned}$$

Finalmente, por la Proposición 14.12 tenemos que

$$r\vec{u} + (\vec{v} - r\vec{u}) = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} + \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

es la descomposición deseada.

Como veremos en el siguiente capítulo, el vector

$$\left(\frac{\vec{v}\cdot\vec{u}}{\|\vec{u}\|^2}\right)\vec{u},$$

conocido como la proyección ortogonal de $\vec{v}$ sobre $\vec{u}$, es el elemento del *span*$\{\vec{u}\}$ "más cercano" a $\vec{v}$. Este vector también juega un papel muy importante en la obtención de dos desigualdades muy famosas.

Proposición 14.14 (Desigualdad de Bunyakovsky-Cauchy-Schwarz) *Si $\vec{u}$ y $\vec{v}$ son vectores de $\mathbb{R}^n$, entonces*

$$|\vec{v}\cdot\vec{u}| \le \|\vec{v}\|\,\|\vec{u}\|$$

y la igualdad se da syss $\vec{u}$ y $\vec{v}$ forman un conjunto LD.

Demostración. Comenzamos con el caso en el cual $\{\vec{u}, \vec{v}\}$ es un conjunto LD. Sin pérdida de generalidad, podemos suponer que existe un escalar c tal que $\vec{v} = c\vec{u}$ (el caso en el cual $\vec{u}$ es un múltiplo escalar de $\vec{v}$ es análogo). Así pues,

$$\|\vec{v}\|\,\|\vec{u}\| = \|c\vec{u}\|\,\|\vec{u}\| = |c|\,\|\vec{u}\|^2 = |(c\vec{u})\cdot\vec{u}| = |\vec{v}\cdot\vec{u}|\,.$$

Ahora bien, si $\{\vec{u}, \vec{v}\}$ es LI, entonces $\vec{u} \ne \vec{0}$ y $\vec{v}$ no es un múltiplo escalar de $\vec{u}$. Por lo tanto,

$$\begin{aligned}
0 &< \left\| \vec{v} - \left(\frac{\vec{v} \cdot \vec{u}}{\|\vec{u}\|^2} \right) \vec{u} \right\|^2 \\
&= \left(\vec{v} - \left(\frac{\vec{v} \cdot \vec{u}}{\|\vec{u}\|^2} \right) \vec{u} \right) \cdot \left(\vec{v} - \left(\frac{\vec{v} \cdot \vec{u}}{\|\vec{u}\|^2} \right) \vec{u} \right) \\
&= \|\vec{v}\|^2 - \frac{2\,(\vec{v} \cdot \vec{u})^2}{\|\vec{u}\|^2} + \frac{(\vec{v} \cdot \vec{u})^2}{\|\vec{u}\|^4} \|\vec{u}\|^2 \\
&= \|\vec{v}\|^2 - \frac{(\vec{v} \cdot \vec{u})^2}{\|\vec{u}\|^2}
\end{aligned}$$

de manera que

$$|\vec{v} \cdot \vec{u}| < \|\vec{v}\| \, \|\vec{u}\| \, .$$

■

Proposición 14.15 (Desigualdad del Triángulo) *Si $\vec{u}$ y $\vec{v}$ son vectores de $\mathbb{R}^n$, entonces*

$$\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$$

Demostración. Utilizando la definición de la norma, las propiedades del producto punto y la Proposición 14.14 tenemos que

$$\begin{aligned}
\|\vec{u} + \vec{v}\|^2 &= (\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v}) \\
&= \vec{u} \cdot \vec{u} + 2\,(\vec{v} \cdot \vec{u}) + \vec{v} \cdot \vec{v} = \|\vec{u}\|^2 + 2\,(\vec{v} \cdot \vec{u}) + \|\vec{v}\|^2 \\
&\leq \|\vec{u}\|^2 + 2\,|\vec{v} \cdot \vec{u}| + \|\vec{v}\|^2 \leq \|\vec{u}\|^2 + 2\,\|\|\vec{v}\| \, \|\vec{u}\| + \|\vec{v}\|^2 \\
&= (\|\vec{u}\| + \|\vec{v}\|)^2 \, .
\end{aligned}$$

Tomando raíces cuadradas de ambos lados de la desigualdad tenemos que

$$\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$$

con lo cual concluye la demostración. ■

La interpretación geométrica de la desigualdad del triángulo es que en un triángulo, la suma de las longitudes de dos de sus lados es mayor o igual a la longitud del tercer lado. Esto se muestra en la Figura 14.5.

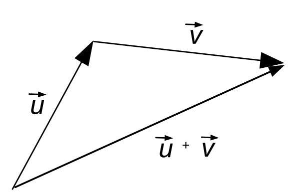
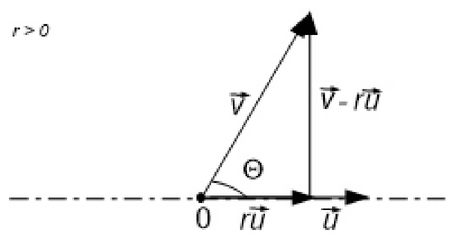


Figura 14.5: Desigualdad del Triángulo.

Medición de ángulos y paralelismo

Dados $\vec{u}$ y $\vec{v}$ vectores no nulos de $\mathbb{R}^2$, podemos medir el ángulo Θ entre ellos examinando los diagramas de la figura 14.6. que dependen del signo del escalar $r = \frac{\vec{v}\cdot\vec{u}}{\|\vec{u}\|^2}$.



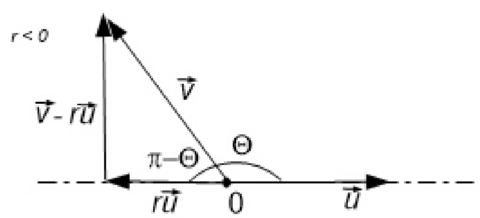


Figura 14.6: Ángulo entre $\vec{u}$ y $\vec{v}$.

En ambos casos tenemos un triángulo rectángulo cuyos lados opuesto, adyacente e hipotenusa son $\vec{v} - r\vec{u}$, $r\vec{u}$ y $\vec{v}$, respectivamente. Además, la igualdad

$$\cos\Theta = \frac{\|r\vec{u}\|}{\|\vec{v}\|} = \frac{|r| \;\; \|\vec{u}\|}{\|\vec{v}\|} = \frac{r\,\|\vec{u}\|}{\|\vec{v}\|}$$

se cumple si $r > 0$ y

$$\cos(\pi - \Theta) = \frac{\|r\vec{u}\|}{\|\vec{v}\|} = \frac{|r| \;\; \|\vec{u}\|}{\|\vec{v}\|} = \frac{-\,r\,\|\vec{u}\|}{\|\vec{v}\|}$$

se satisface cuando $r < 0$. Sin embargo, como $\cos(\pi - \Theta) = -\cos\Theta$, esta última igualdad puede escribirse como

$$-\cos\Theta = \cos(\pi - \Theta) = \frac{-\,r\,\|\vec{u}\|}{\|\vec{v}\|},$$

es decir, independientemente del signo de r hemos mostrado que

$$\cos\Theta = \frac{r\,\|\vec{u}\|}{\|\vec{v}\|} = \left(\frac{\vec{v}\cdot\vec{u}}{\|\vec{u}\|^2}\right)\frac{\|\vec{u}\|}{\|\vec{v}\|} = \frac{\vec{v}\cdot\vec{u}}{\|\vec{v}\|\,\|\vec{u}\|} \qquad (\bigstar)$$

La Expresión ($\bigstar$) nos permite definir el ángulo entre cualquier pareja de vectores no nulos en $\mathbb{R}^n$ como sigue.

Definición 14.16 *Si $\vec{x}, \vec{y} \in \mathbb{R}^n$ son distintos del vector cero, el ángulo entre ellos se define como el único $\Theta \in [0, \pi]$ tal que*

$$\cos\Theta = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \, \|\vec{y}\|}.$$

La definición anterior de ángulo está sustentada en la Desigualdad de Bunyakovsky-Cauchy-Schwarz, que nos garantiza que

$$-1 \leq \cos\Theta \leq 1.$$

Ejemplo 14.17 *Dados*

$$\vec{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix} \in \mathbb{R}^5,$$

podemos calcular el ángulo Θ entre ellos como sigue. Debido a que

$$\|\vec{x}\| = \sqrt{2}, \ \|\vec{y}\| = 1 \ y \ \ \vec{x} \cdot \vec{y} = -1,$$

se tiene que

$$\cos\Theta = \frac{-1}{\sqrt{2} \times 1} = \frac{-1}{\sqrt{2}} = \frac{-\sqrt{2}}{2}$$

y $\Theta = \frac{3\pi}{4}$.

Las Definiciones 14.7 y 14.16 son coherentes en el sentido de que si $\vec{x}, \vec{y} \in \mathbb{R}^n$ son vectores no nulos, entonces el ángulo entre ellos es $\frac{\pi}{2}$ syss $\vec{x} \perp \vec{y}$ (puesto que $\cos\left(\frac{\pi}{2}\right) = 0$). Adicionalmente, dado que $\cos(0) = 1$ y $\cos\pi = -1$ resulta natural introducir el siguiente concepto de paralelismo de acuerdo a la Definición 14.16.

Definición 14.18 *Sean $\vec{x}$ y $\vec{y}$ vectores no nulos de $\mathbb{R}^n$. Decimos que $\vec{x}$ y $\vec{y}$ son **paralelos** syss*

$$|\vec{x} \cdot \vec{y}| = ||\vec{x}||||\vec{y}||.$$

Ejemplo 14.19 *Los vectores $\vec{x} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$ y $\vec{y} = \begin{bmatrix} -2 \\ -10 \end{bmatrix}$ son paralelos pues $\vec{x} \cdot \vec{y} = -52$ y $||\vec{x}||||\vec{v}|| = \sqrt{26}\sqrt{104} = \sqrt{2704} = 52$.*

Una forma más simple de comprobar que los vectores $\vec{x}$ y $\vec{y}$ son paralelos es apelando al siguiente resultado y al hecho de que $\vec{y} = -2\vec{x}$.

Proposición 14.20 *Sean $\vec{x}$ y $\vec{y}$ vectores no nulos de $\mathbb{R}^n$. Entonces $\vec{x}$ y $\vec{y}$ son paralelos syss $\{\vec{x}, \vec{y}\}$ es un conjunto LD.*

Demostración. Basta reformular la segunda parte de la Proposición 14.14 en términos de paralelismo. ■

Definición 14.21 *Sea $\vec{x} \in \mathbb{R}^n$. Decimos que $\vec{x}$ es un vector* ***unitario*** *syss $||\vec{x}|| = 1$.*

Ejemplo 14.22 *Los vectores $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ son claramente unitarios. También el vector*

$$\frac{\sqrt{2}}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

es unitario pues su norma es igual a $\frac{\sqrt{2}}{2}\sqrt{2} = 1$.

La siguiente proposición, cuya demostración se deja como ejercicio para el lector, dice que todo vector distinto de cero tiene asociado, de forma natural, un vector unitario.

Proposición 14.23 *Si $\vec{x} \in \mathbb{R}^n$ es un vector distinto de cero, entonces el vector $\frac{1}{||\vec{x}||}\vec{x}$ es unitario y satisface que $\vec{x}$ y $\frac{1}{||\vec{x}||}\vec{x}$ son paralelos.*

Al vector $\frac{1}{||\vec{x}||}\vec{x}$ se le conoce como la **normalización** de $\vec{x}$ y se ilustra en la Figura 14.7 para el caso $||\vec{x}|| > 1$.

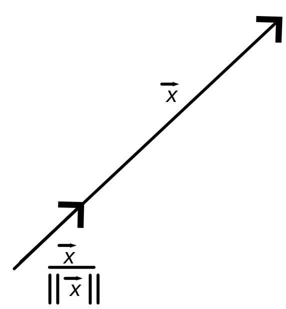


Figura 14.7: Normalización de un vector.

Ejemplo 14.24 *Dado el vector*

$$\vec{v} = \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix},$$

como

$$||\vec{v}|| = \sqrt{9 + 16} = 5,$$

su normalización es el vector,

$$\begin{bmatrix} -\frac{3}{5} \\ \frac{4}{5} \\ 0 \end{bmatrix}.$$

Conjuntos y bases ortogonales

Definición 14.25 *Sea C un subconjunto de vectores de $\mathbb{R}^n$. Decimos que C es un* ***conjunto ortogonal*** *syss para cualesquiera dos vectores distintos $\vec{u}$ y $\vec{v}$ de C se cumple que $\vec{u}$ y $\vec{v}$ son ortogonales esto es, $\vec{u} \cdot \vec{v} = \vec{0}$.*

Por vacuidad, el vacío así como cualquier conjunto de vectores con exactamente un elemento son conjuntos ortogonales. También en el Ejemplo 14.9 ya hemos notado que la base estándar $\{\vec{e}_1, \ldots, \vec{e}_n\}$ de $\mathbb{R}^n$ es un conjunto ortogonal. En el resto de este libro aparecerán otros conjuntos ortogonales menos conocidos.

Ejemplo 14.26 *Si*

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \; y \; \vec{u}_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

entonces el conjunto $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es ortogonal pues

$$\vec{u}_1 \cdot \vec{u}_2 = \vec{u}_1 \cdot \vec{u}_3 = \vec{u}_2 \cdot \vec{u}_3 = 0.$$

Pero, el conjunto $\{\vec{u}_1, \vec{u}_2, \vec{u}_4\}$ no es ortogonal pues $\vec{u}_1 \cdot \vec{u}_4 = 2 \neq 0$.

Proposición 14.27 *Si $\{\vec{v}_1, \ldots, \vec{v}_m\}$ es un subconjunto ortogonal de $\mathbb{R}^n$ cuyos elementos no son el vector cero, entonces se cumplen las siguientes afirmaciones.*

(a) $\{\vec{v}_1, \ldots, \vec{v}_m\}$ es LI.

(b) Si $\vec{v} \in span\{\vec{v}_1, \ldots, \vec{v}_m\}$, entonces

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m$$

donde

$$c_1 = \frac{\vec{v} \cdot \vec{v}_1}{||\vec{v}_1||^2},\; c_2 = \frac{\vec{v} \cdot \vec{v}_2}{||\vec{v}_2||^2}, \ldots,\; c_m = \frac{\vec{v} \cdot \vec{v}_m}{||\vec{v}_m||^2}.$$

Demostración. En el caso de (a) supongamos que

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_m\vec{v}_m = \vec{0}.$$

Debemos probar que todos los α's están forzados a ser cero. Nuestra hipótesis implica que

$$(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_m\vec{v}_m) \cdot \vec{v}_1 = \vec{0} \cdot \vec{v}_1 = 0.$$

y como
$$\vec{v}_1 \cdot \vec{v}_2 = \cdots = \vec{v}_1 \cdot \vec{v}_m = 0,$$
se cumple
$$\begin{aligned}&(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_m\vec{v}_m) \cdot \vec{v}_1 \\ &= \alpha_1\vec{v}_1 \cdot \vec{v}_1 + \alpha_2\vec{v}_2 \cdot \vec{v}_1 + \cdots + \alpha_m\vec{v}_m \cdot \vec{v}_1 = \alpha_1||\vec{v}_1||^2,\end{aligned}$$
obteniendo
$$\alpha_1||\vec{v}_1||^2 = 0.$$
Ahora bien, como $||\vec{v}_1||^2 > 0$ (pues $\vec{v}_1$ no es el vector cero), necesariamente $\alpha_1 = 0$. Análogamente puede verificarse que los otros α_i's también son iguales a cero.

Para el inciso (b) notamos que, gracias al inciso anterior, $\{\vec{v}_1, \ldots, \vec{v}_m\}$ es una base para $span\{\vec{v}_1, \ldots, \vec{v}_m\}$. Por lo tanto, si $\vec{v} \in span\{\vec{v}_1, \ldots, \vec{v}_m\}$, entonces existen escalares únicos $c_1, c_2, \ldots, c_m$ tales que
$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m.$$
Ahora bien, si realizamos el producto punto con $\vec{v}_1$ en ambos lados de esta iguladad, obtenemos
$$\begin{aligned}\vec{v} \cdot \vec{v}_1 &= (c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m) \cdot \vec{v}_1 \\ &= c_1\vec{v}_1 \cdot \vec{v}_1 + \vec{0} + \cdots + \vec{0} = c_1||\vec{v}_1||^2,\end{aligned}$$
en donde la última igualdad se sigue del hecho que $\vec{v}_1 \perp \vec{v}_2, \ldots, \vec{v}_1 \perp \vec{v}_m$. De aquí, $c_1 = \frac{\vec{v}\cdot\vec{v}_1}{||\vec{v}_1||^2}$. Análogamente, puede verificarse que para toda $i \in \{1, \ldots, m\}$ se tiene que $c_i = \frac{\vec{v}\cdot\vec{v}_i}{||\vec{v}_i||^2}$. ■

Ejemplo 14.28 *En el Ejemplo 14.26 ya hemos notado que si*
$$\vec{u}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}, \vec{u}_2 = \begin{bmatrix}-1\\1\\0\end{bmatrix} \text{ y } \vec{u}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix},$$
entonces $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es un conjunto ortogonal de vectores distintos de cero. El teorema anterior asegura que $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es LI y, por lo tanto una base para $\mathbb{R}^3$. Supongamos ahora que deseamos encontrar escalares c_1, c_2 y c_3 tales que si
$$\vec{u} = \begin{bmatrix}2\\-2\\0\end{bmatrix},$$
entonces
$$\vec{u} = c_1\vec{u}_1 + c_2\vec{u}_2 + c_3\vec{u}_3.$$
Por supuesto que una forma de hacerlo es resolviendo el sistema cuya matriz de coeficientes tiene por columnas a $\vec{u}_1$, $\vec{u}_2$ y $\vec{u}_3$ y cuyo vector

de términos independientes es el vector $\vec{u}$. Otra forma de hacerlo es invocando la Proposición 14.27, según la cual

$$c_1 = \frac{\vec{u} \cdot \vec{u}_1}{||\vec{u}_1||^2} = 0, \; c_2 = \frac{\vec{u} \cdot \vec{u}_2}{||\vec{u}_2||^2} = -\frac{4}{2} = -2 \text{ y } c_3 = \frac{\vec{u} \cdot \vec{u}_3}{||\vec{u}_3||^2} = \frac{0}{1} = 0.$$

Ejemplo 14.29 *Es sencillo verificar que si*

$$\vec{v}_1 = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -1 \\ 4 \\ 0 \\ 0 \end{bmatrix} \text{ y } \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 0 \end{bmatrix},$$

entonces

$$\vec{v}_1 \cdot \vec{v}_2 = \vec{v}_1 \cdot \vec{v}_3 = \vec{v}_2 \cdot \vec{v}_3 = 0,$$

por lo que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es un conjunto ortogonal (de vectores distintos de cero). La Proposición 14.27 asegura que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es LI y, por lo tanto, una base para

$$S = span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}.$$

Por supuesto, S es igual al subespacio que consiste de aquellos vectores de $\mathbb{R}^4$ cuya cuarta coordenada es igual a cero.

Supongamos ahora que deseamos encontrar escalares c_1, c_2 y c_3 tales que si

$$\vec{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \\ 0 \end{bmatrix},$$

entonces $\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3$. Una forma de hacerlo es resolviendo el sistema cuya matriz de coeficientes tiene por columnas a $\vec{v}_1$, $\vec{v}_2$ y $\vec{v}_3$ y cuyo vector de términos independientes es el vector $\vec{v}$. Otra forma de hacerlo es invocando la Proposición 14.27, según la cual

$$c_1 = \frac{\vec{v} \cdot \vec{v}_1}{||\vec{v}_1||^2} = \frac{11}{17}, \; c_2 = \frac{\vec{v} \cdot \vec{v}_2}{||\vec{v}_2||^2} = \frac{-7}{17} \text{ y } c_3 = \frac{\vec{v} \cdot \vec{v}_3}{||\vec{v}_3||^2} = \frac{4}{4} = 1.$$

Definición 14.30 *Sea S un subespacio de $\mathbb{R}^n$ y sea $\{\vec{v}_1, \ldots, \vec{v}_m\}$ un subconjunto de S. Decimos que $\{\vec{v}_1, \ldots, \vec{v}_m\}$ es una* ***base ortogonal*** *para S syss $\{\vec{v}_1, \ldots, \vec{v}_m\}$ es un conjunto ortogonal y $span\{\vec{v}_1, \ldots, \vec{v}_m\} = S$.*

Ejemplo 14.31 *La base estándar $\{\vec{e}_1, \ldots, \vec{e}_n\}$ es una base ortogonal para $\mathbb{R}^n$ y el conjunto $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ que definimos previamente en el ejemplo 14.29 es una base ortogonal para el subespacio que consiste de aquellos vectores de $\mathbb{R}^4$ cuya cuarta coordenada es igual a cero.*

Ejercicios

Ejercicio 14.1 *Dados los vectores*

$$\vec{v}_1 = \begin{bmatrix} -1 \\ -2 \end{bmatrix}, \ \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ \vec{v}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix},$$

$$\vec{w}_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}, \ \vec{w}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \ \vec{w}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

calcular:

a. $\|\vec{v}_1\|, \|\vec{v}_2\|, \|\vec{v}_3\|, \|\vec{v}_1 + \vec{v}_2\|, \|-2\vec{v}_2\|, \|\vec{w}_1\|, \|3\vec{w}_2\|, \|-\vec{w}_3\|.$

b. *La distancia entre* $\vec{v}_1$ *y* $\vec{v}_2$.

c. *La distancia entre* $\vec{w}_2$ *y* $\vec{w}_3$.

d. *Los productos* $\vec{v}_1 \cdot \vec{v}_2$, $(\vec{v}_1 + \vec{v}_2) \cdot \vec{v}_3$, $\vec{v}_1 \cdot (-2\vec{v}_2)$, $(-\vec{v}_1 + \vec{v}_2) \cdot \vec{v}_3$.

e. *Los productos* $\vec{w}_1 \cdot \vec{w}_2$, $\vec{w}_1 \cdot (a\vec{w}_2)$ *(con* $a \in \mathbb{R}$*),* $-\vec{w}_1 \cdot (\vec{w}_2 + 2\vec{w}_3)$, $(2\vec{w}_1 - \vec{w}_2) \cdot \vec{w}_3$.

Ejercicio 14.2 *Sean* $\vec{x}, \vec{y} \in \mathbb{R}^n$. *Probar que*

$$\|\vec{x} + \vec{y}\|^2 + \|\vec{x} - \vec{y}\|^2 = 2\|\vec{x}\|^2 + 2\|\vec{y}\|^2.$$

Esta igualdad se conoce como la ***ley del paralelogramo****.*

Ejercicio 14.3 *Sean* $\vec{v}, \vec{w} \in \mathbb{R}^n$ *dos vectores ortogonales. Si* $\|\vec{v}\| = 3$ *y* $\|\vec{w}\| = 4$, *encontrar* $\|\vec{v} + \vec{w}\|$ *y* $\|\vec{v} - \vec{w}\|$.

Ejercicio 14.4 *Considerar los vectores*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \vec{v}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \ \vec{v}_3 = \begin{bmatrix} 0 \\ -3 \\ 4 \end{bmatrix}, \ \vec{v}_4 = \begin{bmatrix} -1 \\ 2 \\ 0 \\ -2 \end{bmatrix}.$$

a. *Expresar al vector* $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ *como la suma de dos vectores, donde uno es un múltiplo de* $\vec{v}_1$ *y el otro es ortogonal a* $\vec{v}_1$.

b. *Expresar al vector* $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ *como la suma de dos vectores, donde uno es un múltiplo de* $\vec{v}_2$ *y el otro es ortogonal a* $\vec{v}_2$.

c. *Expresar al vector* $\begin{bmatrix} 1 \\ -10 \\ 5 \end{bmatrix}$ *como la suma de dos vectores, donde uno es un múltiplo de* $\vec{v}_3$ *y el otro es ortogonal a* $\vec{v}_3$.

d. *Expresar al vector* $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}$ *como la suma de dos vectores, donde uno es un múltiplo de* $\vec{v}_4$ *y el otro es ortogonal a* $\vec{v}_4$.

Ejercicio 14.5 *Dados* $\vec{x}, \vec{y} \in \mathbb{R}^n$ *con* $\vec{x} \neq \vec{0}$, *se define la función*

$$f : \mathbb{R} \longrightarrow \mathbb{R}^+$$

como

$$f(r) = \|\vec{y} - r\vec{x}\|^2 .$$

Utilizando la primera y la segunda derivada, mostrar que f *toma su valor mínimo cuando*

$$r = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\|^2},$$

es decir, para este valor de r *se minimiza el cuadrado de la distancia (por lo tanto la distancia) entre* $\vec{y}$ *y* $r\vec{x}$.

Ejercicio 14.6 *Considerar los vectores* $\vec{x}, \vec{y} \in \mathbb{R}^n$ *y* θ *el ángulo entre ellos. Encontrar* $\vec{x} \cdot \vec{y}$ *para los siguientes casos:*

a. $\|\vec{x}\| = 1, \|\vec{y}\| = 2,\ \theta = \frac{\pi}{4}$.

b. $\|\vec{x}\| = \frac{1}{2}, \|\vec{y}\| = 2,\ \theta = \frac{2\pi}{3}$.

c. $\|\vec{x}\| = 666, \|\vec{y}\| = 1,\ \theta = \pi$.

Ejercicio 14.7 *Encontrar el ángulo entre las siguientes parejas de vectores:*

a. $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ *y* $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

b. $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ *y* $\begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$.

c. $\begin{bmatrix} \sqrt{3} \\ 1 \\ 0 \end{bmatrix}$ *y* $\begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix}$.

d. $\begin{bmatrix} 1 \\ 2 \\ -1 \\ -2 \end{bmatrix}$ *y* $\begin{bmatrix} -1 \\ -2 \\ 1 \\ 2 \end{bmatrix}$.

Ejercicio 14.8 *Demostrar la Proposición 14.23.*

Ejercicio 14.9 *Probar que si $\vec{x} \perp \vec{y}$ y α, β son escalares, entonces $\alpha\vec{x} \perp \beta\vec{y}$.*

Ejercicio 14.10 *Determinar si los siguientes conjuntos de vectores son ortogonales:*

$$a) \left\{ \begin{bmatrix} \pi \\ \pi \end{bmatrix}, \begin{bmatrix} -666 \\ 666 \end{bmatrix} \right\},\ b) \left\{ \begin{bmatrix} \pi \\ \pi \end{bmatrix}, \begin{bmatrix} 666 \\ 666 \end{bmatrix} \right\},$$
$$c) \left\{ \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\},\ d) \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -4 \\ 1 \end{bmatrix} \right\},$$
$$e) \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Ejercicio 14.11 *Dados*

$$\mathcal{B}_1 = \left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\},$$
$$\mathcal{B}_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix} \right\}.$$

a. *Utilizando la Proposición 14.27, expresar al vector* $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *como combinación lineal de los elementos de $\mathcal{B}_1$.*

b. *Utilizando la Proposición 14.27, expresar al vector* $\begin{bmatrix} 0 \\ -3 \\ 6 \end{bmatrix}$ *como combinación lineal de los elementos de $\mathcal{B}_2$.*

c. *Un conjunto* ***ortonormal*** *de vectores es un conjunto ortogonal en el cual todos sus vectores tienen norma unitaria. Encontrar conjuntos ortonormales $\{\vec{v}_1, \vec{v}_2\} \subset \mathbb{R}^2$ y $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\} \subset \mathbb{R}^3$ de forma que*

$$span\,\{\vec{v}_1, \vec{v}_2\} = span \mathcal{B}_1,$$
$$span\,\{\vec{w}_1, \vec{w}_2, \vec{w}_3\} = span \mathcal{B}_2.$$

Capítulo 15

Bases ortogonales y mínimos cuadrados

Introducción

Al final del capítulo anterior exhibimos una de las ventajas de contar con bases ortogonales para un subespacio. Recordemos que si $\{\vec{v}_1, \ldots, \vec{v}_m\}$ es una base ortogonal de un subespacio S de $\mathbb{R}^n$, entonces para todo $\vec{v} \in S$ se tiene que

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_m\vec{v}_m,$$

en donde

$$c_1 = \frac{\vec{v} \cdot \vec{v}_1}{||\vec{v}_1||^2},\ c_2 = \frac{\vec{v} \cdot \vec{v}_2}{||\vec{v}_2||^2}, \ldots,\ c_m = \frac{\vec{v} \cdot \vec{v}_m}{||\vec{v}_m||^2}.$$

Previamente, ya habíamos mostrado que, para toda $i = 1, 2, \ldots, m$, los vectores $c_i\vec{v}_i$ son tales que

$$\vec{v} = c_i\vec{v}_i + (\vec{v} - c_i\vec{v}_i)$$

y adicionalmente,

$$(\vec{v} - c_i\vec{v}_i) \perp \vec{v}_i.$$

Estas aplicaciones (y otras más que veremos más adelante) justifican la necesidad de introducir el siguiente concepto.

Definición 15.1 *Sean $\vec{u}$ y $\vec{v}$ vectores en $\mathbb{R}^n$ con $\vec{u} \neq \vec{0}$. La* ***proyección ortogonal*** *de $\vec{v}$ sobre $\vec{u}$, se denota por $proj_{\vec{u}}\vec{v}$ y se define como el vector*

$$proj_{\vec{u}}\vec{v} = \left(\frac{\vec{v} \cdot \vec{u}}{||\vec{u}||^2}\right)\vec{u}.$$

Ejemplo 15.2 *Si* $\vec{u} = \begin{bmatrix} -2 \\ 5 \\ 1 \end{bmatrix}$ *y* $\vec{v} = \begin{bmatrix} 7 \\ 14 \\ 4 \end{bmatrix}$, *entonces*

$$proj_{\vec{u}}\vec{v} = \left(\frac{\vec{v} \cdot \vec{u}}{||\vec{u}||^2}\right)\vec{u} = \left(\frac{-14+70+4}{30}\right)\begin{bmatrix} -2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 10 \\ 2 \end{bmatrix}.$$

Así, tenemos la siguiente formulación del resultado con el cual comenzamos esta introducción.

Observación 15.3 *Si $\{\vec{v}_1, \ldots \vec{v}_m\}$ es una base ortogonal para S y $\vec{v} \in S$, entonces*

$$\begin{aligned} \vec{v} &= \frac{\vec{v} \cdot \vec{v}_1}{||\vec{v}_1||^2}\vec{v}_1 + \frac{\vec{v} \cdot \vec{v}_2}{||\vec{v}_2||^2}\vec{v}_2 + \ldots + \frac{\vec{v} \cdot \vec{v}_m}{||\vec{v}_m||^2}\vec{v}_m \\ &= proj_{\vec{v}_1}\vec{v} + proj_{\vec{v}_2}\vec{v} + \ldots + proj_{\vec{v}_m}\vec{v}. \end{aligned}$$

Por ejemplo, en el contexto de $\mathbb{R}^3$, tenemos que la base estándar $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ es una base ortogonal para $\mathbb{R}^3$ y cualquier vector $\vec{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ puede expresarse como

$$\vec{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = proj_{\vec{e}_1}\vec{v} + proj_{\vec{e}_2}\vec{v} + proj_{\vec{e}_3}\vec{v} = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ b \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ c \end{bmatrix}.$$

Proyección ortogonal

Hasta el momento hemos visto como obtener la proyección ortogonal de un vector $\vec{v} \in \mathbb{R}^n$ sobre un vector $\vec{u} \neq 0$ o más bien, sobre el subespacio generado por $\vec{u}$ (el cual, por supuesto, tiene dimensión uno). Ahora extenderemos esta noción para subespacios de cualquier dimensión positiva, siempre y cuando éstos posean una base ortogonal (más adelante veremos que éste siempre es el caso).

Definición 15.4 *Sea S un subespacio de $\mathbb{R}^n$ y supongamos que el conjunto $\{\vec{v}_1, \ldots, \vec{v}_m\}$ es una base ortogonal para S. Si $\vec{v} \in \mathbb{R}^n$ (independientemente de si $\vec{v} \in S$ o $\vec{v} \notin S$), entonces definimos la* ***proyección ortogonal*** *de $\vec{v}$ sobre S como*

$$proj_S\vec{v} = \frac{\vec{v} \cdot \vec{v}_1}{||\vec{v}_1||^2}\vec{v}_1 + \frac{\vec{v} \cdot \vec{v}_2}{||\vec{v}_2||^2}\vec{v}_2 + \ldots + \frac{\vec{v} \cdot \vec{v}_m}{||\vec{v}_m||^2}\vec{v}_m$$

$$= proj_{\vec{v}_1}\vec{v} + proj_{\vec{v}_2}\vec{v} + \ldots + proj_{\vec{v}_m}\vec{v}.$$

Gracias a la Observación 15.3, es claro que si $\vec{v} \in S$, entonces

$$\vec{v} = proj_S\vec{v}.$$

Observación 15.5 *Si $\vec{v}$ y $\vec{v}_1$ son vectores de $\mathbb{R}^n$ con $\vec{v}_1 \neq \vec{0}$ y $S = span\{\vec{v}_1\}$, entonces $\{\vec{v}_1\}$ es una base ortogonal para S. En tal caso,*

$$proj_S\vec{v} = proj_{\vec{v}_1}\vec{v}.$$

En la figura 15.1 se ilustra el caso en el cual el subespacio S, de la definición anterior, es un plano.

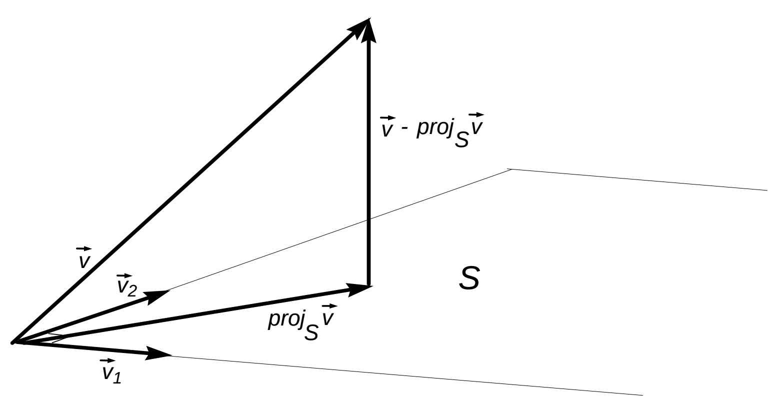


Figura 15.1: Proyección ortogonal de $\vec{v}$ sobre un plano S.

Proposición 15.6 *Con la notación anterior, si $\vec{s} \in S$, entonces*

$$\vec{s} \perp (\vec{v} - proj_S\vec{v}),$$

es decir, $\vec{v} - proj_S\vec{v}$ es ortogonal a todos los vectores de S.

Demostración. *Dado que $\{\vec{v}_1, \ldots \vec{v}_m\}$ es una base ortogonal para S, es suficiente probar, en virtud del Ejercicio 15.3, que para toda $i = 1, 2, \ldots, m$ se cumple*

$$\vec{v}_i \perp (\vec{v} - proj_S\vec{v}).$$

Pero,

$$\begin{aligned}
\vec{v}_i \cdot (\vec{v} - proj_S\vec{v}) &= \vec{v}_i \cdot \vec{v} - \vec{v}_i \cdot proj_S\vec{v} \\
&= \vec{v}_i \cdot \vec{v} - \vec{v}_i \cdot \left(\frac{\vec{v} \cdot \vec{v}_1}{||\vec{v}_1||^2}\vec{v}_1 + \cdots + \frac{\vec{v} \cdot \vec{v}_m}{||\vec{v}_m||^2}\vec{v}_m \right) \\
&= \vec{v}_i \cdot \vec{v} - \left(\frac{\vec{v} \cdot \vec{v}_i}{||\vec{v}_i||^2} (\vec{v}_i \cdot \vec{v}_i) \right) = \vec{v}_i \cdot \vec{v} - \vec{v}_i \cdot \vec{v} = 0,
\end{aligned}$$

lo que concluye la demostración. ■

Ejemplo 15.7 *Sea S el subespacio de $\mathbb{R}^4$ que tiene como base al conjunto $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$, donde*

$$\vec{w}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \ \vec{w}_2 = \begin{bmatrix} -1 \\ 2 \\ -1 \\ 2 \end{bmatrix}, \ \vec{w}_3 = \begin{bmatrix} -3 \\ -3 \\ 1 \\ 2 \end{bmatrix}$$

y supongamos que queremos encontrar la proyección ortogonal de $\vec{u}$ sobre S, donde

$$\vec{u} = \begin{bmatrix} 6 \\ 6 \\ -2 \\ -4 \end{bmatrix}.$$

Claramente $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ es un conjunto ortogonal y como

$$\begin{gathered}
||\vec{w}_1||^2 = 3, ||\vec{w}_2||^2 = 10, ||\vec{w}_3||^2 = 23, \\
\vec{u} \cdot \vec{w}_1 = \vec{u} \cdot \vec{w}_2 = 0 \quad y \quad \vec{u} \cdot \vec{w}_3 = 46,
\end{gathered}$$

tenemos que por la Definición 15.4,

$$\begin{aligned}
proj_S\vec{u} &= proj_{\vec{w}_1}\vec{u} + proj_{\vec{w}_2}\vec{u} + proj_{\vec{w}_3}\vec{u} \\
&= \frac{\vec{u} \cdot \vec{w}_1}{||\vec{w}_1||^2}\vec{w}_1 + \frac{\vec{u} \cdot \vec{w}_2}{||\vec{w}_2||^2}\vec{w}_2 + \frac{\vec{u} \cdot \vec{w}_3}{||\vec{w}_3||^2}\vec{w}_3 \\
&= \frac{0}{||\vec{w}_1||^2}\vec{w}_1 + \frac{0}{||\vec{w}_2||^2}\vec{w}_2 + \left(-\frac{46}{23}\right)\vec{w}_3 = -2 \begin{bmatrix} -3 \\ -3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ -2 \\ -4 \end{bmatrix}.
\end{aligned}$$

Observamos que $proj_S\vec{u} = \vec{u}$, lo cual tiene sentido pues $\vec{u} = -2\vec{w}_3 \in S$.

Ejemplo 15.8 *Sean*

$$S = span \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \right\} \quad y \quad \vec{v} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}.$$

Supongamos que deseamos calcular $proj_S\vec{v}$*. Para ello, primero encontramos una base ortogonal para* S *como sigue. Tomamos* $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ *y*

$$\begin{aligned} \vec{w}_2 &= \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} - proj_{\vec{v}_1}\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} -4 \\ -1 \\ 5 \end{bmatrix}. \end{aligned}$$

Así pues, los vectores $\vec{v}_1$ *y* $\vec{w}_2$ *son ortogonales, están en el subespacio* S *y gracias a la Proposición 14.27, constituyen una base ortogonal para* S*. Definiendo* $\vec{v}_2 = 3\vec{w}_2$ *y apelando al Ejercicio 14.9, tenemos que*

$$\{\vec{v}_1, \vec{v}_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -4 \\ -1 \\ 5 \end{bmatrix} \right\}$$

también es una base ortogonal para S*. Procedemos ahora a calcular* $proj_S\vec{v}$ *observando que*

$$\begin{aligned} &\|\vec{v}_1\|^2 = 3, \|\vec{v}_2\|^2 = 42, \\ &\vec{v}\cdot\vec{v}_1 = 6 \quad y \quad \vec{v}\cdot\vec{v}_2 = 18, \end{aligned}$$

de manera que, de acuerdo a la Definición 15.4,

$$\begin{aligned} proj_S\vec{v} &= proj_{\vec{v}_1}\vec{v} + proj_{3\vec{v}_2}\vec{v} \\ &= \frac{\vec{v}\cdot\vec{v}_1}{\|\vec{v}_1\|^2}\vec{v}_1 + \frac{\vec{v}\cdot\vec{v}_2}{|\,\|\vec{v}_2\|^2}\vec{v}_2 \\ &= \frac{6}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{18}{42}\begin{bmatrix} -4 \\ -1 \\ 5 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 2 \\ 11 \\ 29 \end{bmatrix}. \end{aligned}$$

El siguiente teorema resume algunas propiedades de la proyección ortogonal.

Teorema 15.9 (Propiedades de la proyección ortogonal) *Sea* S *un subespacio de* $\mathbb{R}^n$ *con base ortogonal* $\{\vec{v}_1, \ldots, \vec{v}_m\}$ *y sea* $\vec{v}$ *un vector cualquiera en* $\mathbb{R}^n$*. Entonces,*

1. *$proj_S\vec{v} = proj_{\vec{v}_1}\vec{v} + proj_{\vec{v}_2}\vec{v} + \cdots + proj_{\vec{v}_m}\vec{v} \in S$.*

2. *$proj_S\vec{v} = \vec{v}$ syss $\vec{v} \in S$.*

3. *$\vec{v} - proj_S\vec{v}$ es ortogonal a todos los elementos de S.*

4. *$proj_S\vec{v}$ es el único vector en S con la propiedad de ser el "más cercano" a $\vec{v}$. Esto es,*

$$||\vec{v} - proj_S\vec{v}|| \leq ||\vec{v} - \vec{z}||$$

   para toda $\vec{z}$ en S y la igualdad se tiene syss $\vec{z} = proj_S\vec{v}$.

5. *$proj_S\vec{v}$ es independiente de la elección de la base ortogonal para S.*

Demostración. Los tres primeros incisos son inmediatos o ya han sido verificados previamente. En el caso de 4) observemos que $proj_S\vec{v} - \vec{z} \in S$ y $\vec{v} - proj_S\vec{v}$ es ortogonal a todos los elementos de S. Por la Proposición 14.8 se tiene

$$\|\vec{v} - proj_S\vec{v}\|^2 + \|proj_S\vec{v} - \vec{z}\|^2 = \|\vec{v} - \vec{z}\|^2,$$

con lo cual se sigue el resultado ya que $\|proj_S\vec{v} - \vec{z}\| \geq 0$ y la igualdad se cumple syss $proj_S\vec{v} = \vec{z}$. Finalmente, para demostrar 5), supongamos que $\{\vec{w}_1, \ldots, \vec{w}_m\}$ es otra base ortogonal de S y $proj'_S\vec{v}$ denota la proyección con respecto a esta base. De acuerdo al inciso anterior, los vectores $proj_S\vec{v}$ y $proj'_S\vec{v}$ son ambos iguales al único elemento de S que minimiza la distancia a $\vec{v}$ y, por lo tanto, coinciden. ■

Ejemplo 15.10 *En el caso del Ejemplo 15.2, la distancia mínima entre $\vec{v}$ y el subespacio generado por $\vec{u}$ es igual a*

$$\begin{aligned}\|\vec{v} - proj_{\vec{u}}\vec{v}\| &= \left\| \begin{bmatrix} 7 \\ 14 \\ 4 \end{bmatrix} - \begin{bmatrix} -4 \\ 10 \\ 2 \end{bmatrix} \right\| \\ &= \left\| \begin{bmatrix} 11 \\ 4 \\ 2 \end{bmatrix} \right\| = \sqrt{121 + 16 + 4} = \sqrt{141}.\end{aligned}$$

En el caso del Ejemplo 15.8, la distancia mínima entre $\vec{v}$ y el subespacio S es igual a

$$\begin{aligned}\|\vec{v} - proj_{\vec{u}}\vec{v}\| &= \left\| \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix} - \frac{1}{7} \begin{bmatrix} 2 \\ 11 \\ 29 \end{bmatrix} \right\| \\ &= \frac{1}{7} \left\| \begin{bmatrix} 12 \\ -18 \\ -6 \end{bmatrix} \right\| = \frac{1}{7}\sqrt{504}.\end{aligned}$$

Finalmente, en el Ejemplo 15.7, la distancia mínima entre $\vec{u}$ y el subespacio S es igual a

$$\|\vec{u} - proj_S\vec{u}\| = \|\vec{u} - \vec{u}\| = 0.$$

Proceso de ortogonalización de Gram-Schmidt

Dado cualquier subespacio S en $\mathbb{R}^n$, es natural preguntarse si siempre existe una base ortogonal para éste. Por ejemplo, supongamos que

$$S = span\{\vec{w}_1, \vec{w}_2, \vec{w}_3\} \subset \mathbb{R}^4,$$

en donde

$$\vec{w}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \vec{w}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \vec{w}_3 = \begin{bmatrix} 0 \\ 2 \\ 4 \\ 7 \end{bmatrix}.$$

Estos vectores forman un conjunto LI puesto que claramente $\vec{w}_2 \notin span\{\vec{w}_1\}$ y $\vec{w}_3 \notin span\{\vec{w}_1, \vec{w}_2\}$, no obstante

$$\vec{w}_1 \cdot \vec{w}_2 = -1, \quad \vec{w}_1 \cdot \vec{w}_3 = 2, \quad \vec{w}_2 \cdot \vec{w}_3 = 2,$$

de manera que ningún par de ellos constituye un conjunto ortogonal. ¿Cómo encontrar vectores ortogonales que generen al mismo subespacio S?

Para encontrar una base ortogonal para S, digamos $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$, es suficiente que se cumplan las siguientes dos propiedades:

$$span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = S \text{ y } \{\vec{v}_1, \vec{v}_2, \vec{v}_3\} \text{ es un conjunto ortogonal,}$$

ya que la Proposición 14.27 garantiza la independencia lineal del conjunto $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Para construir dicha base procedemos como sigue:

Sea $\vec{v}_1$ cualquier vector de la base original, digamos $\vec{w}_1$, de manera que $\vec{v}_1 = \vec{w}_1$ y $||\vec{v}_1||^2 = ||\vec{w}_1||^2 = 5$. Como $\vec{w}_2$ no es ortogonal a $\vec{w}_1$, lo reemplazamos por

$$\begin{aligned} \vec{v}_2 &= \vec{w}_2 - proj_{\vec{v}_1}\vec{w}_2 = \vec{w}_2 - \frac{\vec{w}_2 \cdot \vec{v}_1}{||\vec{v}_1||^2}\vec{v}_1 \\ &= \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \left(-\frac{1}{5}\right) \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} -3 \\ 6 \\ 0 \\ 0 \end{bmatrix} \end{aligned}$$

y verificamos que, de acuerdo al Teorema 15.9, se tiene que

$$\vec{v}_2 \perp \vec{v}_1.$$

Adicionalmente,

$$\vec{v}_2 \in span\{\vec{v}_1, \vec{w}_2\} \subset S.$$

En efecto,

$$\begin{bmatrix} -\frac{3}{5} \\ \frac{6}{5} \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} = 0.$$

La intuición geométrica del procedimiento anterior es que $\vec{w}_2$ se descompone como la suma de dos vectores ortogonales, concretamente,

$$\vec{w}_2 = \underbrace{proj_{\vec{v}_1}\vec{w}_2}_{\text{paralelo a } \vec{v}_1} + \underbrace{(\vec{w}_2 - proj_{\vec{v}_1}\vec{w}_2)}_{\text{ortogonal a } \vec{v}_1}$$

y conservamos únicamente la componente que es ortogonal al primer vector de la base. Finalmente, el vector $\vec{v}_3$ se construye a partir de $\vec{w}_3$ de forma similar. Si $R = span\{\vec{v}_1, \vec{v}_2\}$, entonces $\{\vec{v}_1, \vec{v}_2\}$ es una base ortogonal para R y definimos

$$\begin{aligned} \vec{v}_3 &= \vec{w}_3 - proj_R\vec{w}_3 = \vec{w}_3 - (proj_{\vec{v}_1}\vec{w}_3 + proj_{\vec{v}_2}\vec{w}_3) \\ &= \vec{w}_3 - \left(\frac{\vec{w}_3 \cdot \vec{v}_1}{||\vec{v}_1||^2}\vec{v}_1 + \frac{\vec{w}_3 \cdot \vec{v}_2}{||\vec{v}_2||^2}\vec{v}_2 \right) \\ &= \begin{bmatrix} 0 \\ 2 \\ 4 \\ 7 \end{bmatrix} - \left(\frac{2}{5} \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \frac{12}{5} \times \frac{5}{9} \begin{bmatrix} -\frac{3}{5} \\ \frac{6}{5} \\ 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 0 \\ 4 \\ 7 \end{bmatrix}. \end{aligned}$$

Una vez más, por el Teorema 15.9, para toda $\vec{s} \in R$ se cumple que

$$\vec{s} \cdot (\vec{w}_3 - proj_R\vec{w}_3) = \vec{s} \cdot \vec{v}_3 = 0;$$

en particular, como $\vec{v}_1, \vec{v}_2 \in S$ se tiene que,

$$\vec{v}_3 \cdot \vec{v}_1 = \vec{v}_3 \cdot \vec{v}_2 = 0.$$

Adicionalmente,

$$\vec{v}_3 \in span\{\vec{v}_1, \vec{v}_2, \vec{w}_3\} \subset S.$$

Tenemos así que el conjunto $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es una base ortogonal para S.

Ahora sistematizamos la discusión anterior con el fin de encontrar bases ortogonales para cualquier subespacio $S \subset \mathbb{R}^n$.

Teorema 15.11 (Proceso de ortogonalización de Gram - Schmidt) *Supongamos que S es un subespacio de $\mathbb{R}^n$ con base $\{\vec{w}_1, \ldots, \vec{w}_m\}$. Definimos al conjunto $\{\vec{v}_1, \ldots \vec{v}_k\}$ con $1 \leq k \leq m$, como sigue:*

$$\begin{aligned}
\vec{v}_1 &= \vec{w}_1, \\
\vec{v}_2 &= \vec{w}_2 - proj_{\vec{v}_1}\vec{w}_2, \\
\vec{v}_3 &= \vec{w}_3 - proj_{\vec{v}_1}\vec{w}_3 - proj_{\vec{v}_2}\vec{w}_3, \\
&\vdots \\
\vec{v}_k &= \vec{w}_k - \sum_{j=1}^{k-1} proj_{\vec{v}_j}\vec{w}_k = \vec{w}_k - \sum_{j=1}^{k-1} \frac{\vec{w}_k \cdot \vec{v}_j}{||\vec{v}_j||^2}\vec{v}_j.
\end{aligned}$$

Entonces, $\{\vec{v}_1, \ldots \vec{v}_m\}$ es una base ortogonal para S. En particular, todo subespacio S de $\mathbb{R}^n$ tiene una base ortogonal.

Demostración. La demostración se realiza por inducción sobre m, es decir, sobre la dimensión del subespacio S. Si $m = 1$, la demostración es obvia ya que si $\{\vec{w}_1\}$ es una base para S, entonces $span\{\vec{w}_1\} = span\{\vec{v}_1\}$ y $\{\vec{v}_1\}$ es una base ortogonal para S. Supongamos ahora que el teorema se cumple para cualquier $m - 1 \geq 1$ y mostraremos que también será válido para m. Sea $\{\vec{w}_1, \ldots, \vec{w}_m\}$ una base para S y consideremos al siguiente subespacio

$$S_{m-1} = span\{\vec{w}_1, \ldots, \vec{w}_{m-1}\},$$

mismo que tiene dimensión $m - 1$ y está contenido en S. Por hipótesis de inducción se tiene que $\{\vec{v}_1, \ldots, \vec{v}_{m-1}\}$ es una base ortogonal para S_{m-1}. Así pues,

$$\vec{v}_m = \vec{w}_m - \sum_{j=1}^{k-1} proj_{\vec{v}_j}\vec{w}_m = \vec{w}_m - proj_{S_{m-1}}\vec{w}_m.$$

es un vector ortogonal a todos los elementos del conjunto ortogonal $\{\vec{v}_1, \ldots, \vec{v}_{m-1}\}$. Como $proj_{S_{m-1}}\vec{w}_m \in S_{m-1} \subsetneq S$, es claro que $\vec{v}_m \in S$ y que $\vec{v}_m \neq \vec{0}$ dado que $\vec{w}_m \notin S_{m-1}$. De esta forma, la Proposición 14.27, garantiza que $\{\vec{v}_1, \ldots \vec{v}_m\}$ es una base ortogonal para S. ■

Observación 15.12 *Si $\{\vec{v}_1, \ldots, \vec{v}_m\}$ es la base ortogonal para S obtenida por el proceso de Gram-Schimdt, normalizando los vectores de esta base obtenemos lo que se conoce como una base ortonormal $\{\vec{u}_1, \ldots, \vec{u}_m\}$ para S; es decir, se tiene que para todo $i, j = 1, 2, \ldots, m$, se cumple*

$$\vec{u}_i \cdot \vec{u}_j = \begin{cases} 0 \ si \ i \neq j, \\ 1 \ si \ i = j. \end{cases}$$

y además $\|\vec{u}_i\| = 1$. *Los vectores de esta base ortonormal están dados por*

$$\vec{u}_i = \frac{1}{\|\vec{v}_i\|}\vec{v}_i.$$

Si S es un subespacio de $\mathbb{R}^n$ y $\vec{v} \in \mathbb{R}^n$, recordemos que en la Definición 15.4 se requería de una base ortogonal de S para encontrar la proyección ortogonal de $\vec{v}$ sobre S. El proceso de ortogonalización de Gram-Schimdt nos proporciona una forma de encontrar tal base, de manera que siempre tiene sentido hablar de $proj_S\vec{v}$.

En la siguiente sección caracterizaremos a la proyección ortogonal cuando ésta se realiza sobre el espacio columna de una matriz. Como una aplicación importante, explicaremos la técnica de "mínimos cuadrados" (o de regresión) que consiste en interpolar de forma óptima un número finito de elementos de $\mathbb{R}^n$.

Mínimos cuadrados

Sea $S = span\{\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_n\} \subset \mathbb{R}^m$. Definimos la matriz

$$A = [\vec{u}_1\ \vec{u}_2 \cdots \vec{u}_n] \in M_{m\times n},$$

cuyas columnas son los vectores que generan a S. Por construcción, el espacio generado por las columnas de A es precisamente S. Por lo tanto, si $\vec{y}$ es cualquier vector en $\mathbb{R}^m$, como $proj_S\vec{y} \in S$ se tiene que existe alguna $\vec{x} \in \mathbb{R}^n$ tal que

$$A\vec{x} = proj_S\vec{y}.$$

Dado que $\vec{u}_i \in S$, por el Teorema 15.9 sabemos que para todo $i = 1, 2, \ldots, n$, se cumple que

$$0 = (\vec{y} - proj_S\vec{y}) \cdot \vec{u}_i = (\vec{y} - A\vec{x}) \cdot \vec{u}_i = (\vec{y} - A\vec{x})^T \vec{u}_i.$$

Estas n igualdades implican que

$$(\vec{y} - A\vec{x})^T A = \vec{0}^T,$$

o bien, tomando transpuestas de ambos lados,

$$A^T(\vec{y} - A\vec{x}) = \vec{0}.$$

Como consecuencia de esta discusión tenemos el siguiente resultado.

Proposición 15.13 *Sean $S = span\{\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_n\}$, $\vec{y}$ cualquier vector en $\mathbb{R}^m$ y*

$$A = [\vec{u}_1\ \vec{u}_2 \cdots \vec{u}_n] \in M_{m\times n}.$$

Entonces existe $\vec{x} \in \mathbb{R}^n$ tal que $A\vec{x} = proj_S\vec{y}$ y tal $\vec{x}$ satisface

$$A^T A\vec{x} = A^T\vec{y}.$$

En particular, si la matriz $A^T A \in M_{n\times n}$ es invertible, entonces

$$\vec{x} = \left(A^T A\right)^{-1} A^T\vec{y}.$$

Con la misma notación de arriba, al vector $proj_S\vec{y}$ se le conoce como la **aproximación de mínimos cuadrados** de $\vec{y}$ sobre S y a $\vec{y}-proj_S\vec{y}$ se le denomina el **vector residual.**

Una aplicación importante de la Proposición 15.13 es que nos da la posibilidad de encontrar la mejor solución posible para un sistema inconsistente de ecuaciones lineales en el siguiente sentido:

Si el sistema lineal $A\vec{x} = \vec{y}$, en donde $A \in M_{m\times n}$ y $\vec{y} \in \mathbb{R}^m$, es inconsistente; entonces se tiene que

$$\|\vec{y} - A\vec{x}\| > 0.$$

Si $S = span\{A^1, \ldots, A^n\}$, la Proposición 15.13 nos proporciona un vector $\vec{x}$ en $\mathbb{R}^n$ tal que $A\vec{x} = proj_S\vec{y}$, concretamente $\vec{x}$ satisface que $A^T A\vec{x} = A^T\vec{y}$. Así, por el inciso 4 del Teorema 15.9, este vector $\vec{x}$ minimiza la distancia entre y y $A\vec{x}$ de manera que es la mejor solución posible al sistema $A\vec{x} = \vec{y}$.

Ejemplo 15.14 *Sea*

$$A\vec{x} = \vec{y}$$

un sistema de ecuaciones con

$$A = \begin{bmatrix} 2 & 1 \\ -2 & 2 \\ 1 & 1 \\ -1 & 1 \end{bmatrix} \quad y \quad \vec{y} = \begin{bmatrix} 33 \\ -33 \\ 44 \\ -66 \end{bmatrix}.$$

La FER de la matriz aumentada de este sistema está dada como

$$\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right],$$

de manera que el sistema es inconsistente. Es fácil verificar que $A^T A$ es invertible y por la Proposición 15.13, la mejor solución posible al sistema queda dada por

$$\begin{aligned}
\vec{x} &= \left(A^T A\right)^{-1} A^T \vec{y} \\
&= \left(\begin{bmatrix} 2 & -2 & 1 & -1 \\ 1 & 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -2 & 2 \\ 1 & 1 \\ -1 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 2 & -2 & 1 & -1 \\ 1 & 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 33 \\ -33 \\ 44 \\ -66 \end{bmatrix} \\
&= \begin{bmatrix} \frac{7}{66} & \frac{1}{33} \\ \frac{1}{33} & \frac{5}{33} \end{bmatrix} \begin{bmatrix} 2 & -2 & 1 & -1 \\ 1 & 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 33 \\ -33 \\ 44 \\ -66 \end{bmatrix} \\
&= \begin{bmatrix} 24 \\ -1 \end{bmatrix}.
\end{aligned}$$

Aproximación lineal

Supongamos que tenemos m observaciones o datos (correspondientes a millones de botellas de vino vendidas en cierta región, por ejemplo) a lo largo del tiempo:

$$(t_1, y_1), (t_2, y_2), \ldots, (t_m, y_m)$$

con $t_1 \leq t_2 \leq \ldots \leq t_m$. Se conjetura que existe una relación lineal del tipo

$$y = ct + d,$$

entre el tiempo y el número de botellas vendidas. Como se muestra en la Figura 15.2, cualquier recta que se proponga pasará, probablemente, por muy pocos puntos.

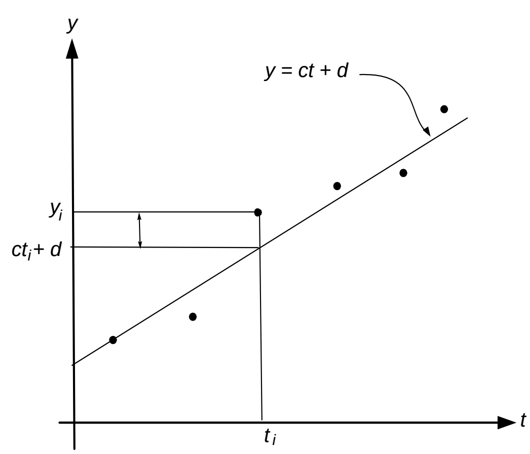


Figura 15.2: Aproximación lineal.

En términos algebraicos, tenemos m valores para t y y y queremos encontrar valores para c y d tales que se cumplan las m ecuaciones dadas por

$$ct_i + d = y_i, \quad i = 1, 2, \ldots, m.$$

La matriz de coeficientes de este sistema (cuyas variables son c y d) está dada por

$$A = \begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{bmatrix}$$

y si $m > 2$ el sistema será inconsistente, a menos que todos los puntos estén perfectamente alineados en una recta (poco probable).

La situación que tenemos es que el vector

$$\vec{y} = \begin{bmatrix} y_i \\ \vdots \\ y_m \end{bmatrix},$$

usualmente no será combinación lineal de las columnas de la matriz A por lo que el sistema

$$A \begin{bmatrix} c \\ d \end{bmatrix} = \vec{y}$$

será inconsistente. Sin embargo, podemos encontrar la solución

$$\vec{x} = \begin{bmatrix} c \\ d \end{bmatrix}$$

que mejor aproxima a este sistema procediendo como en el Ejemplo 15.14.

Utilizando la Proposición 15.13 sabemos que $\vec{x}$ es tal que

$$A^T A\vec{x} = A^T \vec{y}. \tag{$\spadesuit$}$$

Dado que el vector $\vec{x}$ minimiza la distancia

$$\|\vec{y} - A\vec{x}\|,$$

también minimiza la expresión

$$\|\vec{y} - A\vec{x}\|^2 = \sum_{i=1}^{m} (y_i - ct_i - d)^2$$

y de aquí el nombre de mínimos cuadrados.

Resulta natural suponer que al menos dos de las observaciones fueron obtenidas en tiempos distintos, de manera que existirán índices i y j tales que $t_i \neq t_j$. En tal caso, las columnas de A forman un conjunto LI y el siguiente lema nos garantiza que en este caso $A^T A$ es invertible. De esta forma, por la Proposición 15.13, $\vec{x}$ puede obtenerse como $\vec{x} = \left(A^T A\right)^{-1} A^T \vec{y}$.

Lema 15.15 *Sea $A \in M_{m\times n}$ y supongamos que las columnas de A forman un conjunto LI, entonces $A^T A \in M_{n\times n}$ es invertible.*

Demostración. Supongamos que $\vec{z} \in \mathbb{R}^n$ es tal que $A^T A\vec{z} = \vec{0}$. Notemos que

$$0 = \vec{z}^T\vec{0} = \vec{z}^T A^T A\vec{z} = (A\vec{z})^T (A\vec{z}) = (A\vec{z}) \cdot (A\vec{z}) = \|A\vec{z}\|^2 .$$

Por lo tanto $A\vec{z} = \vec{0}$ y como las columnas de A forman un conjunto LI, se tiene que $\vec{z} = \vec{0}$ y en consecuencia $A^T A$ es invertible (ver, por ejemplo, la Proposición 8.49). ■

Corolario 15.16 *La solución al problema de encontrar la recta $y = ct + d$ que mejor aproxime a un conjunto dado de puntos*

$$\{(t_1, y_1), (t_2, y_2), \ldots, (t_m, y_m)\}$$

está dada por

$$\vec{x} = \begin{bmatrix} c^* \\ d^* \end{bmatrix} = (A^T A)^{-1} A^T \vec{y},$$

en donde $A = \begin{bmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{bmatrix}$ *y* $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$.

Ejemplo 15.17 *Supongamos que los datos observados para los millones de botellas de vino vendidas en los años $t = 1, 2, 3, 4$ son $(1, 2)$, $(2, 3)$, $(3, 5)$ y $(4, 7)$. En este caso,*

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} 2 \\ 3 \\ 5 \\ 7 \end{bmatrix}.$$

Por lo tanto,

$$A^T A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 30 & 10 \\ 10 & 4 \end{bmatrix}, \; (A^T A)^{-1} = \frac{1}{20} \begin{bmatrix} 4 & -10 \\ -10 & 30 \end{bmatrix}$$

y se tiene que

$$\vec{x} = \begin{bmatrix} c \\ d \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 4 & -10 \\ -10 & 30 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 1.7 \\ 0 \end{bmatrix}.$$

Así, la recta que mejor predice la venta de vino en el futuro es

$$y = ct + d = 1.7t.$$

Ejemplo 15.18 *Uno de los primeros matemáticos en utilizar el método de mínimos cuadrados fue K.F.Gauss, quien en 1801 aproximó la órbita del asteroide Ceres. Las órbitas de los cuerpos celestes del sistema solar siempre siguen una curva cónica. En coordenadas polares, la distancia r y el ángulo θ con respecto al sol satisfacen la ecuación*

$$r = a - e(r\cos\theta),$$

en donde e representa la excentricidad de la órbita y $a > 0$. Recordemos que $e = 0$ en un círculo, $0 < e < 1$ en una elipse, $e = 1$ en una parábola y $e > 1$ en una hipérbola.

Supongamos que las observaciones de un cometa arrojan el siguiente conjunto de datos para (r, θ) :

$$\{(1.2, 0.3), (2.1, 1.2), (4.1, 2.6), (6.3, 3.8)\}.$$

Para encontrar la órbita aproximada debemos encontrar a y e que mejor aproximen el sistema lineal

$$\begin{aligned} a - e(1.2\cos 0.3) &= 1.2, \\ a - e(2.1\cos 1.2) &= 2.1, \\ a - e(4.1\cos 2.6) &= 4.1, \\ a - e(6.3\cos 3.8) &= 6.3, \end{aligned}$$

mismo que puede reescribirse como

$$\begin{aligned} a - 1.146e &= 1.2, \\ a - 0.761e &= 2.1, \\ a + 3.513e &= 4.1, \\ a + 4.983e &= 6.3. \end{aligned}$$

De aquí

$$A = \begin{bmatrix} 1 & -1.146 \\ 1 & -0.761 \\ 1 & 3.513 \\ 1 & 4.983 \end{bmatrix}, \ \vec{y} = \begin{bmatrix} 1.2 \\ 2.1 \\ 4.1 \\ 6.3 \end{bmatrix}.$$

Por lo tanto,

$$\vec{x} = \begin{bmatrix} a \\ e \end{bmatrix} = \left(A^T A\right)^{-1} A^T \vec{y} = \begin{bmatrix} 2.242\,3 \\ 0.718\,05 \end{bmatrix}.$$

Concluimos que la órbita cónica que mejor aproxima al conjunto de datos observados es

$$r = 2.242\,3 - 0.718\,05(r\cos\theta)$$

lo cual corresponde a una órbita elíptica.

Ejercicios

Ejercicio 15.1 *Mostrar que la matriz* $\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ *representa la transformación lineal que proyecta (ortogonalmente) a cualquier vector de* $\mathbb{R}^2$ *sobre la recta* $y = x$.

Ejercicio 15.2 *Sea* $S \subset \mathbb{R}^n$. *El* ***complemento ortogonal*** *de* S, *denotado por* $S^\perp$, *se define como el conjunto*

$$S^\perp = \{\vec{v} \in \mathbb{R}^n \mid \vec{v} \cdot \vec{x} = 0 \text{ para todo } \vec{x} \in S\}.$$

a. *Demostrar que* $S^\perp$ *es un subespacio de* $\mathbb{R}^n$.

b. *Si* $S = \{\vec{0}\} \subset \mathbb{R}^n$, *encontrar* $S^\perp$.

c. *Si* $S = \mathbb{R}^n$, *encontrar* $S^\perp$.

d. *Si* $S = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\} \subset \mathbb{R}^2$, *encontrar* $S^\perp$.

e. *Si* $S = \left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\} \subset \mathbb{R}^3$, *encontrar* $S^\perp$.

f. *Si* $S = \left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\} \subset \mathbb{R}^3$, *encontrar* $S^\perp$.

g. *Si* $S = span\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\} \subset \mathbb{R}^3$, *encontrar* $S^\perp$

Ejercicio 15.3 *Probar que si* $S = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$, *entonces*

$$S^\perp = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}^\perp.$$

Ejercicio 15.4 *Encontrar* $proy_S\vec{v}$ *en los siguientes casos.*

a. $S = span\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$, $\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

b. $S = span\left\{\begin{bmatrix} -1 \\ 3 \end{bmatrix}\right\}$, $\vec{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

c. $S = span\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$, $\vec{v} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

d. $S = span\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}-1\\0\\1\end{bmatrix}\right\}, \vec{v} = \begin{bmatrix}1\\2\\3\end{bmatrix}.$

e. $S = span\left\{\begin{bmatrix}0\\-1\\0\\1\end{bmatrix}, \begin{bmatrix}2\\1\\0\\1\end{bmatrix}, \begin{bmatrix}-2\\2\\0\\2\end{bmatrix}\right\}, \vec{v} = \begin{bmatrix}0\\1\\1\\0\end{bmatrix}.$

Ejercicio 15.5 *Si $\vec{v}$ y S son los mismos de los incisos del ejercicio anterior, expresar a cada $\vec{v}$ como $\vec{v} = proy_S\vec{v} + \vec{w}$, en donde $\vec{w} \in S^{\perp}$.*

Ejercicio 15.6 *Utilizar el proceso de ortogonalización de Gram-Schmidt para encontrar una base ortogonal de los siguientes subespacios.*

a. $S = span\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}\right\}.$

b. $S = span\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}-1\\2\\2\end{bmatrix}\right\}.$

c. $S = span\left\{\begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right\}.$

d. $S = span\left\{\begin{bmatrix}1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\-1\end{bmatrix}\right\}.$

e. $S = span\left\{\begin{bmatrix}1\\2\\1\\2\end{bmatrix}, \begin{bmatrix}2\\4\\0\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\2\\0\end{bmatrix}\right\}.$

Ejercicio 15.7 *Encontrar bases ortonormales para los subespacios de los incisos a) y c) del ejercicio anterior.*

Ejercicio 15.8 *Dadas las siguientes matrices simétricas,*

$$A = \begin{bmatrix}1 & 4 & 3\\4 & 1 & 0\\3 & 0 & 1\end{bmatrix}, \; B = \begin{bmatrix}1 & 1 & 1\\1 & 1 & 1\\1 & 1 & 1\end{bmatrix},$$

a. *Encontrar todos sus eigenvalores y eigenvectores.*

b. *Verificar que los eigenvectores correspondientes a eigenvalores distintos son ortogonales.*

c. *Encontrar dos bases ortonormales para* $\mathbb{R}^3$ *consistentes de eigenvectores de* A *y* B*, respectivamente. Sugerencia: si* $Eig(\lambda)$ *tiene dimensión mayor a uno para algún eigenvalor* λ*, utilizar el proceso de Gram - Schmidt para encontrar una base ortogonal para* $Eig(\lambda)$ *y normalizar los vectores.*

d. *Como consecuencia del inciso anterior,* A *y* B *son diagonalizables de manera que existen* $D_1, D_2 \in M_{3\times 3}$*, diagonales y* $P_1, P_2 \in M_{3\times 3}$*, invertibles, tales que*

$$D_1 = P_1^{-1} A P_1, \ D_2 = P_2^{-1} A P_2.$$

Las matrices cuyos vectores columna forman un conjunto ortonormal se denominan ***matrices ortogonales****, por lo que* P_1 *y* P_2 *son dos matrices ortogonales.*

e. *Probar que si* P *es una matriz ortogonal, entonces* $P^T P = I$ *(por lo que* $P^{-1} = P^T$*)*

Ejercicio 15.9 *Utilizar la aproximación de mínimos cuadrados para encontrar la recta que mejor aproxime al siguiente conjunto de puntos en el plano:* $\{(1,1),(2,1),(3,2)\}$.

Ejercicio 15.10 *Considerar el siguiente sistema de ecuaciones*

$$\begin{aligned} x + y &= 1, \\ 2x - y &= 2, \\ -x - y &= 0. \end{aligned}$$

a. *Demostrar que el sistema es inconsistente.*

b. *Sean* $A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -1 & -1 \end{bmatrix}$ *la matriz de coeficientes y* $\vec{b} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ *el vector de términos independientes. Encontrar* $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$ *de tal forma que* $\left\| \vec{b} - A\vec{v} \right\|$ *sea mínima. Observar que* $\vec{v}$ *es la solución que "mejor" aproxima al sistema de ecuaciones.*

Apéndice A

Soluciones a los ejercicios

Capítulo 1

1.1. c) Sea $x \in A \cup (A \cap B)$, entonces $x \in A$ o $x \in A \cap B$. Como $A \cap B \subset A$, entonces en ambas instancias $x \in A$ y se tiene

$$A \cup (A \cap B) \subset A.$$

Ahora bien, si $x \in A$, evidentemente $x \in A \cup (A \cap B)$ y tenemos que

$$A \subset A \cup (A \cap B).$$

Concluimos así que

$$A \cup (A \cap B) = A.$$

e) Sea $x \in A \cap (B \cup C)$, entonces se sigue que $x \in A$ y $x \in B \cup C$, es decir, "$x \in A$" y "$x \in B$ o $x \in C$". Entonces "$x \in A$ y $x \in B$" o "$x \in A$ y $x \in C$" lo cual es simplemente $x \in (A \cap B) \cup (A \cap C)$. Se tiene así que

$$A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C).$$

La otra contención es similar.

1.2. Utilizar que $\mathbb{R} = \mathbb{Q} \cup \mathbb{I}$.

1.3. a) Sea $a \in \mathbb{Z}$ par de manera que $a = 2n$ para algún $n \in \mathbb{Z}$. Entonces

$$a + 1 = 2n + 1$$

por lo que $a+1$ es impar. Similarmente, si $a+1$ es impar se tiene que $a + 1 = 2n + 1$ para algún entero n. Claramente

$$a = a + 1 - 1 = 2n + 1 - 1 = 2n,$$

por lo que a es par.

1.4. Supongamos que $rs \notin \mathbb{I}$, por lo tanto $rs \in \mathbb{Q}$ y existen enteros p, q con $q \neq 0$ tales que $rs = \frac{p}{q}$. Ahora bien, como $r \in \mathbb{Q}$, $r \neq 0$, existen enteros a, b, ambos diferentes del cero, de manera que $r = \frac{a}{b}$. Se tiene así que

$$rs = \frac{a}{b}s = \frac{p}{q}$$

por lo que $s = \frac{bp}{aq} \in \mathbb{Q}$. Esto último contradice la hipótesis de que s es un irracional y, por lo tanto, debe tenerse $rs \in \mathbb{I}$.

1.5. b) Sea $x \in \mathbb{R}$, $x \neq 1$. La igualdad se satisface para $n = 1$ ya que

$$(x + 1)(x - 1) = x^2 - 1$$

y, por lo tanto,

$$1 + x = \frac{x^2 - 1}{x - 1}.$$

Supongamos que la igualdad es válida para $n - 1$. Entonces

$$\begin{aligned}\left(1 + x + x^2 + \cdots + x^{n-1}\right) + x^n &= \frac{x^n - 1}{x - 1} + x^n \\ \frac{x^n - 1 + x^n(x-1)}{x - 1} &= \frac{x^{n+1} - 1}{x - 1}.\end{aligned}$$

Capítulo 2

2.1. a) La matriz aumentada, la matriz de coeficientes y el vector de términos independientes son respectivamente

$$\left[\begin{array}{cccc|r} 6 & 1 & 0 & 0 & 8 \\ 1 & \pi & 1 & 0 & -1 \\ 1 & 2 & 1 & 1 & 1 \end{array}\right], \quad \begin{bmatrix} 6 & 1 & 0 & 0 \\ 1 & \pi & 1 & 0 \\ 1 & 2 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 8 \\ -1 \\ 1 \end{bmatrix}.$$

2.2. Las matrices A, D y F no están en FE, las matrices B y C están en FER, la matriz E está en FE pero no en FER.

2.3. Las FER's de las matrices B y C son,

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

.

2.4. En el caso de a) las variables x, y y z son restringidas.

2.5. Los conjuntos solución de estos sistemas son como sigue:

a) $S = \{(2, -3)\}$.
b) $S = \{(0, 0)\}$.
c) $S = \{(5, -1, 4)\}$.
d) $S = \{(0, 0, 0)\}$.
e) $S = \emptyset$.
f) $S = \{(2y, y) \mid y \in \mathbb{R}\}$.
g) $S = \{(-3 + 2y, y) \mid y \in \mathbb{R}\}$.
h) $S = \{(\frac{1}{4}, \frac{-3}{4} + z, \frac{3}{4} - 2z, z) \mid z \in \mathbb{R}\}$.
i) $S = \{(0, z, -2z, z) \mid z \in \mathbb{R}\}$.

2.6. Para la matriz A, x_1 y x_3 son variables restringidas y x_2 y x_4 son libres. El conjunto solución es

$$S = \left\{ \begin{bmatrix} 1 + 2x_2 + x_4 \\ x_2 \\ \pi - 2x_4 \\ x_4 \end{bmatrix} \mid x_2, x_4 \in \mathbb{R} \right\} \subset \mathbb{R}^4.$$

Para la matriz C,Ambas variables son restringidas y $x_1 = -1, x_2 = -2$.

2.7. Los enunciados a), b), c), d), f) y h) son falsos mientras los enunciados e), g), i) y j) son verdaderos.

2.8. En el caso de ii), el sistema no tiene solución syss $c = 5$ y $k \neq -1$, tiene solución única syss $c \neq 5$ y tiene infinitas soluciones syss $c = 5$ y $k = -1$.

2.9. Para los cruceros A, B,C y D se tiene que el tráfico entrante es igual al saliente, es decir

$$\begin{aligned} 700 + 300 &= w + x, \\ w + y &= 900 + 200, \\ x + z &= 200 + 400, \\ 300 + 400 &= y + z. \end{aligned}$$

Este sistema se puede reescribir como

$$\begin{aligned} w + x &= 1000, \\ w + y &= 1100, \\ x + z &= 600, \\ y + z &= 700. \end{aligned}$$

De aquí se concluye que

$$\begin{aligned} z &\text{ es libre,} \\ w &= 400 + z, \\ x &= 600 - z, \\ y &= 700 - z. \end{aligned}$$

Como no tiene sentido tener flujos negativos, z no podrá ser enteramente libre y debe cumplirse $0 \leq z \leq 600$. Por ejemplo, si el flujo $z = 100$, entonces se tendría que $w = x = 500$ y $y = 600$.

2.10 El sistema queda descrito por las siguientes tres ecuaciones:

$$\begin{aligned} c &= 500 + \frac{1}{8}\left(I - \frac{I}{4}\right), \\ p &= \frac{I}{4} - 500 \\ I - \frac{I}{4} - \frac{1}{8}\left(I - \frac{I}{4}\right) &= 2100, \end{aligned}$$

Se deduce que $I = \$3200, c = \$800$ y $p = \$300$. En conclusión, el vendedor ambulante paga \$1100 semanales en cuotas, Si este vendedor fuese formal y pagara impuestos, ¡estas cuotas equivaldrían a un impuesto del 34 % sobre su ingreso bruto!

Capítulo 3

3.2. Se sugiere recordar que $c\vec{v} = \vec{0}$ syss $c = 0$ o $\vec{v} = \vec{0}$.

3.4. a) $\vec{v}_4 \in span\{\vec{v}_1, \vec{v}_2\}$. Más aún, $x_1 = 2$ y $x_2 = 3$.

b) $\vec{v}_4 \notin span\{\vec{v}_1, \vec{v}_3\}$.

c) $\vec{v}_4 \in span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Más aún, $x_1 = 2 - 5x_3$, $x_2 = 3$ y $x_3 \in R$.

d) $\vec{v}_2 \notin span\{\vec{v}_1\}$.

e) $\vec{v}_3 \in span\{\vec{v}_1\}$. Más aún, $x_1 = 5$.

3.5. a) $\vec{v}_3 \notin span\{\vec{v}_1, \vec{v}_2\}$.

b) $\vec{v}_4 \in span\{\vec{v}_1, \vec{v}_2\}$. Más aún, $x_1 = 2$ y $x_2 = -3$.

c) $\vec{v}_5 \in span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Más aún, $x_1 = -\frac{1}{11}$, $x_2 = \frac{3}{11}$ y $x_3 = \frac{1}{11}$.

d) $\vec{v}_5 \notin span\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$.

3.6. $a \in \mathbb{R}$, $b \in \mathbb{R}$, $c \in \mathbb{R}$.

3.7. $a - \frac{2c}{5} - 3b = 0$.

3.9.

$$A\vec{t}_1 = \frac{1}{2}\begin{bmatrix}2\\4\\6\end{bmatrix} - 1\begin{bmatrix}1\\1\\1\end{bmatrix} + \frac{1}{5}\begin{bmatrix}5\\10\\15\end{bmatrix} + 0\begin{bmatrix}-7\\2\\7\end{bmatrix} = \begin{bmatrix}1\\3\\5\end{bmatrix}.$$

3.10.

$$a)\ \begin{bmatrix}-2 & 6\\ 3 & 9\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}-22\\-21\end{bmatrix},\quad c)\ \begin{bmatrix}1 & 2 & 3 & 4\\ 5 & 6 & 7 & 8\\ 3 & 3 & 6 & 9\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} = \begin{bmatrix}1\\2\\3\end{bmatrix}.$$

3.11.

$$a)\ x_1\begin{bmatrix}-2\\3\end{bmatrix} + x_2\begin{bmatrix}6\\9\end{bmatrix} = \begin{bmatrix}-22\\-21\end{bmatrix}.$$

$$b)\ x_1\begin{bmatrix}2\\1\\2\end{bmatrix} + x_2\begin{bmatrix}1\\1\\1\end{bmatrix} + x_3\begin{bmatrix}3\\1\\-1\end{bmatrix} = \begin{bmatrix}21\\8\\5\end{bmatrix}.$$

3.12. a) $span\{A^1, A^2\} = \mathbb{R}^2$. b) $span\{B^1, B^2, B^3\} = \mathbb{R}^2$.
c) $span\{C^1, C^2, C^3\} \subsetneq \mathbb{R}^2$. d) $span\{D^1\} \subsetneq \mathbb{R}^2$.

3.13. a) $span\{A^1, A^2, A^3\} \subsetneq \mathbb{R}^3$. b) $span\{B^1, B^2, B^3\} = \mathbb{R}^3$.
c) $span\{C^1, C^2\} \subsetneq \mathbb{R}^3$.

3.15. Evidentemente, $\vec{v}_1 + \vec{v}_2$ y $\vec{v}_1 - \vec{v}_2$ pertenecen a $span\{v_1, v_2\}$, por lo que el Ejercicio 3.14 asegura que $span\{\vec{v}_1 + \vec{v}_2, \vec{v}_1 - \vec{v}_2\} \subset span\{\vec{v}_1, \vec{v}_2\}$. La otra contención se puede argumentar de forma similar. Exprese pues a v_1 y v_2 como combinaciones lineales de $\vec{v}_1 + \vec{v}_2$ y $\vec{v}_1 - \vec{v}_2$.

3.16. Se sugiere volver a utilizar el Ejercicio 3.14.

Capítulo 4

4.1. a) $\{\vec{v}_1\}$ es LI y $span\{\vec{v}_1\} = \{a\vec{v}_1 \mid a \in \mathbb{R}\}$.
b) $\{\vec{v}_1, \vec{v}_2\}$ es LD y $span\{\vec{v}_1, \vec{v}_2\} = span\{\vec{v}_1\}$.
c) $\{\vec{v}_1, \vec{v}_3\}$ es LI y $span\{\vec{v}_1, \vec{v}_3\} = \mathbb{R}^2$.
d) $\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$ es LD y $span\{\vec{v}_1, \vec{v}_2, \vec{v}_4\} = \mathbb{R}^2$.
e) $\{\vec{v}_1, \vec{v}_3, \vec{v}_4\}$ es LD y $span\{\vec{v}_1, \vec{v}_3, \vec{v}_4\} = \mathbb{R}^2$.
f) $\{\vec{v}_1, \vec{v}_2, \vec{v}_5\}$ es LD y $span\{\vec{v}_1, \vec{v}_2, \vec{v}_5\} = span\{\vec{v}_1\}$.

4.2. a) $\{\vec{v}_1, \vec{v}_2\}$ es LD y $span\{\vec{v}_1, \vec{v}_2\} = span\{\vec{v}_1\} = \{a\vec{v}_1 \mid a \in R\}$.
b) $\{\vec{v}_1, \vec{v}_3\}$ es LI y

$$span\{\vec{v}_1, \vec{v}_3\} = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \mid 17a - 8b + c = 0 \right\}.$$

c) $\{\vec{v}_1, \vec{v}_3, \vec{v}_4\}$ es LD y $span\{\vec{v}_1, \vec{v}_3, \vec{v}_4\} = span\{\vec{v}_1, \vec{v}_3\}$.
d) $\{\vec{v}_1, \vec{v}_3, \vec{v}_5\}$ es LI y $span\{\vec{v}_1, \vec{v}_3, \vec{v}_5\} = \mathbb{R}^3$.
e) $\{\vec{v}_1, \vec{v}_3, \vec{v}_6\}$ es LI y $span\{\vec{v}_1, \vec{v}_3, \vec{v}_6\} = \mathbb{R}^3$.

4.3. a) $\{\vec{v}_2, \vec{v}_3, \vec{v}_4, \vec{v}_5\}$ es LD. b) $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_5, \vec{v}_6\}$ es LD. c) $\{\vec{v}_1, \vec{v}_2, \vec{v}_4, \vec{v}_5\}$ es LI.

4.4. a) Las columnas de A forman un conjunto LI syss $h \neq 20$.
b) Las columnas de B forman un conjunto LD syss $i = \frac{21}{2}$.
c) Las columnas de C forman un conjunto LI syss $j \in \mathbb{R}$.

4.5. Los enunciados a), c), d) y e) son falsos mientras que los enunciados b) f) y g) son verdaderos.

4.6. Se sugiere suponer primero que α, β y γ son escalares tales que $\alpha(\vec{v}_1 + \vec{v}_2 + \vec{v}_3) + \beta(\vec{v}_2 + \vec{v}_3) + \gamma\vec{v}_3 = \vec{0}$ y luego probar que α, β y γ están forzadas a ser 0.

4.9. Una solución particular del sistema (no homogéneo) es $\vec{s}_p = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, por lo que el conjunto solución es

$$S_H + s_p = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \mid s \in \mathbb{R} \right\} \subset \mathbb{R}^3.$$

Capítulo 5

5.1. T_1 y T_3 no son transformaciones lineales. T_2 sí es una transformación lineal y la matriz estándar que la representa es

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

5.2. S sí es una transformación lineal y la matriz estándar que la representa es

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

U no es una transformación lineal.

5.3. Las matrices estándar que representan a F, G y H son

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \; [1\ 1] \text{ y } \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

respectivamente.

5.4. No. Se sugiere observar que $\begin{bmatrix} 4 \\ -1 \end{bmatrix} = 2\begin{bmatrix} 3 \\ -1 \end{bmatrix} + \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

5.5. Si $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \in \mathbb{R}^2$ y $r \in \mathbb{R}$, se tiene

$$\begin{aligned}
\Psi\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) &= \Psi\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right) \\
&= \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \\ 0 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ 0 \end{bmatrix} \\
&= \Psi\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) + \Psi\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}, \\
\Psi\left(r\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) &= \Psi\left(\begin{bmatrix} rx_1 \\ ry_1 \end{bmatrix}\right) = \begin{bmatrix} rx_1 \\ ry_1 \\ 0 \end{bmatrix} \\
&= r\begin{bmatrix} x_1 \\ y_1 \\ 0 \end{bmatrix} = r\Psi\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right)
\end{aligned}$$

y Ψ es lineal.

5.7. La matriz que representa a T es

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

5.8. La matriz estándar que representa a U es

$$\begin{bmatrix} -4 & -2 \\ 2 & 6 \end{bmatrix}.$$

5.9. La matriz estándar que representa a S es

$$\begin{bmatrix} 4 & -3 \\ 5 & 2 \end{bmatrix}.$$

5.11. Una de estas mil transformaciones puede ser

$$T_1\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \pi & 1 \\ 666 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

con $T_1\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

5.12. Una posibilidad muy sencilla es

$$L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 27 & 0 \\ \frac{\sqrt{2}}{3} & 0 \\ 0 & 0 \\ -3 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

5.13. a) $Rango(L_1) = \mathbb{R}^2$ y L_1 es suprayectiva.

b) $Rango(L_2) = span\left\{\begin{bmatrix} -3 \\ \frac{1}{2} \end{bmatrix}\right\} \subsetneq \mathbb{R}^2$ y L_2 no es suprayectiva.

c) $Rango(L_3) = span\left\{\begin{bmatrix} -3 \\ \frac{1}{2} \end{bmatrix}\right\} \subsetneq \mathbb{R}^2$ y L_3 no es suprayectiva.

d) $Rango(L_4) = \mathbb{R}^3$y L_4 es suprayectiva.

e) $Rango(L_5) = \left\{\begin{bmatrix} a \\ b \\ c \end{bmatrix} \mid c - 3a - 9b = 0\right\} \subsetneq \mathbb{R}^3$ y L_5 no es suprayectiva.

f) $Rango(L_6) = span\left\{\begin{bmatrix} 5 \\ -2 \\ -3 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \\ 6 \end{bmatrix}\right\} \subsetneq \mathbb{R}^3$ y L_6 no es suprayectiva.

g) $Rango(L_7) = span \left\{ \begin{bmatrix} 5 \\ -2 \\ -3 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \\ 6 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 3 \\ 1 \end{bmatrix} \right\} \subsetneq \mathbb{R}^4$ y L_7 no es suprayectiva.

5.14. Las únicas transformaciones inyectivas son L_1, L_4, L_5 y L_7.

5.15. Se sugiere utilizar que si A es la matriz estándar que representa a T, entonces el sistema homogeneo $A\vec{x} = \vec{0}$ tiene un número infinito de soluciones.

5.16. Se sugiere meditar sobre el Ejercicio 5.11.

5.17. Recuerde que debe probarse que existen escalares $c_1, c_2, \ldots, c_p$, no todos cero, tales que

$$c_1 T(\vec{v}_1) + c_2 T(\vec{v}_2) + \ldots + c_p T(\vec{v}_p) = \vec{0}.$$

5.18. Los enunciados b), c), d) y f) son falsos mientras los enunciados a), e) y g) son verdaderos.

5.19. Aplica las transformaciones a los vértices e ilustra las figuras resultantes en el plano.

5.20. a) La matriz que representa a esta reflexión es $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. c) La matriz que representa a esta contracción es $\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}$.

Capítulo 6

6.1. a) $DC = \begin{bmatrix} 5 & 7 \\ 7 & 1 \end{bmatrix}$, $DC(F-G) = \begin{bmatrix} -10 & -14 \\ -14 & -2 \end{bmatrix}$, $DC - 4I = \begin{bmatrix} 1 & 7 \\ 7 & -3 \end{bmatrix}$.

b) $CD = \begin{bmatrix} 1 & -4 & 9 \\ 35 & 5 & 0 \\ 21 & 3 & 0 \end{bmatrix}$, $C^T D$ no está definido.

c) $BE = \begin{bmatrix} -28 & -27 \\ 16 & 25 \end{bmatrix}$, $BE + 2F^T = \begin{bmatrix} -16 & -37 \\ 28 & 41 \end{bmatrix}$.

d) DE, $C + E$ no están definidos.

e) $(DA)^T = \begin{bmatrix} 7 & 1 \\ 24 & -9 \\ -20 & -27 \end{bmatrix}$.

f) $HL = \begin{bmatrix} -2 & 7 \\ 0 & 0 \\ -2 & 7 \end{bmatrix}$.

g) $BJ = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

h) $GK = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

6.2. En el caso de $(S \circ S)(\vec{x})$ tenemos

$$\begin{aligned}(S \circ S)\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &= S\left(\begin{bmatrix} -4x_1 + x_2 \\ x_1 - x_2 \end{bmatrix}\right) \\ &= \begin{bmatrix} -4(-4x_1 + x_2) + (x_1 - x_2) \\ -4x_1 + x_2 - (x_1 - x_2) \end{bmatrix} \\ &= \begin{bmatrix} 17x_1 - 5x_2 \\ -5x_1 + 2x_2 \end{bmatrix}.\end{aligned}$$

6.3. a) $A = \begin{bmatrix} 6 & -7 \\ 0 & 7 \end{bmatrix}$, $B = \begin{bmatrix} -1 & -\pi \\ 1 & 9 \end{bmatrix}$.

b) $C = \begin{bmatrix} -1 & -\pi \\ 1 & 9 \end{bmatrix}$, $D = \begin{bmatrix} 6 & -7 \\ 0 & 7 \end{bmatrix}$.

c) $E = F = \begin{bmatrix} 6 & -7 \\ 0 & 7 \end{bmatrix}$.

d) $G = \begin{bmatrix} -1 & -\pi \\ 1 & 9 \end{bmatrix}$, $H = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

6.5. $(A - A^T)^T = A^T - A = -\left(A - A^T\right)$ y $\left(A + A^T\right)^T = A^T + A = A + A^T$. La descomposición deseada es

$$A = \underbrace{\frac{1}{2}\left(A + A^T\right)}_{B} + \underbrace{\frac{1}{2}\left(A - A^T\right)}_{C}.$$

6.6. Para demostrar P4, si r es un escalar y

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \ \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

son vectores en $\mathbb{R}^n$, se tiene que

$$\begin{aligned}\underbrace{(rx_1)y_1 + \cdots + (rx_n)y_n}_{(r\vec{x})\cdot\vec{y}} &= \underbrace{r(x_1y_1) + \cdots + r(x_ny_n)}_{r(\vec{x}\cdot\vec{y})} \\ &= \underbrace{x_1\,(ry_1) + \cdots + x_n\,(ry_n)}_{\vec{x}\cdot(r\vec{y})}.\end{aligned}$$

6.8. a) Notamos que si $A, B \in M_{m\times n}$ y $C \in M_{n\times p}$, entonces

$$((A+B)C)_{ij} = (A+B)_i \cdot C^j = (A_i + B_i)\cdot C^j$$

$$= A_i \cdot C^j + A_i \cdot C^j = (AC)_{ij} + (BC)_{ij} = (AC+BC)_{ij},$$

por lo que las matrices $A(B+C)$ y $AB+AC$ son iguales (pues coinciden entrada por entrada).

c) Observamos que

$$(I_mA)_{ij} = \begin{bmatrix} 0 & \ldots & 1 & \ldots & 0 \end{bmatrix} \begin{bmatrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{mj} \end{bmatrix} = 1(a_{ij}) = a_{ij},$$

por lo que las matrices I_mA y A son iguales (pues coinciden entrada por entrada).

6.10. Los enunciados b), c), d) y h) son falsos mientras los enunciados a), e), f) y g) son verdaderos.

6.11. C, D y E son elementales, A, B y F no lo son.

6.12. $$\tilde{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -8 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{32} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} A.$$

6.13 a) $$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}.$$

Capítulo 7

7.1.

$$A^{-1} = -\frac{1}{3}\begin{bmatrix} -3 & -2 \\ 6 & 5 \end{bmatrix},\ C^{-1} = \begin{bmatrix} \frac{7}{10} & -\frac{1}{20} & \frac{1}{4} \\ -\frac{1}{10} & \frac{1}{40} & -\frac{1}{8} \\ -1 & \frac{1}{4} & -\frac{1}{4} \end{bmatrix},$$

$$D^{-1} = \begin{bmatrix} -\frac{7}{8} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{8} & 0 & 0 \end{bmatrix},\ E^{-1} = \begin{bmatrix} -\frac{1}{2} & \frac{3}{2} & \frac{3}{10} \\ \frac{1}{2} & -1 & -\frac{1}{5} \\ -\frac{1}{2} & \frac{1}{2} & \frac{3}{10} \end{bmatrix},$$

$$G^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & -\frac{3}{2} & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Las matrices B, F, H, I, J no tienen inversas.

7.2.

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A^{-1}\vec{a} = \begin{bmatrix} 8 \\ -17 \end{bmatrix},\quad \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = C^{-1}\vec{c} = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = D^{-1}\vec{d} = \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}.$$

7.3. $(A^{-1})^T = \left(A^T\right)^{-1} = A^{-1}$.

7.4. Utilizar el inciso (e) del ejercicio 6.7.

7.5. Observar que $B^{-1} = \frac{1}{7}A$.

7.7. Los enunciados a), b), c), e) y g) son falsos mientras los enunciados d), f) y h) son verdaderos.

7.8. a) $X = A^{-1}B$. b) $X = (A+C)^{-1}B$. c) $X = B(A-D)^{-1}$. d) $X = (C^{-1}B-A)^{-1}$. e) $X = (2I-A)^{-1}D$. f) $X = AB^T(I-B^T)^{-1}$.

7.9.

$$S^{-1}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{2}x + \frac{1}{2}y \\ \frac{1}{2}x - \frac{1}{2}y \end{bmatrix},\ T^{-1}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -2x + 3y \\ 3x - 4y \end{bmatrix}.$$

7.10. $E(1)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -666 \\ 0 & 0 & 1 \end{bmatrix},\ E(3)^{-1} = \begin{bmatrix} \frac{1}{\pi} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$

7.11.

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

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 2 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Capítulo 8

8.3. Sugerencia: si n es impar se deben hacer $\frac{n-1}{2}$ intercambios de renglones y si n es par se deben hacer $\frac{n}{2}$.

8.4. a) -6. b) 3. c) 3. e) $3^n \times 12$.

8.5. Como $A = -A^T$ y n es impar,

$$\det(A) = \det(-A^T) = (-1)^n \det(A^T) = -\det(A^T) = -\det(A),$$

por lo tanto, $\det(A) = 0$.

8.6. $\det(A) = -2$ (A se obtiene de I intercambiando los renglones 1 y 3 y multiplicando el segundo renglón por 2), $\det(C) = 0$ (el tercer renglón es suma del los otros dos).

8.7. $\det(A) = 12$, $\det(C) = 3$.

8.8. $\det(A) = 4$, $\det(C) = 12$, $\det(F) = 0$, $\det(G) = -30$, $\det(J) = -x^2 - y^2 - z^2$.

8.9. b) $\lambda = 1$.

8.10. $\det\left(\begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 2 & 4 & 1 \end{bmatrix}\right) = 1 \times 1 \times 1 \times (-1) = -1.$

8.11. *a*) $adj(B) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$

b) $\det(B) = 1$.

c) $B^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$

8.12. a) $\det\left(\begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & -1 \\ 1 & 2 & -3 \end{bmatrix}\right) = -3$, $\det\left(\begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 2 & -3 \end{bmatrix}\right) =$ -3, $x = \frac{-3}{-3} = 1$, similarmente $y = z = 1$.

8.13. $y = \det\left(\begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 3 & -1 \end{bmatrix}\right) \div \det\left(\begin{bmatrix} 2 & 1 & 2 \\ 1 & -1 & 1 \\ 0 & 2 & -1 \end{bmatrix}\right) = \frac{1}{3}$.

8.14. a) Falso. b) Verdadero. c) Falso. d) Verdadero. e) Verdadero. f) Falso. g) Verdadero. h) Falso. i) Falso. j) Falso.

8.16. En la demostración de la Proposición 8.55 evaluar la expresión $\sum_{k=1}^{n} a_{kj}C_{ki}$ para $i = j$ e $i \neq j$. En este caso, definir la matriz $A^{'}$ como aquella que se obtiene de A sustituyendo la columna i de A por la columna j de A y evaluar $\det(A^{'})$ a lo largo de la $i-\acute{e}sima$ columna.

Capítulo 9

9.1. $C = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 7 & -1 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 7 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 0 & 5 \end{bmatrix}$.

9.2. Sea $\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, entonces

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 7 & -3 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ 7x' - 3y' + z' \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix},$$

por lo que $x' = 0$, $y' = 2$ y $z' = 10$. Se tiene así,

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + 2y + 3z \\ 5y + 6z \\ 5z \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 10 \end{bmatrix}$$

y resolviendo, $z = 2$, $y = -2$, $x = -2$.

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

9.4. a) Verdadero. b) Falso. c) Verdadero. d) Falso.

Capítulo 10

10.1. No, por ejemplo $\frac{\pi}{2}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} \pi \\ \pi \end{bmatrix} \notin \mathbb{Z}^2$.

10.2. S_1 es subespacio de $\mathbb{R}^2$ mientras que S_2 y S_3 no son subespacios de $\mathbb{R}^2$.

10.3. S_5 es subespacio de $\mathbb{R}^3$ mientras que S_4 no es subespacio de $\mathbb{R}^3$.

10.4. S_6 y S_7 son ambos subespacios de $\mathbb{R}^4$.

10.5. Algunas posibilidades son

$$A = \begin{bmatrix} 1 & -2 \end{bmatrix},\ B = \begin{bmatrix} \pi & 666 & -7 \end{bmatrix},\ C = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$
$$D = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Notamos también que no existe una matriz cuyo espacio nulo sea igual al conjunto S_2 simplemente porque S_2 no es un subespacio.

10.6. Usando la FER de estas tres matrices tenemos que

$$Null(E) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\},\ Null(F) = span\left\{ \begin{bmatrix} -4 \\ 1 \end{bmatrix} \right\},$$
$$Null(G) = span\left\{ \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

10.7. Usando la FER de las matrices que representan a estas transformaciones tenemos que

$$Null(T) = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\} \text{ y } Null(U) = span\left\{ \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

10.8. Dos de ellos pueden ser $Z_1 = span\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ y $Z_2 = span\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$.

10.9. $\{\vec{0}\}$ y $\mathbb{R}$.

10.10. Se sugiere utilizar el segundo ejercicio del Capítulo 3.

10.11. Para probar que $W_1 \subset W_1 + W_2$ basta notar que si $\vec{u} \in W_1$, entonces $\vec{u} = \vec{u} + \vec{0}_n \in W_1 + W_2$ (pues $\vec{0}_n \in \mathbb{R}^n$ es un elemento del subespacio W_2). De forma análoga se prueba que $W_2 \subset W_1 + W_2$.

10.12.

$$W_1 + W_2 = \left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \mid x, y \in \mathbb{R} \right\} = span \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

10.13. De nuevo,

$$W_1 + W_2 = \left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \mid x, y \in \mathbb{R} \right\} = span \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

10.14. Los enunciados a) y d) son falsos mientras los enunciados b), c), e) y f) son verdaderos.

10.15. $M \cap N = span \left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\}$.

10.16. Si $V \cap W = V$, entonces $V \subset W$ y es inmediato verificar que $V + W = W$. Si $V + W = W$, entonces dado $\vec{v} \in V$ se cumple

$$\vec{v} + \vec{0}_n \in W.$$

Por lo tanto, $V \subset W$ y se tiene que $V \cap W = V$.

Capítulo 11

11.1. a) Sí, b) No, c) No, d) Sí, e) No, f) Sí, g) Sí, h) Sí, i) Sí, j) No.

11.2. Dos de ellas serían: $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right\}$.

11.3. a) $\left\{ \begin{bmatrix} 666 \\ 666 \end{bmatrix} \right\}$, b) $\left\{ \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$, c) $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$.

11.4. a) $\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ b) $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ c) $\varnothing$.

11.5. a) 2, b) 2, c) 2, d) 2, e) 4.

11.6. Dos posibilidades son:

$$a)\ \left\{\begin{bmatrix}\pi\\\pi\\\pi\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}1\\0\\0\end{bmatrix}\right\},$$

$$b)\ \left\{\begin{bmatrix}1\\0\\-1\end{bmatrix},\begin{bmatrix}2\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\0\end{bmatrix}\right\}.$$

11.7. a) $span\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_{n-2}\}$, b) $S_1 = span\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\0\end{bmatrix}\right\}$, $S_2 =$ $span\left\{\begin{bmatrix}0\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\}$.

11.8. b) Basta notar que si $a_1, a_2, a_3 \in \mathbb{R}$ y

$$\begin{aligned}\vec{0} &= a_1\vec{u} + a_2(\vec{u}+\vec{v}) + a_3\left(\vec{u}+\vec{v}+\vec{w}\right)\\ &= (a_1+a_2+a_3)\vec{u} + (a_2+a_3)\vec{v} + a_3\vec{w},\end{aligned}$$

necesariamente $a_1 = a_2 = a_3 = 0$ ya que $\{\vec{u},\vec{v},\vec{w}\}$ es LI. Por lo tanto, $\{\vec{u},\vec{u}+\vec{v},\vec{u}+\vec{v}+\vec{w}\}$ también es LI y como $\dim(S) = 3$, constituye una base para S.

11.9. a) $\left\{\begin{bmatrix}1\\0\\-1\end{bmatrix},\begin{bmatrix}0\\1\\-\frac{1}{2}\end{bmatrix}\right\}$, b) $\left\{\begin{bmatrix}1\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\}$.

11.10.

$$a)\ \left\{\begin{bmatrix}1\\0\\-1\end{bmatrix},\begin{bmatrix}-4\\4\\2\end{bmatrix}\right\}, b)\ \left\{\begin{bmatrix}1\\0\\1\\0\end{bmatrix},\begin{bmatrix}1\\1\\1\\1\end{bmatrix},\begin{bmatrix}1\\-1\\1\\1\end{bmatrix}\right\},$$

$$c)\ \left\{\begin{bmatrix}1\\-1\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\1\\0\end{bmatrix},\begin{bmatrix}1\\1\\1\\1\end{bmatrix}\right\}, d)\ \left\{\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}1\\-1\\0\end{bmatrix}\right\}.$$

11.11. Los enunciados a), c), e) y h) son falsos mientras los enunciados b), d), f) y g) son verdaderos.

11.12. Proceder por contradicción suponiendo que $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ no es una base y, por lo tanto, debe ser un conjunto LD.

Capítulo 12

12.1. A) 1) $\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$, 2) $\left\{\begin{matrix} 1 & 0 \\ 0, & 1 \\ 2 & 4 \end{matrix}\right\}$, 3) $\left\{\begin{bmatrix} -2 \\ -4 \\ 1 \end{bmatrix}\right\}$, 4) $rank(A) = 2$, 5) $nulidad(A) = 1$, $2 + 1 = 3$.

B) 1) $\left\{\begin{matrix} 1 \\ 0 \\ 1 \end{matrix}\right\}$, 2) $\left\{\begin{bmatrix} 1 \\ \pi \end{bmatrix}\right\}$, 3) $\left\{\begin{bmatrix} \pi \\ -1 \end{bmatrix}\right\}$, 4) $rank(B) = 1$, 5) $nulidad(B) = 1$.

C) 1) $\left\{\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\right\}$, 2) $\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right\}$, 3) $\varnothing$, 4) $rank(C) = 3$, 5) $rank(C) = 0$.

E) 1) $\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$. 2) $\left\{\begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$,

3) $\left\{\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right\}$, 4) $rank(E) = 3$, 5) $nulidad(E) = 1$.

H) 1) $\left\{\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$, 2) $\left\{\begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$,

3) $\left\{\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right\}$, 4) $rank(H) = 3$, 5) $nulidad(H) = 1$.

12.2. $a \neq -2$, $b = -1$.

12.3. Se sugiere comenzar probando que si $\vec{x}$ está en el espacio nulo de B, entonces $\vec{x}$ está en el espacio nulo del producto AB.

12.4. Se sugiere comenzar probando que el espacio columna de AB está contenido en el espacio columna de A.

12.5. Se sugiere comenzar notando que $(AB)\vec{x} = A(B\vec{x})$.

12.6. Se sugiere utilizar el ejercicio anterior junto con el hecho de que $rank(E) = rank(E^T)$.

12.7. Se sugiere utilizar dos veces el teorema de la dimensión.

12.8. Los enunciados a), c), f), g), y h) son falsos mientras los enunciados b), d), e) i) j) y k) son verdaderos.

Capítulo 13

13.1. a) $p_A(\lambda) = \lambda(\lambda - 5)$ y los eigenvalores de A son $\lambda_1 = 0$ y $\lambda_2 = 5$. Además,

$$\left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\} \text{ y } \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$$

son bases para $Eig(\lambda_1)$ y $Eig(\lambda_2)$, respectivamente.

b) $p_B(\lambda) = \lambda^2 - 2\lambda - 3$ y los eigenvalores de B son $\lambda_1 = 3$ y $\lambda_2 = -1$. Además,

$$\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\} \text{ y } \left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\}$$

son bases para $Eig(\lambda_1)$ y $Eig(\lambda_2)$, respectivamente.

c) $p_C(\lambda) = \lambda^2 - \lambda + 13$ y C no tiene eigenvalores (reales).

d) $p_D(\lambda) = (\lambda - 1)^2$ y D tiene un único eigenvalor $\lambda_1 = 1$. Además,

$$\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

es base para $Eig(\lambda_1)$.

e) $p_E(\lambda) = \lambda^2$ y E tiene un único eigenvalor $\lambda_1 = 0$. Además, $\{\vec{e}_1, \vec{e}_2\}$ es base para $Eig(\lambda_1)$.

f) $p_F(\lambda) = (1 - \lambda)(2 - \lambda)(3 - \lambda)$ y los eigenvalores de F son $\lambda_1 = 1$ y $\lambda_2 = 2$ y $\lambda_3 = 3$. Además,

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\} \text{ y } \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$$

son bases para $Eig(\lambda_1)$, $Eig(\lambda_2)$ y $Eig(\lambda_3)$, respectivamente.

g) $p_G(\lambda) = -\lambda(2-\lambda)^2$ y los eigenvalores de G son $\lambda_1 = 0$ y $\lambda_2 = 2$. Además,

$$\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\} \text{ y } \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

son bases para $Eig(\lambda_1)$ y $Eig(\lambda_2)$, respectivamente.

h) $p_H(\lambda) = (\lambda+1)(\lambda-2)^2(\lambda-3)$ y los eigenvalores de H son $\lambda_1 = -1$ y $\lambda_2 = 2$ y $\lambda_3 = 3$. Además,

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} -5 \\ -4 \\ 0 \\ 3 \end{bmatrix} \right\} \text{ y } \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

son bases para $Eig(\lambda_1)$, $Eig(\lambda_2)$ y $Eig(\lambda_3)$, respectivamente.

13.2. Se sugiere utilizar las propiedades de determinantes.

13.3. Se sugiere probar que si $\vec{x}$ es un eigenvector de A con respecto a λ, entonces $\vec{x}$ es un eigenvector de A^{-1} con respecto a $\frac{1}{\lambda}$.

13.4. Utilizar la definición de eigenvector.

13.5. Los enunciados a), c), f) y g) son falsos mientras los enunciados b), d) y e) son verdaderos.

13.6. a) A es diagonalizable. Una posibilidad es $D = P^{-1}AP$, donde $P = \begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix}$ y $D = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix}$.

b) B es diagonalizable. Una posibilidad es $D = P^{-1}BP$, donde $P = \begin{bmatrix} 1 & 1 \\ 2 & -2 \end{bmatrix}$ y $D = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$.

c) C no es diagonalizable.

d) D no es diagonalizable.

e) E es diagonalizable. Una posibilidad es $D = P^{-1}EP$, donde $P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ y $D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

f) F es diagonalizable. Una posibilidad es $D = P^{-1}FP$, donde $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ y $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

g) G es diagonalizable. Una posibilidad es $D = P^{-1}GP$, donde

$$P = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \text{ y } D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

h) H no es diagonalizable.

13.7. Se sugiere notar que si $D \in M_{n\times n}$ es una matriz diagonal y $P \in M_{n\times n}$ son tales que $D = P^{-1}AP$, entonces $D^T = D$.

13.8. Se sugiere notar que si $D \in M_{n\times n}$ es una matriz diagonal y $P \in M_{n\times n}$ son tales que $D = P^{-1}AP$, entonces D es invertible (es producto de tres matrices invertibles) y D^{-1} es también una matriz diagonal.

13.9. Ver la demostración de la Proposición 13.38 y utilizar el hecho de que $\lambda I = P^{-1}\lambda IP$.

13.10. b) A es diagonalizable y puede expresarse como

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}_{P}\underbrace{\begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}}_{D}\underbrace{\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}}_{P^{-1}}.$$

Por lo tanto,

$$\begin{aligned} A^{(11)} &= PD^{(11)}P^{-1} \\ &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3^{11} & 0 \\ 0 & (-1)^{11} \end{bmatrix}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \\ &= \begin{bmatrix} 88573 & 88574 \\ 88574 & 88573 \end{bmatrix}. \end{aligned}$$

Cadenas de Markov

13.11. a) sí, b) no, c) no, d) sí.

13.12. a) sí, b) no, c) sí.

13.13. $P^2 = \begin{bmatrix} \frac{5}{9} & \frac{4}{9} \\ \frac{4}{9} & \frac{5}{9} \end{bmatrix}$

13.14. a) no, b) sí, c) no, d) sí.

13.15. $\begin{bmatrix} 0 & \frac{1}{3} \\ 1 & \frac{2}{3} \end{bmatrix}$, por ejemplo.

13.16. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ y $\vec{x}_0 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}$.

13.17. $\begin{bmatrix} \frac{3}{19} \\ \frac{4}{19} \\ \frac{12}{19} \end{bmatrix}$ es el vector estacionario atractor.

13.19. a) $\begin{bmatrix} 0.8 & 0.75 \\ 0.2 & 0.25 \end{bmatrix}$ para la CdMx y $\begin{bmatrix} 0.3 & 0.1 \\ 0.7 & 0.9 \end{bmatrix}$ para Edimburgo.

b) Con un 20 % de probabilidad estará nublado mañana en la CdMx.

c) Con un 88 % de probabilidad seguirá nublado en Edimburgo en dos días.

d) $\begin{bmatrix} 0.79 \\ 0.21 \end{bmatrix}$ para CdMx y $\begin{bmatrix} 0.125 \\ 0.875 \end{bmatrix}$ para Edinburgo.

13.20. a) $\begin{bmatrix} 0.9 & 0.4 \\ 0.1 & 0.6 \end{bmatrix}$.

b) $\begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix}$.

c) El 20 %.

Capítulo 14

14.1. a) $\sqrt{5}, \sqrt{5}, 1, 0, 2\sqrt{5}, \sqrt{3}, 3\sqrt{6}, 1$. b) $\sqrt{20}$, c) $\sqrt{3}$ d) -5, 0, 10, $\frac{6}{\sqrt{2}}$ e) 0, 0, -2, 0.

14.2. Desarrollar el lado derecho de la siguiente igualdad:

$$\|\vec{x}+\vec{y}\|^2 + \|\vec{x}-\vec{y}\|^2 = (\vec{x}+\vec{y})\cdot(\vec{x}+\vec{y}) + (\vec{x}-\vec{y})\cdot(\vec{x}-\vec{y})\,.$$

14.3. $\|\vec{v}+\vec{w}\| = \|\vec{v}-\vec{w}\| = 5$.

14.4. a) $\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}$,

b) $\begin{bmatrix} -1 \\ 2 \end{bmatrix} = -\begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}$,

c) $\begin{bmatrix} 1 \\ -10 \\ 5 \end{bmatrix} = 2\begin{bmatrix} 0 \\ -3 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ -4 \\ -3 \end{bmatrix}$,

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

14.5. Reescribir a la función f como sigue:

$$\begin{aligned} f(r) &= \|\vec{y} - r\vec{x}\|^2 = (\vec{y} - r\vec{x}) \cdot (\vec{y} - r\vec{x}) \\ &= \|\vec{y}\|^2 - 2r\,(\vec{x} \cdot \vec{y}) + r^2\,\|\vec{x}\|^2 \end{aligned}$$

y obtener la primera y la segunda derivadas con respecto a r.

14.6. a) $\frac{2}{\sqrt{2}}$, b) $-\frac{1}{2}$, c) -666.

14.7. a) $\frac{\pi}{2}$, b) $\frac{3\pi}{4}$ c) $\frac{\pi}{3}$ d) π.

14.9. Utilizar las propiedades del producto punto para probar que $(\alpha\vec{x}) \cdot (\beta\vec{y}) = 0$.

14.10. a) sí, b) no, c) no, d) sí, e) no.

14.11. a) $\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$,

b) $\begin{bmatrix} 0 \\ -3 \\ 6 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}$,

c)

$$\begin{aligned} span\mathcal{B}_1 &= span\left\{\begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}\right\}, \\ span\mathcal{B}_2 &= span\left\{\begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{18}} \\ \frac{4}{\sqrt{18}} \\ -\frac{1}{\sqrt{18}} \end{bmatrix}\right\}. \end{aligned}$$

Capítulo 15

15.1. Visualizar la transformación dada en $\mathbb{R}^2$.

15.2. a) $S^\perp$ cumple las propiedades de la Definición 10.1. Por ejemplo, $\vec{0} \in S^\perp$ pues dado cualquier $\vec{s} \in S$ se cumple $\vec{0} \cdot \vec{s} = 0$. Si $\vec{x}, \vec{y} \in S^\perp$, entonces dado $\vec{s} \in S$, se cumple

$$(\vec{x} + \vec{y}) \cdot \vec{s} = \vec{x} \cdot \vec{s} + \vec{y} \cdot \vec{s} = 0 + 0 = 0,$$

con lo cual $\vec{x} + \vec{y} \in S^\perp$. La verificación de que $S^\perp$ es cerrado bajo multiplicación por escalares es similar.

b) $\mathbb{R}^n$,

c) $\{\vec{0}\}$,

d) $S^{\perp} = span\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$,

e) $S^{\perp} = span\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$,

f) $S^{\perp} = span\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$,

g) $S^{\perp} = span\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$.

15.3. Si $\vec{x} \in S^{\perp}$, entonces $\vec{x} \cdot \vec{s} = \vec{0}$ para toda $\vec{s} \in S$; en particular $\vec{x} \cdot \vec{v}_i = \vec{0}$ para toda $i = 1, \ldots, m$, por lo que $\vec{x} \in \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}^{\perp}$. Sea $\vec{v} \in \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}^{\perp}$, entonces $\vec{v} \cdot \vec{v}_i = \vec{0}$ para toda $i = 1, \ldots, m$. Dado $\vec{s} \in S$ se tiene que existen escalares $a_1, \ldots, a_m$ tales que $\vec{s} = a_1\vec{v}_1 + \cdots + a_m\vec{v}_m$, por lo tanto

$$\vec{v} \cdot \vec{s} = \vec{v} \cdot (a_1\vec{v}_1 + \cdots + a_m\vec{v}_m) = 0,$$

por lo que $\vec{v} \in S^{\perp}$. Concluimos que $S^{\perp} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}^{\perp}$.

15.4. a) $\begin{bmatrix} \frac{1}{5} \\ \frac{2}{5} \end{bmatrix}$, b) $\begin{bmatrix} -\frac{3}{10} \\ \frac{9}{10} \end{bmatrix}$, c) $\begin{bmatrix} \frac{5}{6} \\ \frac{2}{6} \\ \frac{1}{6} \end{bmatrix}$, d) $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, e) $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$.

15.5. a) $\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{5} \\ \frac{2}{5} \end{bmatrix} + \begin{bmatrix} \frac{4}{5} \\ -\frac{2}{5} \end{bmatrix}$,

b) $\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{3}{10} \\ \frac{9}{10} \end{bmatrix} + \begin{bmatrix} \frac{3}{10} \\ \frac{1}{10} \end{bmatrix}$,

c) $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{5}{6} \\ \frac{2}{6} \\ \frac{1}{6} \end{bmatrix} + \begin{bmatrix} \frac{1}{6} \\ -\frac{2}{6} \\ -\frac{1}{6} \end{bmatrix}$,

d) $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$,

e) $\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$.

15.6. a) $S = span\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} \frac{6}{5} \\ -\frac{3}{5} \end{bmatrix}\right\}$,

b) $S = span\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -\frac{3}{2} \\ 2 \\ \frac{3}{2} \end{bmatrix}\right\}$,

c) $S = span\left\{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}\right\}$,

d) $S = span\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \\ -1 \end{bmatrix}\right\}$,

e) $S = span\left\{\begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} -\frac{4}{5} \\ \frac{2}{5} \\ \frac{8}{5} \\ -\frac{4}{5} \end{bmatrix}\right\}$.

15.7. a) $S = span\left\{\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} \frac{2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{bmatrix}\right\}$,

c) $S = span\left\{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}\right\}$.

15.8. a) La matriz A tiene eigenvalores $1, -4$ y 6 con eigenvectores $\begin{bmatrix} 0 \\ -3 \\ 4 \end{bmatrix}$, $\begin{bmatrix} -5 \\ 4 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}$, respectivamente. La matriz B tiene eigenvalores 0 y 3 con eigenvectores $\left\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right\}$ y $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, respectivamente.

b) Utilizar el producto punto entre los vectores.

c) $\left\{\begin{bmatrix} 0 \\ -\frac{3}{5} \\ \frac{4}{5} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{4}{5\sqrt{2}} \\ \frac{3}{5\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{4}{5\sqrt{2}} \\ \frac{3}{5\sqrt{2}} \end{bmatrix}\right\}$ es una base ortonormal de eigenvectores de A y $\left\{\begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}\right\}$ es una base ortonormal de eigenvectores de B.

$$\text{d) } P_1 = \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{3}{5} & \frac{4}{5\sqrt{2}} & \frac{4}{5\sqrt{2}} \\ \frac{4}{5} & \frac{3}{5\sqrt{2}} & \frac{3}{5\sqrt{2}} \end{bmatrix},\ P_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

15.9.

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix},\ \vec{y} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix},$$

por lo tanto, $(A^T A)^{-1} A^T \vec{y} =$

$$\left(\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{3} \end{bmatrix}.$$

y se tiene que la recta de mínimos cuadrados es $y = \frac{1}{2}x + \frac{1}{3}$.

15.10. b) Para encontrar $\vec{v}$, observar que las columnas de A son ortogonales y si $S = \operatorname{col} A$, entonces $\left\| \vec{b} - A\vec{v} \right\|$ se minimiza cuando

$$A\vec{v} = proy_S \vec{b} = \begin{bmatrix} \frac{1}{2} \\ 2 \\ -\frac{1}{2} \end{bmatrix}$$

y $\vec{v} = \begin{bmatrix} \frac{5}{6} \\ -\frac{1}{3} \end{bmatrix}$.

Índice alfabético