Apuntes del programa completo de Álgebra Lineal

Ricardo Ceballos Sebastián

6 de mayo de 2019

Índice general

Índice de figuras

Índice de tablas

Capítulo 1

Sistemas de Ecuaciones Lineales

La resolución de sistemas de ecuaciones lineales constituye uno de los temas fundamentales del álgebra lineal. Estos sistemas de ecuaciones aparecen de manera natural cuando se modelan diversas situaciones en áreas como las ciencias, las ingenierías e incluso la economía. En esta unidad enfocaremos nuestro esfuerzo en obtener soluciones de estos sistemas de manera sistemática, cuando dichas soluciones existan.

1.1. Sistemas de Ecuaciones Lineales

En esta sección se estudiarán los sistema de ecuaciones lineales. Se comenzará con el análisis de los sistemas más sencillos, con los cuales los estudiantes se encuentran familiarizados, y posteriormente se concluirá con el análisis de un sistema de ecuaciones en general.

1.1.1. Ecuaciones lineales con dos incógnitas

Una recta en el plano xy puede representarse algebraicamente mediante una ecuación de la forma,

$$ax + by = c, \tag{1.1}$$

donde a, b y c son las constantes y; x y y representan a las variables. La ecuación 1.1 se conoce como ecuación lineal en las variales x y y.

Una pareja de números s_1 y s_2 es una solución de la ecuación 1.1, si dicha ecuación se satisface al hacer la sustitución $x = s_1$ y $y = s_2$. Al conjunto de todas las soluciones de la ecuación 1.1 se le conoce como conjunto solución.

Ejemplo 1.1.1 *Encuentre el conjunto solución de la siguiente ecuación lineal,*

$$4x - 3y = 1. \tag{1.2}$$

Solución: El procedimiento consiste en despejar cualquiera de las variables: Si despejamos la variable x obtenemos,

$$x = \frac{3}{4}y + \frac{1}{4},$$

Es posible expresar el conjunto solución en términos de un único parámetro. por ejemplo, si se hace $y = t$, entonces el conjunto solución se expresa como:

$$\begin{cases} x &= \frac{3}{4}t + \frac{1}{4}, \\ y &= t, \end{cases}$$

lo cual significa que para cualquier valor real de t se obtiene una pareja de números reales x y y que forman una solución de la ecuación 1.2. Por ejemplo, si $t = 0$, entonces

$$\begin{cases} x &= \frac{1}{4}, \\ y &= 0, \end{cases}$$

es una solución de la ecuación 1.2.

Despejando la variable y en la ecuación 1.2, y haciendo la parametrización $x = t$, se puede expresar el conjunto solución de una forma aparentemente diferente; sin embargo, puede probarse que cuando t asume todos los valores reales posibles, se obtiene el mismo conjunto solución.

1.1.2. Ecuaciones lineales con tres incógnitas

Un plano en el espacio tridimensional se representa por una ecuación de la forma

$$ax + by + cz = d, \tag{1.3}$$

donde a, b, c y d son las constantes y; x, y y z representan a las variables. La ecuación 1.3 se conoce como ecuación lineal en las variales x, y y z .

Una terna de números s_1, s_2 y s_3 es una solución de la ecuación 1.3, si dicha ecuación se satisface al hacer la sustitución $x = s_1$, $y = s_2$ y $z = s_3$.

1.1.3. Sistema de m ecuaciones lineales con n incógnitas

Definición 1.1.1 *En general se define una ecuación lineal en las* n *variables* $x_1, x_2, ..., x_n$ *como:*

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b, \tag{1.4}$$

donde $a_1, a_2, ...a_n$ y b son constantes reales y; $x_1, x_2, ..., x_n$ representan a las variables.

Ejemplo 1.1.2 *Ejemplos de ecuaciones lineales son:*

$$\begin{aligned} 3x + 2y &= 5, \\ 2x - y + z &= 0, \\ x_1 - x_2 + x_3 - x_4 &= 7. \end{aligned}$$

Se observa que una ecuación lineal no comprende productos o raíces de variables y todas las variables se presentan a la primera potencia y no aparecen como argumento de funciones trigonómetricas, logarítmicas o exponenciales.

Ejemplo 1.1.3 *Ejemplos de ecuaciones no lineales son:*

$$\begin{aligned} x^2 + 2xy &= 5, \\ lnx + \operatorname{sen} y &= 7, \\ \sqrt{x_1} + \frac{x_2}{x_3} &= 20. \end{aligned}$$

Definición 1.1.2 *Una sucesión de números $s_1, s_2, ..., s_n$ es una solución de la ecuación 1.4 si ésta se satisface al hacer la sustitución $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$. .*

Definición 1.1.3 (Sistema de ecuaciones lineales) *Un conjunto finito de ecuaciones lineales en las variables $x_1, x_2, \ldots, x_n$ se conoce como sistema de ecuaciones lineales o sistema lineal.*

Un sistema arbitrario de m ecuaciones lineales en n incógnitas se puede representar de la siguiente manera:

$$\left.\begin{array}{r} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \vdots \qquad\qquad\quad \vdots \qquad\qquad \vdots, \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{array}\right\} . \tag{1.5}$$

El primer subíndice del coeficiente a_{ij} en el sistema representado por la ecuación 1.5 indica la ecuación en la que se encuentra; y el segundo, indica la incógnita que multiplica. Por ejemplo, a_{12} se encuentra en la primera ecuación y multiplica a la incógnita x_2.

Definición 1.1.4 (Solución de un sistema de ecuaciones lineales) *Una sucesión de números $s_1, s_2, ..., s_n$ es una solución de un sistema de ecuaciones como el representado por la ecuación 1.5 si al hacer la sustitución $x_1 = s_1, x_2 = s_2, ..., x_n = s_n$, cada una de las ecuaciones del sistema se satisfacen* [1].

Definición 1.1.5 *Al conjunto de todas las soluciones de un sistema de ecuaciones lineales se le conoce como conjunto solución.*

Definición 1.1.6 (Sistemas consistentes e inconsistentes) *Un sistema de ecuaciones lineales se denomina consistente si tiene al menos una solución. Si el sistema no tiene soluciones se le denomina inconsistente.*

El siguiente sistema lineal constituye un sistema inconsistente,

$$\left.\begin{array}{r} 2x + y = 3 \\ 2x + y = 2 \end{array}\right\} \tag{1.6}$$

Es claro que la afirmación de una de las ecuaciones niega la ecuación restante.

Un sistema general de dos ecuaciones lineales con dos incógnitas puede reprentarse mediante,

$$\left.\begin{array}{r} ax + by = c \\ dx + ey = f \end{array}\right\} \tag{1.7}$$

Las posibilidades que pueden presentarse al resolver el sistema representado por la ecuación 1.7 se muestran de forma gráfica en la figura 1.1.

En la figura 1.1a, las rectas l_1 y l_2 se intersecan en un punto; es decir, el sistema tiene solución única.

En la figura 1.1b, las rectas l_1 y l_2 no se intersecan en ningún punto; el sistema no tiene soluciones.

En la figura 1.1c, las rectas l_1 y l_2 se encuentran superpuestas; en este caso, el sistema tiene infinitas soluciones.

Los resultados anteriores(como demostraremos en las siguientes secciones) son de carácter general: esto significa que todo sistema de ecuaciones lineales tiene solución única, tiene infinitas soluciones o no tiene soluciones.

En los cursos básicos, debachillerato, se enseña a resolver casos explícitos del sistema de euaciones representado por la ecuación 1.7. El resultado general puede hallarse mediante la aplicación de dos propiedades algebraicas elementales:

[1] Bastará con que una ecuación no se satisfaga para que la sucesión no sea una solución del sistema.

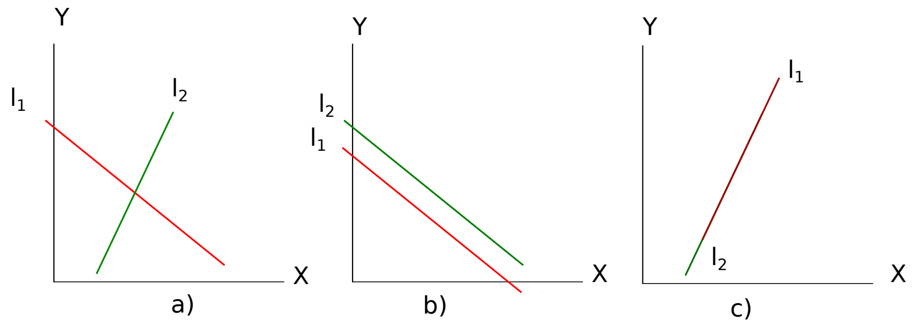

Figura 1.1: Posibles soluciones de un sistema de dos ecuaciones con dos incógnitas

1. Si $a = b$ y $c = d$, entonces $a + c = b + d$.
2. si $c \neq 0$ y $a = b$, entonces $ca = cb$

Ejemplo 1.1.4 *Mediante la aplicación de las propiedades 1 y 2 resuelva el siguiente sistema representado por la ecuación 1.7*

$$\begin{array}{rcl} ax + by &=& c \\ dx + ey &=& f \end{array}$$

Solución:

1. Propiedad 2: Al multipliar la primera euación por $-d$ y la segunda ecuación por a obtenemos,

$$\begin{array}{rcl} -dax - dby &=& -dc, \\ adx + aey &=& af, \end{array}$$

2. Propiedad 1: Al sumar las dos últimas ecuaciones se obtiene,

$$(-dax - dby) + (adx + aey) = -dc + af,$$

de donde se sigue,

$$(-da + ad)x + (ae - db)y = af - dc,$$

$$0x + (ae - db)y = af - dc.$$

Aquí se debe ser muy cuidadoso, ya que para obtener el valor de y es necerario dividir entre, $ae - db$, lo que solo es posible si, $(ae - db) \neq 0$, en este caso,

$$y = \frac{af - dc}{ae - db}. \tag{1.8}$$

Finalmente, se despeja x en la primera ecuación del sistema, posteriormente se sustituye el valor de y, con lo cual se obtiene después de reducciones algebraicas,

$$x = \frac{ce - bf}{ae - db}. \tag{1.9}$$

Se concluye que si, $(ae - db) \neq 0$, entonces la solución general del sistema de dos ecuaciones con dos variables está determinado por

$$\begin{cases} x = \dfrac{ce - bf}{ae - db} \\ y = \dfrac{af - dc}{ae - db}. \end{cases} \tag{1.10}$$

Matriz aumentada y matriz de coeficientes

El arreglo de números que se obtiene al eliminar los signos $+$, las variables x y los signos $=$, en el sistema de ecuaciones lineales representado por la ecuación 1.5, se conoce como matriz aumentada, como se muestra a continuación,

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right]. \tag{1.11}$$

Como es estudiará en la sección 1.2, una matriz es un arreglo rectangular de números que consta de renglones(líneas horizontales) y de columnas(líneas verticales). La última columna se conoce como la columna de los términos constantes[2].

Si en la matriz anterior se elimina la columna de las constantes, obten-

[2]La barra vertical entre las dos últimas columnas de la matriz aumentada, se utiliza para para remarcar el hecho de que la última columna corresponde al de los términos constantes.

dremos lo que se conoce como matriz de coeficientes,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}. \tag{1.12}$$

Se observa que la matriz de coeficientes solo contiene los coeficientes de las variables de la ecuación 1.5 [3].

Ejemplo 1.1.5 *Escriba la matriz aumentada y la matriz de coeficientes del siguientes sistema de ecuaciones,*

$$\left.\begin{array}{r} 2x_1 + 3x_2 - x_3 = 0 \\ -x_1 + 2x_2 - 3x_3 = -3 \\ 3x_1 + 5x_2 - 7x_3 = 5 \end{array}\right\}.$$

Solución: La matriz aumentada es

$$\left[\begin{array}{rrr|r} 2 & 3 & -1 & 0 \\ -1 & 2 & -3 & -3 \\ 3 & 5 & -7 & 5 \end{array}\right].$$

La matriz de coeficientes es

$$\begin{bmatrix} 2 & 3 & -1 \\ -1 & 2 & -3 \\ 3 & 5 & -7 \end{bmatrix}.$$

1.1.4. Eliminación de Gauss y de Gauss-Jordan con pivoteo

El método básico para resolver sistemas de ecuaciones lineales consiste en reemplazar el sistema dado por uno nuevo que tenga el mismo conjunto solución, pero que sea más fácil de resolver. En general, este nuevo sistema se obtiene realizando una serie de pasos, en cada uno de los cuales se aplica una de las siguientes operaciones permitidas, conocidas como operaciones elementales, a fin de eliminar sistemáticamente las incógnitas.

Operaciones elementales sobre las ecuaciones

[3] Las variables en el sistema de ecuaciones siempre deberán aparecer en el mismo orden en todas las ecuaciones, ya que de este modo evitaremos cometer errores al escribir la matriz aumentada y la matriz de coeficientes del sistema.

1. Multiplicar una de las ecuaciones por una constante diferente de cero.
2. Intercambiar dos ecuaciones.
3. Sumar un múltiplo de una ecuación a otra ecuación.

Como se puede apreciar, se trata de las operaciones que se emplearon en el ejemplo 1.1.4.

Como los renglones de la matriz aumentada corresponden a las ecuaciones del sistema, entonces estas tres operaciones sobre las ecuaciones del sistema tienen sus equivalentes sobre los renglones de la matriz aumentada.

Operaciones elementales sobre los renglones

1. Multiplicar uno de los renglones por una constante diferente de cero.
2. Intercambiar dos renglones.
3. Sumar un múltiplo de uno de los renglones a otro renglón.

Definición 1.1.7 (Sistemas equivalentes $\sim$) *Dos sistemas de ecuaciones como los representados por las ecuaciones 1.13 y 1.14 se conocen como equivalentes, si tienen el mismo conjunto solución.*

$$\left.\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array}\right\} \tag{1.13}$$

$$\left.\begin{array}{c} c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n = d_1 \\ c_{21}x_1 + c_{22}x_2 + \ldots + c_{2n}x_n = d_2 \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ c_{m1}x_1 + c_{m2}x_2 + \ldots + c_{mn}x_n = d_m \end{array}\right\} \tag{1.14}$$

Teorema 1.1.1 *Si un sistema de ecuaciones lineales se obtiene de otro por medio de operaciones elementales, entonces los sistemas son equivalentes.*

Demostración: La demostración que se presenta, supone que se aplica una sola operación elemental, ya que el caso general se obtiene de ésta aplicando el mismo argumento sucesivamente.

Consideremos el sistema general de m ecuaciones con n incógnitas.

$$\left.\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n = b_j \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array}\right\} \tag{1.15}$$

Si el nuevo sistema se obtiene aplicando las operaciones elementales (1) y (2), entonces en esencia se sigue teniendo el mismo sistema de ecuaciones y por lo tanto, las soluciones del sistema no se afectan. Sin embargo, si el nuevo sistema se obtiene aplicando la operación elemental(3), entonces debemos asegurarnos que el conjunto solución no se modifica.

Supóngase que el nuevo sistema se obtiene sumando a veces la ecuación (i) a la ecuación (j), entonces los sistemas diferirán en el renglón (j), ya que para el nuevo sistema la ecuación(j) estará determinada por,

$$(aa_{i1} + a_{j1})x_1 + (aa_{i2} + a_{j2})x_2 + \ldots + (aa_{in} + a_{jn})x_n = ab_i + b_j, \tag{1.16}$$

de manera que el nuevo sistema se representa por

$$\left.\begin{array}{lllllllll} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \\ a_{i1}x_1 & + & a_{i2}x_2 & + & \ldots & + & a_{in}x_n & = & b_i \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \\ \multicolumn{7}{l}{(aa_{i1} + a_{j1})x_1 + (aa_{i2} + a_{j2})x_2 + \ldots + (aa_{in} + a_{jn})x_n} & = & ab_i + b_j \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array}\right\}. \tag{1.17}$$

Sean S y S' los conjuntos solución correspondientes a los sistemas de ecuaciones representados por las ecuaciones 1.15 y 1.17, respectivamente, se demostrará que $S = S'$.
$S \subseteq S'$) Si $s_1, s_2, ..., s_n$ es una solución del sistema representado por la ecuación 1.15, entonces en particular se satisfacen las ecuaciones (i) y (j); es decir,

$$a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n = b_i, \tag{1.18}$$

$$a_{j1}s_1 + a_{j2}s_2 + \ldots + a_{jn}s_n = b_j. \tag{1.19}$$

Si se multiplica la ecuación 1.18 por a y se suma a la ecuación 1.19 se obtiene,

$$(aa_{i1} + a_{j1})s_1 + (aa_{i2} + a_{j2})s_2 + \ldots + (aa_{in} + a_{jn})s_n = ab_i + b_j. \quad (1.20)$$

La ecuación 1.20 implica que $s_1, s_2, ..., s_n$ es una solución de la ecuación 1.16, y, por lo tanto, del nuevo sistema representado por la ecuación 1.15; es decir, $S \subseteq S'$.

$S' \subseteq S$) Recíprocamente, si $s_1, s_2, ..., s_n$ es una solución del sitema representado por la ecuación 1.17, entonces se cumple que,

$$a_{k1}s_1 + a_{k2}s_2 + \ldots + a_{kn}s_n = b_k, \quad \forall k \neq j \quad (1.21)$$

además,

$$(aa_{i1} + a_{j1})s_1 + (aa_{i2} + a_{j2})s_2 + \ldots + (aa_{in} + a_{jn})s_n = ab_i + b_j. \quad (1.22)$$

La ecuación 1.21 se cumple de manera particular para $k = i$. En este caso se tiene,

$$a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n = b_i. \quad (1.23)$$

Si se resta a veces la ecuación 1.23 a la ecuación 1.22, entonces se obtiene,

$$a_{j1}s_1 + a_{j2}s_2 + \ldots + a_{jn}s_n = b_j, \quad (1.24)$$

es decir, $s_1, s_2, ..., s_n$ es solución del sistema original representado por la ecuación 1.15, por lo tanto, $S' \subseteq S$), con lo cual concluimos finalmente que $S = S'$. ∎

Definición 1.1.8 (Matriz escalonada en los renglones reducida) *Se dice que una matriz se encuentra en la forma escalonada en los renglones reducida si tiene las siguientes propiedades:*

1. *Si un renglón no consta completamente de ceros, entonces el primer número diferente de cero en el renglón es un uno. A éste se le denomina uno principal(1-principal)*[4].

2. *Si existen renglones que consten completamente de ceros, éstos se agrupan en la parte inferior de la matriz.*

3. *Si dos renglones sucesivos no constan completamente de ceros, entonces el 1-principal del renglón inferior se encuentra más a la derecha que el 1-principal del renglón superior.*

[4] Algunos autores prefieren usar el término entrada principal en lugar de uno principal.

4. *Cada columna que contenga un uno principal tiene ceros en las demás posiciones.*

Definición 1.1.9 (Matriz escalonada en los renglones) *Una matriz que solo cumple con las propiedades 1,2 y 3 se conoce como matriz escalonada en los renglones.*

Ejemplo 1.1.6 *Las siguientes matrices son ejemplos de matrices escalonadas en los renglones reducidas.*

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & 5 \end{array}\right], \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & \pi & 0 & 1 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Ejemplo 1.1.7 *Las siguientes matrices son ejemplos de matrices escalonadas en los renglones.*

$$\left[\begin{array}{ccr|c} 1 & 4 & 3 & 7 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & a \end{array}\right], \qquad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & b & 3 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

En las matrices anteriores a y b representan números reales.

Cuando la matriz aumentada de un sistema de ecuaciones es llevada a la forma escalonada en los renglones reducida, entonces es posible obtener la solución del sistema por simple inspección. Este método se conoce como **eliminación de Gauss-Jordan**. Si la matriz aumentada es llevada a una forma escalonada en los renglones y se aplica sustitución hacia atrás para obtener las soluciones del sistema, entonces estamos aplicando el método de **eliminación de Gauss**.

Notación: Las operaciones elementales realizadas para obtener una matriz equivalente las denotaremos de la siguiente manera:

1. $\underset{ar_i \longrightarrow r_i}{\sim}$ Esta operación indica que se multitlicó el i-ésimo renglón por una constante $a \neq 0$ para generar una matriz equivalente en los renglones. El resultado se aplica al renglón i-ésimo de la nueva matriz.

2. $\underset{r_i \longleftrightarrow r_j}{\sim}$ Esta operación indica que se intercambiaron los renglones r_i y r_j. Ésta es la única operación elemental que al aplicarse de manera aislada modifica dos renglones.

3. $\underset{ar_i+r_j\longrightarrow r_j}{\sim}$ Esta operación indica que se sumó a-veces el renglón i al renglón j. El resultado se aplica al renglón j-ésimo de la nueva matriz.

Antes demostrar el teorema de Gauss-Jordan se muestran algunos ejemplos del uso de las operaciones elementales sobre los renglones de la matriz aumentada, para llevarla a una forma escalonada en los renglones reducida, o simplemente, a su forma escalonada en los renglones.

Ejemplo 1.1.8 *Resuelva el siguiente sistema de ecuaciones mediante eliminación de Gauss-Jordan.*

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

Solución: Considérese la matriz aumentada del sistema,

$$\left[\begin{array}{rrr|r} 2 & -1 & 4 & -3 \\ 1 & 2 & -3 & 1 \\ 5 & 3 & 1 & -2 \end{array}\right].$$

Esta matriz será llevada a su forma escalonada en los renglones reducida, ya que de este modo se cumple con el requisito de resolver el sistema mediante el método de Gauss-Jordan.

1. Se comienza por colocar el primer 1-principal. Esto puede lograrse multiplicando el primer renglón por (1/2) o simplemente intercambiando los renglones (2) y (3). Ente caso optaremos por la segunda opción, ya que con esto evitaremos introducir fracciones desde el comienzo.

$$\underset{r_1\longleftrightarrow r_2}{\sim}\left[\begin{array}{rrr|r} 1 & 2 & -3 & 1 \\ 2 & -1 & 4 & -3 \\ 5 & 3 & 1 & -2 \end{array}\right].$$

   Una vez colocado el primer 1-principal, se usará éste para realizar operaciones que permitan hacer ceros a los elementos restantes de su columna. Este proceso se conoce como pivoteo. Para comenzar súmese (-2) veces el renglón 1 al renglón 2,

$$\underset{-2r_1+r_2\longrightarrow r_2}{\sim}\left[\begin{array}{rrr|r} 1 & 2 & -3 & 1 \\ 0 & -5 & 10 & -5 \\ 5 & 3 & 1 & -2 \end{array}\right].$$

Nuevamente se usa el renglón 1 como pivote. Súmese (-5) veces el renglón 1 al renglón 3 ,

$$\underset{-5r_1+r_3 \longrightarrow r_3}{\sim} \left[\begin{array}{rrr|r} 1 & 2 & -3 & 1 \\ 0 & -5 & 10 & -5 \\ 0 & -7 & 16 & -7 \end{array}\right].$$

2. Para continuar se coloca un 1-principal en el siguiente renglón. En este caso se multiplica el renglón 2 por (-1/5).

$$\underset{-1/5r_2 \longrightarrow r_2}{\sim} \left[\begin{array}{rrr|r} 1 & 2 & -3 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & -7 & 16 & -7 \end{array}\right].$$

Ahora se usa el renglón 2 como pivote para hacer ceros los elementos restantes de la columna del segundo 1-principal.

Súmese (-2) veces el renglón 2 al renglón 1 ,

$$\underset{-2r_2+r_1 \longrightarrow r_1}{\sim} \left[\begin{array}{rrr|r} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 1 \\ 0 & -7 & 16 & -7 \end{array}\right].$$

Súmese (7) veces el renglón 2 al renglón 3 ,

$$\underset{7r_2+r_3 \longrightarrow r_3}{\sim} \left[\begin{array}{rrr|r} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 2 & 0 \end{array}\right].$$

3. En seguida cse coloca un 1-principal en el siguiente renglón. En este caso se multiplica el renglón 3 por (1/2).

$$\underset{1/2r_3 \longrightarrow r_3}{\sim} \left[\begin{array}{rrr|r} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

Finalmente, utilícese el renglón 3 como pivote para hacer ceros los elementos restantes de la columna del tercer 1-principal.

Reste el renglón 3 al renglón 1.

$$\underset{-r_3+r_1 \longrightarrow r_1}{\sim} \left[\begin{array}{rrr|r} 1 & 0 & 0 & -1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

Súmese (2) veces el renglón 3 al renglón 2 ,

$$\underset{2r_3+r_2\longrightarrow r_2}{\sim}\left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

La matriz anterior se encuentra en su forma escalonada en los renglones reducida. El sistema de ecuaciones que corresponde a la matriz anterior es:

$$\begin{cases} x_1 = -1, \\ x_2 = 1, \\ x_3 = 0. \end{cases}$$

Lo cual constituye la solución del sistema.

Ejemplo 1.1.9 *Resuelva el siguiente sistema de ecuaciones usando eliminación gaussiana.*

$$\left.\begin{array}{r} 2x - y + z = 5 \\ x + y + z = 2 \\ 2x - 3y - z = 3 \end{array}\right\}.$$

Solución: Se llevará la matriz aumentada del sistema a su forma escalonada en los renglones, mediante operaciones elementales sobre éstos. De esta manera se garantiza que se está empleando el método de eliminación de Gauss.

La matriz aumentada del sistema es

$$\left[\begin{array}{rrr|r} 2 & -1 & 1 & 5 \\ 1 & 1 & 1 & 2 \\ 2 & -3 & -1 & 3 \end{array}\right].$$

1. En primer lugar se coloca un 1-principal en el primer renglón, lo cual se consegue intercambiando los renglones 1 y 2.

$$\underset{r_1\longleftrightarrow r_2}{\sim}\left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 2 & -1 & 1 & 5 \\ 2 & -3 & -1 & 3 \end{array}\right].$$

Úsese este 1-principal como pivote para hacer cero todos los elementos inferiores de su columna.

Súmese (-2) el renglón 1 al renglón 2.

$$\underset{-2r_1+r_2\longrightarrow r_2}{\sim}\left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 0 & -3 & -1 & 1 \\ 2 & -3 & -1 & 3 \end{array}\right].$$

Súmese (-2) el renglón 1 al renglón 3.

$$\underset{-2r_1+r_2\longrightarrow r_3}{\sim}\left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 0 & -3 & -1 & 1 \\ 0 & -5 & -3 & -1 \end{array}\right].$$

2. Para continuar, se coloca un segundo 1-principal en el segundo renglón. Multiplíquese el segundo renglón por (-1/3,)

$$\underset{-1/3r_2\longrightarrow r_2}{\sim}\left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & -5 & -3 & -1 \end{array}\right].$$

De manera similar, se usará este 1-principal como pivote para hacer cero todos los elementos inferiores de su columna. En este caso debemos sumar 5 veces el renglón 2 al renglón 3.

$$\underset{5r_2+r_3\longrightarrow r_3}{\sim}\left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & -\frac{4}{3} & -\frac{8}{3} \end{array}\right].$$

3. Finalmente, multiplíquese el renglón 3 por (-3/4)

$$\underset{-3/4r_3\longrightarrow r_3}{\sim}\left[\begin{array}{rrr|r} 1 & 1 & 1 & 2 \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 1 & 2 \end{array}\right].$$

La matriz anterior es una matriz escalonada. El sistema de ecuaciones correspondiente es:

$$\left.\begin{array}{r} x+y+z=2 \\ y+\frac{1}{3}z=-\frac{1}{3} \\ z=2 \end{array}\right\}.$$

Despejando las variables y y x, y realizando sustitución hacia atrás se obtiene la solución.

$$\left\{\begin{array}{l} x=1 \\ y=-1 \\ z=2. \end{array}\right.$$

Definición 1.1.10 (Variables principales y variables libres) *Si una matriz aumentada de un sistema de ecuaciones se lleva a su forma escalonada en los renglones reducida, entonces se conocen como variables principales a las variables asociadas con los 1-principales en dicha matriz. El resto de las variables se conocen como variables libres.*

Ejemplo 1.1.10 *Resuelva el siguiente sistema de ecuaciones mediante eliminación de Gauss-Jordan.*

$$\left.\begin{array}{rrrrrcr} x_1 & & +4x_3 & +3x_4 & -x_5 & = & 1 \\ -5x_1 & +2x_2 & -2x_3 & -x_4 & -x_5 & = & -2 \\ -3x_1 & +2x_2 & +2x_3 & +2x_4 & -2x_5 & = & -1 \\ -x_1 & +2x_2 & +14x_3 & +11x_4 & -5x_5 & = & 2 \end{array}\right\}$$

Solución: La matriz aumentada del sistema es

$$\left[\begin{array}{rrrrr|r} 1 & 0 & 4 & 3 & -1 & 1 \\ -5 & 2 & -2 & -1 & -1 & -2 \\ -3 & 2 & 2 & 2 & -2 & -1 \\ -1 & 2 & 14 & 11 & -5 & 2 \end{array}\right].$$

Para reducir la matriz aumentada a su forma escalonada en los renglones reducida realizaremos los siguientes pasos:

1. Súmese (5) veces el renglón(1) al renglón(2).

$$\underset{5r_1+r_2\longrightarrow r_2}{\sim}\left[\begin{array}{rrrrr|r} 1 & 0 & 4 & 3 & -1 & 1 \\ 0 & 2 & 18 & 14 & -6 & 3 \\ -3 & 2 & 2 & 2 & -2 & -1 \\ -1 & 2 & 14 & 11 & -5 & 2 \end{array}\right].$$

Súmese (3) veces el renglón(1) al renglón(4).

$$\underset{3r_1+r_3\longrightarrow r_3}{\sim}\left[\begin{array}{rrrrr|r} 1 & 0 & 4 & 3 & -1 & 1 \\ 0 & 2 & 18 & 14 & -6 & 3 \\ 0 & 2 & 14 & 11 & -5 & 2 \\ -1 & 2 & 14 & 11 & -5 & 2 \end{array}\right].$$

Súmese el renglón(1) al renglón(4).

$$\underset{r_1+r_4\longrightarrow r_4}{\sim}\left[\begin{array}{rrrrr|r} 1 & 0 & 4 & 3 & -1 & 1 \\ 0 & 2 & 18 & 14 & -6 & 3 \\ 0 & 2 & 14 & 11 & -5 & 2 \\ 0 & 2 & 18 & 14 & -6 & 3 \end{array}\right].$$

2. Multiplíquese el renglón(2) por (1/2).

$$\underset{1/2r_2 \longrightarrow r_2}{\sim} \left[\begin{array}{ccccc|c} 1 & 0 & 4 & 3 & -1 & 1 \\ 0 & 1 & 9 & 7 & -3 & \frac{3}{2} \\ 0 & 2 & 14 & 11 & -5 & 2 \\ 0 & 2 & 18 & 14 & -6 & 3 \end{array}\right].$$

Súmese (-2) veces el renglón(2) al renglón(3).

$$\underset{-2r_2+r_3 \longrightarrow r_3}{\sim} \left[\begin{array}{ccccc|c} 1 & 0 & 4 & 3 & -1 & 1 \\ 0 & 1 & 9 & 7 & -3 & \frac{3}{2} \\ 0 & 0 & -4 & -3 & 1 & -1 \\ 0 & 2 & 18 & 14 & -6 & 3 \end{array}\right].$$

Súmese (-2) veces el renglón(2) al renglón(4).

$$\underset{-2r_2+r_4 \longrightarrow r_4}{\sim} \left[\begin{array}{ccccc|c} 1 & 0 & 4 & 3 & -1 & 1 \\ 0 & 1 & 9 & 7 & -3 & \frac{3}{2} \\ 0 & 0 & -4 & -3 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

3. Multiplíquese el renglón(3) por $(-1/4)$.

$$\underset{-1/4r_3 \longrightarrow r_3}{\sim} \left[\begin{array}{ccccc|c} 1 & 0 & 4 & 3 & -1 & 1 \\ 0 & 1 & 9 & 7 & -3 & \frac{3}{2} \\ 0 & 0 & 1 & \frac{3}{4} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Súmese (-4) veces el renglón(3) al renglón(1).

$$\underset{-4r_3+r_1 \longrightarrow r_1}{\sim} \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 9 & 7 & -3 & \frac{3}{2} \\ 0 & 0 & 1 & \frac{3}{4} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Súmese (-9) veces el renglón(3) al renglón(2).

$$\underset{-9r_3+r_2 \longrightarrow r_2}{\sim} \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{3}{4} & -\frac{3}{4} \\ 0 & 0 & 1 & \frac{3}{4} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

La matriz anterior se encuentra en su forma escalonada en los renglones reducida. Podemos observar que los 1-principales corresponden a las variables x_1, x_2 y x_3, de este modo, las variables restantes corresponden a las variables libres; es decir, x_4 y x_5 son las variables libres del sistema.

El sistema de ecuaciones que corresponde a la matriz anterior es:

$$\begin{aligned} x_1 &= 0, \\ x_2 + \frac{1}{4}x_4 - \frac{3}{4}x_5 &= -\frac{3}{4}, \\ x_3 + \frac{3}{4}x_4 - \frac{1}{4}x_5 &= \frac{1}{4}. \end{aligned}$$

Despejando la variables principales del sistema se obtiene,

$$\begin{cases} x_1 = 0, \\ x_2 = -\frac{1}{4}x_4 + \frac{3}{4}x_5 - \frac{3}{4}, \\ x_3 = -\frac{3}{4}x_4 + \frac{1}{4}x_5 + \frac{1}{4}. \end{cases}$$

Si se hace la parametrización de las variables libres mediante, $x_4 = r$ y $x_5 = s$, entonces las soluciones del sistema en términos de estos parámetros son:

$$\begin{cases} x_1 = 0, \\ x_2 = -\frac{1}{4}r + \frac{3}{4}s - \frac{3}{4}, \\ x_3 = -\frac{3}{4}r + \frac{1}{4}s + \frac{1}{4}, \\ x_4 = r, \\ x_5 = s. \end{cases}$$

Como se puede apreciar, con cada pareja de números reales r y s, es posible asociar una solución particular para el sistema de ecuaciones. Por ejemplo, si $r = s = 0$, entonces

$$\begin{cases} x_1 &= 0, \\ x_2 &= -3/4, \\ x_3 &= 1/4, \\ x_4 &= 0, \\ x_5 &= 0. \end{cases}$$

es una solución particular del sistema.

Teorema 1.1.2 (Método de eliminación de Gauss-Jordan) *Toda matriz puede ser llevada a su forma escalonada en los renglones reducida, mediante operaciones elementales sobre sus renglones.*

Demostración: Si la matriz consta completamente de ceros, entonces no habría nada que probar, pues la matriz ya se encontraría en la forma deseada. Supóngase entonces que se parte de una matriz A que no es idénticamente cero. Supóngase que j es el menor entero para el cual $a_{ij} \neq 0$ para algún i y que el primer elemento no cero de la columna j se en cuentra en el primer renglón(en caso contrario se procedería con un intercambio de renglones para cumplir esta condición), entonces la matriz tiene la siguiente forma,

$$A = \begin{bmatrix} 0 & \dots & 0 & a_{1j} & \dots & a_{1n} \\ 0 & \dots & 0 & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & a_{mj} & \dots & a_{mn} \end{bmatrix}, \tag{1.25}$$

para obtener el primer 1-principal se realiza la operación elemental,

$$(a_{1j})^{-1} r_1 \longrightarrow r_1,$$

posteriormente, para obtener ceros debajo del 1-principal se procede con la secuencia de operaciones

$$-a_{ij} r_1 + r_i \longrightarrow r_i, \qquad \text{para todo } i \neq 1.$$

con esto se obtiene la matriz equivalente,

$$A_1 = \begin{bmatrix} 0 & \dots & 0 & 1 & b_{1(j+1)} & \dots & b_{1n} \\ 0 & \dots & 0 & 0 & b_{2(j+1)} & \dots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & b_{m(j+1)} & \dots & b_{mn} \end{bmatrix}. \tag{1.26}$$

Ahora considérese lo submatriz B,

$$B = \begin{bmatrix} b_{2(j+1)} & \dots & b_{2n} \\ \vdots & \vdots & \vdots \\ b_{m(j+1)} & \dots & b_{mn} \end{bmatrix}. \tag{1.27}$$

Si la submatriz representada por la ecuación 1.27 consta completamente de ceros, entonces se habrá terminado. En caso contrario, existirá un k tal que $b_{ki} \neq 0$ para algún $i > 1$. Sea k el menor entero tal que $b_{ki} \neq 0$ para algún

i, además supondremos que $b_{k2} \neq 0$ (en caso contrario se procedería con un intercambio de renglones de tal manera que se satisfaga la condición).

Por lo expuesto anteriormente, la matriz A_1 tendrá la forma,

$$A_1 = \begin{bmatrix} 0 & \ldots & 0 & 1 & b_{1(j+1)} & \ldots & b_{1(k-1)} & b_{1k} & \ldots & b_{1n} \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & b_{2k} & \ldots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & b_{mk} & \ldots & b_{mn} \end{bmatrix}. \tag{1.28}$$

Nuevamente, para obtener el segundo 1-principal se realiza la operación elemental,

$$(b_{2k})^{-1} r_2 \longrightarrow r_2,$$

posteriormente, para obtener ceros debajo de este 1-principal se procede con la secuencia de operaciones,

$$-b_{ik} r_2 + r_i \longrightarrow r_i, \qquad \text{para todo } i \neq 2.$$

con esto obtenemos la matriz equivalente,

$$A_2 = \begin{bmatrix} 0 & \ldots & 1 & b_{1(j+1)} & \ldots & b_{1(k-1)} & 0 & c_{1(k+1)} & \ldots & c_{1n} \\ 0 & \ldots & 0 & 0 & \ldots & 0 & 1 & c_{2(k+1)} & \ldots & c_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & c_{m(k+1)} & \ldots & c_{mn} \end{bmatrix}. \tag{1.29}$$

Se observa que las primeras columnas de A_2 están en la forma escalonada en los renglones reducida. Continuando el proceso se obtendrá una matriz escalonada en los renglones reducida mediante un número finito de pasos. ∎

Soluciones de un sistema de ecuaciones

El análisis de las posibles soluciones de un sistema de ecuaciones puede realizarse de una manera muy sencilla mediante la forma escalonada en los renglones reducida de la matriz aumentada del sistema. De manera más precisa, sea A la matriz aumentada de un sistema de m ecuaciones en n incógnitas, como se muestra a continuación:

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

además, sea R la forma escalonada de la matriz A, entonces existen las condiciones necesarias para comenzar el análisis de las posibles soluciones del sistema.

Sistema inconsistente

Si el último renglón diferente de cero en la matriz R tiene su 1-principal en la columna de los términos independientes; es decir, en la columna $n+1$, entonces el sistema es inconsistente, ya que este renglón representa la ecuación,

$$0x_1 + 0x_2 + \ldots + 0x_n = 1,$$

lo cual representa una condición imposible de satisfacer para cualesquiera valores de $x_1, x_2, \ldots, x_n$.

Sistema consistente: infinitas soluciones

Si el último renglón distinto de cero en la matriz R tiene su 1-principal en la columna $j < n+1$, entonces el sistema es consistente. De manera más precisa, si R tiene l renglones diferentes de cero y sean $k_1, k_2, ...k_l$ las columnas que contienen los 1-principales. Sean además, $x_{k1}, x_{k2}, ..., x_{kl}$ las variables principales y $u_1, u_2, ..., u_{n-l}$ las variables libres del sistema representado por R, entonces los renglones de R representan las ecuaciones,

$$\begin{aligned} x_{k_1} + \sum_{j=1}^{n-l} c_{1j} u_j &= d_1 \\ x_{k_2} + \sum_{j=1}^{n-l} c_{2j} u_j &= d_2 \\ \vdots \; + \qquad \vdots &= \vdots \\ x_{k_l} + \sum_{j=1}^{n-l} c_{lj} u_j &= d_l \end{aligned}$$

Podemos despejar las variables principales en términos de las variables libres para obtener las soluciones, así que debido a la presencia de variables libres el sistema tiene infinitas soluciones.

Sistema consistente: solución única

Del análisis del caso anterior, se concluye que: el sistema tendrá solución única si no existen variables libres en el sistema representado por R. Esto significa que cada renglón solo contiene un elemento distinto de cero antes de la columna $n+1$. Como hay n variables, entonces debe haber n renglones diferentes de cero en R, cada uno con un 1-principal en las posiciones R_{ii}.

Ejemplo 1.1.11 *Determine para qué valores del parámetro a el siguiente sistema de ecuaciones,*

a) Tiene solo una solución.

b) Tiene infinitas soluciones.

c) No tiene soluciones.

$$\left.\begin{array}{rcl} x+2y-3z & = & 4 \\ 3x-y+5z & = & 2 \\ 4x+y+(a^2-14)z & = & a+2 \end{array}\right\}.$$

Solución: El procedimiento consiste en llevar la matriz aumentada del sistema a su forma escalonada en los renglones reducida, y, a partir de ésta, realizar el análisis de las soluciones.

$$\left[\begin{array}{rrr|r} 1 & 2 & -3 & 4 \\ 3 & -1 & 5 & 2 \\ 4 & 1 & (a^2-14) & a+2 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 2 & -3 & 4 \\ 0 & -7 & 14 & -10 \\ 0 & -7 & (a^2-2) & a-14 \end{array}\right] \sim$$

$$\left[\begin{array}{rrr|r} 1 & 2 & -3 & 4 \\ 0 & 1 & -2 & 10/7 \\ 0 & -7 & (a^2-2) & a-14 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 0 & 1 & 8/7 \\ 0 & 1 & -2 & 10/7 \\ 0 & 0 & (a^2-16) & a-4 \end{array}\right].$$

En este punto es muy importante hacer notar que solo será posible colocar un 1-principal en el tercer renglón, si $a^2-16 \neq 0$, es decir, si $a \neq \pm 4$. En este caso se puede seguir reduciendo la matriz hasta obtener,

$$\sim \left[\begin{array}{rrr|r} 1 & 0 & 1 & 8/7 \\ 0 & 1 & -2 & 10/7 \\ 0 & 0 & 1 & (a+4)^{-1} \end{array}\right] \sim \left[\begin{array}{rrr|c} 1 & 0 & 0 & \frac{8a+25}{7(a+4)} \\ 0 & 1 & 0 & \frac{10a+54}{7(a+4)} \\ 0 & 0 & 1 & \frac{1}{(a+4)} \end{array}\right].$$

Se concluye que el sistema tiene solución única cuando $a \neq \pm 4$.

Para finalizar se analizan los casos en que $a = \pm 4$. Esto se logra sustituyendo cada valor en la matriz aumentada del sistema y posteriormente reduciendo la matriz. Sin embargo, es posible y conveniente retomar la matriz obtenida hasta antes de la posible indeterminación.

Si $a = 4$, entonces

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 8/7 \\ 0 & 1 & -2 & 10/7 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

La matriz escalonada en los renglones reducida tiene dos 1-principales que se asocian con variables principales. La tercera variable será libre. La existencia de esta variable libre implica que para $a = 4$ el sistema tiene infinitas soluciones.

Si $a = -4$, entonces

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 8/7 \\ 0 & 1 & -2 & 10/7 \\ 0 & 0 & 0 & -8 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 8/7 \\ 0 & 1 & -2 & 10/7 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

El sistema es inconsistente, ya que el último renglón distinto de cero, en la matriz escalonada en los renglones reducida, tiene su 1-principal en la columna de las constantes.

Ejemplo 1.1.12 *Suponga que la siguiente matriz corresponde a la matriz aumentada de un sistema de ecuaciones. Determine para qué valores de λ el sistema,*

i) tiene solución única,

ii) no tiene solución,

iii) tiene infinitas soluciones.

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

Solución: Se reduce la matriz dada mediante operaciones elementales para obtener conclusiones sobre sus soluciones.

$$\left[\begin{array}{ccc|c} 1 & 0 & \lambda & 1 \\ 0 & \lambda & \lambda & 0 \\ 0 & 2 & 1-\lambda^2 & -1-\lambda \end{array}\right]$$

En este punto, solo es posible seguir reduciendo la matriz, si $\lambda \neq 0$. En este caso obtenemos,

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

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

Nuevamente se requiere tener cuidado, ya que la matriz podrá seguir reduciéndose si y solo si, $\lambda^2+1 \neq 0$, pero ésta es una condición que se cumple para todos los números reales.

$$\sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1+\lambda}{\lambda^2+1} \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1-\lambda}{\lambda^2+1} \\ 0 & 1 & 0 & -\frac{1+\lambda}{\lambda^2+1} \\ 0 & 0 & 1 & \frac{1+\lambda}{\lambda^2+1} \end{array}\right]$$

Por lo anterior, el sistema tendrá solución única dentro de los números reales si $\lambda \neq 0$.

Considérese ahora el caso en que $\lambda = 0$. Sustituyendo este valor en la matriz aumentada obtenemos,

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

Esta matriz se encuentra en la forma escalonada en los renglones reducida. Se aprecia que el sistema admite una variable libre; por lo tanto, tiene infinitas soluciones. Concluimos que el sistema siempre es consistente dentro de los reales.

1.1.5. Sistemas homogéneos

Definición 1.1.11 (Sistemas homogéneos) *Un sistema de ecuaciones homogéneo es aquél en el cual todas sus ecuaciones tienen términos independientes iguales a cero.*

La matriz aumentada de un sistema homogéneo con n incógnitas, tiene únicamente ceros en la columna $n+1$, como se muestra a continuación,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 \\ a_{m1} & a_{m2} & \dots & a_{mn} & 0 \end{bmatrix} \tag{1.30}$$

Los sistemas homogéneos tienen la particularidad de ser sistemas consistentes, ya que, $x_1 = 0, x_2 = 0, \dots x_n = 0$, es una solución del sistema. A esta solución se le conoce como **la solución trivial**.

Teorema 1.1.3 *Dado un sistema homogéneo de m ecuaciones en n incógnitas se cumple lo siguiene: Si $m < n$; es decir, si se tienen más incógnitas que ecuaciones, entonces el sistema tiene infinitas soluciones.*

Demostración: Un sistema de ecuaciones en n variables tiene una matriz aumentada A cuyas primeras n columnas son diferentes de cero, ya que una columna de ceros implicaría que la variable asociada con dicha columna sería irrelevante para el sistema. Ahora, sea R la matriz en la forma escalonada en los renglones reducida equivalente con A, entonces cada una de las primeras n columnas de R contiene al menos un elemento diferente de cero, ya que una columna diferente de cero de la matriz A no puede reducirse a una columna cero mediante operaciones elementales. Además, R tendrá a lo más m renglones distintos de cero. Supóngase que $m < n$,es decir, el número de renglones no ceros en R es menor que el número de variables n. En este punto se tiene un problema equivalente al *problema de las casillas*. En el problema de las casillas, se requiere repartir n objetos(en este caso n variables) en m casillas(en este caso renglones). Un caso particular sería el de repartir tres monedas diferentes en dos monederos diferentes, sin que sobren monedas. Es claro que para cualquier distribución de las monedas, uno de los monederos debe contener más de una moneda. De esta manera concluimos que al menos uno de los renglones debe contener dos o más elementos diferentes de cero, uno de los cuales será un 1-principal, que generará una variable principal, los otros elementos generarán variables libres. El hecho de tener variables libres implica que el sistema tiene infinitas soluciones. ∎

1.2. Matrices

1.2.1. Definición de matriz

Definición 1.2.1 (Matriz y elementos) *Una matriz en un arreglo rectangular de números. Los números que conforman el arreglo se conocen como elementos o entradas de la matriz. En este trabajo solo considerarán matrices cuyos elementos son números reales.*

Denotaremos a las matrices mediante letras mayúsculas y a sus elementos mediante letras minúsculas; por ejemplo[5],

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

[5]De acuerdo con la definición, A no es una matriz, sino la representación de una matriz

En general, una matriz consta de renglones(líneas horizontales) y columnas(líneas verticales). En el ejemplo anterior, a y b son elementos del primer renglón y, c y d son elementos del segundo renglón. De la misma manera a y c son elementos de la primera columna y, b y d son elementos de la segunda columna.

Definición 1.2.2 (Orden o tamaño) *El orden o tamaño de una matriz se especifica mediante el número de renglones y el número de columna de la matriz. Si n y m son enteros positivos, se dice que una matriz A es de tamaño mxn si ésta consta de m renglones y n columnas, en cuyo caso se escribirá $(A)_{mxn}$[6] para denotar esta característica de la matriz.*

Una matriz general A de tamaño mxn puede representarse de manera sencilla mediante el siguiente arreglo,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \tag{1.31}$$

Notación Anteriormente se discutió el elemento a_{ij} como aquél que se encuentra en el renglón i y en la columna j de la matriz A. En este caso,

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

.

Definición 1.2.3 (Vectores columna y vectores renglón) *Una matriz A de tamaño 1xn se comoce como vector renglón. De manera semejante una matriz A de tamaño mx1 se conoce como vector columna.*

La razón por la cual estas matrices se conocen como vectores se discutirá en la unidad 2.

Definición 1.2.4 (Matriz cuadrada) *S dice que una matriz es cuadrada y de tamaño nxn si ésta consta de n renglones y de n columnas. Para matrices cuadradas, los elementos $a_{11}, a_{22}, \ldots, a_{nn}$ se conocen como elementos de la diagonal principal.*

[6]Observe que en general una matriz de orden mxn es diferente que una de orden nxm.

Definición 1.2.5 (Delta de Kronecker) *La delta de Kronecker es una función de dos variables, que vale 1 si éstas son iguales, y 0 si son diferentes. Se representa con el símbolo δ_{ij} como se muestra a continuación,*

$$\delta_{ij} = \begin{cases} 1, & si\ i = j, \\ 0, & si\ i \neq j. \end{cases} \tag{1.32}$$

Si i y j asumen valores enteros desde 1 hasta n, entonces $\delta_{11} = \delta_{22} = ...\delta_{nn} = 1$, siendo cero en cualquier otro caso, por ejemplo, $\delta_{12} = 0$.

Definición 1.2.6 (Matriz diagonal) *Se dice que una matriz cuadrada es diagonal, si los elementos de su diagonal principal son todos diferentes de cero, siendo cero todos sus elementos restantes.*

Una matriz diagonal tiene el siguiente aspecto,

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}. \tag{1.33}$$

Los elementos de una matriz diagonal pueden expresarse en términos de la delta de Kronecker como,

$$[A]_{ij} = a_{ij}\delta_{ij}. \tag{1.34}$$

Definición 1.2.7 (Matriz identidad I_n) *La matriz identidad I_n es una matriz diagonal de tamaño nxn, en la cual los elementos de su diagonal principal son todos iguales a la unidad.*

La matriz identidad I_n tiene la siguiente apariencia,

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{1.35}$$

En términos de la delta de Kronecker los elementos de la matriz identidad pueden expresarse como,

$$[I_n]_{ij} = \delta_{ij}. \tag{1.36}$$

Definición 1.2.8 (Matriz cero) *Una matriz de tamaño mxn se conoce como matriz cero y se denota por $(0)_{mxn}$, si todos sus elementos son iguales a cero.*

A continuación se muestra la matriz cero de tamaño 2x3,

$$(0)_{2X3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

1.2.2. Álgebra matricial

Definición 1.2.9 (Matrices iguales y matrices diferentes) *Se dice que dos matrices A y B son son iguales, si son del mismo tamaño y sus elementos correspondientes son iguales, en este caso se escribirá $A = B$. En caso contrario se dice que las matrices son diferentes y se escribe $A \neq B$.*

Ejemplo 1.2.1 *Consideremos las siguientes matrices*

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

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

Las matrices A y B son iguales, ya que son del mismo tamaño y sus elementos correspondientes son iguales; así que $A = B$.
Las matrices A y C son diferentes, ya que no son matrices del mismo tamaño, en este caso $A \neq C$
Las matrices A y D son diferentes, ya que $a_{23} \neq d_{23}$. En este caso también $A \neq D$.

Definición 1.2.10 (Suma de matrices) *Si A y B son dos matrices ambas de tamaño mxn, entonces la suma de A y B, denotado por $A + B$, es la matriz de tamaño mxn cuyos elementos se obtienen sumando los elementos correspondientes de las matrices A y B; es decir, si a_{ij} y b_{ij} son los elementos correspondientes de las matrices A y B, respectivamente, entonces*

$$[A + B]_{ij} = [A]_{ij} + [B]_{ij} = a_{ij} + b_{ij}. \tag{1.37}$$

Ejemplo 1.2.2 *Obtenga la suma de las matrices A y B, donde A y B son las siguientes matrices,*

$$A = \begin{bmatrix} 2 & 3 & -2 \\ -1 & 4 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 3 & 5 \end{bmatrix}.$$

Solución: *Como A y B son matrices de tamaño 2x3, entonces A+B también será de tamaño 2x3.*

$$A + B = \begin{bmatrix} 5 & 2 & -2 \\ -2 & 7 & 9 \end{bmatrix}.$$

Definición 1.2.11 (Multiplicación por un escalar) *Si A es una matriz de tamaño mxn y k es un escalar, entonces la multiplicación de la matriz A por el escalar k, denotado por kA, es la matriz de tamaño mxn cuyos elementos se obtienen multiplicando cada elemento de la matriz A por el escalar k; es decir,*

$$[kA]_{ij} = k[A]_{ij} = ka_{ij}. \tag{1.38}$$

Ejemplo 1.2.3 *Obtenga las matrices $2A$ y $(-1)A$, donde A es la matriz que se muestra a continuación,*

$$A = \begin{bmatrix} 5 & 2 \\ -2 & 7 \\ 3 & 4 \end{bmatrix}.$$

Solución: *Como A es una matriz de tamaño 3x2, entonces $2A$ y $(-1)A$ son las matrices de tamaño 3x2 que se muestran a continuación:*

$$2A = \begin{bmatrix} 10 & 4 \\ -4 & 14 \\ 6 & 8 \end{bmatrix}, \qquad (-1)A = \begin{bmatrix} -5 & -2 \\ 2 & -7 \\ -3 & -4 \end{bmatrix}.$$

Definición 1.2.12 (Inverso aditivo) *Si A es una matriz de tamaño mxn, entonces el inverso aditivo de A, denotado por -A, es la matriz de tamaño mxn definida como,*

$$-A = (-1)A,$$

es decir, si $[A]_{ij} = a_{ij}$ entonces

$$[-A]_{ij} = (-1)[A]_{ij} = -a_{ij}. \tag{1.39}$$

Definición 1.2.13 (Resta o sustración de matrices) *Si A y B son matrices de tamaño mxn, entonces la resta o sustracción de A y B, denotada por $A - B$, es la matriz que se define de la siguiente manera,*

$$A - B = A + (-B),$$

es decir, si $[A]_{ij} = a_{ij}$ y $[B]_{ij} = b_{ij}$, entonces

$$[A - B]_{ij} = a_{ij} + (-b_{ij}) = a_{ij} - b_{ij}. \tag{1.40}$$

Ejemplo 1.2.4 *Obtenga la resta de las matrices A y B, donde A y B son las siguientes matrices,*

$$A = \begin{bmatrix} 2 & 3 & -2 \\ -1 & 4 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 3 & 5 \end{bmatrix}.$$

Solución: Como A y B son matrices de tamaño 2x3, entonces $A-B$, también será de tamaño 2x3.

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

Puede observarse que para hallar $A-B$ se ha sumado a la matriz A el inverso aditivo de la matriz B. El mismo resultado se obtiene si se restan de manera directa los elementos de la matriz B a los elementos correspondientes de la matriz A.

Definición 1.2.14 (Producto de matrices) *Si A y B son matrices de tamaños mxr y rxn,* [7] *respectivamente, entonces el producto matricial de A y B, denotado por AB, es la matriz de tamaño mxn cuyos elementos se determinan de la siguiente manera,*

$$[AB]_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj} \tag{1.41}$$

donde a_{ik} *y* b_{kj} *representan elementos de las matrices A y B, respectivamente.*

Para comprender mejor el producto de matrices considérense las matrices A y B, y obsérvense el renglón i de la matriz A y la columna j de la matriz B, como se muestra a continuación,

$$[AB]_{ij} = \left[\left[\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1r} \\ \vdots & \vdots & \cdots & \vdots \\ \hline a_{i1} & a_{i2} & \cdots & a_{ir} \\ \hline \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mr} \end{array} \right] \left[\begin{array}{cc|c|cc} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ b_{r1} & \cdots & b_{rj} & \cdots & b_{rn} \end{array} \right] \right]_{ij}$$

$$[AB]_{ij} = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{ir} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{rj} \end{bmatrix} = \sum_{k=1}^{r} a_{ik} b_{kj}. \tag{1.42}$$

Se observa que el elemento $[AB]_{ij}$ se obtiene al multiplicar los elementos del renglón i en la matriz A con los elementos correspondientes de la columna j en la matriz B y posteriormente sumar los productos.

[7] Es importante notar que los índices interiores deben coincidir, de otro modo la multiplicación no estará definida.

Ejemplo 1.2.5 *Obtenga el producto AB, donde A y B son las matrices,*

$$A = \begin{bmatrix} 2 & 3 & -2 \\ -1 & 4 & 4 \\ 3 & 2 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & -1 \\ -1 & 3 \\ 5 & 2 \end{bmatrix}.$$

Solución: Como la matriz A es de tamaño 2x3 y la matriz B es de tamaño 3x2, la matriz AB será de tamaño 2x2.

$$AB = \begin{bmatrix} 2 & 3 & -2 \\ -1 & 4 & 4 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -1 & 3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} -7 & 3 \\ 13 & 21 \\ 12 & 5 \end{bmatrix}.$$

Ejemplo 1.2.6 *Obtenga el producto AB, donde A y B son las matrices,*

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

Solución: Como la matriz A es de tamaño 2x3 y la matriz B es de tamaño 3x1, el producto es posible debido a que los índices internos coinciden, de manera que la matriz AB será de tamaño 2x1.

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

Este ejemplo nos muestra que la multiplicación de matrices tiene propiedades diferentes a las que conocemos para la multiplicación de números reales. Como puede apreciarse, $AB = 0$, aun cuando ninguna de las matrices que aparecen en la multiplicación es la matriz cero. Puede observarse también que el producto BA no está definido, ya que al considerar los tamaños de las matrices, 3x1 y 2x3, los índices internos no coinciden.

Teorema 1.2.1 (Aritmética matricial) *Suponiendo que los tamaños de las matrices son tales que es posible efectuar las operaciones indicadas, entonces son válidas las siguientes reglas de la aritmética matricial:*

a) $A + B = B + A$,

b) $A + (B + C) = (A + B) + C$,

c) $A(BC) = (AB)C$,

d) $A(B + C) = AB + AC$,

e) $(B + C)A = BA + CA$,

f) $A(B - C) = AB - AC$,

g) $(B - C)A = BA - CA$,

h) $k(B + C) = kB + kC$,

i) $k(B - C) = kB - kC$,

j) $(k + l)A = kA + lA$,

k) $(k - l)A = kA - lA$,

l) $(kl)A = k(lA) = l(kA)$,

m) $k(AB) = (kA)B = A(kB)$.

Demostración: Solo se demostrarán los incisos a) y c). La demostración de los incisos restantes se deja como ejercicio para los estudiantes.

a) Sean A y B matrices de tamaños mxn, con elementos correspondientes a_{ij} y b_{ij}, respectivamente, entonces $A + B$ y $B + A$ son matrices de tamaño mxn. Aún queda por demostrar que los elementos correspondientes de estas matrices son iguales.
Por definición,

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

como las entradas de las matrices A y B son números reales, entonces éstos conmutan respecto a la suma, así que

$$[A + B]_{ij} = a_{ij} + b_{ij} = b_{ij} + a_{ij} = [B + A]_{ij}. \tag{1.43}$$

Como puede apreciarse en la ecuación 1.43, las entradas correspondientes de las matrices son iguales, entonces se concluye que $A + B = B + A$, ya que las matrices también son del mismo tamaño.

c) Sean A, B y C matrices de tamaño mxr, rxp y pxn respectivamente, entoncces, A(BC) y (AB)C son de tamaños mxn. Ahora verifiquemos que los elementos correspondientes de estas matrices sean iguales. sean $[A]_{ij} = a_{ij}$, $[B]_{ij=b_{ij}}$ y $[C]_{ij} = c_{ij}$ los elementos correspondientes de las matrices A, B y C respectivamente, entonces aplicando la definición de producto de matrices se obtiene,

$$
\begin{aligned}
[A(BC)]_{ij} &= \sum_{k=1}^{r} a_{ik}[BC]_{kj}, \\
&= \sum_{k=1}^{r} a_{ik}\left(\sum_{l=1}^{p} b_{kl}c_{lj}\right), \\
&= \sum_{k=1}^{r}\sum_{l=1}^{p} a_{ik}b_{kl}c_{lj}, \\
&= \sum_{l=1}^{p}\sum_{k=1}^{r} a_{ik}b_{kl}c_{lj} \\
&= \sum_{l=1}^{p}\left(\sum_{k=1}^{r} a_{ik}b_{kl}\right)c_{lj}, \\
&= \sum_{l=1}^{p} [AB]_{il}c_{lj}, \\
&= [(AB)C]_{ij}. \qquad (1.44)
\end{aligned}
$$

De la última igualdad, representada por la ecuación 1.44, se puede ver que las entradas correspondientes de las matrices $A(BC)$ y $(AB)C$ son iguales, lo que aunado al hecho de tener tamaños iguales, permite concluir que, $A(BC) = (AB)C$. ∎

1.2.3. Transpuesta de una matriz

Definición 1.2.15 (Transpuesta de una matriz) *Si A es una matriz cualquiera de tamaño mxn, entonces la transpuesta de A, denotada por A^T, se define como la matriz de nxm cuya primera columna es el primer renglón de A, su segunda columna es el segundo renglón de A, y así sucesivamente, su última columna corresponde al último renglón de A. En este caso, si $[A]_{ij} = a_{ij}$, entonces $[A^T]_{ij} = a_{ji}$.*

Ejemplo 1.2.7 *Halle las transpuestas de las siguientes matrices,*

$$
A = \begin{bmatrix} 1 & 4 \\ 2 & -3 \\ 5 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}.
$$

Solución:

$$
A^T = \begin{bmatrix} 1 & 2 & 5 \\ 4 & -3 & 0 \end{bmatrix}, \qquad B^T = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}.
$$

Teorema 1.2.2 *Si A y B son matrices tales que las operaciones indicadas pueden realizarse, entonces*

a) $(A^T)^T = A$,

b) $(A+B)^T = A^T + B^T$,

c) $(kA)^T = kA^T$, *donde k es un escalar,*

d) $(AB)^T = B^T A^T$.

Demostración: Solo se demostrará el inciso d), el resto se deja como ejercicio para el estudiante.

d) Sean $(A)_{mxp}$ y B_{pxn}, entonces $(AB)_{mxn}$ y $((AB)^T)_{nxm}$. Además, $(A^T)_{pxm}$, $(B^T)_{nxp}$, luego $(B^T A^T)_{nxm}$. Se ha probado con esto que las matrices involucradas en la igualdad son del mismo tamaño. Todavía falta probar que los elementos correspondientes son iguales. Sean $[A]_{ij} = a_{ij}$ y $[B]_{ij} = b_{ij}$, entonces

$$\begin{aligned}
[(AB)^T]_{ij} &= [AB]_{ji}, \\
&= \sum_{k=1}^{p} a_{jk} b_{ki}, \\
&= \sum_{k=1}^{p} [A^T]_{kj} [B^T]_{ik}, \\
&= \sum_{k=1}^{p} [B^T]_{ik} [A^T]_{kj}, \\
&= [B^T A^T]_{ij}.
\end{aligned}$$

Por lo tanto, $(AB)^T = B^T A^T$. ▮

1.2.4. Representación matricial de un sistema de ecuaciones

Considérense las matrices $(A)_{mxn}$, $(x)_{nx1}$ y $(b)_{mx1}$ como se muestra a continuación,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Multiplicando A por x se obtiene,

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}.$$

Además, de acuerdo con la definición para la igualdad de matrices, $Ax = b$,

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}. \tag{1.45}$$

Lo anterior significa que una sola ecuación matricial, $Ax = b$, puede representar un sistema de m ecuaciones en n variables.

Teorema 1.2.3 *Suponiendo que los tamaños de las matrices son tales que pueden efectuarse las operaciones indicadas, entonces las siguientes reglas de la aritmética matricial son válidas,*

a) $A + 0 = 0 + A = A$,

b) $A - A = 0$,

c) $A0 = 0A = 0$.

Demostración: Solo se demostrará el inciso a), el resto se deja como ejercicio para el estudiante.

a) Si A y 0 son matrices de tramaño mxn, entonces A+0 y 0+A son matrices de tamaño mxn, así que

$$[A + 0]_{ij} = [A]_{ij} + [0]_{ij} = a_{ij} + 0 = a_{ij} = [A]_{ij},$$

luego,

$$A + 0 = A,$$

además, por el inciso a) del teorema 1.2.1 se concluye el resultado. ∎

Teorema 1.2.4 *Todo sistema de ecuaciones lineales tiene una de las siguientes propiedades:*

i) No tiene soluciones.

ii) Tiene solución única.

iii) Tiene infinitas soluciones.

Demostración: Considérese el sistema de ecuaciones representado por la ecuación matricial $Ax = b$. Ya se han estudiado los casos en que el sistema no tiene soluciones y también el caso en el cual el sistema tiene solución única. Nuestro interés se centra en investigar si un sistema de ecuaciones puede tener solamente dos soluciones o si el hecho de tener más de una solución implica que el sistema tiene infinitas soluciones. Para tal fin supóngase que x_1 y x_2 son soluciones de la ecuación $Ax = b$; es decir,

$$Ax_1 = b,$$

$$Ax_2 = b,$$

restando las ecuaciones anteriores se obtiene,

$$Ax_2 - Ax_1 = b - b,$$

$$A(x_2 - x_1) = 0.$$

Ahora, sea $x_o = x_2 - x_1$, entonces

$$Ax_o = 0.$$

Por otro lado, sea k un escalar, entonces $kx_o + x_1$ es también solución del sistema $Ax = b$, es decir,

$$A[kx_o + x_1] = A(kx_0) + Ax_1 = kAx_o + b == k0 + b = 0 + b = b.$$

Como esto es válido para cada valor de k, entonces el sistema tiene infinitas soluciones. ∎

Teorema 1.2.5 *Si A y B son matrices de tamaños nxn y mxn, respectivamente, entonces los siguientes enunciados son válidos,*

i) $AI_n = I_nA = A$,

ii) $I_mB = B$,

iii) $BI_n = B$.

Demostración: Solo se demostrará el inciso ii), el resto se deja como ejercicio para el alumno. Como B es una matriz de tamaño mxn, es claro que I_mB es también una matriz de tamaño mxn, además,

$$
\begin{aligned}
[I_m B]_{ij} &= \sum_{k=1}^{k=m} [I]_{ik}[B]_{kj}, \\
&= \sum_{k=1}^{k=m} \delta_{ik} b_{kj}, \\
&= b_{ij}, \\
&= [B]_{ij}.
\end{aligned} \tag{1.46}
$$

La ecuación 1.46 muestra que los elementos correspondientes de las matrices $I_m B$ y B son iguales; por lo tanto, $I_m B = B$. ∎

1.3. Inversa de una matriz

Definición 1.3.1 (Inversa de una matriz) *Si A es una matriz cuadrada cualquiera y si es posible hallar una matriz B tal que, AB=BA=I, entonces se dice que A es invertible y B se conoce como su inversa. Las matrices invertibles también se conocen como matrices no singulares.*

Ejemplo 1.3.1 *Compruebe que B es la inversa de la matriz A, donde A y B son las siguientes matrices,*

$$
A = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}.
$$

Solución: Para realizar la comprobación se debe calcular AB y BA.

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

de la misma manera,

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

Como se observa que, $AB = BA = I$, por lo tanto A es invertible y su inversa es B.

Ejemplo 1.3.2 *Compruebe que para cualquier valor del ángulo θ, B es la inversa de la matriz A, donde A y B son las siguientes matrices,*

$$A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}, \qquad B = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

Solución: Consideremos los productos AB y BA.

$$\begin{aligned} AB &= \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\theta + \operatorname{sen}^2\theta & 0 \\ 0 & \cos^2\theta + \operatorname{sen}^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

De la misma manera,

$$\begin{aligned} BA &= \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\theta + \operatorname{sen}^2\theta & 0 \\ 0 & \cos^2\theta + \operatorname{sen}^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Por lo anterior, $AB = BA = I$, lo que prueba que A es invertible y su inversa es B.

Teorema 1.3.1 (Unicidad de la inversa) *Si tanto B como C son inversas de la matriz A, entonces B=C.*

Demostración: Como B es inversa de A, entonces

$$AB = I,$$

multiplicando la ecuación anterior por la matriz C se obtiene,

$$C(AB) = CI,$$

por los teoremas 1.2.1 y 1.2.5,

$$(CA)B = C,$$

además, como C es también inversa de A, entonces $CA = I$, luego

$$\begin{aligned} IB &= C, \\ B &= C. \end{aligned}$$

Se ha demostrado, con este teorema, que la inversa de una matriz invertible A es única. De aquí en adelante se le denotará por A^{-1}.

Teorema 1.3.2 *Si A y B son matrices invertibles del mismo tamaño, entonces AB es invertible y su inversa es*[8]*,*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Demostración: Como A y B son invertibles, entonces

$$AA^{-1} = A^{-1}A = I,$$
$$BB^{-1} = B^{-1}B = I.$$

Ahora, AB será invertible si existe una matriz C tal que $C(AB) = (AB)C = I$, en este caso C será la inversa de AB. En efecto éste es el caso, ya que si se hace $C = B^{-1}A^{-1}$, se verifica lo anterior, como se muestra a continuación.

$$\begin{aligned} (B^{-1}A^{-1})(AB) &= (B^{-1})(A^{-1}(AB)), \\ &= (B^{-1})((A^{-1}A)B), \\ &= (B^{-1})(IB) = (B^{-1})(B), \\ &= I. \quad \blacksquare \end{aligned}$$

Se deja como ejercicio para el estudiante verificar que $(AB)(B^{-1}A^{-1}) = I$.

Un resultado más general puede enunciarse de la siguiente manera: si $A_1, A_2, \ldots, A_k$ son matrices invertibles del mismo tamaño, entonces el producto $A_1A_2\ldots A_k$ es invertible y su inversa es

$$(A_1A_2\ldots A_k)^{-1} = A_k^{-1}A_{k-1}^{-1}\ldots A_1^{-1}. \tag{1.47}$$

Este resultado puede demostrarse por el método de inducción matemática, se deja como ejercicio al estudiante.

Definición 1.3.2 (Potencia de matrices) *Si A es una matriz cuadrada y n es un entero positivo, se define A^n como,*

$$A^n = \underbrace{AA\ldots A}_{n-factores} \tag{1.48}$$

además,

$$A^0 = I \tag{1.49}$$

Por otro lado, si A es una matriz invertible y n es un entero positivo, entonces definimos A^{-n} como,

$$A^{-n} = \underbrace{A^{-1}A^{-1}\ldots A^{-1}}_{n-factores} \tag{1.50}$$

[8]Observe el orden en que aparece el producto de las inversas.

Teorema 1.3.3 *Si A es una matriz invertible, entonces*

a) A^{-1} *es invertible y* $(A^{-1})^{-1} = A$,

b) A^n *es invertible y* $(A^n)^{-1} = (A^{-1})^n$, *para* $n = 0, 1, 2, \ldots$

c) Para cualquier escalar $k \neq 0$, kA *es invertible y* $(kA)^{-1} = \frac{1}{k}A^{-1}$.

Demostración: Solo se demostrará el inciso a), el resto se deja como ejercicio para el estudiante.

Se sabe que A es invertible, así que

$$AA^{-1} = A^{-1}A = I.$$

Ahora, A^{-1} será invertible si existe una matriz B tal que,

$$BA^{-1} = A^{-1}B = I,$$

En efecto, dicha matriz existe, ya que haciendo $B = A$ se obtiene la condición deseada. Así que A es la inversa de A^{-1} o de manera equivalente,

$$(A^{-1})^{-1} = A. \quad \blacksquare$$

1.3.1. Matrices elementales y matrices equivalentes a la matriz identidad

Definición 1.3.3 (Matriz elemental) *Se dice que una matriz de nxn es una matriz elemental si ésta puede obtenerse de la matriz identidad realizando una sola operación sobre sus renglones.*

Ejemplo 1.3.3 *A continuación se listan cuatro matrices elementales y las operaciones que las producen:*

i) $\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$ Multiplique el segundo renglón de I_2 por -3

ii) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ Intercambie los renglones 2 y 4 de I_4

iii) $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ Sume 3 veces el tercer renglón de I_3 al primero.

iv) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ Multiplique el primer renglón de I_3 por 1.

Teorema 1.3.4 *Si la matriz elemental E resulta al efectuar cierta operación elemental sobre los renglones de I_m y A es una matriz de tamaño mxn, entonces el producto EA es la matriz que resulta al efectuar la misma operación elemental sobre los renglones de A.*

Demostración: Las operaciones elementales sobre los renglones generan tres tipos de matrices elementales, a saber,

i) E_{ij}: Es la matriz elemental que se obtiene al intercambiar los renglones i y j de la matriz identidad.

ii) $E_i(k)$: Es la matriz elemental que resulta al multiplicar el renglón i de la matriz identidad por un escalar $k \neq 0$.

iii) $E_{ij}(k)$: Es la matriz elemental que resulta al sumar k veces el renglón i al renglón j en la matriz identidad.

Solo se demostrará el inciso iii), los incisos restantes se dejan como ejercicio para el estudiante.

Sea B la matriz que se obtiene al realizar la operación $kr_i + r_j \longrightarrow r_j$ sobre los renglones de la matriz A, además supongamos que $1 \leq i < j \leq m$, entonces

$$A = \begin{array}{cc} & \begin{array}{ccccc} \phantom{a_{11}} & & p & & \phantom{a_{1n}} \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} & \begin{pmatrix} a_{11} & \dots & a_{1p} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{ip} & \dots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j1} & \cdots & a_{jp} & \dots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mp} & \dots & a_{mn} \end{pmatrix} \end{array}$$

en consecuencia,

$$B = \begin{array}{c c} & \begin{array}{ccccc} \phantom{ka_{i1}+a_{j1}} & & p & & \phantom{ka_{in}+a_{jn}} \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} & \begin{pmatrix} a_{11} & \dots & a_{1p} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{ip} & \dots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ ka_{i1} + a_{j1} & \cdots & ka_{ip} + a_{jp} & \dots & ka_{in} + a_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mp} & \dots & a_{mn} \end{pmatrix} \end{array}$$

Por otro lado, la matriz $E_{ij}(k)$ de tamaño mxm definida anteriormente, se expresa como,

$$E = E_{ij}(k) = \begin{array}{c c} & \begin{array}{ccccccc} & & i & & j & & \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} & \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & k & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{array}$$

Ahora considérese el producto EA,

$$EA = \begin{array}{c c} & \begin{array}{ccccccc} & & i & & j & & \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} & \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & k & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{array} \begin{array}{c c} & \begin{array}{ccccc} & & p & & \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} & \begin{pmatrix} a_{11} & \dots & a_{1p} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{ip} & \dots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j1} & \cdots & a_{jp} & \dots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mp} & \dots & a_{mn} \end{pmatrix} \end{array}$$

$$EA = \begin{array}{cc} & \begin{array}{ccccc} \phantom{ka_{i1}+a_{j1}} & & p & & \phantom{ka_{in}+a_{jn}} \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} & \begin{pmatrix} a_{11} & \cdots & a_{1p} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{ip} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ ka_{i1} + a_{j1} & \cdots & ka_{ip} + a_{jp} & \cdots & ka_{in} + a_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mp} & \cdots & a_{mn} \end{pmatrix} \end{array}$$

Como puede obervarse, ésta es la matriz que se obtiene si se realiza directamente la operación $kr_i + r_j \longrightarrow r_j$ sobre los renglones de la matriz A; es decir, $EA = B$, con lo cual queda demostrado el teorema. ▌

1.3.2. La inversa de una matriz como el producto de matrices elementales

Si E es la matriz elemental que se obtiene al realizar una operación elemental sobre los renglones de la matriz identidad I, entonces existe una segunda operación elemental que realizada sobre los renglones de E nos devuelve la matriz identidad. En la siguiente tabla se muestran estas operaciones.

Tabla 1.1: Operaciones elementales

Operación sobre los renglones de I que produce E	Operación sobre los renglones de E que produce I
Multiplíquese el renglón i por $k \neq 0$	Multiplíquese el renglón i por $\frac{1}{k}$
Intercámbiense los renglones i y j	Intercámbiense los renglones i y j
Súmese k veces el renglón i al renglón j	Súmese $-k$ veces el renglón i al renglón j

Teorema 1.3.5 *Toda matriz elemental es invertible y la inversa también es una matriz elemental.*

Demostración: Supóngase que se realiza una operación elemental sobre los renglones I (como las que se indican en primera columna la tabla 1.1 para obtener la matriz elemental E y que se aplica la operación elemental inversa corresponciente(como las que se indican en la segunda columna de la tabla

1.1) sobre los rengones de I, para obtener la matriz E_o. Ahora, por el teorema 1.3.4 , se tiene que,

$$E_o E = I,$$

es decir, al realizar la operación inversa sobre los renglones de E nos da la identidad; pero esto equivale a multiplicar la matriz E por la matriz elemental E_o (teorema 1.3.4). Con argumentos similares se establece,

$$EE_o = I,$$

es decir, E es invertible y,

$$E^{-1} = E_0. \quad \blacksquare$$

Definición 1.3.4 (Matrices equivalentes respecto a los renglones $\sim$) *Son aquellas matrices que pueden obtenerse una de la otra mediante una sucesión finita de operaciones elementales sobre los renglones.*

Teorema 1.3.6 *Si A es una matriz de tamaño nxn, entonces las siguientes proposiciones son equivalentes,*

a) A es invertible.

b) $Ax = 0$, *solo tiene la solución trivial.*

c) A es equivalente respecto a los renglones a I_n.

Demostración: Se demostrará el teorema mediante la línea de implicaciones $a) \Rightarrow b) \Rightarrow c) \Rightarrow a)$.

$a) \Rightarrow b)$ Si A es invertible, entonces $AA^{-1} = A^{-1}A = I$. Sea x una solución de,

$$Ax = 0,$$

entonces multiplicando por A^{-1} se obtiene,

$$\begin{aligned} A^{-1}(Ax) &= A^{-1}0, \\ (A^{-1}A)x &= 0, \\ Ix &= 0, \\ x &= 0. \end{aligned} \tag{1.51}$$

La ecuación 1.51 nos muestra que la homogénea $Ax = 0$ solo tiene la solución trivial.

b) $\Rightarrow$ *c*) Si $Ax = 0$ solo tiene la solución trivial $x = 0$, entonces, reescribiendo las matrices aumentadas de estos dos sistemas de ecuaciones se tiene,

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 \end{array}\right],$$

el sistema de ecuaciones correspondiente a $Ix = x = 0$ es

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

Lo anterior muestra que la matriz A puede ser llevada mediante operaciones elementales sobre sus renglones a la matriz identidad; es decir, $A \sim I_n$.

c) $\Rightarrow$ *a*) Se sabe que $A \sim I_n$; es decir, mediante una sucesión finita de operaciones elementales sobre los renglones de A es posible obtener la matriz identidad. En otras palabras, por el teorema 1.3.4, existe una sucesión de matrices elementales E_1, E_2, ...E_k tales que,

$$E_k...E_2E_1A = I_n.$$

Cada una de las matrices elementales es invertible, así que multiplicando por las inversas de cada matriz elemental obtenemos,

$$A = E_1^{-1}E_2^{-1}...E_k^{-1}I_n.$$

Como A es el producto de matrices invertibles, entonces por el teorema 1.3.2, A es invertible y su inversa es

$$A^{-1} = I_nE_k...E_2E_1. \qquad \blacksquare$$

Se observa que la sucesión de operaciones elementales que llevan a la matriz A hacia la identidad, son aquellas que en la misma secuencia, pero actuando sobre la identidad nos generan la inversa de la matriz A. Esto constituye un método básico para determinar la inversa de una matriz, como se muestra en el siguiente esquema.

$$
\begin{array}{rcl}
A & | & I_n \\
E_1A & | & E_1I_n \\
\vdots & | & \vdots \\
E_k\ldots E_1A & | & E_k\ldots E_1I_n \\
I_n & | & A^{-1}.
\end{array}
$$

Paja fijar las ideas expuestas anteriormente se considerarán los siguientes ejemplos.

Ejemplo 1.3.4 *Mediante operaciones elementales sobre los renglones, hallar la inversa de la matriz A.*

$$
A = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix}.
$$

Solución: Se agrega la matriz identidad I_3 a la derecha de la matriz A, para llevar acabo el método descrito en teorema anterior.

$$
\left[\begin{array}{rrr|rrr} 2 & -1 & 4 & 1 & 0 & 0 \\ 1 & 2 & -3 & 0 & 1 & 0 \\ 5 & 3 & 1 & 0 & 0 & 1 \end{array}\right].
$$

A continuación se procede a reducir esta matriz mediante operaciones elementales. El estudiante debe ser capaz de reconocer las operaciones elementales que se realizan en cada paso, de lo contrario, se le recomienda repasar la sección 1.1.4.

$$
\left[\begin{array}{rrr|rrr} 2 & -1 & 4 & 1 & 0 & 0 \\ 1 & 2 & -3 & 0 & 1 & 0 \\ 5 & 3 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|rrr} 1 & 2 & -3 & 0 & 1 & 0 \\ 2 & -1 & 4 & 1 & 0 & 0 \\ 5 & 3 & 1 & 0 & 0 & 1 \end{array}\right]
$$

$$
\sim \left[\begin{array}{rrr|rrr} 1 & 2 & -3 & 0 & 1 & 0 \\ 0 & -5 & 10 & 1 & -2 & 0 \\ 0 & -7 & 16 & 0 & -5 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|rrr} 1 & 2 & -3 & 0 & 1 & 0 \\ 0 & 1 & -2 & -\frac{1}{5} & \frac{2}{5} & 0 \\ 0 & -7 & 16 & 0 & -5 & 1 \end{array}\right]
$$

$$
\sim \left[\begin{array}{rrr|rrr} 1 & 0 & 1 & \frac{2}{5} & \frac{1}{5} & 0 \\ 0 & 1 & -2 & -\frac{1}{5} & \frac{2}{5} & 0 \\ 0 & 0 & 2 & -\frac{7}{5} & -\frac{11}{5} & 1 \end{array}\right] \sim \left[\begin{array}{rrr|rrr} 1 & 0 & 1 & \frac{2}{5} & \frac{1}{5} & 0 \\ 0 & 1 & -2 & -\frac{1}{5} & \frac{2}{5} & 0 \\ 0 & 0 & 1 & -\frac{7}{10} & -\frac{11}{5} & \frac{1}{2} \end{array}\right]
$$

$$
\sim \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & \frac{11}{10} & \frac{13}{10} & -\frac{5}{10} \\ 0 & 1 & 0 & -\frac{16}{10} & -\frac{18}{5} & \frac{10}{10} \\ 0 & 0 & 1 & -\frac{7}{10} & -\frac{11}{10} & \frac{5}{10} \end{array}\right].
$$

De lo anterior se tiene que la inversa de la matriz A es

$$A^{-1} = \begin{bmatrix} \frac{11}{10} & \frac{13}{10} & -\frac{5}{10} \\ -\frac{16}{10} & -\frac{18}{10} & \frac{10}{10} \\ -\frac{7}{10} & -\frac{11}{10} & \frac{5}{10} \end{bmatrix}. = \frac{1}{10}\begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix}.$$

Con esto queda resuelto el ejercicio. Sin embargo, se deja al estudiante determinar la inversa de la matriz,

$$\begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix}.$$

Ejemplo 1.3.5 *Determine la inversa de la siguiente matriz,*

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

Solución: Nuevamente se agrega la matriz identidad I_3 a la derecha de la matriz A y se reduce mediante operaciones elementales.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 9 & 5 & -2 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 9 & 5 & -2 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right].$$

Luego la matriz inversa de la matriz A es

$$A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}.$$

Teorema 1.3.7 *Si A y B son matrices cuadradas tales que,*

a) AB=I, entonces $B = A^{-1}$.

b) BA=I, entonces $B = A^{-1}$.

Demostración: a) Se sabe que $AB = I$. Considérese ahora la ecuación homogénea, $Bx = 0$. Multiplicando la ecuación anterior por la matriz A se obtiene,

$$\begin{aligned} A(Bx) &= A0, \\ (AB)x &= 0, \\ Ix &= 0, \\ x &= 0. \end{aligned} \tag{1.52}$$

La ecuación 1.52 muestra que la ecuación homogénea, $Bx = 0$, tiene únicamente la solución trivial. Por el teorema 1.3.6 la matriz B es invertible; es decir, existe una única matriz B^{-1}, tal que $B^{-1}B = BB^{-1} = I$, así que multiplicando la ecuación, $AB = I$, por B^{-1} (por la derecha) se obtiene,

$$\begin{aligned} AB &= I, \\ (AB)B^{-1} &= IB^{-1}, \\ A(BB^{-1}) &= B^{-1}, \\ AI &= B^{-1}, \\ A &= B^{-1}. \end{aligned} \tag{1.53}$$

La matriz A es invertible, ya que $BA = BB^{-1} = I$, y por hipótesis $AB = I$. Se concluye que $A^{-1} = B$.

b) Se sabe que $BA = I$. Considérese la ecuación homogénea $Ax = 0$. Multiplicando la ecuación anterior por la matriz B se obtiene,

$$\begin{aligned} B(Ax) &= B0, \\ (BA)x &= 0, \\ Ix &= 0, \\ x &= 0. \end{aligned} \tag{1.54}$$

La ecuación 1.54 muestra que la ecuación homogénea, $Ax = 0$, tiene únicamente la solución trivial, por el teorema 1.3.6, la matriz A es invertible; es decir, existe una matriz única A^{-1}, tal que, $A^{-1}A = AA^{-1} =$

I, así que multiplicando la ecuación, $BA = I$, por A^{-1} por la derecha obtenemos,

$$\begin{aligned} BA &= I, \\ (BA)A^{-1} &= IA^{-1}, \\ B(A^{-1}A) &= A^{-1}, \\ BI &= A^{-1}, \\ B &= A^{-1}. \quad \blacksquare \end{aligned} \tag{1.55}$$

$\blacksquare$

Teorema 1.3.8 *Si A es una matriz de tamaño nxn, entonces las proposiciones siguientes son equivalentes,*

a) A es invertible.

b) $Ax = 0$, tiene únicamente la solución trivial $x = 0$.

c) A es equivalente respecto a los renglones a I_n.

d) $Ax = b$, es consistente para toda matriz b de nx1.

Demostración: Únicamente se ha agregado el inciso d) al teorema 1.3.6, por lo tanto, solo se demostrará la equivalencia entre d) y cualquiera de los otros incisos. En particular se demostrará la equivalencia entre a) y d).

$a) \Rightarrow d)$ Como A es invertible, entonces $A^{-1}A = AA^{-1} = I$. Considérese ahora la ecuación matricial,

$$Ax = b,$$

multiplíquese la ecuación anterior por A^{-1}, así que

$$\begin{aligned} Ax &= b, \\ A^{-1}(Ax) &= A^{-1}b, \\ (A^{-1}A)x &= A^{-1}b, \\ Ix &= A^{-1}b, \\ x &= A^{-1}b. \end{aligned} \tag{1.56}$$

Se ha probado que $Ax = b$ es consistente y tiene solución $x = A^{-1}b$ para toda matriz b de $nx1$.

$d) \Rightarrow a)$. Si $Ax = b$ es consistente para toda matriz b de tamaño $nx1$, entonces en particular es consistente en los siguientes casos,

$$Ax = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, Ax = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, Ax = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Sean

$$x_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix}, x_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix}, \ldots, x_n = \begin{bmatrix} x_1n \\ x_2n \\ \vdots \\ x_nn \end{bmatrix},$$

Sean $x_1, x_2, \ldots x_n$ las soluciones correspondientes de las n ecuaciones matriciales anteriores; es decir,

$$Ax_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, Ax_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, Ax_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Ahora formemos la matriz C con las n matrices x_1,x_2,...,x_n; es decir,

$$C = \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{nn} \end{bmatrix} = \left[\; x_1 \mid x_2 \mid \ldots \mid x_n \;\right].$$

Finalmente, resulta claro que AC es la matriz cuyas columnas son Ax_1, Ax_2, ..., Ax_n.

$$AC = \left[\; Ax_1 \mid Ax_2 \mid \ldots \mid Ax_n \;\right] = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix},$$

es decir,se ha construido una matriz C, tal que,

$$AC = I$$

por el teorema 1.3.7, A es invertible y su inversa es C, con lo cual el teorema queda demostrado. ▮

1.3.3. Solución de un sistema de ecuaciones lineales usando la inversa de la matriz de coeficientes

Para finalizar esta sección se resolverán dos ejemplos de sistema de ecuaciones mediante la inversa de la matriz de coeficientes.

Ejemplo 1.3.6 *Resuelva el siguiente sistema de ecuaciones lineales usando la inversa de la matriz de coeficientes.*

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

Solución: La matriz de coeficientes del sistema es

$$A = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix}.$$

La inversa de esta matriz fue determinada anteriormente, (ver ejemplo 1.3.4), de modo que

$$A^{-1} = \frac{1}{10}\begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix}.$$

Por el teorema 1.3.8, $x = A^{-1}b$, luego

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix}\begin{bmatrix} -3 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Por lo anterior, la solución del sistema es

$$\left\{\begin{array}{rcr} x_1 & = & -1 \\ x_2 & = & 1 \\ x_3 & = & 0 \end{array}\right.$$

Este ejemplo ya había sido resuelto mediante el método de Gauss-Jordan. (ver ejemplo 1.1.8).

Ejemplo 1.3.7 *Resuelva el siguiente sistema de ecuaciones lineales,*

$$\left.\begin{array}{lcr} x+2y+3z & = & 1 \\ 2x+5y+3z & = & 2 \\ x+8z & = & -1 \end{array}\right\}.$$

Solución: La representación matricial del sistema anterior es

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix},$$

La matriz de coeficientes del sistema es

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

La inversa de esta matriz fue determinada en el ejemplo 1.3.5, retomaremos el resultado obtenido.

$$A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}.$$

De acuerdo con el teorema 1.3.8, la solución del sistema está dado por,

$$x = A^{-1}b,$$

en este caso,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}b,$$

sustituyendo obtenemos,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -17 \\ 6 \\ 2 \end{bmatrix}.$$

La solución del sistema es

$$\begin{cases} x &= -17 \\ y &= 6 \\ z &= 2 \end{cases}$$

Este ejercicio había sido resuelto anteriormente mediante el método de eliminación de Gauss (ver ejemplo 1.1.9).

1.4. Determinantes

En esta sección se estudia una de las aplicaciones más importantes definidas sobre las matrices, la función determinante. Aun cuando la función determinante tiene una amplia gama de aplicaciones, en este punto el objetivo principal es desarrollar sus propiedades y aplicarlas en la resolución de sistemas de ecuaciones mediante la regla de Cramer.

Definición 1.4.1 (Producto elemental) *Si A es una matriz cuadrada, entonces se conoce como producto elemental tomado de A, a cualquier producto de n elementos tomados de A, sin que dos cualesquiera de ellos provengan del mismo renglón o de la misma columna.*

Ejemplo 1.4.1 *Determine los productos elementales de la matriz,*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Solución: Los productos elementales deben provenir de renglones diferentes, por lo tanto, los productos elementales son de la forma $a_{1j_1}a_{2j_2}a_{3j_3}$, donde $j_1j_2j_3$ es una permutación del conjunto $\{1, 2, 3\}$. Se tendrán entonces tantos productos elementales como permutaciones del conjunto $\{1, 2, 3\}$; es decir, 3!=6 productos elementales, como se muestra en tabla 1.2.

Tabla 1.2: Productos elementales para una matriz de 3x3

Permutación	Producto elemental
123	$a_{11}a_{22}a_{33}$
132	$a_{11}a_{23}a_{32}$
213	$a_{12}a_{21}a_{33}$
231	$a_{12}a_{23}a_{31}$
312	$a_{13}a_{21}a_{32}$
321	$a_{13}a_{22}a_{31}$

Definición 1.4.2 (Producto elemental con signo) *Dada una matriz cuadrada de tamaño nxn, se denomina producto elemental con signo tomado de A a un producto elemental $a_{1j_1}a_{2j_2}\ldots a_{nj_n}$ multiplicado por $sgn(\sigma)$, donde $\sigma = j_1j_2\ldots j_n$ es una permutación de S_n.*

Definición 1.4.3 (Determinante) *Dada una matriz cuadrada A de tamaño nxn, la función determinante, denotada por $detA$ o $|A|$, se define como la suma de todos los productos elementales con signos tomados de A.*

$$detA = \sum_{\sigma \in S_n} sgn(\sigma) a_{1j_1} a_{2j_2} \dots a_{nj_n},$$

donde S_n es el conjunto de todas las permutaciones del conjunto $\{1, 2, \dots, n\}$.

Definición 1.4.4 *Si A es una matriz de tamaño nxn, entonces se dice que $detA$ es un determinante de orden n.*

Ejemplo 1.4.2 *Sea A la matriz general de tramaño 3x3, hallar $detA$.*

Solución: Considerando las tablas 1.2 y 1.3 se tiene,

Tabla 1.3: Productos elementales con signo para una matriz de 3x3

Permutación	paridad	Producto elemental	Producto elemental con signo
123	par	$a_{11}a_{22}a_{33}$	$a_{11}a_{22}a_{33}$
132	impar	$a_{11}a_{23}a_{32}$	$-a_{11}a_{23}a_{32}$
213	impar	$a_{12}a_{21}a_{33}$	$-a_{12}a_{21}a_{33}$
231	par	$a_{12}a_{23}a_{31}$	$a_{12}a_{23}a_{31}$
312	par	$a_{13}a_{21}a_{32}$	$a_{13}a_{21}a_{32}$
321	impar	$a_{13}a_{22}a_{31}$	$-a_{13}a_{22}a_{31}$

Por lo anterior,

$$\begin{aligned} detA &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + \\ & \quad a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned} \tag{1.57}$$

Ejemplo 1.4.3 *Hallar $detA$, donde A es la matriz general de tamaño 2x2,*

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Tabla 1.4: Productos elementales con signo para una matriz de 2x2

Permutación	paridad	Producto elemental	Producto elemental con signo
12	par	$a_{11}a_{22}$	$a_{11}a_{22}$
21	impar	$a_{12}a_{21}$	$-a_{12}a_{21}$

Solución: En la siguiente tabla se muestran los productos elementales con signo,

De la tabla se tiene

$$\begin{aligned} detA &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \\ &= a_{11}a_{22} - a_{12}a_{21} \end{aligned} \tag{1.58}$$

1.4.1. Propiedades y cálculo de determinantes

Teorema 1.4.1 *Si A es una matriz cuadrada, entonces $detA = detA^T$.*

Demostración: Considérese una matriz cuadrada A con elementos a_{ij}; es decir,

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

de modo que los elementos de su transpuesta son,

$$[A^T]_{ij} = b_{ij} = a_{ji}.$$

Además, el determinante de la matriz A^T es

$$\begin{aligned} |A^T| &= \sum_{\sigma \in S_n} sgn(\sigma) b_{1j_1} b_{2j_2} \dots b_{nj_n}, \\ &= \sum_{\sigma \in S_n} sgn(\sigma) a_{j_1 1} a_{j_2 2} \dots a_{j_n n}. \end{aligned}$$

Sea τ la permutación inversa de σ, entonces de acuerdo con el teorema A.1.11 (ver el ap[endice A), $sgn(\sigma) = sgn(\tau)$, además

$$a_{j_1 1} a_{j_2 2} \dots a_{j_n n} = a_{1\tau(1)} a_{2\tau(2)} \dots a_{n\tau(n)}.$$

Por lo anterior,

$$|A^T| = \sum_{\sigma \in S_n} sgn(\tau) a_{1\tau(1)} a_{2\tau(2)} \dots a_{n\tau(n)}.$$

Como σ corre sobre todas la permutaciones de S_n, entonces τ también correrá sobre todas las permutaciones de S_n, así que $|A^T| = |A|$. ∎

Teorema 1.4.2 *Si B es la matriz que se obtiene al intercambiar dos renglones o dos columnas de A, entonces $detB = -detA$.*

Demostración: Sea B la matriz que se obtiene al intercambiar dos columnas de la matriz A y sea τ la transposición que intercambia los dos índices; si además, los a_{ij} son las entradas de la matriz A, entonces $b_{ij} = a_{i\tau(j)}$. Por lo anterior tenemos que,

$$a_{1\sigma(1)}a_{2\sigma(2)}\ldots a_{n\sigma(n)} = b_{1\tau\sigma(1)}b_{2\tau\sigma(2)}\ldots b_{n\tau\sigma(n)},$$

por los teoremas A.1.10 y A.1.5,(ver el ap[endice A), se tiene

$$\begin{aligned} sgn(\tau\circ\sigma) &= sgn(\tau)sgn(\sigma), \\ &= -sgn(\sigma). \end{aligned}$$

Por lo tanto,

$$\begin{aligned} |A| &= \sum_{\sigma\in S_n} sgn(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\ldots a_{n\sigma(n)}, \\ &= -\sum_{\sigma\in S_n} sgn(\tau\circ\sigma)b_{1\tau\sigma(1)}b_{2\tau\sigma(2)}\ldots b_{n\tau\sigma(n)}. \end{aligned}$$

Como σ corre sobre todas las permutaciones de S_n, entonces $\tau\circ\sigma$ también corre sobre todas las permutaciones de S_n(teorema A.1.6)(ver el ap[endice A). Con lo cual se concluye $|A| = -|B|$. ∎

Teorema 1.4.3 *Si A es una matriz cualquiera que contiene un renglón de ceros o una columna de ceros, entonces*

$$detA = 0.$$

Demostración: Cada producto elemental tomado de A debe contener un factor (o elemento) que provenga del renglón de ceros, por lo que el producto en sí mismo vale cero, luego

$$detA = \sum_{\sigma\in S_n} sgn(\sigma)a_{1j_1}a_{2j_2}\ldots a_{nj_n} = 0. \quad ∎$$

Definición 1.4.5 (Matrices triangulares) *Sea A una matriz cuadrada de nxn con entradas $[A]_{ij} = a_{ij}$, entonces se dice que A es*

a) triangular superior, si $a_{ij} = 0, \quad \forall i > j$,

b) triangular inferior, si $a_{ij} = 0, \quad \forall i < j$,

c) triangular, si es triangular superior o inferior.

Las matrices triangulares tienen los siguientes aspectos

a) Matriz triangular superior.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix}. \tag{1.59}$$

b) Matriz triangular inferior.

$$\begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}. \tag{1.60}$$

Teorema 1.4.4 *Si A es una matriz triangular de tamaño nxn, entonces el determinante de A es el producto de los elementos de su diagonal principal; es decir,*

$$detA = a_{11}a_{22}\dots a_{nn}.$$

Demostración: Considérese una matriz triangular inferior, en este caso, probaremos que el único producto elemental diferente de cero es justamente el que surge de la diagonal. Se sabe que los productos elementales pueden expresarse de manera general como,

$$a_{1j_1}a_{2j_2}a_{3j_3}\dots a_{nj_n}.$$

El único elemento del primer renglón diferente cero es a_{11}, por lo tanto, j_1 solo puede valer uno. Así que los únicos productos elementales diferentes de cero son de la forma,

$$a_{11}a_{2j_2}a_{3j_3}\dots a_{nj_n}.$$

Para el renglón 2 tenemos dos elementos diferentes de cero $\{a_{21}, a_{22}\}$, pero como a_{12} está en la misma columna que a_{11}, entonces el único elemento a considerar es a_{22}, es decir, j_2 solo puede valer 2, de este modo, los únicos productos elementales diferentes de cero son de la forma,

$$a_{11}a_{22}a_{3j_3}\ldots a_{nj_n}.$$

Para el tercer renglón los únicos elementos diferentes de cero son $\{a_{31}, a_{32}, a_{33}\}$. Sin embargo ya no es posible considerar elementos de la primera y la segunda columna, por lo tanto, el único elemento admisible en el producto elemental es a_{33}; es decir, $j_3 = 3$. Continuando con el proceso se concluye que el único producto elemental diferente de cero es $a_{11}a_{22}a_{33}\ldots a_{nn}$. Este producto elemental corresponde a la permutación $j_1j_2\ldots j_n = 12\ldots n$, la cual constituye una permutación par, por lo tanto,

$$|A| = a_{11}a_{22}a_{33}\ldots a_{nn}.$$

Por último, si A es una matriz trangular superior, entonces por el teorema 1.4.1, $|A| = |A^T|$, donde A^T es una matriz triangular inferior cuyos elementos de la diagonal son los mismo que los de la matriz A, por lo tanto,

$$|A| = |A^T| = a_{11}a_{22}a_{33}\ldots a_{nn}. \quad \blacksquare$$

Teorema 1.4.5 *Si una matriz A de tamaño nxn contiene dos renglones(o dos columnas) idénticas, entonces* $detA = 0$.

Demostración: Supóngase que los renglones i y j de la matriz A son idénticos. Ahora, si B es la matriz que se obtiene al intercambiar estos renglones, por un lado se tiene, $B = A$, luego, $|B| = |A|$, además, por el teorema 1.4.2 , $|B| = -|A|$. De lo anterior concluimos que, $|A| = -|A|$. Esta igualdad solo puede satisfacerce cuando $|A| = 0$. $\blacksquare$

Teorema 1.4.6 (Operaciones elementales sobre los determinantes) *Sea A una matriz de nxn,*

a) Si B es la matriz que se obtiene cuando un renglón de la matriz A se multiplica por una constante $k \neq 0$*, entonces* $detB = kdetA$*,*

b) Si B es la matriz que se obtiene al intencambiar dos renglones de la matriz A, entonces $detB = -detA$*,*

c) Si B es la matriz que se obtiene al sumar un múltiplo de uno de los renglones de A a otro renglón, entonces $detB = detA$

Demostración: a) Supóngase que el renglón i de la matriz A se multiplica por $k \neq 0$, entonces

$$\begin{aligned}
detB &= \sum_{\sigma \in S_n} sgn(\sigma) b_{1j_1} b_{2j_2} \dots b_{ij_i} \dots b_{nj_n}, \\
&= \sum_{\sigma \in S_n} sgn(\sigma) a_{1j_1} a_{2j_2} \dots ka_{ij_i} \dots a_{nj_n}, \\
&= k \sum_{\sigma \in S_n} sgn(\sigma) a_{1j_1} a_{2j_2} \dots a_{ij_i} \dots a_{nj_n}, \\
&= kdetA.
\end{aligned}$$

b) Se demostró en el teorema 1.4.2

c) Supongamos que B se obtiene de A al sumar k veces el renglón l al renglón m (con $l < m$), entonces

$$\begin{aligned}
detB &= \sum_{\sigma \in S_n} sig(\sigma) b_{1j_1} \dots b_{lj_l} \dots b_{mj_m} \dots b_{nj_n}, \\
&= \sum_{\sigma \in S_n} sig(\sigma) a_{1j_1} \dots a_{lj_l} \dots (ka_{lj_l} + a_{mj_m}) \dots a_{nj_n}, \\
&= k \sum_{\sigma \in S_n} sig(\sigma) a_{1j_1} \dots a_{lj_l} \dots a_{lj_l} \dots a_{nj_n} \\
&\quad + \sum_{\sigma \in S_n} sig(\sigma) a_{1j_1} \dots a_{lj_l} \dots a_{mj_m} \dots a_{nj_n}.
\end{aligned}$$

La primera sumatoria corresponde al determinante de una matriz cuyos renglones l y m son idénticos, así que por el teorema 1.4.5, este determinante vale cero. Finalmente, la segunda suma corresponde al determinante de la matriz A, por lo tanto, $detB = detA$. ∎

El teorema anterior proporciona una herramienta poderosa para reducir el determinante de una matriz cualquiera al determinante de una matriz triangular. Antes de mostrar un ejemplo concreto considérese el ejemplo de una matriz general de 2x2, para ilustrar el uso de las operaciones elementales.

Para indicar las operaciones sobre el determinante recurriremos a la notación descrita en la sección 1.1.4, con la variante que, cuando las operaciones se realicen sobre las columnas, cambiaremos r por c.

i) Multiplicación del renglón 1 por una constante $k \neq 0$.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \underset{kr_1 \longrightarrow r_1}{=} \frac{1}{k} \begin{vmatrix} ka_{11} & ka_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Aquí es necesario hacer la siguiente aclaración: si originalmente se tiene una matriz en la cual uno de sus renglones es múltiplo de una constante k, como la matriz del lado derecho, entonces este factor puede simplemente sacarse como un factor de todo el determinante; sin embargo, la operación sobre el renglón sería $(\frac{1}{k})r_1 \longrightarrow r_1$; es decir,

$$\begin{vmatrix} ka_{11} & ka_{12} \\ a_{21} & a_{22} \end{vmatrix} \underset{(\frac{1}{k})r_1 \longrightarrow r_1}{=} k \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

ii) Intercambio de renglones.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \underset{r_1 \longrightarrow r_2}{=} - \begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix}.$$

iii) Se suma un múltiplo de un renglón a otro.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \underset{kr_2 + r_1 \longrightarrow r_1}{=} \begin{vmatrix} ka_{21} + a_{11} & ka_{22} + a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Ejemplo 1.4.4 *Hallar el determinante de la siguiente matriz realizando operaciones elementales hasta obtener el determinante de una matriz triangular.*

$$A = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix}.$$

Solución:

1. Sume (-2) veces el renglón 2 al renglón 1.

$$|A| = \begin{vmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{vmatrix} \underset{(-2)r_2 + r_1 \longrightarrow r_1}{=} \begin{vmatrix} 0 & -5 & 10 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{vmatrix}$$

2. Sume (-5) veces el renglón 2 al renglón 3

$$\begin{vmatrix} 0 & -5 & 10 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{vmatrix} \underset{(-5)r_2 + r_3 \longrightarrow r_3}{=} \begin{vmatrix} 0 & -5 & 10 \\ 1 & 2 & -3 \\ 0 & -7 & 16 \end{vmatrix}$$

3. Intercambie los renglones 1 y 2

$$\begin{vmatrix} 0 & -5 & 10 \\ 1 & 2 & -3 \\ 0 & -7 & 16 \end{vmatrix} \underset{r_2 \longleftrightarrow r_3}{=} -\begin{vmatrix} 1 & 2 & -3 \\ 0 & -5 & 10 \\ 0 & -7 & 16 \end{vmatrix}$$

4. Factorice (-5) del renglón 2.

$$-\begin{vmatrix} 1 & 2 & -3 \\ 0 & -5 & 10 \\ 0 & -7 & 16 \end{vmatrix} \underset{(-\frac{1}{5})r_2 \longrightarrow r_2}{=} -(-5)\begin{vmatrix} 1 & 2 & -3 \\ 0 & 1 & -2 \\ 0 & -7 & 16 \end{vmatrix}$$

5. Sume 7 veces el renglón 2 al renglón 3.

$$-(-5)\begin{vmatrix} 1 & 2 & -3 \\ 0 & 1 & -2 \\ 0 & -7 & 16 \end{vmatrix} \underset{(-7)r_2 + r_3 \longrightarrow r_3}{=} 5\begin{vmatrix} 1 & 2 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{vmatrix}.$$

6. El último determinante corresponde a una matriz triangular, cuyo determinante es el producto de los elementos de la diagonal (teorema 1.4.4), por lo tanto, $|A| = 5(2) = 10$.

Teorema 1.4.7 *Sean $E_i(k)$, E_{ij} y $E_{ij}(k)$ las matrices elementales definidas en el teorema 1.3.4, entonces*

$$|E_i(k)| = k, \qquad |E_{ij}| = -1, \qquad |E_{ij}(k)| = 1.$$

Demostración: i) $E_i(k)$ es la matriz que se obtiene al multiplicar el i-ésimo renglón de la matriz identidad I_n por k, por el teorema anterior,

$$|E_i(k)| = k|I_n|.$$

Ahora, I_n es una matriz triangular cuyos elementos de la diagonal son todos iguales a 1, por el teorema 1.4.4, $|I_n| = 1$, luego

$$|E_i(k)| = k|I_n| = k.$$

ii) E_{ij} es la es la matriz que se obtiene al intercambiar los renglones i y j de la matriz identidad I_n, por el teorema anterior,

$$|E_{ij}| = -|I_n| = -1.$$

iii) $E_{ij}(k)$ es la matriz que se obtiene al sumar k veces el renglón i al renglón j en la matriz I_n, por el teorema anterior,

$$|E_{ij}(k)| = |I_n| = 1. \quad \blacksquare$$

Teorema 1.4.8 *Si E y A son matrices de tamaño nxn, donde E es una matriz elemental, entonces*

$$|EA| = |E||A|.$$

Demostración: La matriz elemental E será considerada por casos.

i) Si $E = E_i(k)$, entonces por el teorema 1.3.4, EA es la matriz que se obtiene al multiplicar el renglón i de la matriz A por k, de manera que, por el teorema 1.4.6, $|EA| = k|A|$, pero $|E| = k$, por lo tanto, $|EA| = |E||A|$.

ii) Si $E = E_{ij}$, entonces por el teorema 1.3.4, EA es la matriz que se obtiene al intercambiar los renglones i y j en la matriz A; por el teorema 1.4.6, tenemos que, $|EA| = -|A|$, además, $|E| = -1$, de manera que, $|EA| = |E||A|$.

iii) Finalmente, si $E = E_{ij}(k)$, entonces por el teorema 1.3.4, EA es la matriz que se obtiene al sumar k veces el renglón i al renglón j en la matriz A, por el teorema 1.4.6, $|EA| = |A|$, como $|E| = 1$, entonces $|EA| = |E||A|$. $\blacksquare$

Puede probarse por inducción un resultado más general: Si E_1,E_2,...,E_k son matrices elementales de tamaño nxn y A es una matriz cuadrada de tamaño nxn, entonces

$$|E_k \ldots E_2E_1A| = |E_k| \ldots |E_2||E_1||A|.$$

La demostración se deja como ejercicio para el estudiante.

Teorema 1.4.9 *Si A es equivalente respecto a los renglones con B ($A \sim B$); es decir, $E_k \ldots E_2E_1A = B$, entonces*

a) $|B| = |E_k| \ldots |E_2||E_1||A|$,

b) $|B| \neq 0 \iff |A| \neq 0$.

Demostración: a) Como $B = E_k \ldots E_2E_1A$, entonces por el teorema anterior,

$$|B| = |E_k \ldots E_2E_1A| = |E_k| \ldots |E_2||E_1||A|.$$

b)Como las matrices elementales tienen determinantes diferentes de cero, entonces

$$|B| \neq 0 \iff |A| \neq 0. \quad \blacksquare$$

Teorema 1.4.10 *A es invertible* $\Longleftrightarrow |A| \neq 0$.

Demostración: Por el teorema 1.3.8 A es invertible si y solo si A es equivalente respecto a los renglones a la identidad; es decir, existen matrices elementales $E_1, E_2, \ldots, E_k$ tales que,

$$E_k \ldots E_2 E_1 A = I_n,$$

de modo que

$$|E_k| \ldots |E_2||E_1||A| = |I_n| = 1 \neq 0,$$

como las matrices elementales son invertibles, entonces

$$|E_k| \ldots |E_2||E_1| \neq 0,$$

finalmente,

$$|A| \neq 0. \quad \blacksquare$$

Teorema 1.4.11 *Si A y B son matrices cuadradas del mismo tamaño, entonces* $|AB| = |A||B|$.

Demostración: Supóngase que A es invertible, entonces A es equivalente respecto a los renglones a la identidad, de manera que existen matrices elementales $\{E_1, E_2, \ldots, E_k\}$, tales que,

$$A = E_k \ldots E_2 E_1 I_n = E_k \ldots E_2 E_1,$$

de manera que,

$$|A| = |E_k| \ldots |E_2||E_1|.$$

Por otra parte,

$$AB = E_k \ldots E_2 E_1 B,$$

así que el determinante AB se expresa de la siguiente manera,

$$\begin{aligned} |AB| &= |E_k \ldots E_2 E_1 B|, \\ &= |E_k| \ldots |E_2||E_1||B|, \\ &= |A||B|. \end{aligned}$$

Si A no es invertible, entonces $|A| = 0$, de manera que $|A||B| = 0$. Por otro lado, la negación del teorema 1.3.2 implica que si A o B son matrices no invertibles; entonces, AB es una matriz no invertible; es decir, $|AB| = 0$; lo tanto, $|A||B| = 0 = |AB|$, de manera que incluso en este caso se cumple que $|AB| = |A||B|$. $\blacksquare$

Corolario 1.4.1 *Si A es invertible, entonces* $det(A^{-1}) = \dfrac{1}{detA}$.

Demostración: Si A es invertible, entonces

$$AA^{-1} = I,$$

por el teorema anterior,

$$\begin{aligned} |AA^{-1}| &= |I|, \\ |A||A^{-1}| &= 1, \\ |A^{-1}| &= \frac{1}{|A|}. \qquad \blacksquare \end{aligned}$$

1.4.2. Desarrollo por cofactores

Definición 1.4.6 (Menor del elemento a_{ij}) *Sea A una matriz cuadrada, entonces el menor del elemento a_{ij}, denotado por M_{ij}, se define como el determinante de la submatriz que resulta al suprimir el renglón i y la columna j en la matriz A.*

Ejemplo 1.4.5 *Dada la matriz,* $A = \begin{bmatrix} 3 & 2 & 1 \\ -1 & 0 & 4 \\ 2 & 3 & -5 \end{bmatrix}$, *determine los menores* M_{22}, M_{32}

Solución:

1. M_{22} es el determinante de la submatriz que se obtiene al eliminar el renglón 2 y la columna 2 en la matriz A.

$$M_{22} = \left| \begin{array}{r|rr} 3 & 2 & 1 \\ -1 & 0 & 4 \\ \hline 2 & 3 & -5 \end{array} \right| = \left| \begin{array}{rr} 3 & 1 \\ 2 & -5 \end{array} \right| = -17.$$

2. M_{32} es el determinante de la submatriz que se obtiene al eliminar el renglón 3 y la columna 2 en la matriz A.

$$M_{32} = \left| \begin{array}{r|rr} 3 & 2 & 1 \\ -1 & 0 & 4 \\ 2 & 3 & -5 \\ \hline \end{array} \right| = \left| \begin{array}{rr} 3 & 1 \\ -1 & 4 \end{array} \right| = 13.$$

Ejemplo 1.4.6 *Considérense los determinantes de orden 5 de las matrices A y A', con elementos correspondientes $[A]_{ij} = a_{ij}$ y $[A']_{ij} = a'_{ij}$, como se muestra a continuación.*

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix},$$

$$|A^*| = \begin{vmatrix} a_{11} & a_{13} & a_{14} & a_{15} & a_{12} \\ a_{21} & a_{23} & a_{24} & a_{25} & a_{22} \\ a_{41} & a_{43} & a_{44} & a_{45} & a_{42} \\ a_{51} & a_{53} & a_{54} & a_{55} & a_{52} \\ a_{31} & a_{33} & a_{34} & a_{35} & a_{32} \end{vmatrix}.$$

Obsérvese que $a'_{55} = a_{32}$. De hecho, la matriz A' se obtiene de la matriz A mediante el intercambio de renglones consecutivos hasta llevar el renglón 3 hasta último renglón, después mediante intercambios consecutivos de columnas, se lleva la columna 2 hasta la última columna. Ahora sean M_{32} y M'_{55} los menores correspondientes a los elementos a_{32} y a'_{55} de las matrices A y A', respectivamente, entonces

i) Determine la relación entre los menores M_{32} y M'_{33}.

ii) Determine la relación existente entre $|A|$ y $|A'|$.

Solución:

i) Eliminando el renglón 3 y la columna 3 en el determinante de $|A'|$ se obtiene,

$$M'_{55} = \begin{vmatrix} a_{11} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{23} & a_{24} & a_{25} \\ a_{41} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{53} & a_{54} & a_{55} \end{vmatrix} = M_{32}.$$

ii) Ahora mediante el intercambio de renglones y columnas podemos relacionar los determinantes de las matrices $|A|$ y $|A'|$ como se muestra a continuación.

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix}$$

$$
\begin{aligned}
&= (-1)\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix} \\
&= (-1)(-1)\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{vmatrix} \\
&= (-1)^2(-1)\begin{vmatrix} a_{11} & a_{13} & a_{12} & a_{14} & a_{15} \\ a_{21} & a_{23} & a_{22} & a_{24} & a_{25} \\ a_{41} & a_{43} & a_{42} & a_{44} & a_{45} \\ a_{51} & a_{53} & a_{52} & a_{54} & a_{55} \\ a_{31} & a_{33} & a_{32} & a_{34} & a_{35} \end{vmatrix} \\
&= (-1)^2(-1)(-1)\begin{vmatrix} a_{11} & a_{13} & a_{14} & a_{12} & a_{15} \\ a_{21} & a_{23} & a_{24} & a_{22} & a_{25} \\ a_{41} & a_{43} & a_{44} & a_{42} & a_{45} \\ a_{51} & a_{53} & a_{55} & a_{52} & a_{55} \\ a_{31} & a_{33} & a_{34} & a_{32} & a_{35} \end{vmatrix} \\
&= (-1)^2(-1)(-1)(-1)\begin{vmatrix} a_{11} & a_{13} & a_{14} & a_{15} & a_{12} \\ a_{21} & a_{23} & a_{24} & a_{25} & a_{22} \\ a_{41} & a_{43} & a_{44} & a_{45} & a_{42} \\ a_{51} & a_{53} & a_{55} & a_{55} & a_{52} \\ a_{31} & a_{33} & a_{34} & a_{35} & a_{32} \end{vmatrix} \\
&= (-1)^2(-1)^3|A'| \\
&= (-1)^{2+3}|A'|.
\end{aligned}
$$

En general, para una matriz A de orden nxn en la cual se desplaza el elemento a_{ij} como se ha indicado para obtener la matriz A^* se tienen $(n-i)$ intercambios de renglones y $(n-j)$ intercambios de columnas, de lo que resulta,

$$M_{ij} = M'_{nn} \tag{1.61}$$

y

$$\begin{aligned}
|A| &= (-1)^{(n-i)+(n-j)}|A'|, \\
&= (-1)^{(2n-i-j)}|A'|, \\
&= (-1)^{-i-j}|A'|, \\
&= (-1)^{i+j}|A'|. \qquad (1.62)
\end{aligned}$$

Definición 1.4.7 (Cofactor del elemento a_{ij}) *Dada una matriz cuadrada A con entradas a_{ij}, se define el cofactor del elemento a_{ij}, denotado por c_{ij} como el número,*

$$c_{ij} = (-1)^{i+j} M_{ij}. \qquad (1.63)$$

Es claro que si $i+j$ es impar, entonces las cantidades c_{ij} y M_{ij} se diferencian en el signo.

Definición 1.4.8 (Matriz de cofactores) *Si A es una matriz cuadrada con entradas a_{ij} y cofactores c_{ij}, entonces la matriz cuadrada C con entradas c_{ij} se conoce como la matriz de cofactores.*

Ejemplo 1.4.7 *Hallar matriz de cofactores para de matriz A, donde*

$$A = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix}.$$

Solución: Determínense las entradas de la matriz de cofactores.

$$\begin{aligned}
c_{11} &= (-1)^2 M_{11} = \begin{vmatrix} 2 & -3 \\ 3 & 1 \end{vmatrix} = 11 \\
c_{12} &= (-1)^3 M_{12} = -\begin{vmatrix} 1 & -3 \\ 5 & 1 \end{vmatrix} = -16 \\
c_{13} &= (-1)^4 M_{13} = \begin{vmatrix} 1 & 2 \\ 5 & 3 \end{vmatrix} = -7 \\
c_{21} &= (-1)^3 M_{21} = -\begin{vmatrix} -1 & 4 \\ 3 & 1 \end{vmatrix} = 13 \\
c_{22} &= (-1)^4 M_{22} = \begin{vmatrix} 2 & 4 \\ 5 & 1 \end{vmatrix} = -18 \\
c_{23} &= (-1)^5 M_{23} = -\begin{vmatrix} 2 & -1 \\ 5 & 3 \end{vmatrix} = -11 \\
c_{31} &= (-1)^4 M_{31} = \begin{vmatrix} -1 & 4 \\ 2 & -3 \end{vmatrix} = -5 \\
c_{32} &= (-1)^5 M_{32} = -\begin{vmatrix} 2 & 4 \\ 1 & -3 \end{vmatrix} = 10 \\
c_{33} &= (-1)^6 M_{33} = \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} = 5
\end{aligned}$$

Finalmente, con las entradas anteriores se construye la matriz de cofactores,

$$C = \begin{bmatrix} 11 & -16 & -7 \\ 13 & -18 & -11 \\ -5 & 10 & 5 \end{bmatrix}.$$

Teorema 1.4.12 *Dada una matriz A de tamaño nxn es posible calcular el determinante de A multiplicando los elementos de cualquier renglón (o columna) por sus cofactores y sumando los productos; es decir, para cada*

$$1 \leq i \leq n \qquad y \qquad 1 \leq j \leq n,$$

a) Desarrollo por cofactores a lo largo del i-ésimo renglón.

$$detA = \sum_{k=1}^{n} a_{ik} c_{ik}. \tag{1.64}$$

b) Desarrollo por cofactores a lo largo de la j-ésima columna.

$$detA = \sum_{k=1}^{n} a_{kj}c_{kj}. \tag{1.65}$$

Demostración: Se sabe que el determinante de la matriz A se expresa de manera general como,

$$detA = \sum_{\sigma \in S_n} (sgn\sigma) a_{1j_1} a_{2j_2} \ldots a_{nj_n}.$$

Considérense los elementos del renglón i-ésimo, $\{a_{i1}, a_{i2}, \ldots a_{1n}\}$. Cada producto elemental del determinante de A contiene solo un elemento del renglón i-ésimo, por lo tanto, es posible expresar el determinante de A como,

$$|A| = a_{i1}p_{i1} + a_{i2}p_{i2} + \ldots + a_{in}p_{in}, \tag{1.66}$$

donde el factor p_{ij} es una sumatoria de productos, ninguno de los cuales contiene factores que provengan del renglón i ni de la columna j. Si se puede probar que estos p_{ij} corresponden a los cofactores correspondientes a cada elemento, entonces se habrá concluido la demostración.

Para comenzar se prueba una propiedad de los factores p_{ij} aquí definidos.

Considérese la matriz A' con entradas a'_{ij}, tal que A' se obtiene de A al reacomodar el elemento a_{ij} mediante intercambios de renglones sucesivos y posteriormente intercambiando columnas sucesivas, hasta llevarlo a la posición "nn", como en el ejemplo 1.4.6. Es importante notar que,

$$a'_{nn} = a_{ij}. \tag{1.67}$$

Los mismos argumentos que permitieron expresar el determinante de la matriz A mediante la ecuación 1.66, permiten expresar el determinante de la matriz A' como,

$$|A'| = a'_{n1}p'_{n1} + a'_{n2}p'_{n2} + \ldots + a'_{nn}p'_{nn}. \tag{1.68}$$

De la ecuación 1.62,

$$|A| = (-1)^{i+j}|A'|. \tag{1.69}$$

Sustituyendo las ecuaciones 1.66, 1.67 y 1.68 en la ecuación anterior se obtiene,

$$\begin{aligned}
|A| &= (-1)^{i+j}|A'|,\\
&= (-1)^{i+j}(a'_{n1}p'_{n1} + a'_{n2}p'_{n2} + \ldots + a'_{nn}p'_{nn}),\\
\sum_{k\neq j}^{n} a_{ik}A_{ik} + a_{ij}A_{ij} &= (-1)^{i+j}\sum_{k=1}^{n-1} a'_{nk}p'_{nk} + a_{ij}((-1)^{i+j}p'_{nn}). \qquad (1.70)
\end{aligned}$$

Los únicos términos que contienen el factor a_{ij} son los que se han dejado explícitamente en la ecuación 1.70, de donde se concluye

$$p_{ij} = (-1)^{i+j}p'_{nn}. \qquad (1.71)$$

Ahora se prueba un caso particular para el cual p_{ij} es el cofactor del elemento a_{ij}. Considérese el término $a_{nn}p_{nn}$, es decir, $i = n$ y $j = n$. Este término corresponde a las permutaciones de S_n para las cuales $\sigma(n) = n$, de manera que,

$$a_{nn}p_{nn} = a_{nn} \sum_{\{\sigma\in S_n:\sigma(n)=n\}} (sgn\sigma)a_{1j_1}a_{2j_2}\ldots a_{(n-1)j_{(n-1)}}$$

Como se demostró en el teorema A.1.8(ver el apéndice A), $sgn(\sigma) = sgn(\sigma^*)$. Además, cuando σ corre sobre todas las permutaciones de S_n para las cuales $\sigma(n) = n$, entonces σ^* corre sobre todas las permutaciones de S_{n-1}, por lo tanto,

$$a_{nn}p_{nn} = a_{nn} \sum_{\sigma^*\in S_{(n-1)}} (sgn\sigma^*)a_{1j_1}a_{2j_2}\ldots a_{(n-1)j_{(n-1)}}$$

La sumatoria corresponde al determinante de la submatriz que se obtiene al eliminar el renglón n y la columna n de la matriz A; es decir,

$$p_{nn} = M_{nn} = (-1)^{n+n}M_{nn} = c_{nn}. \qquad (1.72)$$

Donde c_{nn} es el cofactor correspondiente al elemento a_{nn}.

Ahora considérese el elemento $a_{ij}p_{ij}$. En este caso se recurre al procedimiento mostrado en el ejemplo 1.4.6. Es posible reacomodar el elemento a_{ij} mediante intercambios de renglones sucesivos y posteriormente intercambiando columnas sucesivas, hasta llevarlo a la posición "nn".

De la ecuación 1.72,

$$p'_{nn} = c'_{nn} = M'_{nn}. \qquad (1.73)$$

De la ecuación 1.61,

$$p'_{nn} = M'_{nn} = M_{ij}. \qquad (1.74)$$

De la ecuación 1.71

$$p'_{nn} = M_{ij} = (-1)^{i+j} p_{ij}. \tag{1.75}$$

Finalmente se concluye,

$$p_{ij} = (-1)^{i+j} M_{ij}. \quad \blacksquare \tag{1.76}$$

Ejemplo 1.4.8 *Calcule el determinante de la siguiente matriz desarrollando por el renglón 1 y por la columna 3.*

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

Solución:

1. Desarrollando por el renglón 1 tenemos:

$$\begin{aligned} |A| &= a_{11}c_{11} + a_{12}c_{12} + a_{13}c_{13} \\ &= 3c_{11} - c_{12} - 2c_{13} \\ &= 3M_{11} + M_{12} - 2M_{12} \\ &= 3\begin{vmatrix} 4 & 0 \\ 3 & 0 \end{vmatrix} + \begin{vmatrix} 4 & 0 \\ 5 & 0 \end{vmatrix} - 2\begin{vmatrix} 4 & 3 \\ 5 & 4 \end{vmatrix} \\ &= -2. \end{aligned}$$

2. Desarrollemos por la columna 3.

$$\begin{aligned} |A| &= a_{13}c_{13} + a_{23}c_{23} + a_{33}c_{33} \\ &= -2c_{13} + 0c_{23} + 0c_{33} \\ &= -2M_{13} \\ &= -2\begin{vmatrix} 4 & 3 \\ 5 & 4 \end{vmatrix} = -2. \end{aligned}$$

El desarrollo por cofactores representa una herramienta valiosa para calcular el determinante de una matriz. El hecho que se pueda elegir cualquier renglón o cualquier columna para hacer nuestro desarrollo, concede una gran ventaja, ya que para desarrollar el determinante, es posible escoger aquel renglón o aquella columna que posea un mayor número de elementos iguales a cero (como se puede observar en el ejemplo anterior). El método de cofactores puede complementarse con las operaciones elementales de tal manera que se puedan colocar ceros en un renglón determinado y de este modo reducir el orden del determinante.

Ejemplo 1.4.9 *Calcular el siguiente determinante combinando operaciones elementales y el desarrollo por cofactores.*

$$A = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix}.$$

Solución: Se realizarán operaciones elementales sobre el determinante hasta reducirlo a una forma simple (de un determinante de 2x2). Para comenzar se usa el 1 de la primer columna, como pivote para hacer cero el resto de los elemento de su columna.

1. Súmese -2 veces el segundo renglón al primero.

$$|A| = \begin{vmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{vmatrix} \underset{-2r_2+r_1 \longrightarrow r_1}{=} \begin{vmatrix} 0 & -5 & 10 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{vmatrix}$$

2. Súmese -5 veces el segundo renglón al tercero.

$$= \begin{vmatrix} 0 & -5 & 10 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{vmatrix} \underset{-5r_2+r_3 \longrightarrow r_3}{=} \begin{vmatrix} 0 & -5 & 10 \\ 1 & 2 & -3 \\ 0 & -7 & 16 \end{vmatrix}$$

3. Use el desarrollo por cofactores mediante la primera columna.

$$= -\begin{vmatrix} -5 & 10 \\ -7 & 16 \end{vmatrix}$$

4. Factorize el (-1) de la primer columna.

$$= -\begin{vmatrix} -5 & 10 \\ -7 & 16 \end{vmatrix} = -(-1)\begin{vmatrix} 5 & 10 \\ 7 & 16 \end{vmatrix}$$

5. Factorice el factor (5) del primer renglón

$$= \begin{vmatrix} 5 & 10 \\ 7 & 16 \end{vmatrix} = 5\begin{vmatrix} 1 & 2 \\ 7 & 16 \end{vmatrix}$$

6. Factorice el factor (2) de la segunda columna.

$$= \quad 5\begin{vmatrix} 1 & 2 \\ 7 & 16 \end{vmatrix} = 5(2)\begin{vmatrix} 1 & 1 \\ 7 & 8 \end{vmatrix} = 10.$$

Finalmente, $|A| = 10$.

Ejemplo 1.4.10 *Calcular el siguiente determinante combinando operaciones elementales y el desarrollo por cofactores.*

$$|A| = \begin{vmatrix} 1 & 2 & -1 & 4 \\ 3 & -1 & 2 & 0 \\ 4 & 5 & -2 & 1 \\ -2 & 3 & 0 & 2 \end{vmatrix}.$$

Solución: Demanera análoga al ejemplo anterior se procederá a reducir el determinante hasta obtener un determinante de tamaño 2x2. Se aprovechará el hecho de que la matriz A tiene un cero en la columna 4; es decir, se usará el 1 de la columna 4 como pivote para hacer cero los elemento restantes de dicha columna. En este ejercicio también se muestra el uso de las operaciones sobre las columnas (Recuérdese que $|A| = |A^T|$).

1. Súmese (-4) veces el renglón 3 al renglón 1.

$$|A| \quad = \begin{vmatrix} 1 & 2 & -1 & 4 \\ 3 & -1 & 2 & 0 \\ 4 & 5 & -2 & 1 \\ -2 & 3 & 0 & 2 \end{vmatrix} \underset{-4r_3+r_1 \longrightarrow r_1}{=} \begin{vmatrix} -15 & -18 & 7 & 0 \\ 3 & -1 & 2 & 0 \\ 4 & 5 & -2 & 1 \\ -2 & 3 & 0 & 2 \end{vmatrix}$$

2. Súmese (-2) veces el renglón 3 al renglón 4

$$= \begin{vmatrix} -15 & -18 & 7 & 0 \\ 3 & -1 & 2 & 0 \\ 4 & 5 & -2 & 1 \\ -2 & 3 & 0 & 2 \end{vmatrix} \underset{-2r_3+r_4 \longrightarrow r_4}{=} \begin{vmatrix} -15 & -18 & 7 & 0 \\ 3 & -1 & 2 & 0 \\ 4 & 5 & -2 & 1 \\ -10 & -7 & 4 & 0 \end{vmatrix}$$

3. Desarrolle por cofactores mediante la cuarta columna.

$$= -\begin{vmatrix} -15 & -18 & 7 \\ 3 & -1 & 2 \\ -10 & -7 & 4 \end{vmatrix}$$

4. Ahora se tiene un determinante de orden 3x3. Para continuar la reducción factorice (-1) de la columna 2.

$$= -\begin{vmatrix} -15 & -18 & 7 \\ 3 & -1 & 2 \\ -10 & -7 & 4 \end{vmatrix} = -(-1)\begin{vmatrix} -15 & 18 & 7 \\ 3 & 1 & 2 \\ -10 & 7 & 4 \end{vmatrix}$$

5. Súmese (-3)veces la columna 2 a la columna 1

$$= \begin{vmatrix} -15 & 18 & 7 \\ 3 & 1 & 2 \\ -10 & 7 & 4 \end{vmatrix} \underset{-3c_2+c_1 \longrightarrow c_1}{=} \begin{vmatrix} -69 & 18 & 7 \\ 0 & 1 & 2 \\ -31 & 7 & 4 \end{vmatrix}$$

6. Súmese (-2)veces la columna 2 a la columna 3

$$= \begin{vmatrix} -15 & 18 & 7 \\ 3 & 1 & 2 \\ -10 & 7 & 4 \end{vmatrix} \underset{-2c_2+c_3 \longrightarrow c_3}{=} \begin{vmatrix} -69 & 18 & -29 \\ 0 & 1 & 0 \\ -31 & 7 & -10 \end{vmatrix}$$

7. Desarrollando por cofactores mediante el segundo renglón se tiene,

$$= \begin{vmatrix} -69 & -29 \\ -31 & -10 \end{vmatrix}$$

8. Factorizando (-1) en cada renglón se tiene,

$$= \begin{vmatrix} 69 & 29 \\ 31 & 10 \end{vmatrix} = -209.$$

Finalmente, $|A| = -209$.

Teorema 1.4.13 *Sea A una matriz cuadrada de tamaño nxn con elementos a_{ij}, entonces*

$$\sum_{k=1}^{n} a_{ik}c_{jk} = |A|\delta_{ij}.$$

Demostración: Si $i = j$ se obtiene el desarrollo por cofactores del determinante de la matriz A mediante el renglón i. Ahora consideremos el caso en que $i \neq j$. Para comenzar constrúyase la matriz B a partir de la matriz A, como se muestra a continuación,

$$B = \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

En esencia solo se ha sustituido el renglón j-ésimo de A por otro renglón i. Como B tiene dos renglones idénticos, entonces por el teorema 1.4.5 $detB = 0$. Por otro lado, el desarrollo por cofactores del determinante de B mediante el renglón j-ésimo nos da,

$$detB = \sum_{k=1}^{n} b_{jk}c_{jk},$$

ya que los cofactores de los elementos del j-ésimo renglón de las matrices A y B son iguales. Además, como $b_{jk} = a_{ik}$, entonces

$$detB = \sum_{k=1}^{n} a_{ik}c_{jk} = 0.$$

Con lo anterior se ha probado que,

$$\sum_{k=1}^{n} a_{ik}c_{jk} = |A|\delta_{ij}. \qquad \blacksquare$$

1.4.3. La inversa de una matriz a través de su adjunta

Definición 1.4.9 (Adjunta de la matriz A) *Dada una matriz cuadrada A con entradas a_{ij}, cuya matriz de cofactores C tiene entradas c_{ij}, entonces la matriz cuadrada C^T se conoce como la adjunta de la matriz A y se denota por $adj(A)$; es decir,*

$$adj(A) = C^T,$$

de manera equivalente,

$$[adj(A)]_{ij} = [C^T]_{ij} = c_{ji}.$$

Teorema 1.4.14 *Si A es una matriz invertible, entonces*

$$A^{-1} = \frac{1}{detA} adj(A).$$

Demostración: Considérese el producto,

$$\begin{aligned} [Aadj(A)]_{ij} &= \sum_{k=1}^{n} [A]_{ik}[adj(A)]_{kj}, \\ &= \sum_{k=1}^{n} a_{ik}[C^T]_{kj}, \\ &= \sum_{k=1}^{n} a_{ik}[C]_{jk}, \\ &= \sum_{k=1}^{n} a_{ik}c_{jk}. \end{aligned}$$

Por el teorema 1.4.13

$$\begin{aligned} [Aadj(A)]_{ij} &= |A|\delta_{ij}, \\ &= |A|[I]_{ij}. \end{aligned}$$

Como A es invertible, entonces $|A| \neq 0$. Dividiendo la última ecuación entre este número se tiene,

$$\left[A\frac{adj(A)}{|A|} \right]_{ij} = [I]_{ij},$$

es decir,

$$A\left[\frac{adj(A)}{|A|}\right] = I.$$

Como A es una matriz cuadrada, entonces por el teorema 1.3.7, A es invertible y su inversa es

$$A^{-1} = \frac{adj(A)}{|A|}. \quad \blacksquare$$

.

Este teorema nos provee de un método comptutacional para determinar la inversa de una matriz.

Ejemplo 1.4.11 *Por el método de la adjunta halle la inversa de la matriz A, donde*

$$A = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix}.$$

Solución: De acuerdo con el teorema 1.4.14 se debe hallar el determinante de la matriz A y también construir su adjunta.

1. El determinante de esta matriz ha sido calculado anteriormente, (ejemplo 1.4.4), el resultado fue, $|A| = 10$.

2. La matriz de cofactores para este caso fue determinado en el ejemplo 1.4.7. La matriz obtenida fue,

$$C = \begin{bmatrix} 11 & -16 & -7 \\ 13 & -18 & -11 \\ -5 & 10 & 5 \end{bmatrix}.$$

3. Ahora constrúyase la matriz adjunta de A. Como se sabe, $Adj(A) = C^T$, luego,

$$Adj(A) = C^T = \begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix}.$$

4. Finalmente,

$$A^{-1} = \frac{Adj(A)}{|A|} = \frac{1}{10}\begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix}.$$

Este ejemplo ya había sido resuelto anteriormente mediante operaciones elementales(ver ejemplo 1.3.4).

Ejemplo 1.4.12 *Mediante la adjunta hallar la inversa de la matriz B, donde*

$$B = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta & 0 \\ -\operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solución: Se procede de la misma manera que en el ejemplo anterior.

1. Mediante el desarrollo por cofactores.

$$|B| = \begin{vmatrix} \cos\theta & \operatorname{sen}\theta & 0 \\ -\operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{vmatrix} = \cos^2\theta + \operatorname{sen}^2\theta = 1.$$

2. En seguida se determinan los elementos de la matriz de cofactores.

$$\begin{aligned}
c_{11} &= (-1)^2 M_{11} = \begin{vmatrix} \cos\theta & 0 \\ 0 & 1 \end{vmatrix} = \cos\theta \\
c_{12} &= (-1)^3 M_{12} = -\begin{vmatrix} -\operatorname{sen}\theta & 0 \\ 0 & 1 \end{vmatrix} = \operatorname{sen}\theta \\
c_{13} &= (-1)^4 M_{13} = \begin{vmatrix} -\operatorname{sen}\theta & \cos\theta \\ 0 & 0 \end{vmatrix} = 0 \\
c_{21} &= (-1)^3 M_{21} = -\begin{vmatrix} \operatorname{sen}\theta & 0 \\ 0 & 1 \end{vmatrix} = -\operatorname{sen}\theta \\
c_{22} &= (-1)^4 M_{22} = \begin{vmatrix} \cos\theta & 0 \\ 0 & 1 \end{vmatrix} = \cos\theta \\
c_{23} &= (-1)^5 M_{23} = -\begin{vmatrix} \cos\theta & \operatorname{sen}\theta \\ 0 & 0 \end{vmatrix} = 0 \\
c_{31} &= (-1)^4 M_{31} = \begin{vmatrix} \operatorname{sen}\theta & 0 \\ \cos\theta & 0 \end{vmatrix} = 0 \\
c_{32} &= (-1)^5 M_{32} = -\begin{vmatrix} \cos\theta & 0 \\ -\operatorname{sen}\theta & 0 \end{vmatrix} = 0 \\
c_{33} &= (-1)^6 M_{33} = \begin{vmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{vmatrix} = 1
\end{aligned}$$

3. Con los coeficientes se construye la matriz de cofactores

$$C = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta & 0 \\ -\operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

4. La matriz adjunta de B es

$$Adj(B) = C^T = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

5. Finalmente,

$$B^{-1} = \frac{Adj(B)}{|B|} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Obsérvese que en este caso la inversa de la matriz B es simplemente su transpuesta.

1.4.4. Regla de Cramer

Teorema 1.4.15 (Regla de Cramer) *Si $Ax = b$ es un sistema de ecuaciones lineales en n incógnitas tal que, $detA \neq 0$, entonces el sistema tiene una solución única. Esta solución es*

$$x_j = \frac{detA_j}{detA}. \qquad 1 \leq j \leq n. \tag{1.77}$$

donde A_j es la matriz que se obtiene al reemplazar los elementos de la j-ésima columna de A por los elementos de la matriz b.

Demostración: Si $detA \neq 0$, entonces el sistema de ecuaciones $Ax = b$ tiene solución única, $x = A^{-1}b$, además, por el teorema anterior,

$$x = \frac{adj(A)}{detA}b.$$

Se sabe x y b son vectores columna; es decir, son matrices de tamaño $nx1$,

$$
\begin{aligned}
x_{j1} &= \frac{1}{detA}[adj(A)b]_{j1} \\
&= \frac{1}{detA}[C^T b]_{j1} \\
&= \frac{1}{detA}\sum_k [C^T]_{jk}[b]_{k1} \\
&= \frac{1}{detA}\sum_k [C]_{kj}b_{k1} \\
&= \frac{1}{detA}\sum_k c_{kj}b_{k1} \\
&= \frac{1}{detA}\sum_k b_{k1}c_{kj}. \qquad (1.78)
\end{aligned}
$$

Ahora considérense las matrices A y A_j,

$$
A = \begin{array}{c}
\begin{matrix} \phantom{a_{11}} & & j & & \phantom{a_{1n}} \end{matrix} \\
\begin{pmatrix}
a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
a_{n1} & \cdots & a_{nj} & \cdots & a_{nn}
\end{pmatrix}
\end{array}
\qquad
A_j = \begin{array}{c}
\begin{matrix} \phantom{a_{11}} & & j & & \phantom{a_{1n}} \end{matrix} \\
\begin{pmatrix}
a_{11} & \cdots & b_1 & \cdots & a_{1n} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
a_{i1} & \cdots & b_i & \cdots & a_{in} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
a_{n1} & \cdots & b_n & \cdots & a_{nn}
\end{pmatrix}
\end{array}
$$

Desarrollando los determinantes de las matrices anteriores mediante la columna j,

$$
\begin{aligned}
|A| &= \sum_k a_{kj}c_{kj}, \\
|A_j| &= \sum_k b_{k1}c_{kj}. \qquad (1.79)
\end{aligned}
$$

De las ecuaciones 1.78 y 1.79 se concluye que,

$$x_j = \frac{detA_j}{detA} \qquad \blacksquare$$

Ejemplo 1.4.13 *Resuelva el siguiente sistema de ecuaciones mediante la regla de Cramer.*

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

Solución: El sistema se expresa como un producto de matrices mediante,

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

Por lo anterior la matriz de coeficientes es

$$A = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix}$$

y la matriz de los términos constantes es

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

Las matrices A_1, A_2 y A_3 se forman sustituyendo la matriz b en las columnas 1, 2 y 3 de la matriz A, respectivamente, como se muestra a continuación.

$$A_1 = \begin{bmatrix} -3 & -1 & 4 \\ 1 & 2 & -3 \\ -2 & 3 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 2 & -3 & 4 \\ 1 & 1 & -3 \\ 5 & -2 & 1 \end{bmatrix} \quad A_3 \begin{bmatrix} 2 & -1 & -3 \\ 1 & 2 & 1 \\ 5 & 3 & -2 \end{bmatrix}.$$

Ahora se calculan los determinantes de las matrices A, A_1, A_2 y A_3. El estudiante debe ser capaz de comprender los pasos realizados para reducir los determinantes, en caso contrario, se recomienda repasar los ejercicios de la sección 1.4.2.

1. El determinante de la matriz A fue calculado en el ejemplo 1.4.9, obteniéndose $|A| = 10$.

2. Calcúlese el determinante de A_1

$$A_1 = \begin{vmatrix} -3 & -1 & 4 \\ 1 & 2 & -3 \\ -2 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 5 & -5 \\ 1 & 2 & -3 \\ 0 & 7 & -5 \end{vmatrix} = -\begin{vmatrix} 5 & -5 \\ 7 & -5 \end{vmatrix} = 5\begin{vmatrix} 5 & 1 \\ 7 & 1 \end{vmatrix} = -10.$$

Así que de acuerdo con la regla de Cramer,

$$x_1 = \frac{-10}{10} = -1.$$

3. Calcúlese el determinante de la matriz A_2.

$$A_2 = \begin{vmatrix} 2 & -3 & 4 \\ 1 & 1 & -3 \\ 5 & -2 & 1 \end{vmatrix} = \begin{vmatrix} 0 & -5 & 10 \\ 1 & 1 & -3 \\ 0 & -7 & 16 \end{vmatrix} = - \begin{vmatrix} -5 & 10 \\ -7 & 16 \end{vmatrix} = 5(2) \begin{vmatrix} 1 & 1 \\ 7 & 8 \end{vmatrix} = 10.$$

De acuerdo con la regla de Cramer,

$$x_2 = \frac{10}{10} = 1.$$

4. Por último calcílese el determinante de la matriz A_3.

$$A_3 = \begin{vmatrix} 2 & -1 & -3 \\ 1 & 2 & 1 \\ 5 & 3 & -2 \end{vmatrix} = \begin{vmatrix} 0 & -5 & -5 \\ 1 & 2 & 1 \\ 0 & -7 & -7 \end{vmatrix} = - \begin{vmatrix} -5 & -5 \\ -7 & -7 \end{vmatrix} = -5(7) \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0.$$

Así que

$$x_3 = \frac{0}{10} = 0.$$

Finalmente, la solución del sistema es

$$\begin{cases} x_1 = -1, \\ x_2 = 1, \\ x_3 = 0. \end{cases}$$

Ejemplo 1.4.14 *Use la regla de Cramer para obtener la solución del siguiente sistema de ecuaciones.*

$$\left.\begin{array}{rrrrcr} x_1 & +2x_2 & -x_3 & +4x_4 & = & 4 \\ 3x_1 & -x_2 & +2x_3 & & = & 10 \\ 4x_1 & +5x_2 & -2x_3 & +x_4 & = & -5 \\ -2x_1 & +3x_2 & & +2x_4 & = & -1 \end{array}\right\}$$

Solución: Hallar los determinantes de las matrices A, A_1, A_2, A_3 y A_4. Mediante la reducción de los renglones en combinación con el desarrollo por cofactores.

1. Determinante de la matriz de coeficientes,

$$|A| = \begin{vmatrix} 1 & 2 & -1 & 4 \\ 3 & -1 & 2 & 0 \\ 4 & 5 & -2 & 1 \\ -2 & 3 & 0 & 2 \end{vmatrix}.$$

Este determinante fue calculado en el ejemplo 1.4.10, resultando $|A| = -209$

2. Determinante A_1.

$$\begin{aligned}
|A_1| &= \begin{vmatrix} 4 & 2 & -1 & 4 \\ 10 & -1 & 2 & 0 \\ -5 & 5 & -2 & 1 \\ -1 & 3 & 0 & 2 \end{vmatrix} = \begin{vmatrix} 24 & -18 & 7 & 0 \\ 10 & -1 & 2 & 0 \\ -5 & 5 & -2 & 1 \\ 9 & -7 & 4 & 0 \end{vmatrix} \\
&= -\begin{vmatrix} 24 & -18 & 7 \\ 10 & -1 & 2 \\ 9 & -7 & 4 \end{vmatrix} = \begin{vmatrix} 24 & 18 & 7 \\ 10 & 1 & 2 \\ 9 & 7 & 4 \end{vmatrix} = \begin{vmatrix} -156 & 18 & -29 \\ 0 & 1 & 0 \\ -61 & 7 & -10 \end{vmatrix} \\
&= \begin{vmatrix} -156 & -29 \\ -61 & -10 \end{vmatrix} = \begin{vmatrix} 156 & 29 \\ 61 & 10 \end{vmatrix} = -209
\end{aligned}$$

Determínese la incógnita x_1. De acuerdo con la regla de Cramer se tiene,

$$x_1 = \frac{|A_1|}{|A|} = \frac{-209}{-209} = 1$$

3. Determinante ahora $|A_2|$.

$$\begin{aligned}
|A_2| &= \begin{vmatrix} 1 & 4 & -1 & 4 \\ 3 & 10 & 2 & 0 \\ 4 & -5 & -2 & 1 \\ -2 & -1 & 0 & 2 \end{vmatrix} = \begin{vmatrix} -15 & 24 & 7 & 0 \\ 3 & 10 & 2 & 0 \\ 4 & -5 & -2 & 1 \\ -10 & 9 & 4 & 0 \end{vmatrix} \\
&= -\begin{vmatrix} -15 & 24 & 7 \\ 3 & 10 & 2 \\ -10 & 9 & 4 \end{vmatrix} = \begin{vmatrix} 15 & 24 & 7 \\ -3 & 10 & 2 \\ 10 & 9 & 4 \end{vmatrix} = \begin{vmatrix} 22 & 24 & 7 \\ -1 & 10 & 2 \\ 14 & 9 & 4 \end{vmatrix} \\
&= \begin{vmatrix} 0 & 244 & 51 \\ -1 & 10 & 2 \\ 0 & 149 & 32 \end{vmatrix} = \begin{vmatrix} 244 & 51 \\ 149 & 32 \end{vmatrix} = 209
\end{aligned}$$

De acuerdo con la regla de Cramer se tiene,

$$x_2 = \frac{|A_2|}{|A|} = \frac{209}{-209} = -1$$

4. Calcúlese el siguiente determinante,

$$\begin{aligned}
|A_3| &= \begin{vmatrix} 1 & 2 & 4 & 4 \\ 3 & -1 & 10 & 0 \\ 4 & 5 & -5 & 1 \\ -2 & 3 & -1 & 2 \end{vmatrix} = \begin{vmatrix} -15 & -18 & 24 & 0 \\ 3 & -1 & 10 & 0 \\ 4 & 5 & -5 & 1 \\ -10 & -7 & 9 & 0 \end{vmatrix} \\
&= -\begin{vmatrix} -15 & -18 & 24 \\ 3 & -1 & 10 \\ -10 & -7 & 9 \end{vmatrix} = \begin{vmatrix} -15 & 18 & 24 \\ 3 & 1 & 10 \\ -10 & 7 & 9 \end{vmatrix} = \begin{vmatrix} -69 & 18 & -156 \\ 0 & 1 & 0 \\ -31 & 7 & -61 \end{vmatrix} \\
&= \begin{vmatrix} -69 & -156 \\ -31 & -61 \end{vmatrix} = \begin{vmatrix} 69 & 156 \\ 31 & 61 \end{vmatrix} = -627
\end{aligned}$$

En este caso,

$$x_3 = \frac{|A_3|}{|A|} = \frac{-627}{-209} = 3.$$

5. Finalmente, $|A_4|$.

$$\begin{aligned}
|A_4| &= \begin{vmatrix} 1 & 2 & -1 & 4 \\ 3 & -1 & 2 & 10 \\ 4 & 5 & -2 & -5 \\ -2 & 3 & 0 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 & 4 \\ 3 & -1 & 2 & 10 \\ 7 & 4 & 0 & 5 \\ -2 & 3 & 0 & -1 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 2 & -1 & 4 \\ 5 & 3 & 0 & 18 \\ 7 & 4 & 0 & 5 \\ -2 & 3 & 0 & -1 \end{vmatrix} = -\begin{vmatrix} 5 & 3 & 18 \\ 7 & 4 & 5 \\ -2 & 3 & -1 \end{vmatrix} = -\begin{vmatrix} -31 & 57 & 18 \\ -3 & 19 & 5 \\ 0 & 0 & -1 \end{vmatrix} \\
&= \begin{vmatrix} -31 & 57 \\ -3 & 19 \end{vmatrix} = -\begin{vmatrix} 31 & 57 \\ 3 & 19 \end{vmatrix} = -418.
\end{aligned}$$

Con esto se determina la cuarta incógnita,

$$x_4 = \frac{|A_4|}{|A|} = \frac{-418}{-209} = 2$$

En resumen, la solución del sistema es

$$\begin{cases} x_1 = 1, \\ x_2 = -1, \\ x_3 = 3, \\ x_4 = 2. \end{cases}$$

Capítulo 2

Espacios Vectoriales

2.1. Espacios Vectoriales

Cantidades físicas como velocidad, aceleración, fuerza, campos eléctricos y magnéticos, entre otras, están íntimamente relacionadas con el concepto de campo vectorial. En esta sección se estudian los conceptos y teoremas más importantes relacionados con los campos vectoriales.

2.1.1. Definición y propiedades básicas

Definición 2.1.1 (Espacio Vectorial) *Sea V un conjunto arbitrario de objetos sobre los cuales se definen dos operaciones: la adición y la multiplicación por un escalar. La adición es una regla que asocia con cada pareja de elementos u y v en V, el elemento $u+v$, llamado la suma de u y v. La multiplicación por un escalar es una regla que asocia con cada elemento v en V y escalar k, el elemento kv, llamado múltiplo escalar de u por k. Se dice que V es un espacio vectorial y a sus elementos se les denomina vectores, siempre y cuando la suma y la multiplicación por un escalar cumplan con los siguientes axiomas.*

Si u, v y w son elementos de V, y, k y l son escalares, entonces

1. *Propiedad de cerradura bajo la suma.*
   $u+v \in V$.

2. *Propiedad conmutativa bajo la suma.*
   $u+v = v+u$.

3. *Propiedad asociativa para la suma.*
   $u+(v+w) = (u+v)+w$.

4. *Existencia del neutro aditivo.*
   Existe un elemento en V, llamado elemento cero, denotado por 0_e o simplemente por 0, tal que, $0 + u = u + 0 = u$, para todo $u \in V$.

5. *Existencia del inverso aditivo.*
   Para todo $u \in V$ existe el elemento $-u \in V$, llamado inverso aditivo de u, tal que, $u + (-u) = (-u) + u = 0$.

6. *Propiedad de cerradura bajo la multiplicación por un escalar.*
   Para todo $u \in V$ y escalar k, $ku \in V$.

7. *Primera ley distributiva.*
   $k(u + v) = ku + kv$.

8. *Segunda ley distributiva.*
   $(k + l)u = ku + lu$.

9. *Ley asociativa de la multiplicación por un escalar.*
   $k(lu) = (kl)u$.

10. $1.u = u$.

Bastará con que uno de los axiomas no se satisfaga para que el conjunto V no sea un espacio vectorial.

El concepto de campo vectorial está íntimamente relacionado con el concepto de campo o cuerpo[1]. En este trabajo se usará el término escalar para hacer referencia a un número real; es decir, se trabajará con el campo de los números reales; además, se denotará a los vectores mediante letras en negritas, para diferenciarlos de los escalares.

2.1.2. Ejemplos de espacios vectoriales de distintos géneros

Ejemplo 2.1.1 (Espacio vectorial $\mathbb{R}^n$) *Si n es un entero positivo, entonces una n-ada ordenada es una sucesión de números reales $(a_1, a_2, \ldots, a_n)$. El conjunto de todas las n-adas ordenadas se conoce como espacio n-dimensional y se denota por $\mathbb{R}^n$; es decir,*

$$\mathbb{R}^n = \{(a_1, a_2, \ldots, a_n) \,:\, a_i \in \mathbb{R}, \quad \textit{para todo} \;\; i = 1, 2, \ldots, n\}$$

Si $u = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$, entonces la entrada n-ésima de u, se denota por $(u)_i$; es decir, $(u)_i = u_i$.

[1]Los cuerpos o campos son estructuras algebraicas que comparten las mismas propiedades con los números reales.

Si $u = (u_1, u_2, \ldots, u_n)$ y $v = (v_1, v_2, \ldots, v_n)$ son elementos de $\mathbb{R}^n$, con entradas $(u)_i = u_i$ y $(v)_i = v_i$, entonces se dice que son iguales y se escribe $u = v$, si y solo si, las entradas correspondientes son iguales; es decir,

$$u = v \Longleftrightarrow u_i = v_i, \quad \textit{para todo} \quad i = 1, 2, \ldots, n$$

Las operaciones usuales de suma y multiplicación por un escalar definidas sobre $\mathbb{R}^n$ son:

Suma

Si $u = (u_1, u_2, \ldots, u_n)$ y $v = (v_1, v_2, \ldots, v_n)$ están en $\mathbb{R}^n$, entonces

$$u + v = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n);$$

es decir,

$$(u + v)_i = u_i + v_i, \quad \textit{para todo} \quad i = 1, 2, \ldots, n$$

Multiplicación por un escalar

Si Si $u = (u_1, u_2, \ldots, u_n) \in \mathbb{R}$ y k es un escalar, entonces

$$ku = (ku_1, ku_2, \ldots, ku_n);$$

es decir,

$$(ku)_i = ku_i \quad \textit{para todo} \quad i = 1, 2, \ldots, n$$

Para probar que $\mathbb{R}^n$ con las operaciones anteriores constituye un espacio vectorial, se debe verificar que se cumplen los diez axiomas de espacio vectorial.

Si u, v y w son elementos de $\mathbb{R}^n$, con entradas correspondientes u_i, v_i, w_i respectivamente, y, k y l son escalares, entonces:

1. Como u y v están en $\mathbb{R}^n$, entonces las entradas correspondientes u_i y v_i son números reales; así que, $u_i + v_i \in \mathbb{R}$, por lo tanto, $u + v \in \mathbb{R}^n$.

2. Considérese la entrada i-ésima de $u + v$,
   $(u+v)_i = u_i + v_i = v_i + u_i = (v + u)_i$,ya que las entradas son números reales y éstos conmutan bajo la suma. Así que, $u + v = v + u$.

3. Considérese la entrada i-ésima de $u + (v + w)$,
   $(u+(v+w))_i = u_i+(v+w)_i = u_i+(v_i+w_i) = (u_i+v_i)+w_i = ((u+v)+w)_i$,ya que las entradas son números reales y éstos son asociativos. Así que, $u + (v + w) = (u + v) + w$.

4. Sea el elemento $0_e \in \mathbb{R}^n$ aquel cuyas entradas son todas iguales a cero; es decir, $(0_e)_i = 0$, entonces

$$(0_e + u)_i = 0 + u_i = u_i + 0 = u_i = (u)_i.$$

por lo anterior, el elemento cero definido de esta manera cumple con la propiedad fundamental

$$0_e + u = u + 0_e = u.$$

5. Si $u \in \mathbb{R}^n$ y $(u)_i = u_i$, defínase el elemento $-u \in \mathbb{R}^n$ como $(-u)_i = -u_i$, entonces

$$(u + (-u))_i = (u)_i + (-u)_i = u_i + (-u_i) = u_i - u_i = 0 = (0_e)_i;$$

es decir,

$$u + (-u) = (-u) + u = 0_e.$$

6. Si $u \in \mathbb{R}^n$, y sea $(u)_i = u_i \in \mathbb{R}$. Si además k es un número real, entonces $k(u)_i = ku_i \in \mathbb{R}$, luego $ku \in \mathbb{R}^n$.
7. $(k(u + v))_i = k(u + v)_i = k(u_i + v_i) = kv_i + kv_i = (ku)_i + (kv)_i = (ku + kv)_i$; es decir, $k(u + v) = ku + kv$.
8. $((k + l)u)_i = (k + l)u_i = ku_i + lu_i = (ku)_i + (lu)_i = (ku + lu)_i$; es decir, $(k + l)u = ku + lu$.
9. $(k(lu))_i = k(lu)_i = k(lu_i) = (kl)u_i = ((kl)u)_i$; es decir, $k(lu) = (kl)u$.
10. $(1.u)_i = 1(u)_i = 1.u_i = u_i = (u)_i$; por lo tanto, $1.u = u$.

Por lo anterior $\mathbb{R}^n$ con las operaciones usuales de suma y multiplicación por un escalar constituye un espacio vectorial.

Ejemplo 2.1.2 (Espacio de las matrices de tamaño mxn) *En la sección 1.2 representó a una matriz A de tamaño mxn como $(A)_{mxn}$ y sus entradas mediante, $[A]_{ij} = a_{ij}$. El conjunto de todas las matrices de tamaño mxn con entradas en los reales se denota por, $M_{mxn}(\mathbb{R})$; es decir,*

$$M_{mxn}(\mathbb{R}) = \{(A)_{mxn} \,:\, a_{ij} \in \mathbb{R}, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n\}$$

Se dice que dos matrices A y B con entradas correspondientes a_{ij} y b_{ij} son iguales si son del mismo tamaño y si sus entradas correspondientes son iguales; es decir, si $a_{ij} = b_{ij}$, para todo i, j. En tal caso se escribe $A = B$. Las operaciones usuales definidas sobre $M_{mxn}(\mathbb{R})$ son las que se estudiaron en la sección 1.2.

Suma
Si A y B están en $M_{mxn}(\mathbb{R})$ y sus entradas correspondientes son, a_{ij} y b_{ij}, entonces la suma de A y B, denotado por $A+B$, es la matriz de tamaño mxn que se obtiene sumando las entradas correspondientes de las matrices A y B; es decir,

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

Multiplicación por un escalar
Si $A \in M_{mxn}(\mathbb{R})$ tiene entradas a_{ij} y k es un escalar, entonces kA es la matriz de tamaño mxn cuyas entras se obtienen multiplicando las entradas de A por k; es decir,

$$[kA]_{ij} = ka_{ij}.$$

Resta verificar que los diez axiomas se satisfcen para para probar que $M_{mxn}(\mathbb{R})$ con las operaciones usuales constituye un espacio vectorial. Si A, B y C son elementos de $M_{mxn}(\mathbb{R})$, con entradas correspondientes a_{ij}, b_{ij}, c_{ij} respectivamente, y, k y l son escalares, entonces

1. Considérese el elemento $[A+B]_{ij}$ de la matriz $A+B$.

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

   Como las entradas de las matrices A y B son números reales, entonces

   $$[A+B]_{ij} = a_{ij} + b_{ij} \in \mathbb{R},$$

   esto significa que, $A+B \in M_{mxn}(\mathbb{R})$.

2. Considérese el elemento $[A+B]_{ij}$ de la matriz $A+B$.

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

   ya que las entradas son números reales y por lo tanto conmutan, así que,

   $$A+B = B+A.$$

3. Considérese el elemento $[A+(B+C)]_{ij}$ de la matriz $A+(B+C)$.

   $$[A+(B+C)]_{ij} = a_{ij}+[B+C]_{ij} = a_{ij}+(b_{ij}+c_{ij}) = (a_{ij}+b_{ij})+c_{ij} = [(A+B)+C]_{ij},$$

   ya que las entradas son números reales, y por lo tanto son asociativos,
   así que,

   $$A+(B+C) = (A+B)+C$$

   .

4. Sea el elemento $0_e \in M_{mxn}(\mathbb{R})$ aquella matriz cuyas entradas son todas iguales a cero, es decir, $(0_e)_{mxn}$ y sea $A \in M_{mxn}(\mathbb{R})$, entonces

$$[A + 0_e]_{ij} = [A]_{ij} + [0_e]_{ij} = a_{ij} + 0 = a_{ij} = [A]_{ij},$$

por lo tanto, la matriz cero de tamaño mxn cumple con la propiedad fundamental,

$$A + 0_e = 0_e + A = A.$$

5. Sea $A \in M_{mxn}(\mathbb{R})$, con entradas a_{ij} y sea $-A$ la matriz de tamaño mxn cuyas entradas se definen como $[-A]_{ij} = -a_{ij}$. Es claro que $-A \in M_{mxn}(\mathbb{R})$ además,

$$[A + (-A)]_{ij} = [A]_{ij} + [-A]_{ij} = a_{ij} + (-a_{ij}) = a_{ij} - a_{ij} = 0 = [0]_{ij},$$

por lo tanto,

$$A + (-A) = (-A) + A = 0_e.$$

6. Sea $A \in M_{mxn}(\mathbb{R})$, con entradas a_{ij} y k un escalar, entonces

$$[kA]_{ij} = ka_{ij} \in \mathbb{R},$$

ya que tanto las entradas de la matriz A como el escalar son números reales. Así que, kA es una matriz de tamaño mxn con entradas reales; es decir, $kA \in M_{mxn}(\mathbb{R})$.

7. Considérese el elemento $[k(A + B)]_{ij}$ de la matriz $k(A + B)$.

$$[k(A+B)]_{ij} = k[A+B]_{ij} = k(a_{ij}+b_{ij}) = ka_{ij}+kb_{ij} = [kA]_{ij}+[kB]_{ij} = [kA+kB]_{ij},$$

por lo anterior,

$$k(A + B) = kA + kB.$$

8. Considérese el elemento $[(k + l)A]_{ij}$ de la matriz $(k + l)A$.

$$[(k + l)A]_{ij} = (k + l)a_{ij} = ka_{ij} + la_{ij} = [kA]_{ij} + [lA]_{ij} = [kA + lA]_{ij},$$

por lo anterior,

$$(k + l)A = kA + lA.$$

9. Considérese el elemento $[k(lA)]_{ij}$ de la matriz $k(lA)$.

$$[k(lA)]_{ij} = k[lA]_{ij} = k(la_{ij}) = (kl)a_{ij} = [(kl)A]_{ij},$$

así que,

$$k(lA) = (kl)A.$$

10. Finalmente,
$$[1.A]_{ij} = 1.a_{ij} = a_{ij} = [A]_{ij},$$
por lo anterior,
$$1.A = A.$$

Ejemplo 2.1.3 (Espacio vectorial de las funciones de valor real) *El conjunto de todas las funciones cuyo dominio y codominio son los reales se denota por $F(-\infty, \infty)$. Diremos que dos funciones $f(x)$ y $g(x)$ son iguales, si $f(x) = g(x)$ para todo $x \in \mathbb{R}$.*

La suma y la multiplicación por un escalar definidos usualmente para $F(-\infty, \infty)$ son:

Suma
Si $f(x)$ y $g(x)$ son funciones de valor real, entonces
$$(f + g)(x) = f(x) + g(x).$$

Multiplicación por un escalar
Si k es un escalar y $f(x) \in F(-\infty, \infty)$, entonces
$$(kf)(x) = kf(x).$$

Pruebe que $F(-\infty, \infty)$ con las operaciones de suma y multiplicación por un escalar aquí definidas forma un espacio vectorial.

Sean $f(x)$, $g(x)$ y $h(x)$ elementos de $F(-\infty, \infty)$ y, k y l escalares entonces:

1. Como $(f+g)(x) = f(x)+g(x)$, de modo que $(f+g)(x)$ representa una función con dominio y codominio en los reales, entonces
$$(f + g)(x) \in F(-\infty, \infty).$$
2. $(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$, $\forall x \in \mathbb{R}$.
3.
$$\begin{aligned} ((f + g) + h)(x) &= (f + g)(x) + h(x) = (f(x) + g(x)) + h(x), \\ &= f(x) + (g(x) + h(x)) = f(x) + (g + h)(x), \\ &= (f + (g + h))(x), \quad \forall x \in \mathbb{R}. \end{aligned}$$
4. Sea $0_e = 0(x) = 0$, $\forall x \in \mathbb{R}$, entonces
$$\begin{aligned} (f + 0_e)(x) &= f(x) + 0(x) \\ &= f(x) + 0 = 0 + f(x) \\ &= 0_e(x) + f(x) \\ &= (0_e + f)(x), \quad \forall x \in \mathbb{R}. \end{aligned}$$

5. Si $f(x) \in F(-\infty,\infty)$ y sea $(-f)(x) = -f(x)\ \forall x \in \mathbb{R}$, entonces

$$\begin{aligned}
(f+(-f))(x) &= f(x)+(-f(x)),\\
&= f(x)-f(x),\\
&= 0,\\
&= -f(x)+f(x),\\
&= (-f+f)(x), \quad \forall x \in \mathbb{R}.
\end{aligned}$$

6. Si k es un escalar y $f(x) \in F(-\infty,\infty)$, entonces $(kf)(x) = kf(x)$ se observa que $kf(x)$ representa una función con dominio y codominio en los reales, por lo tanto,

$$(kf)(x) \in F(-\infty,\infty).$$

7.
$$\begin{aligned}
(k(f+g))(x) &= k(f(x)+g(x)),\\
&= kf(x)+kg(x) = (kf)(x)+(kg)(x),\\
&= (kf+kg)(x), \quad \forall x \in \mathbb{R}.
\end{aligned}$$

8.
$$\begin{aligned}
((k+l)f)(x) &= (k+l)f(x) = kf(x)+lf(x),\\
&= (kf)(x)+(lf)(x),\\
&= (kf+lf)(x), \quad \forall x \in \mathbb{R}.
\end{aligned}$$

9.
$$(k(lf))(x) = k(lf)(x) = k(lf(x)) = (kl)f(x) = ((kl)f)(x), \quad \forall x \in \mathbb{R}.$$

10.
$$(1.f)(x) = 1.f(x) = f(x), \quad \forall x \in \mathbb{R}.$$

Se ha probado que se cumplen todos los axiomas de espacio vectorial, por lo tanto, $F(-\infty,\infty)$ con las operaciones usuales constituye un espacio vectorial.

Teorema 2.1.1 *Si V es un espacio vectorial, $\boldsymbol{u}$ un vector en V y k un escalar, entonces* *a)* $0\boldsymbol{u} = \boldsymbol{0}$,

b) $k\boldsymbol{0} = \boldsymbol{0}$,

c) $-1.\boldsymbol{u} = -\boldsymbol{u}$.

d) Si $k\boldsymbol{u} = \boldsymbol{0}$, *entonces* $k = 0$ *o bien* $\boldsymbol{u} = \boldsymbol{0}$.

Demostración: a)

$$\begin{aligned}
0\mathbf{u} + 0\mathbf{u} &= (0+0)\mathbf{u}, && \text{axioma 8,}\\
0\mathbf{u} + 0\mathbf{u} &= 0\mathbf{u}, && \text{propiedad del cero,}\\
(0\mathbf{u} + 0\mathbf{u}) + (-0\mathbf{u}) &= 0\mathbf{u} + (-0\mathbf{u}), && \text{axioma 5,}\\
0\mathbf{u} + (0\mathbf{u} + (-0\mathbf{u})) &= 0\mathbf{u} + (-0\mathbf{u}), && \text{axioma 3,}\\
0\mathbf{u} + \mathbf{0} &= \mathbf{0}, && \text{axioma 5,}\\
0\mathbf{u} &= \mathbf{0}, && \text{axioma 4.}
\end{aligned}$$

b)

$$\begin{aligned}
0\mathbf{u} &= \mathbf{0}, && \text{inciso a,}\\
k(0\mathbf{u}) &= k\mathbf{0}, && \text{multiplicando por k,}\\
(k0)\mathbf{u} &= k\mathbf{0}, && \text{axioma 9.}\\
0\mathbf{u} &= k\mathbf{0}, && \text{propiedad de los reales,}\\
\mathbf{0} &= k\mathbf{0}, && \text{inciso a.}
\end{aligned}$$

c)

$$\begin{aligned}
\mathbf{u} + (-1.\mathbf{u}) &= 1.\mathbf{u} + (-1.\mathbf{u}), && \text{axioma 10,}\\
\mathbf{u} + (-1.\mathbf{u}) &= (1. + (-1))\mathbf{u}, && \text{axioma 8,}\\
\mathbf{u} + (-1.\mathbf{u}) &= (0)\mathbf{u}, && \text{propiedad de los reales,}\\
\mathbf{u} + (-1.\mathbf{u}) &= \mathbf{0}, && \text{inciso a.}
\end{aligned}$$

De lo anterior se concluye que $-\mathbf{u} = -1.\mathbf{u}$.

d) Supóngase que $\mathbf{u} \neq \mathbf{0}$, ahora, si $k\mathbf{u} = \mathbf{0}$ y si $k \neq 0$, entonces $\mathbf{u} = \dfrac{1}{k}\mathbf{0} = \mathbf{0}$ lo que constituye una contradicción, ya que de inicio se supuso $\mathbf{u} \neq \mathbf{0}$. Con lo cual se concluye que si $k\mathbf{u} = \mathbf{0}$, entonces $\mathbf{u} = \mathbf{0}$ o $k = 0$. ∎

2.2. Subespacios

2.2.1. Definición y propiedades básicas

Definición 2.2.1 (Subespacios) *Un subconjunto W de un espacio vectorial V es un subespacio de V, si W es un espacio vectorial bajo las operaciones de adición y de multiplicación por un escalar definidas sobre V.*

Teorema 2.2.1 *Si W es un conjunto de uno o más vectores de un espacio vectorial V, entonces W es un subespacio de V si y solo si se cumplen las condiciones siguientes,*

a) Cerradura bajo la suma

Si $\boldsymbol{u}$ *y* $\boldsymbol{v}$ *están en W, entonces* $\boldsymbol{u}+\boldsymbol{v}$ *también está en W.*

b) Cerradura bajo la multiplicación por un escalar

Si k es un escalar cualquiera y $\boldsymbol{u}$ *es cualquier vector en W, entonces* $k\boldsymbol{u}$ *también está en W.*

Demostración: $\Rightarrow$) Supóngase que W es un subespacio de V, en este caso W es un espacio vectorial bajo las operaciones definidas sobre V; es decir, los elementos de W satisfacen los 10 axiomas de espacio vectorial, en particular satisfacen los axiomas 1 y 6 que corresponden a las condiciones a y b de este teorema.
($\Leftarrow$ Se sabe que se satisfacen las condiciones a y b que equivalen a los axiomas 1 y 6; además, como $W \subseteq V$ entonces los elementos de W heredan las propiedades de V contenidos en los axiomas 2, 3, 7 8, 9, 10. Así que solo resta probar que los elementos de W satisfacen los axiomas 4 y 5. De hecho como se cumple la condición b); es decir, si k es cualquier escalar y $\mathbf{w} \in W$, entonces $k\mathbf{w} \in W$, en particular si $k = -1$, tenemos que, $-1.\mathbf{u} = -\mathbf{u} \in W$ (teorema 2.1.1c), de manera que se satisface el axioma 5. Finalmente, si k=0, entonces $0.\mathbf{u} = \mathbf{0} \in W$ (teorema 2.1.1a), por lo tanto se satisface el axioma 4. Se ha demostrado que W es un espacio vectorial bajo las operaciones definidas en V; con lo cual se concluye la demostración. ∎

2.2.2. Ejemplos de subespacios vectoriales de distintos géneros

Ejemplo 2.2.1 *Pruebe que las rectas que pasan por el origen son subespacios de* $\mathbb{R}^3$*.*

Solución: En general la recta que pasa por los punto $A(a_o, b_o, c_o)$ y $B(a, b, c)$ se expresa como,

$$\mathbf{r} = t\mathbf{r}_B + (1 - t)\mathbf{r}_A \tag{2.1}$$

donde $\mathbf{r}_A$, $\mathbf{r}_B$ y $\mathbf{r}$ son los vectores de posición correspondientes a los puntos $A(a_o, b_o, c_o)$, $B(a, b, c)$ y $P(x, y, z)$ respectivamente, y t es cualquier número real.

Si la recta pasa por el origen, entonces $A(a_o, b_o, c_o) = A(0, 0, 0)$ en este caso $\mathbf{r}_A = \mathbf{0}$. Por lo tanto, la ecuación de la recta que pasa por el origen es,

$$\mathbf{r} = t\mathbf{r}_B = t(a, b, c) \tag{2.2}$$

Es claro que B(a,b,c) es un punto distinto del origen. Por lo anterior, la recta que pasa por el origen es el lugar geométrico que se describe a continuación,

$$W = \{(x, y, z) = t(a, b, c) : t \in \mathbb{R}\}. \tag{2.3}$$

Se procede entonces a demostrar que el conjunto W es un subespacio de $\mathbb{R}^3$.

a) Considérese que $\mathbf{u}$ y $\mathbf{v}$ están en W, en este caso, existen escalares t_1 y t_2 tales que, $\mathbf{u} = t_1(a, b, c)$ y $\mathbf{v} = t_2(a, b, c)$, así que,

$$\mathbf{u} + \mathbf{v} = t_1(a, b, c) + t_2(a, b, c) = (t_1 + t_2)(a, b, c)$$

como $t_1 + t_2 \in \mathbb{R}$, entonces $\mathbf{u} + \mathbf{v} \in W$

b) Sea k un número real y $\mathbf{u}$ un elemento de W , entonces

$$k\mathbf{u} = k(t_1(a, b, c)) = (kt_1)(a, b, c)$$

Como $kt_1 \in \mathbb{R}$, entonces $k\mathbf{u} \in W$.

Concluimos que W es un subespacio de $\mathbb{R}^3$, ya que se satisfacen las propiedades a) y b).

Ejemplo 2.2.2 *Demuestre que las matrices de tamaño 2x2 que tienen ceros en la diagonal principal forman un subespacio de* $M_{2X2}(\mathbb{R})$

Solución: En este caso,

$$W = \left\{ \begin{bmatrix} 0 & a_{12} \\ a_{21} & 0 \end{bmatrix} : a_{12}, a_{21} \in \mathbb{R} \right\}.$$

Verifiquemos que se satisfacen las propiedades a) y b) del teorema anterior.

a) Sean A y B elementos de W, donde

$$A = \begin{bmatrix} 0 & a_1 \\ a_2 & 0 \end{bmatrix} \quad y \quad B = \begin{bmatrix} 0 & b_1 \\ b_2 & 0 \end{bmatrix}$$

con a_1, a_2, b_1, b_2 dentro de los reales, entonces

$$A + B = \begin{bmatrix} 0 & a_1 \\ a_2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & b_1 \\ b_2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & a_1 + b_1 \\ a_2 + b_2 & 0 \end{bmatrix}$$

como $a_1 + b_1$ y $a_2 + b_2$ son números reales y la matriz $A + B$ tienes ceros en su diagonal, entonces $A + B \in W$.

b) Ahora sean k un escalar y A un elemento de W, entonces

$$kA = k \begin{bmatrix} 0 & a_1 \\ a_2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & ka_1 \\ ka_2 & 0 \end{bmatrix}$$

nuevamente obtuvimos una matriz cuyos elementos de la diagonal son ceros, además, ka_1 y ka_2 son elementos de los reales, por lo tanto $kA \in W$.

Finalmente se concluye que W es un subespacio de M_{2x2}.

Ejemplo 2.2.3 *Demuestre que todo espacio vectorial V tiene al menos dos subespacios.*

Solución: Todo espacio vectorial se contiene a sí mismo, por lo tanto $W_1 = V$ es un subestacio de V además, todo espacio vectorial contiene al vector cero, en este caso $W_2 = \{\mathbf{0}\}$ es cerrado bajo la suma y bajo la multiplicación por un escalar, de modo que W_2 es un subespacio.

Ejemplo 2.2.4 *Sea $Ax = 0$ un sistema homogéneo de m ecuaciones con n incógnitas, demuestre que el conjunto solución es un subespacio de $\mathbb{R}^n$.*

Solución: En este caso es necesario precisar que los elementos de $\mathbb{R}^n$ pueden ser considerados como matrices de tamaño nx1, es decir,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

De manera que las soluciones del sistema de ecuaciones representado por la ecuación homogénea $Ax = 0$ serán elementos de $\mathbb{R}^n$. Si el sistema tiene solución única, entonces $\{\mathbf{0}\}$ es un subespacio de $\mathbb{R}^n$. Supóngase por lo tanto, que el sistema tiene infinitas soluciones, en este caso,

$$W = \{x \in \mathbb{R}^n : Ax = 0\}.$$

Verifiquemos que se satisfacen las condiciones de cerradura.

a) Si x y y son elementos de W, entonces, $Ax = 0$ y $Ay = 0$, por la propiedad distributiva tenemos que,

$$\begin{aligned} A(x+y) &= Ax + Ay \\ &= 0 + 0 \\ &= 0. \end{aligned}$$

Concluimos que $x + y$ es una solución de la homogénea, por lo tanto $x + y \in W$

b) Sea k un escalar y x un elemento de W, entonces

$$A(kx) = k(A(x)) = k(0) = 0.$$

Lo anterior significa que kx es una solución de la homogénea, por lo tanto, $kx \in W$

Se concluye que W es un subespacio de $\mathbb{R}^n$.

Ejemplo 2.2.5 (Las funciones continuas) *Las funciones continuas en todos los reales se denotan por $C(-\infty, \infty)$. Este conjunto constituye un subespacio del espacio vectorial de las funciones de valor real; es decir, de $F(-\infty, \infty)$. De los cursos de cálculo aprendemos que la suma de dos funciones continuas es otra función continua; y que el producto de un escalar por una función continua es también una función continua. No se abundará en esta parte en el concepto de límite para realizar una demostración formal.*

Ejemplo 2.2.6 (Las funciones diferenciables) *Las funciones diferenciables en todos los reales se denotan mediante $C^1(-\infty, \infty)$. En los cursos de cálculo se prueba que las funciones diferenciables son continuas, así que, al sumar dos funciones diferenciables se obtiene una función continua y de la misma manera al multiplicar una función diferenciable por un escalar se obtiene una función continua. De manera que $C^1(-\infty, \infty)$ constituye un subespacio de $C(-\infty, \infty)$.*

El conjunto de las funciones con derivadas continuas de orden n en todos los reales se denota por $C^n(-\infty, \infty)$. El subespacio de las funciones con derivadas continuas de todos los órdenes se denota por $C^\infty(-\infty, \infty)$, la función $f(x) = e^x$ pertenece al conjunto anterior. Tanto $C^n(-\infty, \infty)$ como $C^\infty(-\infty, \infty)$ son subespacios de $C^1(-\infty, \infty)$.

Ejemplo 2.2.7 (Los polinomios) *El conjunto de los polinomios de grado menor o igual con n, donde n es un entero positivo se denotan por P_n; es decir,*

$$P_n = \{f(x) = a_0 + a_1x + a_2x^2 + \ldots a_nx^n : a_i \in \mathbb{R}, \quad \forall i\}.$$

Demuestre que P_n es un subespacio de $F(-\infty, \infty)$

Solución: Si p y q son elementos de P_n y k es un escalar entonces:

a) p y q pueden expresarse de la siguiente manera,

$$\begin{aligned} p(x) &= a_0 + a_1x + a_2x^2 + \ldots a_nx^n, \\ q(x) &= b_0 + b_1x + b_2x^2 + \ldots b_nx^n, \end{aligned}$$

donde tanto los a_i como los b_i son números reales para todo i.

Ahora considérese la suma $p + q$.

$$\begin{aligned} (p+q)(x) &= p(x) + q(x), \\ &= (a_0 + a_1x + \ldots a_nx^n) + (a_o + b_1x + \ldots b_nx^n), \\ &= (a_0 + b_0) + (a_1 + b_1)x + \ldots + (a_n + b_n)x^n, \\ &= c_0 + c_1x + c_2x^2 + \ldots + c_nx^n, \\ &\quad \text{donde,} \quad c_i = a_i + b_i, \quad \forall i = 1, 2, \ldots, n. \end{aligned}$$

Como $c_i = a_i + b_i \in \mathbb{R} \quad \forall i$, entonces $(p+q)(x)$ es un polinomio de grado menor o igual a n.

b) Si k es un escalar y $p(x) \in P_n$, entonces

$$\begin{aligned} (kp)(x) &= k(a_0 + a_1x + a_2x^2 \ldots a_nx^n) + (a_0 + b_1x + \ldots b_nx^n), \\ &= ka_0 + ka_1x + ka_2x^2 \ldots ka_nx^n, \\ &= d_0 + d_1x + d_2x^2 + \ldots d_nx^n, \\ &\quad \text{donde,} \quad d_i = ka_i, \quad \forall i = 1, 2, \ldots, n. \end{aligned}$$

Nuevamente, como $d_i = ka_i \in \mathbb{R} \quad \forall i$, entonces $(kp)(x)$ es un polinomio de grado menor o igual a n.

Se concluye que P_n es un subespacio de $F(-\infty, \infty)$.

2.3. Combinaciones lineales

Definición 2.3.1 (Combinaciones lineales) *Sean $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r$ vectores en un espacio vectorial V, entonces un vector $\boldsymbol{v}$ en V es una combinación lineal de los vectores $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r$ si existen escalares $k_1, k_2, \ldots, k_r$ tales que la ecuación*

$$\boldsymbol{v} = k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2, \ldots + k_r\boldsymbol{v}_r, \tag{2.4}$$

se satisface.

Ejemplo 2.3.1 *Sean $\boldsymbol{v}_1 = (2, 4, 5), \boldsymbol{v}_2 = (1, 3, 4)$ y $\boldsymbol{v}_3 = (0, -1, 2)$ vectores en $\mathbb{R}^3$. Determine si $\boldsymbol{v} = (3, 2, 1)$ es combinación lineal de los vectores anteriores.*

Solución: Aplicando la ecuación 2.4, se obtiene

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 = \mathbf{v}. \tag{2.5}$$

Sustituyendo los vectores dados se obtiene,

$$k_1(2, 4, 5) + k_2(1, 3, 4) + k_3(0, -1, 2) = (3, 2, 1). \tag{2.6}$$

Si es posible obtener los escalares k_1, k_2 y k_3 que satisfagan la ecuación anterior, entonces $\mathbf{v}$ será combinación lineal de los vectores $\mathbf{v}_1, \mathbf{v}_2$ y $\mathbf{v}_3$.

Desarrollando las operaciones indicadas en la ecuación 2.6 se tiene el siguiente sistema de ecuaciones,

$$\left.\begin{array}{ll} 2k_1 + k_2 & = 3 \\ 4k_1 + 3k_2 - k_3 & = 2 \\ 5k_1 + 4k_2 + 2k_3 & = 1 \end{array}\right\},$$

El sistema anterior puede expresarse en forma matricial como,

$$\begin{bmatrix} 2 & 1 & 0 \\ 4 & 3 & -1 \\ 5 & 4 & 2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}.$$

Si el sistema es consistente, entonces será posible hallar los escalares k_1, k_2 y k_3. Como en este caso particular la matriz de coeficientes es cuadrada, hallaremos el determinante.

$$detA = \begin{vmatrix} 2 & 1 & 0 \\ 4 & 3 & -1 \\ 5 & 4 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 0 \\ 4 & 3 & -1 \\ 13 & 10 & 0 \end{vmatrix} = -(-1)\begin{vmatrix} 2 & 1 \\ 13 & 10 \end{vmatrix} = 7 \neq 0.$$

De lo anterior se deduce que el sistema tiene solución única, la cual puede determinarse por cualquiera de los métodos estudiados en el capítulo 1. Lo esencial en este punto es que los escalares existen y por lo tanto, $\mathbf{v}$ es combinación lineal de los vectores $\mathbf{v}_1, \mathbf{v}_2$ y $\mathbf{v}_3$. A continuación se resuelve un sistema de ecuaciones lineales mediante la regla de Cramer.

$$detA_1 = \begin{vmatrix} 3 & 1 & 0 \\ 2 & 3 & -1 \\ 1 & 4 & 2 \end{vmatrix} = \begin{vmatrix} 3 & 1 & 0 \\ 2 & 3 & -1 \\ 5 & 10 & 0 \end{vmatrix} = -(-1)\begin{vmatrix} 3 & 1 \\ 5 & 10 \end{vmatrix} = 30 - 5 = 25.$$

De la misma manera,

$$detA_2 = \begin{vmatrix} 2 & 3 & 0 \\ 4 & 2 & -1 \\ 5 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 3 & 0 \\ 4 & 2 & -1 \\ 13 & 5 & 0 \end{vmatrix} = -(-1)\begin{vmatrix} 2 & 3 \\ 13 & 5 \end{vmatrix} = 10 - 39 = -29.$$

Nuevamente,

$$detA_3 = \begin{vmatrix} 2 & 1 & 3 \\ 4 & 3 & 2 \\ 5 & 4 & 1 \end{vmatrix} = \begin{vmatrix} -13 & -11 & 0 \\ -6 & -5 & 0 \\ 5 & 4 & 1 \end{vmatrix} = \begin{vmatrix} -13 & -11 \\ -6 & -5 \end{vmatrix} = \begin{vmatrix} 13 & 11 \\ 6 & 5 \end{vmatrix} = 65-66 = -1.$$

luego,

$$\begin{cases} k_1 = \dfrac{detA_1}{detA} = \dfrac{25}{7} \\[2ex] k_2 = \dfrac{detA_2}{detA} = -\dfrac{29}{7} \\[2ex] k_3 = \dfrac{detA_3}{detA} = -\dfrac{1}{7} \end{cases}$$

Finalmente, sustituyendo estos valores en la ecuación 2.5 se tiene

$$\mathbf{v} = \frac{25}{7}\mathbf{v}_1 - \frac{29}{7}\mathbf{v}_2 - \frac{1}{7}\mathbf{v}_3.$$

Ejemplo 2.3.2 *Determine si el vector* $\boldsymbol{v} = (4, -3, 1)$*, es combinación lineal de los vectores* $\boldsymbol{v}_1 = (2, 1, 5)$ *y* $\boldsymbol{v}_2 = (-1, 2, 3)$*.*

Solución: La ecuación 2.4 en este caso se convierte en,

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 = \mathbf{v}. \tag{2.7}$$

Sustituyendo los vectores en la ecuación 2.7 se obtiene,

$$k_1(2, 1, 5) + k_2(-1, 2, 3) = (4, -3, 1). \tag{2.8}$$

Desarrollando las operaciones indicadas en la ecuación 2.8 se obtiene el siguiente sistema de ecuaciones,

$$\left.\begin{aligned} 2k_1 - k_2 &= 4 \\ k_1 + 2k_2 &= -3 \\ 5k_1 + 3k_2 &= 1 \end{aligned}\right\}.$$

El sistema matricial equivalente es,

$$\begin{bmatrix} 2 & -1 \\ 1 & 2 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}$$

En este caso reduciremos la matriz aumentada del sistema mediante el método de eliminación de Gauss.

$$\left[\begin{array}{rr|r} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{array}\right] \sim \left[\begin{array}{rr|r} 1 & 2 & -3 \\ 2 & -1 & 4 \\ 5 & 3 & 1 \end{array}\right] \sim \left[\begin{array}{rr|r} 1 & 2 & -3 \\ 0 & -5 & 10 \\ 0 & -7 & 16 \end{array}\right]$$

$$\sim \left[\begin{array}{rr|r} 1 & 2 & -3 \\ 0 & 1 & -2 \\ 0 & -7 & 16 \end{array}\right] \sim \left[\begin{array}{rr|r} 1 & 2 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{array}\right].$$

El tercer renglón de la última matriz tiene su entrada principal en la columna de las constante, lo cual constituye una inconsistencia. Por lo anterior, no existen escalares k_1 y k_2 que satisfagan la ecuación 2.8, con lo cual concluimos que $\mathbf{v}$ no es combinación lineal de los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$.

2.3.1. Espacio generado

Definición 2.3.2 (Espacio generado) *Sea $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ un conjunto de r vectores en un espacio vectorial V, entonces el conjunto de todos los vectores en V, que son expresables como combinaciones lineales de los vectores en S, se conoce como espacio lineal generado o espacio generado y se denota por gen(S) o $\mathcal{L}(S)$.*

Teorema 2.3.1 *Sea $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ un conjunto de r vectores en un espacio vectorial V, entonces*

a) $\mathcal{L}(S)$ es un subespacio de V.

b) $\mathcal{L}(S)$ es el menor subespacio que contiene a S, en el sentido de que cualquier otro subespacio que contenga a S, también contiene a $\mathcal{L}(S)$.

Demostración: 1. Si $\mathbf{u}$ y $\mathbf{w} \in \mathcal{L}(S)$, entonces

$$\mathbf{u} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_r\mathbf{v}_r,$$

y

$$\mathbf{v} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \ldots + b_r\mathbf{v}_r,$$

donde las a_i y b_i son escalares. Además,

$$\mathbf{u}+\mathbf{v} = (a_1 + b_1)\mathbf{v}_1 + (a_2 + b_2)\mathbf{v}_2 + \ldots + (a_r + b_r)\mathbf{v}_r,$$

como las a_i y $b_i \in \mathbb{R}$, entonces $a_i + b_i \in \mathbb{R}$ $\forall i = 1, 2, \ldots r$. Así que $\mathbf{u}+\mathbf{v} \in \mathcal{L}(S)$.

2. Si $\mathbf{u} \in \mathcal{L}(S)$ y k es un escalar, entonces

$$k\mathbf{u} = (ka_1)\mathbf{v}_1 + (ka_2)\mathbf{v}_2 + \ldots + (ka_r)\mathbf{v}_r,$$

como k y $a_i \in \mathbb{R}$, entonces $ka_i \in \mathbb{R}$, $\forall i = 1, 2, \ldots, r$, así que, $k\mathbf{u} \in \mathcal{L}(S)$. Como $\mathcal{L}(S)$ es cerrado bajo la suma y la multiplicación por un escalar definidos sobre V, entonces $\mathcal{L}(S)$ es un subespacio de V.

b) En primer lugar verifiquemos que $S \subset \mathcal{L}(S)$.
Cada vector en S puede expresarse como,

$$\mathbf{v}_i = 0\mathbf{v}_1 + \ldots + 1.\mathbf{v}_i + \ldots + 0\mathbf{v}_r \quad \forall i = 1, 2, \ldots, r.$$

Por lo anterior, cada vector $\mathbf{v}_i \in \mathcal{L}(S)$, de manera que, $S \subset \mathcal{L}(S)$. Ahora, si W es un subespacio que contiene a S, entonces W contiene a todas las combinaciones lineales de los vectores en S, ya que, W es cerrado bajo la suma y la multiplicación por un escalar; es decir, $\mathcal{L}(S) \subseteq W$. ∎

2.3.2. Dependencia e independencia lineal

Definición 2.3.3 (Dependencia e independencia lineal) *Si $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ es un conjunto de r vectores en un espacio vectorial V, entonces S es linealmente independiente si la ecuación vectorial,*

$$k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \ldots + k_r\boldsymbol{v}_r = \boldsymbol{0}, \tag{2.9}$$

solo tiene la solución trivial; es decir, si $k_1 = k_2 = \ldots = k_r = 0$, es la única solución posible. En caso contrario, diremos que S es linealmente dependiente.

Ejemplo 2.3.3 *Sea $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ un conjunto de vectores en $\mathbb{R}^3$, donde $\boldsymbol{v}_1 = (1, -2, 3), \boldsymbol{v}_2 = (5, 6, -1)$ y $\boldsymbol{v}_3 = (3, 2, 1)$. Determine si S es linealmente independiente o linealmente dependiente.*

Slolución: Aplicando la ecuación 2.9 se tiene

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 = \mathbf{0}.$$

Sustituyendo los vectores se tiene

$$k_1(1,-2,3) + k_2(5,6,-1) + k_3(3,2,1) = (0,0,0). \tag{2.10}$$

Realizando las operaciones indicadas se obtiene el siguiente sistema de ecuaciones.

$$\begin{aligned} k_1 + 5k_2 + 3k_3 &= 0 \\ -2k_1 + 6k_2 + 2k_3 &= 0 \\ 3k_1 - k_2 + k_3 &= 0 \end{aligned}$$

El sistema matricial correspondiente es,

$$\begin{bmatrix} 1 & 5 & 3 \\ -2 & 6 & 2 \\ 3 & -1 & 1 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

La ecuación matricial $Ax = 0$ solo tiene la solución trivial si A es inversible; es decir, si $detA \neq 0$. Hallemos entonces $detA$.

$$detA = \begin{vmatrix} 1 & 5 & 3 \\ -2 & 6 & 2 \\ 3 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 5 & 3 \\ 0 & 16 & 8 \\ 0 & -16 & -8 \end{vmatrix} = \begin{vmatrix} 1 & 5 & 3 \\ 0 & 16 & 8 \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

Lo anterior implica que el sistema homogéneo tiene soluciones no triviales; por lo tanto, la ecuación 2.10 admite soluciones no triviales y el conjunto de vectores es linealmente dependiente.

Teorema 2.3.2 (El wronskiano) *Si $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n \in C^{n-1}(-\infty, \infty)$; es decir, las funciones tienen derivadas continuas de orden $(n-1)$, y si el wronskiano de estas funciones no es idénticamente cero sobre los reales, entonces las funciones son linealmente independiente. Donde el wronskiano es el siguiente determinante.*

$$w = \begin{vmatrix} \mathbf{f}_1(x) & \mathbf{f}_2(x) & \ldots & \mathbf{f}_n \\ \mathbf{f}_1'(x) & \mathbf{f}_2'(x) & \ldots & \mathbf{f}_n' \\ \mathbf{f}_1''(x) & \mathbf{f}_2''(x) & \ldots & \mathbf{f}_n'' \\ \vdots & \vdots & \ldots & \vdots \\ \mathbf{f}_1^{(n-1)}(x) & \mathbf{f}_2^{(n-1)}(x) & \ldots & \mathbf{f}_n^{(n-1)}. \end{vmatrix}.$$

Demostración: Supóngase que las funciones son linealmente dependientes, por lo tanto, existen escalares $k_1, k_2 \ldots, k_n$ no todos ceros tales que la ecuación,

$$k_1\mathbf{f}_1(x) + k_2\mathbf{f}_2(x) + \ldots + k_n\mathbf{f}_n(x) = \mathbf{0}, \tag{2.11}$$

se satisface, para todo $x \in \mathbb{R}$.
La ecuación anterior, junto con sus derivadas sucesivas hasta el orden $(n-1)$ nos proporcionan el siguiente sistema de ecuaciones.

$$\begin{aligned}
k_1\mathbf{f}_1(x) + k_2\mathbf{f}_2(x) + \ldots + k_n\mathbf{f}_n(x) &= \mathbf{0},\\
k_1\mathbf{f}_1^{'}(x) + k_2\mathbf{f}_2^{'}(x) + \ldots + k_n\mathbf{f}_n^{'}(x) &= \mathbf{0},\\
k_1\mathbf{f}_1^{''}(x) + k_2\mathbf{f}_2^{''}(x) + \ldots + k_n\mathbf{f}_n^{''}(x) &= \mathbf{0},\\
\vdots \qquad \vdots \qquad \vdots \qquad \vdots &= \mathbf{0},\\
k_1\mathbf{f}_1^{(n-1)}(x) + k_2\mathbf{f}_2^{(n-1)}(x) + \ldots k_n\mathbf{f}_n^{(n-1)}(x) &= \mathbf{0}.
\end{aligned}$$

En forma matricial el sistema se representa mediante,

$$\begin{bmatrix}
\mathbf{f}_1(x) & \mathbf{f}_2(x) & \ldots & \mathbf{f}_n(x)\\
\mathbf{f}_1^{'}(x) & \mathbf{f}_2^{'}(x) & \ldots & \mathbf{f}_n^{'}(x)\\
\mathbf{f}_1^{''}(x) & \mathbf{f}_2^{''}(x) & \ldots & \mathbf{f}_n^{''}(x)\\
\vdots & \vdots & \ddots & \vdots\\
\mathbf{f}_1^{(n-1)}(x) & \mathbf{f}_2^{(n-1)}(x) & \ldots & \mathbf{f}_n^{(n-1)}(x).
\end{bmatrix}
\begin{bmatrix} k_1\\ k_2\\ k_3\\ \vdots\\ k_n \end{bmatrix}
=
\begin{bmatrix} 0\\ 0\\ 0\\ \vdots\\ 0 \end{bmatrix}.$$

Como este sistema de ecuaciones tiene infinitas soluciones, entonces el determinante de la matriz de coeficientes debe valer cero para todo $x \in \mathbb{R}$; es decir,

$$w = \begin{vmatrix}
\mathbf{f}_1(x) & \mathbf{f}_2(x) & \ldots & \mathbf{f}_n\\
\mathbf{f}_1^{'}(x) & \mathbf{f}_2^{'}(x) & \ldots & \mathbf{f}_n^{'}\\
\mathbf{f}_1^{''}(x) & \mathbf{f}_2^{''}(x) & \ldots & \mathbf{f}_n^{''}\\
\vdots & \vdots & \ddots & \vdots\\
\mathbf{f}_1^{(n-1)}(x) & \mathbf{f}_2^{(n-1)}(x) & \ldots & \mathbf{f}_n^{(n-1)}.
\end{vmatrix} = 0, \qquad \forall x \in \mathbb{R}.$$

Por lo anterior, las funciones serán linealmente independientes si w no es idénticamente cero sobre todos los reales. ∎

Ejemplo 2.3.4 *Determine si las funciones senx y cosx son linealmente independientes o linealmente dependientes.*

Solución: Considérese el wronskiano de las funciones,

$$w = \begin{vmatrix} senx & cosx\\ cosx & -senx \end{vmatrix} = -\operatorname{sen}^2 x - \cos^2 x = -1 \neq 0.$$

Como el wronskiano no es idénticamente cero, entonces las funciones son linealmente independientes.

Teorema 2.3.3 *Si $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ es un conjunto de vectores en $\mathbb{R}^n$, con $r > n$, entonces S es linealmente dependientes.*

Demostración: Los vectores en S pueden expresarse como,
$\mathbf{v}_i = (a_{1i}, a_{2i}, \ldots, a_{ni})$, para todo $1 \leq i \leq r$.

La ecuación 2.9 en este caso se convierte en,

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r = \mathbf{0}. \tag{2.12}$$

Sustituyendo los vectores dados y realizando las operaciones se obtiene el siguiente sistema de ecuaciones,

$$\left.\begin{array}{lcl} a_{11}k_1 + a_{12}k_2 + \ldots + a_{1r}k_r & = & 0 \\ a_{21}k_1 + a_{22}k_2 + \ldots + a_{2r}k_r & = & 0 \\ \vdots \qquad\qquad \vdots \qquad\qquad\quad \vdots & = & \vdots \\ a_{n1}k_1 + a_{n2}k_2 + \ldots + a_{nr}k_r & = & 0 \end{array}\right\}.$$

En forma matricial,

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nr} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

El sistema homogéneo anterior contiene n ecuaciones y r incógnitas, como $r > n$, entonces el sistema tiene más incógnitas que ecuaciones. Anteriormente probamos que un sistema homogéneo con más incógnitas que ecuaciones tiene infinitas soluciones (teorema 1.1.3); de manera que la ecuación 2.12 tiene soluciones no triviales; por lo tanto, S es un conjunto linealmente dependiente. ∎

2.4. Bases de un Espacio Vectorial

Definición 2.4.1 (Base) *Si V es un espacio vectorial cualquiera y $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ es un conjunto finito de vectores en V, entonces S se denomina base para V si,*

a) *S es linealmente independiente,*

b) *S genera a V.*

Ejemplo 2.4.1 (Base canónica de $\mathbb{R}^n$) *El conjunto $S = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ se conoce como base canónica de $\mathbb{R}^n$, donde*

$$\begin{aligned} \boldsymbol{e}_1 &= (1, 0, \ldots, 0), \\ \boldsymbol{e}_2 &= (0, 1, 0, \ldots, 0), \\ \vdots &= \vdots \\ \boldsymbol{e}_n &= (0, 0, \ldots, 1). \end{aligned}$$

Pruebe que efectivamente S es una base para $\mathbb{R}^n$.

Solución:

a) Verifiquemos que S es linealmente independiente mediante la ecuación,

$$k_1\mathbf{e}_1 + k_2\mathbf{e}_2 + \ldots k_n\mathbf{e}_n = \mathbf{0}. \tag{2.13}$$

Sustituyendo los vectores,

$$k_1(1, 0, 0, \ldots, 0) + k_2(0, 1, 0 \ldots, 0) + \ldots + k_n(0, 0, 0, \ldots, 1) = (0, 0, 0, \ldots, 0).$$

Realizando las operaciones,

$$(k_1, k_2, \ldots, k_n) = (0, 0, \ldots, 0).$$

Por lo anterior, $k_1 = k_2 = \ldots = k_n = 0$. Debido a que la ecuación 2.13 solo admite la solución trivial, entonces S es linealmente independiente.

b) Ahora considérese un elemento cualquiera de $\mathbb{R}^n$. Si $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, se requiere investigar si $\mathbf{x}$ se puede expresar como combinación lineal de los vectores en S, en cuyo caso, S generaría a $\mathbb{R}^n$. Considérese la ecuación,

$$k_1\mathbf{e}_1 + k_2\mathbf{e}_2 + \ldots k_n\mathbf{e}_n = \mathbf{x}. \tag{2.14}$$

Sustituyendo los vectores en la ecuación anterior,

$$k_1(1, 0, \ldots, 0) + k_2(0, 1 \ldots, 0) + \ldots + k_n(0, 0, \ldots, 1) = (x_1, x_2, \ldots, x_n).$$

Realizando las operaciones indicadas se obtiene,

$$(k_1, k_2, \ldots, k_n) = (x_1, x_2, \ldots, x_n).$$

La anterior significa que la ecuación 2.14 tiene la solución $k_1 = x_1, k_2 = x_2, \ldots, k_n = x_n$; por lo tanto, $\mathbf{x}$ es combinación lineal de los vectores en S, de manera que S genera a $\mathbb{R}^n$.

En conclusión, como S es linealmente independiente y genera a $\mathbb{R}^n$, entonces es una base para $\mathbb{R}^n$.

Ejemplo 2.4.2 *Pruebe que* $S = \{M_1, M_2, M_3, M_4\}$ *es una base para* $M_{2x2}(\mathbb{R})$ *donde,*

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, M_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad y \quad M_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Solución:

i) Verifiquemos que S es linealmente independiente considerando la ecuación,

$$k_1M_1 +_2 M_2 + k_3M_3 + k_4M_4 = 0. \tag{2.15}$$

Si esta ecuación solo admite la solución trivial, entonces el conjunto S será linealmente independiente.
Sustituyendo las matrices del conjunto se obtiene,

$$k_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + k_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + k_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + k_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

realizando las operaciones,

$$\begin{bmatrix} k_1 & k_2 \\ k_3 & k_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Lo anterior implica que la única solución posible para la ecuación 2.15 es la solución trivial $k_1 = k_2 = k_3 = k_4 = 0$; por lo tanto, el conjunto es linealmente independiente.

ii) Considérese una matriz cualquiera $A \in M_{2x2}$. En este caso,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Si A es combinación lineal de las matrices en S, entonces S genera a M_{2x2}. Para verificar lo anterior, considérese la ecuación,

$$k_1M_1 +_2 M_2 + k_3M_3 + k_4M_4 = A. \tag{2.16}$$

Si existen escalares que satisfagan la ecuación anterior, entonces A será combinación lineal de los vectores en S. Sustituyendo las matrices en la ecuación 2.16 se obtiene,

$$k_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + k_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + k_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + k_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Realizando las operaciones se tiene,

$$\begin{bmatrix} k_1 & k_2 \\ k_3 & k_4 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Lo anterior significa que $k_1 = a_{11}, k_2 = a_{12}, k_3 = a_{21}$ y $k_4 = a_{22}$ es una solución de la ecuación 2.16. De modo que A es combinación lineal de los vectores en S, y por lo tanto, S genera a M_{2x2}.

En conclusión como S genera a M_{2x2} y es linealmente independiente, entonces S es una base para M_{2x2}.

Ejemplo 2.4.3 (Base canónica de P_n) *El conjunto $S = \{1, x, x^2, \ldots, x^n\}$ se conoce como base canónica de P_n, Pruebe que efectivamente S es una base para P_n.*

Solución:

a) Para probar que el conjunto S es linealmente independiente recurriremos al wronskiano (teorema 2.3.2).

$$w = \begin{vmatrix} 1 & x & x^2 & \ldots & x^n \\ 0 & 1! & 2x & \ldots & nx^{n-1} \\ 0 & 0 & 2! & \ldots & n(n-1)x^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & n!. \end{vmatrix} = 1!2!\ldots n!$$

Como el determinante no es idénticamente cero, entonces S es linealmente independiente.

b) Verifiquemos que S genere a V. Considérese un elemento $p(x) \in P_n$; es decir,

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n, \quad \text{donde} \quad a_i \in \mathbb{R}, \quad \forall i = 0, \ldots, n. \quad (2.17)$$

Ahora considérese la ecuación,

$$k_0(1) + k_1 x + k_2 x^2 \ldots + k_n x^n = p(x), \quad (2.18)$$

si existen los escalares $k_i \in \mathbb{R}, \quad \forall i = 1, 2, \ldots, n$, entonces $p(x)$ será combinación lineal de los vectores en S, y por lo tanto, S generará a P_n.

Sustituyendo la ecuación 2.17 en la 2.18 se obtiene,

$$k_0(1) + k_1x + k_2x^2 \ldots + k_nx^n = a_0 + a_1x + \ldots + a_nx^n, \qquad (2.19)$$

Agrupando términos se tiene,

$$(k_0 - a_0) + (k_1 - a_1)x + (k_2 - a_2)x^2 \ldots + (k_n - a_n)x^n = 0. \qquad (2.20)$$

La ecuación 2.20 es idénticamente cero para todo $x \in \mathbb{R}$ si y solo si los coeficientes son idénticamente cero, ya que los vectores en S son linealmente independientes, por lo tanto, $k_0 = a_0, k_1 = a_1, k_2 = a_2, \ldots, k_n = a_n$. Por lo anterior, $p(x)$ es combinación lineal de los vectores en S, de manera que S genera a P_n.
Se concluye que S es una base para P_n.

2.4.1. Dimensión de un espacio vectorial

Definición 2.4.2 (Espacios de dimensión finita) *Se dice que un espacio vectorial V distinto del espacio cero, es de dimensión finita, si contiene un conjunto finito de vectores* $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ *que conforman una base para este espacio. Si no existe un conjunto de este tipo, se dice que V es de dimensión infinita. El espacio vectorial cero se considera de dimensión finita.*

Teorema 2.4.1 *Si* $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ *es una base para un espacio vectorial V, entonces todo conjunto en V con más de n vectores es linealmente dependiente.*

Demostración: Sea $W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ un conjunto de m vectores en V tales que $m > n$, dado que S es una base para V, entonces genera a V. En particular los vectores en W pueden expresarse como,

$$\mathbf{w}_i = a_{1i}\mathbf{v}_1 + a_{2i}\mathbf{v}_2 + \ldots + a_{ni}\mathbf{v}_n, \qquad 1 \le i \le m. \qquad (2.21)$$

Considérese la ecuación,

$$k_1\mathbf{w}_1 + k_2\mathbf{w}_2 + \cdots + k_m\mathbf{w}_m = \mathbf{0}. \qquad (2.22)$$

Sustituyendo las ecuaciones 2.21 en la ecuación 2.22 se tiene lo siguiente,

$$\begin{aligned}
k_1\mathbf{w}_1 + k_2\mathbf{w}_2 + \cdots + k_m\mathbf{w}_m &= k_1(\,a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + \ldots + a_{n1}\mathbf{v}_n) \\
&\quad +k_2(\,a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \ldots + a_{n2}\mathbf{v}_n) \\
&\quad \vdots \qquad\qquad\qquad\qquad\qquad \vdots \\
&\quad +k_m(\,a_{1m}\mathbf{v}_1 + a_{2m}\mathbf{v}_2 + \ldots + a_{nm}\mathbf{v}_n) \\
&= \mathbf{0}.
\end{aligned}$$

Reagrupando términos,

$$(\,k_1a_{11} + k_2a_{12} + \ldots + k_ma_{1m}\,)\mathbf{v}_1 + (\,k_1a_{21} + k_2a_{22} + \ldots + k_ma_{2m}\,)\mathbf{v}_2 + \ldots$$

$$+(\,k_1a_{n1} + k_2a_{n2} + \ldots + k_ma_{nm}\,)\mathbf{v}_n = \mathbf{0}.$$

Como los vectores en S son linealmente independientes, entonces los escalares en la ecuación anterior son todos iguales a cero; es decir,

$$\begin{aligned} k_1a_{11} + k_2a_{12} + \ldots + k_ma_{1m} &= 0 \\ k_1a_{21} + k_2a_{22} + \ldots + k_ma_{2m} &= 0 \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \quad &= \vdots \\ k_1a_{n1} + k_2a_{n2} + \ldots + k_ma_{nm} &= 0 \end{aligned}$$

En forma matricial,

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Lo anterior representa un sistema homogéneo con m incógnitas y n ecuaciones. Como $m > n$, entonces tenemos más incógnitas que ecuaciones, así que, de acuerdo con el teorema 1.1.3, el sistema tiene infinitas soluciones. Finalmente, la ecuación 2.22 admite soluciones no triviales; por lo tanto, W es linealmente dependiente. ∎

Teorema 2.4.2 *Dos bases cualesquiera para un espacio vectorial de dimensión finita tienen el mismo número de vectores.*

Demostración: Sean $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots,\ \mathbf{v}_n\}$ y $S' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ dos bases del mismo espacio vectorial V, entonces

a) Por el teorema anterior, como S es linealmente independiente y S' es una base para V, entonces $n \leq m$.

b) Por el teorema anterior, como S' es linealmente independiente y S es una base para V, entonces $m \leq n$.

La única posibilidad de que las dos relaciones anteriores se cumplan simultáneamente es que $m = n$. ∎

Definición 2.4.3 (Dimensión) *La dimensión de un espacio vectorial V de dimensión finita se define como el número de vectores en una base para V. Por definición el espacio vectorial cero es de dimensión cero. Denotaremos la dimensión de un espacio vectorial V mediante dim(V).*

Por el teorema anterior sabemos que cualquier base de $\mathbb{R}^n$ debe contener n vectores, ya que es el número de vectores en su base canónica. De la misma manera, toda base para P_n debe contener exactamente $n+1$ vectores.

Teorema 2.4.3 *Sea $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \ \boldsymbol{v}_r\}$, con $r \geq 2$, un conjunto de vectores en un espacio vectorial V, entonces los vectores en S son linealmente dependientes, si y solo si, al menos uno de ellos es combinación lineal de los vectores restantes. De manera más precisa, si $\boldsymbol{v}_1$ no es el vector cero, entonces los vectores son linealmente dependientes, si y solo si, para algún $j > 1$, $\boldsymbol{v}_j$ es combinación lineal de los vectores previos.*

Demostración: $\Rightarrow$) Si los vectores en S son linealmente dependiente, entonces la ecuación,

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r \ \mathbf{v}_r = \mathbf{0},$$

admite soluciones no triviales; es decir, algunos de los escalares son diferentes de cero. Se l el mayor índice tal que $k_l \neq 0$, entonces

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_l \ \mathbf{v}_l = \mathbf{0}.$$

Despejando $\mathbf{v}_l$ se tiene,

$$\mathbf{v}_l = -(\frac{k_1}{k_l})\mathbf{v}_1 - (\frac{k_2}{k_l})\mathbf{v}_2 \ldots - (\frac{k_{l-1}}{k_l})\mathbf{v}_{l-1}.$$

La ecuación anterior prueba que $\mathbf{v}_l$ es combinación lineal de los vectores que le anteceden.
$\Leftarrow$) Si $\mathbf{v}_i$ es combinación lineal del resto de los vectores; es decir,

$$\mathbf{v}_i = \sum_{j \neq i} c_j\mathbf{v}_j,$$

por lo anterior,

$$\sum_{j \neq i} c_j\mathbf{v}_j - \mathbf{v}_i = \mathbf{0}.$$

Finalmente, la sucesión de escalares, $c_1, \ldots, c_{i-1}, -1, c_{i+1}, \ldots, c_r$ es una solución no trivial de la ecuación,

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r \ \mathbf{v}_r = \mathbf{0},$$

por lo tanto, S es linealmente dependiente. ∎

Teorema 2.4.4 *Si $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \ \boldsymbol{v}_r\}$, genera un espacio vectorial V y si $\boldsymbol{v}_i$ es combinación lineal del resto de los vectores en S, entonces $S' = S - \{\boldsymbol{v}_i\}$ también genera a V. De manera más precisa, $\mathcal{L}(S) = \mathcal{L}(S - \{\boldsymbol{v}_i\}) = \mathcal{L}(S')$*

Demostración: Sean $V = \mathcal{L}(S)$ y $V' = \mathcal{L}(S')$ probaremos que $V = V'$.
$V \subseteq V'$) Si $\mathbf{v} \in V$, entonces

$$\begin{aligned} \mathbf{v} &= \sum_{j=1}^{r} k_j \mathbf{v}_j, \\ &= \sum_{j \neq i} k_j \mathbf{v}_j + k_i \mathbf{v}_i. \end{aligned} \tag{2.23}$$

Por otro lado,

$$\mathbf{v}_i = \sum_{j \neq i} c_j \mathbf{v}_j. \tag{2.24}$$

Sustituyendo la ecuación 2.24 en la ecuación 2.23 se obtiene,

$$\begin{aligned} \mathbf{v} &= \sum_{j \neq i} k_j \mathbf{v}_j + k_i \sum_{j \neq i} c_j \mathbf{v}_j, \\ &= \sum_{j \neq i} (k_j + k_i c_j) \mathbf{v}_j \end{aligned} \tag{2.25}$$

Si se hace $l_j = k_j + k_i c_j$ y se sustituye en la ecuación 2.25 se obtiene,

$$\mathbf{v} = \sum_{j \neq i} l_j \mathbf{v}_j. \tag{2.26}$$

La ecuación 2.26 prueba que $\mathbf{v}$ es combinación lineal de los vectores en S', por lo tanto, $\mathbf{v} \in V'$, luego $V \subseteq V'$.
$V \subseteq V'$) Si $\mathbf{v}' \in V'$, entonces

$$\mathbf{v}' = \sum_{j \neq i} k_j \mathbf{v}_j,$$

luego,

$$\mathbf{v}' = \sum_{j \neq i} k_j \mathbf{v}_j + 0 \mathbf{v}_i. \tag{2.27}$$

La ecuación 2.27 muestra que $\mathbf{v}'$ es combinación lineal de los vectores en S, es decir, $\mathbf{v}' \in V$, luego, $V' \subseteq V$. Finalmente se concluye que $V = V'$. ∎

Corolario 2.4.1 *Si $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots,\ \boldsymbol{v}_r\}$ genera un espacio vectorial V y $\boldsymbol{w} \in V$, entonces $S' = \{\boldsymbol{w}, \boldsymbol{v}_1, \boldsymbol{v}_2, \dots,\ \boldsymbol{v}_r\}$ también genera a V.*

Demostración: Sea $V' = \mathcal{L}(S')$, entonces, como $\mathbf{w}$ es combinación lineal del resto de los vectores en S', por el teorema anterior,

$$V' = \mathcal{L}(S') = \mathcal{L}(S' - \{\mathbf{w}\}) = \mathcal{L}(S) = V. \quad \blacksquare$$

Teorema 2.4.5 *Si $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots,\ \boldsymbol{v}_r\}$ genera un espacio vectorial V de dimensión n y si $S' = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots,\ \boldsymbol{w}_k\}$ es un subconjunto de V linealmente independiente, entonces $k \leq r$.*

Demostración: Considérese el conjunto $\{\mathbf{w}_1, \mathbf{v}_1, \mathbf{v}_2, \dots,\ \mathbf{v}_r\}$. Como $\mathbf{w}_1$ es combinación lineal de los vectores restantes en el conjunto, entonces este conjunto es linealmente dependiente (teorema 2.4.3). Además, este conjunto sigue generando a V (corolario 2.4.4). Como este conjunto es linealmente dependiente, el teorema 2.4.3 garantiza que existe un $i > 1$, tal que $\mathbf{v}_i$ es combinación lineal de los vectores que le preceden. Podemos extraer este vector del conjunto y el conjunto que nos queda sigue generando a V (teorema 2.4.4); es decir, $\{\mathbf{w}_1, \mathbf{v}_1, \dots, \mathbf{v}_{i-1},\ \mathbf{v}_{i+1}, \dots,\ \mathbf{v}_r\}$ sigue generando a V. Considérese ahora el conjunto $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{v}_1, \dots, \mathbf{v}_{i-1},\ \mathbf{v}_{i+1}, \dots,\ \mathbf{v}_r\}$. Argumentando de la misma manera que en la primera parte de esta demostración, concluimos que este conjunto es linealmente dependiente y que genera a V. Además, podemos extraer un vector $\mathbf{v}_j$ y el conjunto que resulta seguirá generando a V. Este proceso de introducir un vector $\mathbf{w}$ y extraer un vector $\mathbf{v}$ puede realizarse de manera sucesiva. Sin embargo, si suponemos que $k > r$, llegaremos a un punto en el cual introduciremos un vector $\mathbf{w}_l$ y no quedarán vectores $\mathbf{v}$ para extraer y el conjunto $\{\mathbf{w}_1, \mathbf{w}_1, \mathbf{w}_2, \dots,\ \mathbf{w}_l\}$ será por lo anteriormente argumentado linealmente dependiente, lo cual constituye una contradicción a la hipótesis inicial. La contradicción se genera al suponer que $k > r$, por lo tanto, $k \leq r$. $\blacksquare$

Corolario 2.4.2 *Si un espacio vectorial V está generado por m elementos y V tiene dimensión n, entonces $n \leq m$.*

Demostración: Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots,\ \mathbf{v}_n\}$ una base para V, como este conjunto es linealmente independiente, entonces, por el teorema anterior debe cumplirse que,

$$n \leq m. \quad \blacksquare$$

Teorema 2.4.6 *Si $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots,\ \boldsymbol{v}_n\}$ es un conjunto de n vectores en un espacio finito de dimensión n, entonces se cumple que:*

a) Si S es linealmente independiente, entonces S es una base para V.

b) Si S genera a V, entonces S es una base para V.

Demostración: a) Dado que S es linealmente independiente, debemos probar que genera a V para que sea una base de este espacio. Supóngase que S no genera a V, en este caso debe existir al menos un vector $\mathbf{w}$, tal que $\mathbf{w}$ no puede ser expresado como combinación lineal de los vectores en S. Ahora considérese el conjunto

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \ \mathbf{v}_n, \mathbf{w}\}$$

Este conjunto es linealmente dependiente, ya que contiene más vectores que cualquier base para V (teorema 2.4.1). El teorema 2.4.3 garantiza que existe un vector en el conjunto que es combinación lineal de los vectores previos. Sin embargo, al observar el conjunto vemos que este resultado no se satisface; es decir, tenemos una clara contradicción. Así que S debe generar a V y por lo tanto, es una base para este espacio.
b)Dado que S genera V, debemos probar que S es linealmente independiente, para que S sea una base para este espacio. Supóngase que S es linealmente dependiente. En este caso existe un $i > 1$, tal que $\mathbf{v}_i$ es combinación lineal de los vectores que le anteceden. Nuevamente, podemos extraer este vector del conjunto y el conjunto que resulta seguirá generando a V. Es decir,

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \ \mathbf{v}_n\}$$

Este conjunto contiene $n - 1$ elementos y genera un espacio de dimensión n. Lo anterior constituye una contradicción al corolario 2.4.2, la cual se origina de suponer que S es linealmente dependiente. Por lo anterior, concluimos que S es linealmente independiente y en consecuencia forma una base para V. ∎

Teorema 2.4.7 *Sea $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \ \boldsymbol{v}_r\}$ un conjunto de vectores que no contiene al vector cero y que genera un espacio vectorial de dimensión n, entonces puede extraerse de S una base para V.*

Demostración: Si S es linealmente independiente, entonces no hay nada que probar pues S sería una base para V. Supóngase que S es linealmente dependiente. en este caso, debe existir un $i > 1$ tal que $\mathbf{v}_i$ es combinación lineal de los vectores previos (teorema 2.4.3). Además por el teorema 2.4.4 podemos extraer este vector y el conjunto

$$\{\mathbf{v}_1, \ldots, \ \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots \ \mathbf{v}_r\}$$

seguirá generando a V. Si este conjunto es linealmente independiente habremos terminado, en caso caso contrario, aplicaremos el proceso anterior de manera sucesiva hasta obtener un conjunto linealmente independiente (como el conjunto original no contiene al vector cero, este punto está garantizado). ∎

Teorema 2.4.8 *Sea $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \ \boldsymbol{v}_r\}$ un conjunto de vectores linealmente independiente en un espacio vectorial V de dimensión n, entonces existe una base de V que contiene S.*

Demostración: Sea $S' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \ \mathbf{w}_n\}$ una base para V, entonces el conjunto

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$$

genera a V y es linealmente dependiente. (teorema 2.4.3) y corolario 2.4.4). Procediendo como en el teorema 2.4.5, al final del proceso obtendremos una base para V que contiene a S. ∎

Teorema 2.4.9 *Si $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \ \boldsymbol{v}_n\}$ es una base de un espacio vectorial V, entonces para todo vector $\boldsymbol{v} \in V$ existe un conjunto único de escalares $a_1, a_2, \ldots, a_n$ tales que,*

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \ldots + a_n\boldsymbol{v}_n.$$

Demostración: Como $\mathbf{v} \in V$ y S es una base para V, entonces existen los escalares $a_1, a_2, \ldots, a_n$, tales que,

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n. \tag{2.28}$$

Para probar que el conjunto de escalares es único supondremos que existe un segundo conjunto de escalares $b_1, b_2, \ldots, b_n$, tales que,

$$\mathbf{v} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \ldots + b_n\mathbf{v}_n. \tag{2.29}$$

Restando las ecuaciones 2.28 y 2.29,

$$\mathbf{0} = (a_1 - b_1)\mathbf{v}_1 + (a_2 - b_2)\mathbf{v}_2 + \ldots + (a_n - b_n)\mathbf{v}_n. \tag{2.30}$$

Como los vectores en S son linealmente independientes, entonces la ecuación 2.30 solo se satisface si los escalares son todos iguales a cero; es decir, $a_1 - b_1 = 0$, $a_2 - b_2 = 0$, $\ldots$, $a_n - b_n = 0$. De lo anterior concluimos que $a_1 = b_1$, $a_2 = b_2$, $\ldots$,$a_n = b_n$, lo cual significa que el conjunto de escalares es único. ∎

2.4.2. Rango y nulidad de una matriz

Espacio de renglores y espacio de columnas

Considérese una matriz general A de tamaño mxn.

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

Definición 2.4.4 (Vectores renglón y espacio de renglones) *Los vectores en $\mathbb{R}^n$ formados con los elementos de los renglones de A se conocen como vectores renglón de A; es decir, los vectores renglón para la matriz A son:*

$$\boldsymbol{r}_i = (a_{i1}, a_{i1}, \dots, a_{in}) \text{ para todo } 1 \leq i \leq m.$$

El conjunto de todos los vectores renglón de A lo denotaremos por S_r. El espacio generado por S_r; es decir, $\mathcal{L}(S_r)$, se conoce como el espacio de renglones.

Definición 2.4.5 (Vectores columna y espacio de columnas) *Los vectores en $\mathbb{R}^m$ formados con los elementos de las columnas de A se conocen como vectores columna de A; es decir, los vectores columna para la matriz A son:*

$$\boldsymbol{r}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \qquad 1 \leq j \leq n.$$

El conjunto de todos los vectores columna de A lo denotaremos por S_c. El espacio generado por S_c; es decir, $\mathcal{L}(S_c)$, se conoce como el espacio de columnas.

Teorema 2.4.10 *Las operaciones elementales sobre los renglones no cambian el espacio de renglones de una matriz.*

Demostración: Considérese que la matriz B se obtiene de la matriz A mediante la realización de una sola operación elemental. Además, sean $S_r = \{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m\}$ y $S'_r = \{\mathbf{r}'_1, \mathbf{r}'_2, \dots, \mathbf{r}'_m\}$ los vectores renglón de A y de B, respectivamente, entonces debemos probar que $\mathcal{L}(S_r) = \mathcal{L}(S'_r)$. Para continuar consideraremos los tres casos de operaciones elementales.
i) Si $B = E_{ij}A$; es decir, si B se obtiene de la matriz A al intercambiar los

renglones i y j, entonces $S_r = S'_r$, ya que simplemente se intercambian de posiciones dos vectores; por lo tanto, $\mathcal{L}(S_r) = \mathcal{L}(S'_r)$.
ii) Si $B = E_i(k)A$; es decir, si B se obtiene al multiplicar el renglón i de la matriz A por el escalar $k \neq 0$.. En este caso

$$S_r = \{\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \ \mathbf{r}_m\}$$

y

$$S'_r = \{\mathbf{r}_1, \ldots, k\mathbf{r}_i, \ldots, \ \mathbf{r}_m\}$$

considérese el conjunto

$$S''_r = \{\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \ \mathbf{r}_m, k\mathbf{r}_i\}$$

Como el último vector en el conjunto es combinación lineal de los vectores que le anteceden, entonces

$$\mathcal{L}(S''_r) = \mathcal{L}(S''_r - \{k\mathbf{r}_i\}) = \mathcal{L}(S_r). \tag{2.31}$$

Reordenando los vectores en S''_r,

$$S''_r = \{\mathbf{r}_1, \ldots, k\mathbf{r}_i, \ldots, \ \mathbf{r}_m, \mathbf{r}_i\}$$

Como el último vector en el conjunto es combinación lineal de los vectores que le anteceden, entonces

$$\mathcal{L}(S''_r) = \mathcal{L}(S''_r - \{\mathbf{r}_i\}) = \mathcal{L}(S'_r). \tag{2.32}$$

De las ecuaciones 2.31 y 2.32 concluimos que $\mathcal{L}(S_r) = \mathcal{L}(S'_r)$.
iii) Si $B = E_{ij}(k)A$; es decir, si B se obtiene de A al sumar k veces el renglón i al renglón j. En este caso

$$S_r = \{\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_j, \ldots, \ \mathbf{r}_m\}$$

y

$$S'_r = \{\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, k\mathbf{r}_i + \mathbf{r}_j, \ldots, \ \mathbf{r}_m\}$$

considérese el conjunto

$$S''_r = \{\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_j, \ldots, \ \mathbf{r}_m, k\mathbf{r}_i + \mathbf{r}_j\}$$

Como el último vector en el conjunto es combinación lineal de los vectores que le anteceden, entonces

$$\mathcal{L}(S''_r) = \mathcal{L}(S''_r - \{k\mathbf{r}_i + \mathbf{r}_j\}) = \mathcal{L}(S_r). \tag{2.33}$$

Reordenando los vectores en S_r'',

$$S_r'' = \{\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, k\mathbf{r}_i + \mathbf{r}_j, \ldots,\ \mathbf{r}_m, \mathbf{r}_j\}.$$

Ahora sea,

$$\mathbf{r}_j' = k\mathbf{r}_i + \mathbf{r}_j \tag{2.34}$$

Despejando $\mathbf{r}_j$ en la ecuación anterior,

$$\mathbf{r}_j = \mathbf{r}_j' - k\mathbf{r}_i \tag{2.35}$$

La ecuación 2.35 muestra que el último vector en S_r'' es combinación lineal de los vectores que le anteceden, así que,

$$\mathcal{L}(S_r'') = \mathcal{L}(S_r'' - \{\mathbf{r}_j\}) = \mathcal{L}(S_r'). \tag{2.36}$$

De las ecuaciones 2.33 y 2.36 concluimos que $\mathcal{L}(S_r) = \mathcal{L}(S_r')$. ∎

Teorema 2.4.11 *Los vectores renglón diferentes de cero en la forma escalonada en los renglones de una matriz forman una base para el espacio de renglones.*

Demostración: Sea A una matriz cualquiera de tamaño mxn con vectores renglón $S_r = \{\mathbf{r}_1, \mathbf{r}_2, \ldots,\ \mathbf{r}_m\}$; además, sea B la forma escalonada en los renglones de la matriz A con vectores renglón $S_r' = \{\mathbf{r}_1', \mathbf{r}_2', \ldots,\ \mathbf{r}_m'\}$. Ahora, sea l el mayor índice tal que, $\mathbf{r}_l' \neq \mathbf{0}$, entonces $S_r'' = \{\mathbf{r}_1', \mathbf{r}_2', \ldots,\ \mathbf{r}_l'\}$ es el conjunto de vectores diferentes de cero en la matriz B. Debemos probar que el conjunto S_r'' es una base para el espacio de renglones de la matriz B. Por el teorema 2.4.10,

$$\mathcal{L}(S_r) = \mathcal{L}(S_r'),$$

además, por el teorema 2.4.4,

$$\mathcal{L}(S_r') = \mathcal{L}(S_r' - \{\mathbf{0}\}) = \mathcal{L}(S_r'').$$

Lo anterior prueba que S_r'' genera al espacio de renglones de la matriz A. Solo resta probar que S_r'' es un conjunto linealmente independiente. Supóngase que S_r'' es linealmente dependiente. En este caso, el teorema 2.4.3 nos garantiza que exite un $1 < i \leq l$, tal que $\mathbf{r}_i$ es combinación lineal de los vectores previos; es decir,

$$\mathbf{r}_i = k_1\mathbf{r}_1 + k_2\mathbf{r}_2 + \ldots + k_{i-1}\mathbf{v}_{li-1}$$

Este vector, que corresponde a un renglón en la matriz B, puede reducirse a un renglón cero mediante la siguiente secuencia de operaciones elementales sobre B.

$$E_{1,i-1}(-k_{i-1}) \ldots E_{2i}(-k_2)E_{1i}(-k_1)B.$$

Lo anterior constituye una contradicción, ya que B se encuentra en la forma escalonada en los renglones y no podemos generar en ella renglones cero mediante operaciones elementales. La contradicción surge de suponer que S_r'' es linealmente dependiente; por lo tanto, concluimos que es linealmente independiente y en consecuencia, es una base para el espacio de renglones de la matriz A. ∎

Ejemplo 2.4.4 *Encuentre una base para el espacio generado por los vectores*

$$\begin{aligned}
\boldsymbol{v}_1 &= (1,0,4,3,-1,1),\\
\boldsymbol{v}_2 &= (-5,2,-2,-1,-1,-2),\\
\boldsymbol{v}_3 &= (-3,2,2,2,-2,-1),\\
\boldsymbol{v}_4 &= (-1,2,14,11,-5,2).
\end{aligned}$$

Solución: Los vectores dados constituyen los renglones de la siguiente matriz A, de manera que el espacio generado por los vectores corresponda al espacio de renglones de la matriz.

$$A = \begin{bmatrix} 1 & 0 & 4 & 3 & -1 & 1 \\ -5 & 2 & -2 & -1 & -1 & -2 \\ -3 & 2 & 2 & 2 & -2 & -1 \\ -1 & 2 & 14 & 11 & -5 & 2 \end{bmatrix}.$$

Mediante operaciones elementales se obtiene la forma escalonada en los renglones (ver el ejemplo 1.1.10),

$$\begin{bmatrix} 1 & 0 & 4 & 3 & -1 & 1 \\ 0 & 1 & 9 & 7 & -3 & \frac{3}{2} \\ 0 & 0 & 1 & \frac{3}{4} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Por el teorema anterior, los vectores renglón diferentes de cero en la forma escalonda, constituyen una base para el espacio de renglones de la matriz A. Estos vectores son:

$$\begin{aligned}
\mathbf{u}_1 &= (1,0,4,3,-1,1),\\
\mathbf{u}_2 &= (0,1,9,7,-3,\frac{3}{2}),\\
\mathbf{u}_3 &= (0,0,1,\frac{3}{4},-\frac{1}{4},\frac{1}{4}).
\end{aligned}$$

Ejemplo 2.4.5 *Encuentre una base para el espacio de columnas de la matriz.*

$$A = \begin{bmatrix} 1 & 2 & -2 \\ 2 & 4 & 6 \\ 0 & 0 & -20 \end{bmatrix}$$

Solución: En este caso, realizaremos operaciones elementales sobre los renglones de matriz transpuesta, ya que éstos corresponden a las columnas de la matriz dada.

$$A^T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ -2 & 6 & -20 \end{bmatrix}.$$

Llevaremos la matriz anterior a su forma escalonada en los renglones,

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ -2 & 6 & -20 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 10 & -20 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 \\ 0 & 10 & -20 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}.$$

La matriz anterior se encuentra en su forma escalonada en los renglones. Para obtener los vectores de la base para el espacio de columnas, debemos obtener la transpuesta de esta matriz.

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

Los vectores columna diferentes del vector cero formarán una base para el espacio de columnas. Finalmente, los vectores de la base son:

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \quad y \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}.$$

Teorema 2.4.12 *Si W es un subespacio de un espacio vectorial V de dimensión finita, entonces $dim(W) \leq dim(V)$.*

Demostración: Sea $dim(V) = n$, entonces V es generado por n vectores que componen cualquiera de sus bases. Si $W = \{\mathbf{0}\}$ por definición $dim(W) = 0$ y la desigualdad se cumple de manera inmediata. Supóngase ahora que W es un subespacio propio de V(subespacios distintos a V y al espacio cero). Sea $\mathbf{v}_1$ un elemento de W, entonces considérese el espacio $W_1 = \mathcal{L}\{\mathbf{v}_1\}$. Si

$W = W_1$, entonces $dim(V) = 1$. Si $W \neq W_1$, entonces existe un vector $\mathbf{v}_2 \in W$, tal que $\{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente independiente. Ahora considérese el espacio $W_2 = \mathcal{L}(\{\mathbf{v}_1, \mathbf{v}_2\})$ si $W = W_2$, entonces $dim(V) = 2$. Si $W \neq W_2$, podemos continuar el proceso hasta obtener un conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ que sea una base para W. En este caso, $dim(W) = k$. Además, como el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es un conjunto linealmente independiente y V es generado por n vectores, entonces por el teorema 2.4.5,

$$k \leq n,$$

o de manera equivalente,

$$din(W) \leq dim(V). \quad \blacksquare$$

Teorema 2.4.13 *Si A es una matriz cualquiera, entonces el espacio de renglones y el espacio de columnas de A tienen la misma dimensión.*

Demostración: Considérese la matriz general A de tamaño mxn.

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

Sea k la dimensión del espacio de renglones; es decir,

$$dim(\mathcal{L}(S_r)) = k$$

Ahora, sea $S = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k\}$ una base para $(\mathcal{L}(S_r)$, entonces cada vector renglón puede expresarse como combinación lineal de los vectores en la base; es decir,

$$\left.\begin{array}{rcl} \mathbf{r}_1 & = & c_{11}\mathbf{b}_1 + c_{12}\mathbf{b}_2 + \ldots + c_{1k}\mathbf{b}_k \\ \mathbf{r}_2 & = & c_{21}\mathbf{b}_1 + c_{22}\mathbf{b}_2 + \ldots + c_{2k}\mathbf{b}_k \\ \vdots & = & \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \mathbf{r}_m & = & c_{m1}\mathbf{b}_1 + c_{m2}\mathbf{b}_2 + \ldots + c_{mk}\mathbf{b}_k \end{array}\right\} \tag{2.37}$$

Además, los vectores en la base pueden expresarse como

$$\begin{array}{rcl} \mathbf{b}_1 & = & (b_{11}, b_{12}, \ldots, b_{1n}) \\ \mathbf{b}_2 & = & (b_{21}, b_{22}, \ldots, b_{2n}) \\ \vdots & = & \vdots \qquad\qquad \vdots \\ \mathbf{b}_k & = & (b_{k1}, b_{k2}, \ldots, b_{kn}) \end{array} \tag{2.38}$$

Sustituyendo las ecuaciones 2.38 en las ecuaciones 2.37,

$$\begin{aligned}
\mathbf{r}_1 &= (a_{11},\ldots,a_{1j},\ldots,a_{1n}) = c_{11}(b_{11},\ldots,b_{1j},\ldots,b_{1n})+ \\
& \quad c_{12}(b_{21},\ldots,b_{2j},\ldots,b_{2n}) + \ldots + c_{1k}(b_{k1},\ldots,b_{kj},\ldots,b_{kn}), \\
\mathbf{r}_2 &= (a_{21},\ldots,a_{2j},\ldots,a_{2n}) = c_{21}(b_{11},\ldots,b_{1j},\ldots,b_{1n})+ \\
& \quad c_{22}(b_{21},\ldots,b_{2j},\ldots,b_{2n}) + \ldots + c_{2k}(b_{k1},\ldots,b_{kj},\ldots,b_{kn}), \\
\vdots &= \qquad \vdots \qquad\qquad \vdots \\
\mathbf{r}_m &= (a_{m1},\ldots,a_{mj},\ldots,a_{mn}) = c_{m1}(b_{11},\ldots,b_{1j},\ldots,b_{1n})+ \\
& \quad c_{m2}(b_{21},\ldots,b_{2j},\ldots,b_{2n}) + \ldots + c_{mk}(b_{k1},\ldots,b_{kj},\ldots,b_{kn}).
\end{aligned}$$

De las ecuaciones anteriores,

$$\begin{aligned}
a_{1j} &= c_{11}b_{1j} + c_{12}b_{2j} + \ldots + c_{1k}b_{kj} \\
a_{2j} &= c_{21}b_{1j} + c_{22}b_{2j} + \ldots + c_{2k}b_{kj} \\
\vdots &= \vdots \qquad\quad \vdots \qquad\quad \vdots \\
a_{mj} &= c_{m1}b_{1j} + c_{m2}b_{2j} + \ldots + c_{mk}b_{kj}
\end{aligned} \tag{2.39}$$

En forma matricial tenemos,

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = b_{1j}\begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{bmatrix} + b_{2j}\begin{bmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{m2} \end{bmatrix} + \ldots + b_{kj}\begin{bmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{mk} \end{bmatrix} \qquad 1 \leq j \leq n$$

La ecuación anterior muestra que los vectores columna están contenidos en un espacio generado por k vectores. Sabemos por el teorema 2.3.1 que el espacio generado por estos k vectores contiene al espacio de columnas; además, por el teorema anterior,

$$dim(\mathcal{L}(S_c)) \leq k,$$

$$dim(\mathcal{L}(S_c)) \leq dim(\mathcal{L}(S_r)). \tag{2.40}$$

Si aplicamos este resultado a la transpuesta de la matriz A, concluimos que:

$$dim(\mathcal{L}(S_c^t)) \leq dim(\mathcal{L}(S_r^t)),$$

$$dim(\mathcal{L}(S_r)) \leq dim(\mathcal{L}(S_c)). \tag{2.41}$$

Las relaciones de orden descritas en las ecuaciones 2.40 y 2.41 solo se satisfacen simultáneamente cuando se cumple la igualdad; es decir,

$$dim(\mathcal{L}(S_r)) = dim(\mathcal{L}(S_c)). \qquad \blacksquare$$

Definición 2.4.6 (Rango de una matriz) *Se conoce como rango de una matriz A, denotado por $\rho(A)$, a la dimensión del espacio de renglones o de columnas.*

Definición 2.4.7 (Nulidad de una matriz) *Se conoce como nulidad de una matriz A, denotado por $\nu(A)$, a la dimensión del espacio nulo; es decir, a la dimensión del espacio de soluciones de la homogénea Ax=0.*

Teorema 2.4.14 *Si A es una matriz de tamaño mxn, entonces*

$$\rho(A) + \nu(A) = n. \tag{2.42}$$

Demostración: Dado que la matriz A tiene n columnas, entonces la ecuación homogénea $Ax = 0$, tendrá n incógnitas o variables, de las cuales r serán variables principales y $(n-r)$ serán variables libres. Las variables principales están relacionadas directamente con los 1-principales en la forma escalonada en los renglones reducida de la matriz A y éstos a su vez forman una base para el espacio de renglones de la matriz A; es decir, $\rho(A) = r$. Si no existen variables libres, la homogénea solo tendría la solución trivial, en este caso, $\nu(A) = 0$, si solo existe una variable libre $\nu(A) = 1$, entonces, en general las variables libres están relacionadas con la dimensión del espacio nulo, por lo anterior,

$$r + (n - r) = n,$$

o

$$\rho(A) + \nu(A) = n. \quad \blacksquare$$

Ejemplo 2.4.6 *Para la matriz del ejemplo 2.4.4, determine la nulidad de la matriz A y una base para el espacio nulo.*

$$A = \begin{bmatrix} 1 & 0 & 4 & 3 & -1 & 1 \\ -5 & 2 & -2 & -1 & -1 & -2 \\ -3 & 2 & 2 & 2 & -2 & -1 \\ -1 & 2 & 14 & 11 & -5 & 2 \end{bmatrix}.$$

Solución: Como la matriz A tiene $n = 6$ columnas y $\rho(A) = 3$ (ver el ejemplo 2.4.4), entonces de la ecuación 2.42 tenemos,

$$\nu(A) = n - \rho(A) = 6 - 3 = 3.$$

La forma escalonada en los rengloes reducida de la matriz A es (ver el ejemplo 1.1.10),

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{3}{4} & -\frac{3}{4} \\ 0 & 0 & 1 & \frac{3}{4} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Debido a que una columna de ceros no puede ser modificada mediante operaciones elementales sobre los renglones, entonces la forma escalonada reducida para la homogénea $Ax = 0$ es simplemente,

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{3}{4} & -\frac{3}{4} & 0 \\ 0 & 0 & 1 & \frac{3}{4} & -\frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Para determinar una base del espacio nulo, debemos resolver el sistema anterior.

Despejando las variables principales en términos de las variables libres se obtiene,

$$\begin{cases} x_1 &= 0 \\ x_2 &= -\frac{1}{4}x_4 + \frac{3}{4}x_5 + \frac{3}{4}x_6, \\ x_3 &= -\frac{3}{4}x_4 + \frac{1}{4}x_5 - \frac{1}{4}x_6. \end{cases}$$

Realizando la parametrización de las variables libres mediante $x_4 = r, x_5 = s$ y $x_6 = t$ se obtiene,

$$\begin{cases} x_1 &= 0, \\ x_2 &= -\frac{1}{4}r + \frac{3}{4}s + \frac{3}{4}t, \\ x_3 &= -\frac{3}{4}r + \frac{1}{4}s - \frac{1}{4}t, \\ x_4 &= r, \\ x_5 &= s, \\ x_6 &= t. \end{cases}$$

El espacio nulo puede expresarse como una matriz de coordenadas mediante,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} \begin{bmatrix} 0 \\ -\frac{1}{4}r + \frac{3}{4}s + \frac{3}{4}t, \\ -\frac{3}{4}r + \frac{1}{4}s - \frac{1}{4}t, \\ r \\ s \\ t \end{bmatrix} = r \begin{bmatrix} 0 \\ -\frac{1}{4} \\ -\frac{3}{4} \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ \frac{3}{4} \\ \frac{1}{4} \\ 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ \frac{3}{4} \\ -\frac{1}{4} \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

De manera equivalente tenemos,

$$x = (x_1, x_2, x_3, x_4, x_5, x_6) = r(0, -\frac{1}{4}, -\frac{3}{4}, 1, 0, 0) + s(0, \frac{3}{4}, \frac{1}{4}, 0, 1, 0) + t(0, \frac{3}{4}, -\frac{1}{4}, 0, 0, 1).$$

Hemos expresado el espacio nulo como la combinación lineal de tres vectores; es decir, estos tres vectores generan el espacio nulo, que como sabemos es de dimensión 3, de manera que, por el teorema 2.4.6, estos vectores forman una base para dicho espacio.

Teorema 2.4.15 *Si A es una matriz de tamaño nxn, entonces los siguientes enunciados son equivalentes.*

a) A es inversible.

b) Ax=0 es solo tiene la solución trivial.

c) A es equivalente respecto a los renglones a I_n.

d) Ax=b tiene solución única para toda matriz b de tamaño nx1.

e) $detA \neq 0$.

f) A tiene rango n.

g) Los vectores renglón de A son linealmente independientes.

h) Los vectores columna de A son linealmente independientes.

Demostración: Ya hemos probado la equivalencia entre los incisos a, b, c, d y e. Para completar la demostración seguiremos la línea demostrativa, $c) \Rightarrow f) \Rightarrow g) \Rightarrow h) \Rightarrow c)$

$c) \Rightarrow f)$ Sabemos que A es equivalente respecto a los renglones con I_n. Esto significa que la forma escalonada en los renglones de la matriz A tiene n renglones diferentes de cero, estos n renglones constituyen una base para el espacio de renglones de la matriz A (teorema 2.4.11).Por lo anterior,

$$dim(\mathcal{L}(S_r)) = n = \rho(A).$$

$f) \Rightarrow g)$ Sabemos que $\rho(A) = n$, es decir, $dim(\mathcal{L}(S_r)) = n$; Además, este espacio es generado por los n renglones de la matriz A; así que, por el teorema 2.4.6, estos n vectores forman una base para $\mathcal{L}(S_r)$ y, por lo lanto, son linealmente independientes.

$g \Rightarrow h)$ Sabemos que los n vectores renglón de A son linealmente independientes y , además, generan al espacio de renglones, entonces estos n vectores conforman una base para $\mathcal{L}(S_r)$. Ahora, por el teorema 2.4.13, $\mathcal{L}(S_r) = \mathcal{L}(S_c) = n$. En este punto, los n vectores columna generan un espacio vectorial de dimensión n, por el teorema(2.4.6), estos n vectores forman una base para $\mathcal{L}(S_c)$ y, por lo lanto, son linealmente independientes.

h) $\Rightarrow$ *c*) Sabemos que los n vectores columna de A son linealmente independientes, así que, $dim(\mathcal{L}(S_c)) = n = dim(\mathcal{L}(S_r))$. Luego, los n vectores renglones de A forman una base para $\mathcal{L}(S_r)$. Además, la forma escalonada en los renglones reducida de la matriz A debe tener n renglones diferentes de cero; por lo tanto, $A \sim I_n$. ∎

Ejemplo 2.4.7 *Determine si los vectores,*

$$\begin{aligned}\boldsymbol{v}_1 &= (1,2,-1,4),\\ \boldsymbol{v}_2 &= (3,-1,2,0),\\ \boldsymbol{v}_3 &= (4,5,-2,1),\\ \boldsymbol{v}_4 &= (-2,3,0,2).\end{aligned}$$

forman una base para $\mathbb{R}^4$

Solución: Para comenzar debemos construir la matriz A cuyos vectores renglón sean los vectores dados; es decir,

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

Por el teorema anterior, los vectores serán linealmente independientes en $\mathbb{R}^4$ si $|A| \neq 0$. Éste es el caso (ver el ejemplo 1.4.14),

$$|A| = \begin{vmatrix} 1 & 2 & -1 & 4\\ 3 & -1 & 2 & 0\\ 4 & 5 & -2 & 1\\ -2 & 3 & 0 & 2 \end{vmatrix} = -209 \neq 0.$$

Por otro lado, 4 vectores linealmente independientes forman una base para $\mathbb{R}^4$ (teorema 2.4.6). En conclusión, los vectores anteriores constituyen una base para $\mathbb{R}^4$.

2.5. Espacios con producto interior

Definición 2.5.1 (Producto interior) *Un producto interior sobre un espacio vectorial V es una función que asocia un número real* $\langle \boldsymbol{u}, \boldsymbol{v}\rangle$ *con cada pareja de vectores* $\boldsymbol{u}$ *y* $\boldsymbol{v}$ *en V, de tal manera que se satisfacen los siguientes axiomas para todos los vectores* $\boldsymbol{u}$,$\boldsymbol{v}$ *y* $\boldsymbol{w}$ *en V y todos los escalares k.*

a) $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \langle \boldsymbol{v}, \boldsymbol{u}\rangle$*, axioma de simetría.*

b) $\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle$, *axioma de aditividad.*

c) $\langle k\boldsymbol{u}, \boldsymbol{v} \rangle = k\langle \boldsymbol{u}, \boldsymbol{v} \rangle$, *axioma de homogeneidad.*

d) $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$ *y* $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = 0$, *si y solo si,* $\boldsymbol{v} = \boldsymbol{0}$, *axioma de positividad.*

Definición 2.5.2 (Espacio de productos interiores) *Un espacio vectorial con producto interior se conoce como espacio de productos interiores.*

Ejemplo 2.5.1 (Producto euclidiano interior) *Para el espacio vectorial* $\mathbb{R}^n$ *el producto euclidiano interior se define de la siguiente manera: Si* $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, *y* $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ *son vectores de* $\mathbb{R}^n$, *entonces*

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Verifiquemos que el producto euclidiano interior satisface los axiomas de anteriores:

a) Simetría:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} y_i x_i = \langle \mathbf{y}, \mathbf{x} \rangle.$$

b) Aditividad:

$$\begin{aligned} \langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle &= \sum_{i=1}^{n} (x_i + y_i) z_i = \sum_{i=1}^{n} (x_i z_i + y_i z_i), \\ &= \sum_{i=1}^{n} x_i z_i + \sum_{i=1}^{n} y_i z_i, = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle. \end{aligned}$$

c) Homogeneidad:

$$\langle k\mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} (k x_i) y_i = k \sum_{i=1}^{n} x_i y_i = k\langle \mathbf{x}, \mathbf{y} \rangle.$$

d) Positividad:

$$\langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^{n} x_i^2 = 0 \Leftrightarrow x_i = 0, \text{ para todo i;}$$

por lo tanto,

$$\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}.$$

Si los vectores en $\mathbb{R}^n$ se escriben como columnas, entonces el producto euclidiano interior puede expresarse como un producto matricial. Sean $\mathbf{x}$ y $\mathbf{y}$ vectores de $\mathbb{R}^n$ tales que:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad y \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

entonces,

$$\langle \mathbf{x}, \mathbf{y} \rangle = [y_1, y_2, \ldots, y_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{y}^T\mathbf{x}.$$

Si A es una matriz de nxn, entonces,

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^T\mathbf{y} \rangle. \tag{2.43}$$

Para verificar este resultado considérese

$$\begin{aligned} \langle A\mathbf{x}, \mathbf{y} \rangle &= \mathbf{y}^T(A\mathbf{x}), \\ &= (\mathbf{y}^T A)\mathbf{x}, \quad \text{Teorema 1.2.1c} \\ &= (A^T\mathbf{y})^T\mathbf{x}, \quad \text{Teorema 1.2.3d} \\ &= \langle \mathbf{x}, A^T\mathbf{y} \rangle. \end{aligned}$$

Teorema 2.5.1 *Si V es un espacio de productos interiores, entonces para todo $\boldsymbol{u}$,$\boldsymbol{v}$ y $\boldsymbol{w}$ en V y todos los escalares k, se cumple que:*

a) $\langle \boldsymbol{0}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{0} \rangle = 0$

b) $\langle \boldsymbol{w}, \boldsymbol{u} + \boldsymbol{v} \rangle = \langle \boldsymbol{w}, \boldsymbol{u} \rangle + \langle \boldsymbol{w}, \boldsymbol{v} \rangle$

c) $\langle \boldsymbol{u}, k\boldsymbol{v} \rangle = k\langle \boldsymbol{u}, \boldsymbol{v} \rangle$

Demostración: a)Sea $\mathbf{u}$ cualquier vector en V, entonces

$$\begin{aligned} \langle \mathbf{0}, \mathbf{v} \rangle &= \langle \mathbf{u} - \mathbf{u}, \mathbf{v} \rangle, && \text{propiedad del inverso aditivo,} \\ &= \langle \mathbf{u}, \mathbf{v} \rangle + \langle \text{-}\mathbf{u}, \mathbf{v} \rangle, && \text{axioma de aditividad,} \\ &= \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{v} \rangle, && \text{axioma de homogeneidad,} \\ &= \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{v} \rangle = 0, && \text{propiedad de los números reales.} \end{aligned}$$

b)

$$\begin{aligned}
\langle \mathbf{w}, \mathbf{u}+\mathbf{v} \rangle &= \langle \mathbf{u}+\mathbf{v}, \mathbf{w} \rangle, && \text{axioma de simetría,}\\
&= \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle, && \text{axioma de aditividad,}\\
&= \langle \mathbf{w}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle, && \text{axioma de simetría.}
\end{aligned}$$

c)

$$\begin{aligned}
\langle \mathbf{u}, k\mathbf{v} \rangle &= \langle k\mathbf{v}, \mathbf{u} \rangle, && \text{axioma de simetría}\\
&= k\langle \mathbf{v}, \mathbf{u} \rangle, && \text{axioma de homogeneidad}\\
&= k\langle \mathbf{u}, \mathbf{v} \rangle, && \text{axioma de simetría.} \quad \blacksquare
\end{aligned}$$

Teorema 2.5.2 (Desigualdad de Cauchy-Schwarz) *Si* $\boldsymbol{u}$ *y* $\boldsymbol{v}$ *son vectores en un espacio de productos interiores V, entonces*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 \leq \langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{v}, \boldsymbol{v} \rangle.$$

Demostración: Si $\mathbf{u} = \mathbf{0}$, entonces la igualdad se cumple de manera inmediata, ya que, $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{0}, \mathbf{v} \rangle = 0$, y $\langle \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{0}, \mathbf{0} \rangle = 0$. Supóngase entonces que $\mathbf{u} \neq \mathbf{0}$ y considérese el vector $\mathbf{w} = x\mathbf{u} + \mathbf{v}$, donde x es cualquier escalar. Por el axioma de positividad tenemos que, $0 \leq \langle \mathbf{w}, \mathbf{w} \rangle$, además, sustituyendo $\mathbf{w}$ se tiene

$$\begin{aligned}
0 &\leq \langle \mathbf{w}, \mathbf{w} \rangle,\\
0 &\leq \langle \mathbf{w}, x\mathbf{u} + \mathbf{v} \rangle,\\
0 &\leq \langle \mathbf{w}, x\mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle,\\
0 &\leq x\langle \mathbf{w}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle,\\
0 &\leq x\langle x\mathbf{u} + \mathbf{v}, \mathbf{u} \rangle + \langle x\mathbf{u} + \mathbf{v}, \mathbf{v} \rangle,\\
0 &\leq x(\langle x\mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle) + \langle x\mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle,\\
0 &\leq x\langle x\mathbf{u}, \mathbf{u} \rangle + x\langle \mathbf{v}, \mathbf{u} \rangle + \langle x\mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle,\\
0 &\leq x^2\langle \mathbf{u}, \mathbf{u} \rangle + x\langle \mathbf{u}, \mathbf{v} \rangle + x\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle,\\
0 &\leq x^2\langle \mathbf{u}, \mathbf{u} \rangle + 2x\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle.
\end{aligned}$$

Si definimos las parámetros $a = \langle \mathbf{u}, \mathbf{u} \rangle$, $b = 2\langle \mathbf{u}, \mathbf{v} \rangle$ y $c = \langle \mathbf{v}, \mathbf{v} \rangle$, entonces la última relación se expresa como,

$$0 \leq ax^2 + bx + c. \tag{2.44}$$

La ecuación 2.44 es válida para cualquier número real x. Es decir, ésta representa una parábola con concavidad positiva y que se ubica por encima de la horizontal, tocándola a lo más en un solo punto, así que las raíces del polinomio son las dos complejas o las dos son reales pero iguales. Lo anterior impone la siguiente condición al discriminante de las raíces de la parábola.

$$b^2 - 4ac \leq 0.$$

luego,

$$b^2 \leq 4ac$$

Sustituyendo los parámetros a, b, y c se tiene

$$\begin{aligned}(2\langle \mathbf{u}, \mathbf{v}\rangle)^2 &\leq 4\langle \mathbf{u}, \mathbf{u}\rangle\langle \mathbf{v}, \mathbf{v}\rangle, \\ 4\langle \mathbf{u}, \mathbf{v}\rangle^2 &\leq 4\langle \mathbf{u}, \mathbf{u}\rangle\langle \mathbf{v}, \mathbf{v}\rangle, \\ \langle \mathbf{u}, \mathbf{v}\rangle^2 &\leq \langle \mathbf{u}, \mathbf{u}\rangle\langle \mathbf{v}, \mathbf{v}\rangle. \quad \blacksquare\end{aligned}$$

Longitud y ángulo en los espacios de productos interiores

Definición 2.5.3 (Norma) *Si V es un espacio de productos interiores, entonces la norma o longitud de un vector $\boldsymbol{v}$, denotado por $\|\boldsymbol{v}\|$, se define por*

$$\|\boldsymbol{v}\| = \langle \boldsymbol{v}, \boldsymbol{v}\rangle^{1/2}.$$

De la definición anterior se tiene, $\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v}\rangle$, de este modo, la desigualdad de Cuachy- Schwarz puede expresarse como,

$$\langle \mathbf{u}, \mathbf{v}\rangle^2 \leq \|\mathbf{u}\|^2\|\mathbf{v}\|^2.$$

Definición 2.5.4 (Distancia) *La distancia entre dos vectores $\boldsymbol{u}$ y $\boldsymbol{v}$, denotado por $d(\boldsymbol{u}, \boldsymbol{v})$, se define como*

$$d(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|.$$

Teorema 2.5.3 *Si V es un espacio de productos interiores, entonces la norma y la distancia satisfacen todas las propiedades que se enlistan en la tabla 2.1.*

Demostración: Solo probaremos la propiedad n_4, el resto de las propiedades se deja como ejercicio.

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle, && \text{Definición} \\ &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}\rangle\langle \mathbf{u}+\mathbf{v}, \mathbf{v}\rangle, && \text{Aditividad} \\ &= \langle \mathbf{u}, \mathbf{u}\rangle + \langle \mathbf{v}, \mathbf{u}\rangle + \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle, && \text{Aditividad} \\ &= \langle \mathbf{u}, \mathbf{u}\rangle + 2\langle \mathbf{u}, \mathbf{v}\rangle + \mathbf{v}, \mathbf{v}\rangle, && \text{Simetría} \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle. && \text{Simetría} \qquad (2.45)\end{aligned}$$

Tabla 2.1: Propiedades de la norma y la distancia

Propiedades básicas de la norma	Propiedades básicas de la distancia
n_1) $\|\mathbf{u}\| \geq 0$	d_1) $d(\mathbf{u}, \mathbf{v}) \geq 0$
n_2) $\|\mathbf{u}\| = 0 \Longleftrightarrow \mathbf{u} = \mathbf{0}$	d_2) $d(\mathbf{u}, \mathbf{v}) = 0 \Longleftrightarrow \mathbf{u} = \mathbf{v}$
n_3) $\|k\mathbf{u}\| = k\|\mathbf{u}\|$	d_3) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$
n_4) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{u}\|$ Desigualdad del triángulo	d_4) $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$

La ecuación 2.45 puede expresarse como,

$$\frac{\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2}{2} = \langle \mathbf{u}, \mathbf{v} \rangle. \tag{2.46}$$

De la desigualdad de Cauchy-Schwarz tenemos,

$$\langle \mathbf{u}, \mathbf{v} \rangle^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2.$$

Extrayendo la raiz cuadrada y recordando que la raíz cuadrada es una función monótonamente creciente,

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\| \tag{2.47}$$

Además, el valor absoluto de un número siempre será mayor o igual que dicho número; es decir,

$$\langle \mathbf{u}, \mathbf{v} \rangle \leq |\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\| \tag{2.48}$$

Sustituyendo la ecuación 2.48 en la ecuación 2.46,

$$\frac{\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2}{2} \leq \|\mathbf{u}\| \|\mathbf{v}\| \tag{2.49}$$

La ecuación 2.49 puede expresarse como,

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 &\leq \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| \\ \|\mathbf{u}+\mathbf{v}\|^2 &\leq (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \end{aligned} \tag{2.50}$$

Finalmente, extrayendo la raíz cuadrada en la ecuación 2.50 se obtiene la desigualdad del triángulo.

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|. \quad \blacksquare \tag{2.51}$$

Ángulo entre vectores

La desigualdad de Cauchy Schwarz puede expresarse como

$$\frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\|\mathbf{u}\|^2 \|\mathbf{v}\|^2} \leq 1,$$

la raíz cuadrada nos da,

$$|\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|}| \leq 1,$$

lo anterior significa que,

$$1 \leq \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1.$$

Como consecuencia de este resultado definimos,

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|}.$$

θ se conoce como el ángulo entre los vectores $\mathbf{u}$ y $\mathbf{v}$. Ahora podemos expresar el producto interior como,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\| \|\mathbf{v}\| \cos\theta.$$

Es claro que si $\theta = 90°$, entonces

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0$$

2.5.1. Bases ortonormales

Definición 2.5.5 (Vectores ortogonales) *Si $\boldsymbol{u}$ y $\boldsymbol{v}$ son vectores en un espacio de productos interiores, entonces diremos que $\boldsymbol{u}$ y $\boldsymbol{v}$ son vectores ortogonales si $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$. Además, si $\boldsymbol{u}$ es ortogonal a cada vector en un conjunto W, diremos que $\boldsymbol{u}$ es ortogonal a W.*

Teorema 2.5.4 (Teorema de Pitágoras generalizado) *Si $\boldsymbol{u}$ y $\boldsymbol{v}$ son vectores ortogonales en un espacio de productos interiores, entonces*

$$\|\boldsymbol{u} + \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2.$$

Demostración:

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \end{aligned}$$

Pero como $\mathbf{u}$ y $\mathbf{v}$ son ortogonales, entonces $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, luego

$$\begin{aligned} \|\mathbf{v} + \mathbf{v}\|^2 &= \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2. \quad \blacksquare \end{aligned}$$

Definición 2.5.6 (Vectores unitarios) *En un espacio de productos interiores, diremos que un vector es unitario si su norma es uno. Si* $\boldsymbol{v}$ *es un vector unitario lo representaremos como* $\widehat{\boldsymbol{v}}$*, para destacar dicha propiedad.*

Si $\mathbf{v}$ es un vector distinto del vector cero, entonces $\dfrac{\mathbf{v}}{\|\mathbf{v}\|}$ define un vector unitario, en este caso escribiremos,

$$\widehat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}. \tag{2.52}$$

Despejando $\mathbf{v}$ de la ecuación anterior se obtiene,

$$\mathbf{v} = \|\mathbf{v}\|\widehat{\mathbf{v}}. \tag{2.53}$$

Definición 2.5.7 (Conjunto ortogonal y ortonormal) *Se dice que un conjunto de vectores en un espacio de productos interiores es ortogonal si todas las parejas de vectores diferentes en el conjunto son ortogonales. Si además, cada vector es unitario, entonces el conjunto se conoce como ortonormal.*

Teorema 2.5.5 *Si* $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ *es un conjunto ortogonal de vectores diferentes del vector cero, en un espacio vectorial* V *de productos interiores y de dimensión finita* n*, entonces* S *es una base para este espacio.*

Demostración: Como S tiene n vectores y V es de dimensión n, de acuerdo con el teorema 2.4.6, bastará con probar que S es linealmente independiente para que sea una base.

Considérese la ecuación vectorial,

$$\begin{aligned}
\mathbf{0} &= k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_n\mathbf{v}_n && (2.54)\\
&= \sum_{j=1}^{n} k_j\mathbf{v}_j && (2.55)
\end{aligned}$$

Si probamos que la única solución de la ecuación anterior corresponde a $k_1 = k_2 = \ldots = k_n = 0$, entonces S será linealmente independiente. Para tal fin considérese el producto interior $\langle \mathbf{0}, \mathbf{v}_i \rangle$.

$$\begin{aligned}
0 &= \langle \mathbf{v}_i, \mathbf{0}\rangle, && \text{teorema 2.5.1}\\
&= \langle \mathbf{v}_i, \sum_{j=1}^{n} k_j\mathbf{v}_j\rangle, && \text{sustitución}\\
&= \sum_{j=1}^{n} \langle \mathbf{v}_i, k_j\mathbf{v}_j\rangle, && \text{aditividad}\\
&= \sum_{j=1}^{n} k_j\langle \mathbf{v}_i, \mathbf{v}_j\rangle, && \text{homogeneidad}\\
&= k_i\langle \mathbf{v}_i, \mathbf{v}_i\rangle, && \text{ortogonalidad}
\end{aligned}$$

Pero $\langle \mathbf{v}_i, \mathbf{v}_i\rangle > 0$, por el axioma de positividad y debido a que $\mathbf{0} \notin S$; por lo tanto, $k_i = 0$, para todo i. Esto significa que S es linealmente independiente. Se concluye que S es una base para V. ∎

Teorema 2.5.6 *Si $S = \{\widehat{\boldsymbol{v}}_1, \widehat{\boldsymbol{v}}_2, \ldots, \widehat{\boldsymbol{v}}_n\}$ es una base ortonormal para un espacio de productos interiores V y $\boldsymbol{u}$ es cualquier vector en V, entonces*

$$\boldsymbol{u} = \langle \boldsymbol{u}, \widehat{\boldsymbol{v}}_1\rangle\widehat{\boldsymbol{v}}_1 + \langle \boldsymbol{u}, \widehat{\boldsymbol{v}}_2\rangle\widehat{\boldsymbol{v}}_2 + \cdots + \langle \boldsymbol{u}, \widehat{\boldsymbol{v}}_n\rangle\widehat{\boldsymbol{v}}_n$$

Demostración: Como S es una base para V y $\mathbf{u} \in V$, entonces por el teorema 2.4.9, existe un conjunto único de escalares $\{a_1, a_2, \ldots, a_n\}$ tales que,

$$\begin{aligned}
\mathbf{u} &= a_1\widehat{\mathbf{v}}_1 + a_2\widehat{\mathbf{v}}_2 + \ldots + a_n\widehat{\mathbf{v}}_n\\
&= \sum_{j=1}^{n} a_j\widehat{\mathbf{v}}_j
\end{aligned} \tag{2.56}$$

Ahora considérese el producto interior $\langle \mathbf{u}, \widehat{\mathbf{v}}_i\rangle$

$$\begin{aligned}
\langle \mathbf{u}, \widehat{\mathbf{v}}_i\rangle &= \langle\sum_{j=1}^{n} a_j\widehat{\mathbf{v}}_j, \widehat{\mathbf{v}}_i\rangle, &&\\
&= \sum_{j=1}^{n}\langle a_j\widehat{\mathbf{v}}_j, \widehat{\mathbf{v}}_i\rangle, && \text{aditividad}\\
&= \sum_{j=1}^{n} a_j\langle\widehat{\mathbf{v}}_j, \widehat{\mathbf{v}}_i\rangle, && \text{homogeneidad}\\
&= \sum_{j=1}^{n} a_j\delta_{ij}, && \text{propiedad de las bases ortonormales}\\
\langle \mathbf{u}, \widehat{\mathbf{v}}_i\rangle &= a_i, && \text{definición de la delta de Kronecker.}
\end{aligned} \tag{2.57}$$

Sustituyendo la ecuación 2.57 en la ecuación 2.56 se tiene

$$\mathbf{u} = \sum_{j=1}^{n} \langle \mathbf{u}, \widehat{\mathbf{v}}_j \rangle \widehat{\mathbf{v}}_j \qquad \blacksquare$$

2.5.2. Proceso de ortonormalización de Gram-Schmidt

Teorema 2.5.7 *Todo espacio de productos interiores diferente del espacio cero y de dimensión finita tiene una base ortonormal.*

Demostración: Este teorema lo demostraremos construyendo una base ortonormal para un espacio vectorial cualquiera de productos interiores. Procederemos de manera explícita para los casos más sencillos.

i) Supóngase para comenzar que V es como se indica en el teorema y de dimensión 1. Así que, sea $\mathbf{v}_1 \in V$ un vector distinto del vector cero tal que, $V = \mathcal{L}\{\mathbf{v}_1\}$, entonces es claro que $\widehat{\mathbf{u}}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$ cumple con las condiciones del teorema.

ii) Considérese ahora el caso en que V es como se indica en el teorema y de dimensión 2; y sea además, $S_2 = \{\mathbf{v}_1, \mathbf{v}_2\}$ una base cualquiera para V, se requiere construir una base $S_2' = \{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2\}$ tal que $\langle \widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2 \rangle = 0$. Para tal fin, primero se hará $\widehat{\mathbf{u}}_1 = \dfrac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$, es claro que $V = \mathcal{L}\{\mathbf{v}_1, \mathbf{v}_2\} = \mathcal{L}\{\widehat{\mathbf{u}}_1, \mathbf{v}_2\}$. Ahora determinaremos un vector $\mathbf{w}_2 \in V$ que sea ortogonal a $\widehat{\mathbf{u}}_1$. Como $\mathbf{w}_2 \in V$, entonces

$$\mathbf{w}_2 = b_2\mathbf{v}_2 + b_1\widehat{\mathbf{u}}_1,$$

para simplificar se hace $b_2 = 1$, en este caso,

$$\mathbf{w}_2 = \mathbf{v}_2 + b_1\widehat{\mathbf{u}}_1. \tag{2.58}$$

Para determinar el escalar b_1 recordemos que $\mathbf{w}_2$ debe ser ortogonal a $\widehat{\mathbf{u}}_1$. En este caso

$$\begin{aligned} 0 &= \langle \widehat{\mathbf{u}}_1, \mathbf{w}_2 \rangle \\ &= \langle \widehat{\mathbf{u}}_1, \mathbf{v}_2 + b_1\widehat{\mathbf{u}}_1 \rangle \\ &= \langle \widehat{\mathbf{u}}_1, \mathbf{v}_2 \rangle + \langle \widehat{\mathbf{u}}_1, b_1\widehat{\mathbf{u}}_1 \rangle \\ &= \langle \widehat{\mathbf{u}}_1, \mathbf{v}_2 \rangle + b_1\langle \widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_1 \rangle \\ 0 &= \langle \widehat{\mathbf{u}}_1, \mathbf{v}_2 \rangle + b_1. \end{aligned} \tag{2.59}$$

Despejando de la última ecuación,

$$b_1 = -\langle \widehat{\mathbf{u}}_1, \mathbf{v}_2 \rangle = -\langle \mathbf{v}_2, \widehat{\mathbf{u}}_1 \rangle \tag{2.60}$$

Sustituyendo la ecuación 2.60 en la ecuación 2.58,

$$\mathbf{w}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1 \tag{2.61}$$

Finalmente, normalizando $\mathbf{w}_2$,

$$\widehat{\mathbf{u}}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} \tag{2.62}$$

$$= \frac{\mathbf{v}_2 - \langle \mathbf{v}_2, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1}{\|\mathbf{v}_2 - \langle \mathbf{v}_2, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1\|} \tag{2.63}$$

La normalización siempre es posible, ya que, en caso contrario; es decir, cuando el denominador se anula, se obtiene una contradicción a la independencia lineal de $\{\widehat{\mathbf{u}}_1, \mathbf{v}_2\}$. Finalmente, $S_2' = \{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2\}$, como se mostró en el teorema anterior, es linealmente independiente y, por tanto, una base para V.

iii) Considérese ahora el caso en que V es como se indica en el teorema y de dimensión n; y sea ahora, $S_n = \{\mathbf{v}_1, \mathbf{v}_2\ \ldots, \mathbf{v}_n\}$ una base cualquiera para V, se busca construir una base $S_n' = \{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2 \ldots, \widehat{\mathbf{u}}_n\}$ tal que $\langle \widehat{\mathbf{u}}_i, \widehat{\mathbf{u}}_j \rangle = \delta_{ij}$ para todo $1 \leq i \leq n$ y $1 \leq j \leq n$. Supóngase que para algún $r > 1$, los conjuntos $S_r = \{\mathbf{v}_1, \mathbf{v}_2\ \ldots, \mathbf{v}_r\}$ y $S_r' = \{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2 \ldots, \widehat{\mathbf{u}}_r\}$ generan el mismo subespacio de V que denotaremos como V_r, además,$\langle \widehat{\mathbf{u}}_i, \widehat{\mathbf{u}}_j \rangle = \delta_{ij}$ para todo $1 \leq i, j \leq r$. En este punto se busca construir una base ortonormal para el subesacio de V generado por el conjunto $S_{r+1} = \{\mathbf{v}_1, \mathbf{v}_2\ \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}\}$. Como $\mathcal{L}(S_r) = \mathcal{L}(S_r')$, entonces $\mathcal{L}(S_r \cup \mathbf{v}_{r+1}) = \mathcal{L}(S_r' \cup \{\mathbf{v}_{r+1}\})$, de este modo, S_{r+1} y

$$S_{r+1}'' = S_r' \cup \{\mathbf{v}_{r+1}\} = \{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2 \ldots, \widehat{\mathbf{u}}_r, \mathbf{v}_{r+1}\},$$

generan el mismo subespacio de V que llamaremos V_{r+1}. Ahora determinemos un vector $\mathbf{w}_{r+1} \in V_{r+1}$, tal que, $\mathbf{w}_{r+1}$ sea ortogonal al conjunto S_r'. Como $\mathbf{w}_{r+1} \in V_{r+1}$ entonces puede expresarse como,

$$\mathbf{w}_{r+1} = b_1\widehat{\mathbf{u}}_1 + b_2\widehat{\mathbf{u}}_2 + \ldots + b_r\widehat{\mathbf{u}}_r + b_{r+1}\mathbf{v}_{r+1}.$$

Para simplificar hágase $b_{r+1} = 1$, con lo cual se obtiene

$$\mathbf{w}_{r+1} = b_1\widehat{\mathbf{u}}_1 + b_2\widehat{\mathbf{u}}_2 + \ldots + b_r\widehat{\mathbf{u}}_r + \mathbf{v}_{r+1} \tag{2.64}$$

Para determinar los escalares considérese el producto $\langle \mathbf{w}_{r+1}, \widehat{\mathbf{u}}_i \rangle$

$$
\begin{aligned}
0 &= \langle \mathbf{w}_{r+1}, \widehat{\mathbf{u}}_i \rangle \\
&= \langle \sum_{j=1}^{r} b_j \widehat{\mathbf{u}}_j + \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_i \rangle \\
&= \langle \sum_{j=1}^{r} b_j \widehat{\mathbf{u}}_j, \widehat{\mathbf{u}}_i \rangle + \langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_i \rangle \\
&= \sum_{j=1}^{r} \langle b_j \widehat{\mathbf{u}}_j, \widehat{\mathbf{u}}_i \rangle + \langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_i \rangle \\
&= \sum_{j=1}^{r} b_j \langle \widehat{\mathbf{u}}_j, \widehat{\mathbf{u}}_i \rangle + \langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_i \rangle \\
&= \sum_{j=1}^{r} b_j \delta_{ij} + \langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_i \rangle \\
&= b_i + \langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_i \rangle \qquad (2.65)
\end{aligned}
$$

Despejando b_i de la última ecuación,

$$b_i = -\langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_i \rangle \qquad (2.66)$$

Sustituyendo la ecuación 2.66 en la ecuación 2.64 se obtiene,

$$
\begin{aligned}
\mathbf{w}_{r+1} &= \mathbf{v}_{r+1} - \langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1 - \langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_2 \rangle \widehat{\mathbf{u}}_2 - \ldots - \langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_r \rangle \widehat{\mathbf{u}}_r \\
&= \mathbf{v}_{r+1} - \sum_{i=1}^{r} \langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_i \rangle \widehat{\mathbf{u}}_i \qquad (2.67)
\end{aligned}
$$

Normalizando el vector $\mathbf{w}_{r+1}$,

$$
\begin{aligned}
\widehat{\mathbf{u}}_{r+1} &= \frac{\mathbf{w}_{r+1}}{\|\mathbf{w}_{r+1}\|} \qquad (2.68) \\
&= \frac{\mathbf{v}_{r+1} - \sum_{i=1}^{r} \langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_i \rangle \widehat{\mathbf{u}}_i}{\|\mathbf{v}_{r+1} - \sum_{i=1}^{r} \langle \mathbf{v}_{r+1}, \widehat{\mathbf{u}}_i \rangle \widehat{\mathbf{u}}_i\|} \qquad (2.69)
\end{aligned}
$$

La normalización siempre es posible, de otro modo, se violaría la independencia lineal de S''_{r+1}.
El conjunto $S'_{r+1} = \{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2 \ldots, \widehat{\mathbf{u}}_{r+1}\}$ es un conjunto ortogonal, por tanto, linealmente independiente (teorema 2.5.5), así que constituye una base para V_{r+1}. Este proceso puede extenderse hasta incluir todos los vectores de la base para V, de manera que al final obtendremos una base ortonormal para dicho espacio. ∎

Ejemplo 2.5.2 *Considere el espacio vectorial* $\mathbb{R}^3$ *con el producto euclidiano interior. Aplique el proceso de Gram-Schmidt para transformar la base* $\boldsymbol{v}_1 = (1,1,1)$*,* $\boldsymbol{v}_2 = (0,1,1)$ *y* $\boldsymbol{v}_3 = (0,0,1)$ *en una base ortonormal.*

Solución:

1. Para comenzar hagamos $\mathbf{w}_1 = \mathbf{v}_1$ y normalicemos dicho vector.

$$\langle \mathbf{w}_1, \mathbf{w}_1 \rangle = \langle \mathbf{v}_1 \mathbf{v}_1 \rangle = \langle (1,1,1), (1,1,1) \rangle = 3,$$

de manera que,

$$\widehat{\mathbf{u}}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = \frac{1}{\sqrt{3}}(1,1,1).$$

2. Ahora determinemos un vector $\mathbf{w}_2$ en V_2 que sea ortogonal a $\widehat{\mathbf{u}}_1$. En este caso el proceso de Gram-Schmidt nos proporciona,

$$\begin{aligned}
\mathbf{w}_2 &= \mathbf{v}_2 - \langle \mathbf{v}_2, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1, \\
&= (0,1,1) - \langle (0,1,1), \frac{(1,1,1)}{\sqrt{3}} \rangle \frac{(1,1,1)}{\sqrt{3}}, \\
&= (0,1,1) - \frac{1}{\sqrt{3}} \frac{1}{\sqrt{3}} \langle (0,1,1), (1,1,1) \rangle (1,1,1), \\
&= (0,1,1) - \frac{2}{3}(1,1,1), \\
&= (-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}), \\
&= \frac{1}{3}(-2,1,1).
\end{aligned}$$

Se procede a normalizar $\mathbf{w}_2$,

$$\begin{aligned}
\widehat{\mathbf{u}}_2 &= \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \frac{\frac{1}{3}(-2,1,1)}{\|\frac{1}{3}(-2,1,1)\|}, \\
&= \frac{(-2,1,1)}{\|(-2,1,1)\|} = \frac{(-2,1,1)}{\sqrt{6}}, \\
&= \frac{1}{\sqrt{6}}(-2,1,1).
\end{aligned}$$

3. Determinemos ahora un vector $\mathbf{w}_3$ en V_3 que sea ortogonal a $\{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2\}$. En este caso el proceso de Gram-Schmidt nos proporciona,

$$
\begin{aligned}
\mathbf{w}_3 &= \mathbf{v}_3 - \langle \mathbf{v}_3, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1 - \langle \mathbf{v}_3, \widehat{\mathbf{u}}_2 \rangle \widehat{\mathbf{u}}_2,\\
&= (0,0,1) - \langle (0,0,1), \frac{1}{\sqrt{3}}(1,1,1) \rangle \frac{1}{\sqrt{3}}(1,1,1) - \langle (0,0,1), \frac{1}{\sqrt{6}}(-2,1,1) \rangle \frac{1}{\sqrt{6}}(-2,1,1),\\
&= (0,0,1) - \frac{1}{3}\langle (0,0,1), (1,1,1) \rangle (1,1,1) - \frac{1}{6}\langle (0,0,1), (-2,1,1) \rangle (-2,1,1),\\
&= (0,0,1) - \frac{1}{3}(1,1,1) - \frac{1}{6}(-2,1,1),\\
&= (0, -\frac{1}{2}, \frac{1}{2}),\\
&= \frac{1}{2}(0,-1,1).
\end{aligned}
$$

Normalizando $\mathbf{w}_3$ se obtiene,

$$
\begin{aligned}
\widehat{\mathbf{u}}_3 &= \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \frac{\frac{1}{2}(0,-1,1)}{\|\frac{1}{2}(0,-1,1)\|},\\
&= \frac{(0,-1,1)}{\|(0,-1,1)\|} = \frac{(0,-1,1)}{\sqrt{2}},\\
&= \frac{1}{\sqrt{2}}(0,-1,1).
\end{aligned}
$$

Finalmente, el conjunto $S = \{\frac{1}{\sqrt{3}}(1,1,1), \frac{1}{\sqrt{6}}(-2,1,1), \frac{1}{\sqrt{2}}(0,-1,1)\}$ constituye una base ortonormal para $\mathbb{R}^3$.

Ejemplo 2.5.3 *Para el conjunto de vectores del ejemplo anterior, construya una base ortonormal considerando el producto interior definido como,*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = u_1 v_1 + 2u_2 v_2 + 3u_3 v_3.$$

Solución:

1. Hagamos $\mathbf{w}_1 = \mathbf{v}_1$ y normalicemos dicho vector, es muy importante considerar el producto interior dado para normalizar los vectores.

$$\langle \mathbf{w}_1, \mathbf{w}_1 \rangle = \langle \mathbf{v}_1, \mathbf{v}_1 \rangle = 1 + 2 + 3 = 6,$$

de manera que,

$$\widehat{\mathbf{u}}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = \frac{1}{\sqrt{6}}(1,1,1).$$

2. Construyamos un vector $\mathbf{w}_2$ ortogonal a $\mathbf{u}_1$.

$$\begin{aligned}
\mathbf{w}_2 &= \mathbf{v}_2 - \langle \mathbf{v}_2, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1, \\
&= (0,1,1) - \left\langle (0,1,1), \frac{1}{\sqrt{6}}(1,1,1) \right\rangle \frac{1}{\sqrt{6}}(1,1,1), \\
&= (0,1,1) - \frac{1}{6}\langle (0,1,1), (1,1,1) \rangle (1,1,1), \\
&= (0,1,1) - \frac{1}{6}(5)(1,1,1), \\
&= (0,1,1) - \frac{5}{6}(1,1,1), \\
&= (-5/6, 1/6, 1/6), \\
&= \frac{1}{6}(-5,1,1).
\end{aligned}$$

Normalicemos el vector $\mathbf{w}_2$.

$$\begin{aligned}
\widehat{\mathbf{u}}_2 &= \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \frac{\frac{1}{6}(-5,1,1)}{\|\frac{1}{6}(-5,1,1)\|} = \frac{(-5,1,1)}{\|(-5,1,1)\|}, \\
&= \frac{1}{\sqrt{30}}(-5,1,1).
\end{aligned}$$

3. Determinemos el vector $\mathbf{w}_3$ ortogonal tanto a $\widehat{\mathbf{u}}_1$ como a $\widehat{\mathbf{u}}_2$.

$$\begin{aligned}
\mathbf{w}_3 &= \mathbf{v}_3 - \langle \mathbf{v}_3, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1 - \langle \mathbf{v}_3, \widehat{\mathbf{u}}_2 \rangle \widehat{\mathbf{u}}_2, \\
&= (0,0,1) - \left\langle (0,0,1), \frac{1}{\sqrt{6}}(1,1,1) \right\rangle \frac{1}{\sqrt{6}}(1,1,1) - \left\langle (0,0,1), \frac{(-5,1,1)}{\sqrt{30}} \right\rangle \frac{(-5,1,1)}{\sqrt{30}} \\
&= (0,0,1) - \frac{1}{6}\langle (0,0,1), (1,1,1) \rangle (1,1,1) - \frac{1}{30}\langle (0,0,1), (-5,1,1) \rangle (-5,1,1), \\
&= (0,0,1) - \frac{1}{6}(3)(1,1,1) - \frac{1}{30}(3)(-5,1,1), \\
&= (0,0,1) - \frac{1}{2}(1,1,1) - \frac{1}{10}(-5,1,1), \\
&= (-1/2, -1/2, 1/2) - \frac{1}{10}(-5,1,1), \\
&= \frac{1}{5}(0,-3,2).
\end{aligned}$$

Normalicemos el vector $\mathbf{w}_3$.

$$\begin{aligned}\widehat{\mathbf{u}}_3 &= \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \frac{\frac{1}{5}(0,-3,2)}{\|\frac{1}{5}(0,-3,2)\|} = \frac{(0,-3,2)}{\|(0,-3,2)\|}, \\ &= \frac{1}{\sqrt{30}}(0,-3,2).\end{aligned}$$

Finalmente, el conjunto $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ forma una base ortonormal para el espacio vectorial V con el producto interior dado.

Ejemplo 2.5.4 (Primeros polinomios normalizados de Legendre) *A partir de la base canónica $\{1, x, x^2\}$ y considerando el producto interior,*

$$\langle \boldsymbol{p}, \boldsymbol{q} \rangle = \int_{-1}^{1} p(x)q(x)dx, \tag{2.70}$$

determine una base ortonormal para P_2.

Solución: Apliquemos el proceso de Gram- Schmidt.

1. Hagamos $w_1(x) = 1$, y normalicemos este vector,

$$\langle \mathbf{w}_1, \mathbf{w}_1 \rangle = \int_{-1}^{1} w_1(x)w_1(x)dx = \int_{-1}^{1} dx = 2,$$

de manera que,

$$\widehat{\mathbf{u}}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = \frac{1}{\sqrt{2}}.$$

2. Construyamos ahora un vector $\mathbf{w}_2$ ortogonal a $\widehat{\mathbf{u}}_1$.

$$\begin{aligned}\mathbf{w}_2 &= \mathbf{v}_2 - \langle \mathbf{v}_2, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1, \\ &= x - \left\langle x, \frac{1}{\sqrt{2}} \right\rangle \frac{1}{\sqrt{2}}, \\ &= x - \frac{1}{2}\int_{-1}^{1} x dx = x.\end{aligned}$$

Normalicemos el vector $\mathbf{w}_2$.

$$\langle \mathbf{w}_2, \mathbf{w}_2 \rangle = \int_{-1}^{1} w_2(x)w_2(x)dx = \int_{-1}^{1} x^2 dx = \frac{2}{3},$$

por lo anterior,

$$\widehat{u}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \sqrt{\frac{3}{2}}x.$$

3. Construyamos un vector $\mathbf{w}_3$ ortogonal a $\{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2\}$

$$\begin{aligned}
\mathbf{w}_3 &= \mathbf{v}_3 - \langle \mathbf{v}_3, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1 - \langle \mathbf{v}_3, \widehat{\mathbf{u}}_2 \rangle \widehat{\mathbf{u}}_2, \\
&= x^2 - \left\langle x^2, \frac{1}{\sqrt{2}} \right\rangle \frac{1}{\sqrt{2}} - \left\langle x^2, \sqrt{\frac{3}{2}} x \right\rangle \sqrt{\frac{3}{2}} x, \\
&= x^2 - \frac{1}{2} \int_{-1}^{1} x^2 dx - \frac{3}{2} x \int_{-1}^{1} x^3 dx, \\
&= x^2 - \frac{1}{3}.
\end{aligned}$$

Finalmente, normalicemos el vector $\mathbf{w}_3$.

$$\langle \mathbf{w}_3, \mathbf{w}_3 \rangle = \int_{-1}^{1} w_3(x) w_3(x) dx = \int_{-1}^{1} (x^2 - \frac{1}{3})^2 dx = \frac{8}{45},$$

de manera que,

$$\widehat{\mathbf{u}}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \sqrt{\frac{45}{8}} (x^3 - \frac{1}{3}) = \sqrt{\frac{5}{8}} (3x^2 - 1).$$

Los vectores $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ forman una base ortonormal para P_2. Estos vectores se conocen como los primeros polinomios normalizados de Legendre.

2.6. Cambio de Base

Supóngase que $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base para un espacio vectorial V de dimensión finita. Dado que S genera a V, entonces todo vector $\mathbf{v} \in V$ puede expresarse de manera única como,

$$\mathbf{v} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \ldots + a_n \mathbf{v}_n$$

Los escalares $a_1, a_2, \ldots, a_n$ se conocen como coordenadas de $\mathbf{v}$ relativas a la base S.

Definición 2.6.1 (Vector de coordenadas del vector v relativo a la base S) *Se denota por $(\boldsymbol{v})_S$ y es el vector en $\mathbb{R}^n$ definido por*

$$(\boldsymbol{v})_S = (a_1, a_2, \ldots, a_n).$$

Definición 2.6.2 (Matriz de coordenadas del vector v relativo a la base S) *Se denota por* $[\boldsymbol{v}]_S$ *y es la matriz de* $nx1$ *definido por*

$$[\boldsymbol{v}]_S = \begin{bmatrix} a_1 \\ a_2 \vdots \\ a_n \end{bmatrix}.$$

Ejemplo 2.6.1 *Sea* $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ *una base para* $\mathbb{R}^3$*, donde* $\boldsymbol{v}_1 = (2, 1, 5)$, $\boldsymbol{v}_2 = (-1, 2, 3)$ *y* $\boldsymbol{v}_3 = (4, -3, 1)$

a) Determine tanto el vector de coordenadas como la matriz de coordenadas para el vector $\boldsymbol{v} = (-3, 1, -2)$.

b) Encuentre el vector $\boldsymbol{v} \in \mathbb{R}^3$ *cuyo vector de coordenadas respecto a la base* S *es* $(\boldsymbol{v})_S = (-3, 1, -2)$.

Solución:

a) Debemos hallar los escalares x_1, x_2 y x_3 tales que

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

Sustituyendo los vectores dados,

$$x_1(2, 1, 5) + x_2(-1, 2, 3) + x_3(4, -3, 1) = (-3, 1, -2),$$

realizando las operaciones indicadas se obtiene el siguiente sistema de ecuaciones,

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

Este sistema fue resuelto en el ejemplo 1.1.8. Los valores obtenidos fueron, $x_= - 1, x_2 = 1$ y $x_3 = 0$. De esta manera tenemos que,

$$(\mathbf{v})_S = (-1, 1, 0). \quad y \quad [\mathbf{v}]_S = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

b) En este caso sabemos que, $(\mathbf{v})_S = (-3, 1, -2)$, así que,

$$\begin{aligned} \mathbf{v} &= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3, \\ &= -3\mathbf{v}_1 + \mathbf{v}_2 - 2\mathbf{v}_3, \\ &= -3(2, 1, 5) + (-1, 2, 3) - 2(4, -3, 1), \\ &= (-15, 5, -14). \end{aligned}$$

2.6.1. Representación de un vector mediante una base ortonormal

Si $S = \{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2, \ldots, \widehat{\mathbf{u}}_n\}$ es una base ortonormal de un espacio vectorial V de productos interiores y $\mathbf{v} \in V$, entonces podemos expresar $\mathbf{v}$ como,

$$\mathbf{v} = \langle \mathbf{v}, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1 + \langle \mathbf{v}, \widehat{\mathbf{u}}_2 \rangle \widehat{\mathbf{u}}_2 + \ldots + \langle \mathbf{v}, \widehat{\mathbf{u}}_n \rangle \widehat{\mathbf{u}}_n.$$

Por lo anterior, la representación de $\mathbf{v}$ respecto a esta base ortonormal es,

$$(\mathbf{v})_S = (\langle \mathbf{v}, \widehat{\mathbf{u}}_1 \rangle, \langle \mathbf{v}, \widehat{\mathbf{u}}_2 \rangle, \ldots, \langle \mathbf{v}, \widehat{\mathbf{u}}_n \rangle).$$

$$[\mathbf{v}]_S = \begin{bmatrix} \langle \mathbf{v}, \widehat{\mathbf{u}}_1 \rangle \\ \langle \mathbf{v}, \widehat{\mathbf{u}}_2 \rangle \\ \vdots \\ \langle \mathbf{v}, \widehat{\mathbf{u}}_n \rangle \end{bmatrix}.$$

Ejemplo 2.6.2 *Cosidere la base ortonormal* $S = \{\widehat{\boldsymbol{u}}_1, \widehat{\boldsymbol{u}}_2, \widehat{\boldsymbol{u}}_3\}$*, donde* $\widehat{\boldsymbol{u}}_1 = \frac{1}{\sqrt{6}}(1,1,1)$, $\widehat{\boldsymbol{u}}_2 = \frac{(-5,1,1)}{\sqrt{30}}$, $\widehat{\boldsymbol{u}}_3 = \frac{(0,-3,2))}{\sqrt{30}}$*. Hallar la representación de* $\boldsymbol{v} = (1,2,3)$ *respecto a* S*.*

Solución: Recordemos que el producto interior para este espacio vectorial fue definido como,

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + 2u_2 v_2 + 3u_3 v_3.$$

Determinemos ahora los productos interiores de $\mathbf{v}$ con los vectores de la base ortonormal.

$$\begin{aligned}
\langle \mathbf{v}, \widehat{\mathbf{u}}_1 \rangle &= \left\langle (1,2,3), \frac{1}{\sqrt{6}}(1,1,1) \right\rangle, \\
&= \frac{1}{\sqrt{6}} \langle (1,2,3), (1,1,1) \rangle, \\
&= \frac{1}{\sqrt{6}}(14) = \frac{14}{\sqrt{6}}.
\end{aligned}$$

$$\begin{aligned}
\langle \mathbf{v}, \widehat{\mathbf{u}}_2 \rangle &= \left\langle (1,2,3), \frac{(-5,1,1)}{\sqrt{30}} \right\rangle, \\
&= \frac{1}{\sqrt{30}} \langle (1,2,3), (-5,1,1) \rangle, \\
&= \frac{1}{\sqrt{30}}(8) = \frac{8}{\sqrt{30}}.
\end{aligned}$$

$$\begin{aligned}
\langle \mathbf{v}, \widehat{\mathbf{u}}_3 \rangle &= \left\langle (1,2,3), \frac{(0,-3,2)}{\sqrt{30}} \right\rangle, \\
&= \frac{1}{\sqrt{30}} \langle (1,2,3), (0,-3,2) \rangle, \\
&= \frac{1}{\sqrt{30}}(6) = \frac{6}{\sqrt{30}}.
\end{aligned}$$

Finalmente se forma la matriz de coordenadas respecto a la base ortogonal S.

$$[\mathbf{v}]_S = \begin{bmatrix} \frac{14}{\sqrt{6}} \\ \frac{8}{\sqrt{30}} \\ \frac{6}{\sqrt{30}}. \end{bmatrix}.$$

Teorema 2.6.1 *Si S es una base ortonormal para un espacio de productos interiores V de dimensión n, y si, $\boldsymbol{u}$ y $\boldsymbol{v}$ son dos vectores tales que,$(\boldsymbol{u})_S = (u_1, u_2, \ldots, u_n)$ y $(\boldsymbol{v})_S = (v_1, v_2, \ldots, v_n)$, entonces*

a) $\|\boldsymbol{u}\| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$.

b) $d(\boldsymbol{u}, \boldsymbol{v}) = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_n - v_n)^2}$.

c) $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$.

Demostración: Solo se demuestra el inciso a), el resto se deja como ejercicio para el estudiante. Considérese que $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$, entonces

$$\mathbf{u} = \sum_i u_i \mathbf{w}_i$$

de manera que,

$$\begin{aligned}
\|\mathbf{u}\|^2 &= \langle \mathbf{u}, \mathbf{u} \rangle, \\
&= \left\langle \sum_i u_i \mathbf{w}_i, \sum_j u_j \mathbf{w}_j \right\rangle, \\
&= \sum_i \sum_j u_i u_j \langle \mathbf{w}_i, \mathbf{w}_j \rangle, \\
&= \sum_i \sum_j u_i u_j \delta_{ij}, \\
&= \sum_i u_i u_i = \sum_i u_i^2.
\end{aligned}$$

Finalmente, extrayendo la raíz cuadrada se obtiene,

$$\|\mathbf{u}\| = \sqrt{\sum_i u_i^2}. \qquad \blacksquare$$

2.6.2. Matriz de cambio de base

Teorema 2.6.2 *Si se cambia la base para un espacio vectorial V, de cierta base dada* $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ *a otra nueva base* $S' = \{\boldsymbol{u}'_1, \boldsymbol{u}'_2, \ldots, \boldsymbol{u}'_n\}$, *entonces la matriz de coordenadas inicial ,* $[\boldsymbol{v}]_S$ *para un vector* $\boldsymbol{v}$ *cualquiera, está relacionada con la nueva matriz de coordenadas,* $[\boldsymbol{v}]'_S$*, por medio de la ecuación*

$$[\boldsymbol{v}]_S = P[\boldsymbol{v}]_{S'}, \tag{2.71}$$

en donde las columnas de P *son las matrices de coordenadas de los vectores base de la nueva base, respecto a la base inicial; es decir,*

$$P = [[\boldsymbol{u}'_1]_S, [\boldsymbol{u}'_2]_S, \ldots, [\boldsymbol{u}'_n]_S], \tag{2.72}$$

esta matriz se conoce como la matriz de transición de la base S' *a la base* S.

Demostración: Como S es una base para V, entonces los vectores en S' pueden expresarse como,

$$\mathbf{v}'_j = \sum_{i=1}^{n} p_{ij}\mathbf{v}_i, \qquad 1 \leq j \leq n, \tag{2.73}$$

así que,

$$[\mathbf{v}'_j]_S = \begin{bmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{nj} \end{bmatrix}, \qquad 1 \leq j \leq n. \tag{2.74}$$

Como tanto S y S' son bases para V, entonces para cualquier vector $\mathbf{v} \in V$ se tiene que,

$$\mathbf{v} = \sum_{j=1}^{n} k_j \mathbf{v}_j, \tag{2.75}$$

de manera que,

$$[\mathbf{v}]_S = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}, \tag{2.76}$$

de la misma manera,

$$\mathbf{v} = \sum_{j=1}^{n} k'_j \mathbf{v}'_j, \tag{2.77}$$

y

$$[\mathbf{v}]_{S'} = \begin{bmatrix} k'_1 \\ k'_2 \\ \vdots \\ k'_n \end{bmatrix}, \qquad 1 \leq i \leq n. \tag{2.78}$$

Sustituyendo la ecuación 2.73 en la ecuación 2.75 se tiene,

$$\begin{aligned} \mathbf{v} &= \sum_{j=1}^{n} k'_j \mathbf{v}'_j & (2.79) \\ &= \sum_{j=1}^{n} k'_j \left(\sum_{i=1}^{n} p_{ij} \mathbf{v}_i \right) \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} k'_j p_{ij} \right) \mathbf{v}'_i \end{aligned}$$

$$\tag{2.80}$$

Comparando las ecuaciones 2.80 y 2.77 se observa que,

$$k_i = \sum_{j=1}^{n} k'_j p_{ij}, \qquad 1 \leq i \leq n \tag{2.81}$$

La representación matricial de la ecuación 2.81 es,

$$\begin{bmatrix} k_1 \\ k_2 \\ \cdots \\ k_n \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} k'_1 \\ k'_2 \\ \vdots \\ k'_n \end{bmatrix} \tag{2.82}$$

Finalmente, haciendo

$$P = [[\mathbf{u}'_1]_S, [\mathbf{u}'_2]_S, \ldots, [\mathbf{u}'_n]_S]$$

concluimos que,

$$[\mathbf{v}]_S = P[\mathbf{v}]'_S. \quad \blacksquare$$

Ejemplo 2.6.3 *Considere las bases $S = \{\boldsymbol{e}_1, \boldsymbol{e}_1, \boldsymbol{e}_3\}$ y $S' = \{\boldsymbol{u}'_1, \boldsymbol{u}'_2, \boldsymbol{u}'_3\}$ para $\mathbb{R}^3$, donde*

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad y \quad \boldsymbol{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

Además,

$$\boldsymbol{u}'_1 = \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}, \boldsymbol{u}'_2 = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} \quad y \quad \boldsymbol{u}'_3 = \begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}.$$

a) Hallar la matriz de transición de S' hacia S.

b) Hallar $[\boldsymbol{v}]_S$, si $[\boldsymbol{v}]_{S'} = \begin{bmatrix} 7 \\ 2 \\ 4 \end{bmatrix}$.

Solución:

a) Para construir la matriz P debemos hallar la matriz de coordenadas respecto a la base S de los vectores en la base S'. Como en este caso S es la base canónica para $\mathbb{R}^3$ entonces por simple inspección tenemos,

$$[\mathbf{u}'_1]_S = \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}, \quad [\mathbf{u}'_2]_S = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}, \quad [\mathbf{u}'_3]_S = \begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}.$$

De manera que la matriz de transición de la base S' a la base S es:

$$P = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix}.$$

b) De la ecuación 2.71 se tiene,

$$[\mathbf{v}]_S = P[\mathbf{v}]'_S = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix} \begin{bmatrix} 7 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 28 \\ -1 \\ 45 \end{bmatrix}.$$

Teorema 2.6.3 *Si P es la matriz de transición de una base S' a otra base S de un espacio vectorial, entonces* *a)* P *es inversible.*

b) P^{-1} *es la matriz de transición de* S *a* S'.

Demostración: Sean $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ y $S' = \{\mathbf{u}'_1, \mathbf{u}'_2, \ldots, \mathbf{u}'_n\}$ dos bases diferentes de V, sea P la matriz de transición de la base S' hacia la base S y sea Q la matriz de transición de la base S hacia la base S'; es decir, si $\mathbf{v}$ es cualquier vector en V, entonces

$$[\mathbf{v}]_S = P[\mathbf{v}]_{S'}, \tag{2.83}$$

y

$$[\mathbf{v}]_{S'} = Q[\mathbf{v}]_S. \tag{2.84}$$

Probaremos que $QP = I$, por lo tanto, P será inversible y su inversa será Q.

Sustituyendo la ecuación 2.83 en la ecuación 2.84,

$$[\mathbf{v}]_{S'} = QP[\mathbf{v}]_{S'}.$$

Sea $C = QP$ entonces,

$$[\mathbf{v}]_{S'} = C[\mathbf{v}]_{S'}. \tag{2.85}$$

Además, podemos expresar C en términos de sus entradas.

$$C = \begin{bmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{bmatrix}.$$

La ecuación 2.85 es válida para todo vector $\mathbf{v} \in V$, en particular se satisface para los vectores de la base S'. En este caso tenemos,

$$[\mathbf{u}'_1]_{S'} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad [\mathbf{u}'_2]_{S'} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad [\mathbf{u}'_n]_{S'} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

De manera que,

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{bmatrix}.$$

De la misma manera, sustituyendo el resto de los vectores del conjunto S' se concluye que,

$$C = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = I.$$

Hemos demostrado que $C = QP = I$, como P es una matriz cuadrada, entonces de acuerdo con el teorema 1.3.7 P es inversible y Q es su inversa. ∎

Ejemplo 2.6.4 *Para el ejemplo 2.6.3 determine la matriz de transición de S a S',*

a) de manera directa,

b) mediante el teorema anterior.

Solución:

a) Para construir la matriz Q de manera directa, debemos hallar la matriz de coordenadas de los vectores en la base S respecto a la base S' ; es decir, debemos resolver,

$$\begin{aligned} Q_{11}\mathbf{u}'_1 + Q_{21}\mathbf{u}'_2 + Q_{31}\mathbf{u}'_3 &= \mathbf{e}_1, \\ Q_{12}\mathbf{u}'_1 + Q_{22}\mathbf{u}'_2 + Q_{32}\mathbf{u}'_3 &= \mathbf{e}_2, \\ Q_{13}\mathbf{u}'_1 + Q_{23}\mathbf{u}'_2 + Q_{33}\mathbf{u}'_3 &= \mathbf{e}_3. \end{aligned}$$

Cada una de las ecuaciones vectoriales anteriores nos conduce a un sistema matricial equivalente, como se muestra a continuación:

$$\begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix} \begin{bmatrix} Q_{11} \\ Q_{21} \\ Q_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix} \begin{bmatrix} Q_{12} \\ Q_{22} \\ Q_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix} \begin{bmatrix} Q_{13} \\ Q_{23} \\ Q_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Tenemos que resolver tres ecuaciones matriciales de la forma $Ax_i = b_i$. El método recomendado es obtener la inversa de la matriz A y posteriormente hallar las x_i mediante multiplicaciones matriciales; es decir, $x_i = A^{-1}b_i$. La inversa de la matriz A fue determinada en el ejemplo 1.3.4, retomaremos el resultado.

$$\frac{1}{10}\begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix}.$$

Por lo argumentado anteriormente,

$$\begin{bmatrix} Q_{11} \\ Q_{21} \\ Q_{31} \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 11 \\ -16 \\ -7 \end{bmatrix},$$

$$\begin{bmatrix} Q_{12} \\ Q_{22} \\ Q_{32} \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 13 \\ 18 \\ -11 \end{bmatrix},$$

$$\begin{bmatrix} Q_{13} \\ Q_{23} \\ Q_{33} \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} -5 \\ 10\ 1 \end{bmatrix}$$

De manera que,

$$Q = \frac{1}{10}\begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix}.$$

b) En este caso sabemos que $Q = P^{-1}$, donde P es la matriz que se determinó en el ejemplo 2.6.3.

$$P = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 2 & -3 \\ 5 & 3 & 1 \end{bmatrix}$$

La inversa de esta matriz es:

$$P^{-1} = \frac{1}{10}\begin{bmatrix} 11 & 13 & -5 \\ -16 & -18 & 10 \\ -7 & -11 & 5 \end{bmatrix} = Q.$$

Teorema 2.6.4 *Si P es la matriz de transición de una base ortonormal hacia otra base ortonormal, para un espacio de productos interiores, entonces* $P^{-1} = P^T$.

Demostración: Sea V un espacio con productos interiores de dimensión finita n y sean $S = \{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2, \ldots, \widehat{\mathbf{u}}_n\}$ y $S' = \{\widehat{\mathbf{u}}'_1, \widehat{\mathbf{u}}'_2, \ldots, \widehat{\mathbf{u}}'_n\}$ dos bases para V, entonces

$$\widehat{\mathbf{u}'}_j = \sum_{i=1}^{n} P_{ij}\widehat{\mathbf{u}}_i = \sum_{i=1}^{n} \langle \widehat{\mathbf{u}}_i, \widehat{\mathbf{u}'}_j \rangle \widehat{\mathbf{u}}_i$$

donde

$$P_{ij} = \langle \widehat{\mathbf{u}}_i, \widehat{\mathbf{u}'}_j \rangle, \tag{2.86}$$

de manera análoga,

$$\widehat{\mathbf{u}}_j = \sum_{i=1}^{n} Q_{ij}\widehat{\mathbf{u}'}_i = \sum_{i=1}^{n} \langle \widehat{\mathbf{u}}'_i, \widehat{\mathbf{u}}_j \rangle \widehat{\mathbf{u}'}_i$$

donde

$$Q_{ij} = \langle \widehat{\mathbf{u}}'_i, \widehat{\mathbf{u}}_j \rangle. \tag{2.87}$$

Por otro lado,

$$\begin{aligned} [P^T]_{ij} &= P_{ji} \\ &= \langle \widehat{\mathbf{u}}_j, \widehat{\mathbf{u}}'_i \rangle \\ &= \langle \widehat{\mathbf{u}}'_i, \widehat{\mathbf{u}}_j \rangle \\ &= Q_{ij} = [Q]_{ij} \end{aligned}$$

Se observa que, $P^T = Q$. y por el teorema anterior, $P^T = Q = P^{-1}$. ∎

Ejemplo 2.6.5 (Rotaciones en el plano) *Considérese el espacio vectorial $\mathbb{R}^2$ con el producto euclidiano interior. Sea $S = \{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ la base estándar para este espacio y sea $S' = \{\boldsymbol{e}'_1, \boldsymbol{e}'_2\}$ la base ortogonal que se obtiene mediante una rotación de los ejes en un ángulo θ, en el sentido contrario a las manecillas del reloj, como se muestra en la figura 2.1. Los vectores de la base S' pueden expresarse como:*

$$\begin{aligned} \boldsymbol{e}'_1 &= \cos\theta\, \boldsymbol{e}_1 + \operatorname{sen}\theta\, \boldsymbol{e}_2, \\ \boldsymbol{e}'_2 &= -\sin\theta\, \boldsymbol{e}_1 + \cos\theta\, \boldsymbol{e}_2, \end{aligned}$$

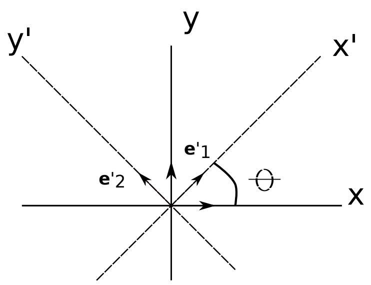

Figura 2.1: Rotación.

de manera que la matriz de transición P de la base S' a S es:

$$P = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

La matriz de transición Q de la base S a S' es:

$$Q = P^T = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

Si $\theta = \frac{\pi}{4}$*, entonces* $\operatorname{sen}\frac{\pi}{4} = \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}$*, por lo tanto, para este ángulo específico, las matrices anteriores se expresan como:*

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Definición 2.6.3 (Matriz ortogonal) *Una matriz es ortogonal si ésta es invertible y si su inversa es igual a su transpuesta.*

Teorema 2.6.5 *Las proposiciones son equivalentes,*

a) A es ortogonal.

b) Los vectores renglón de A forman un conjunto ortonormal en $\mathbb{R}^n$, con el producto euclidiano interior.

c) Los vectores columna de A forman un conjunto ortonormal en $\mathbb{R}^n$*, con el producto euclidiano interior.*

Demostración: Seguiremos la línea demostrativa $a) \Rightarrow b) \Rightarrow c) \Rightarrow a)$.

$a) \Rightarrow b)$ En este caso suponemos que A es una matriz ortogonal; es decir, $A^{-1} = A^T$ y debemos probar que los renglones de A son ortonormales bajo el producto euclidiano interior. Como $A^{-1} = A^T$, entonces $AA^T = I$, luego,

$$\begin{aligned}
[AA^T]_{ij} &= [I]_{ij}, \\
\sum_{k=1}^{n} [A]_{ik}[A^T]_{kj} &= \delta_{ij}, \\
\sum_{k=1}^{n} [A]_{ik}[A]_{jk} &= \delta_{ij}, \\
\langle \mathbf{r}_i, \mathbf{r}_j \rangle &= \delta_{ij}.
\end{aligned}$$

$b) \Rightarrow c)$ La matriz A^T también es ortogonal ya que $(A^T)^{-1} = A = (A^T)^T$. Por lo tanto, para los renglones de A^T también se cumple que,

$$\begin{aligned}
\langle \mathbf{r}_i^T, \mathbf{r}_j^T \rangle &= \delta_{ij}, \\
\langle \mathbf{c}_i, \mathbf{c}_j \rangle &= \delta_{ij},
\end{aligned}$$

$c) \Rightarrow d)$ En este caso sabemos que los vectores columna de la matriz A son ortonormales; es decir,

$$\begin{aligned}
\langle \mathbf{c}_i, \mathbf{c}_j \rangle &= \delta_{ij}, \\
\langle \mathbf{r}_i^t, \mathbf{r}_j^t \rangle &= \delta_{ij}, \\
\sum_{k=1}^{n} [A^T]_{ik}[A^T]_{jk} &= \delta_{ij}, \\
\sum_{k=1}^{n} [A^T]_{ik}[A]_{kj} &= \delta_{ij}.
\end{aligned}$$

$$[A^T A]_{ij} = [I]_{ij}.$$

De lo anterior, $A^T A = I$, como A es una matriz cuadrada, entonces por el teorema 1.3.7, $A^{-1} = A^T$, por lo tanto, A es ortogonal. ∎

Ejemplo 2.6.6 *Hallar la inversa de la matriz B, donde*

$$B = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solución: Debido a que los renglones de esta matriz forman un conjunto ortonormal, entonces la matriz es ortogonal y su inversa es simplemente su transpuesta,

$$B^{-1} = B^T = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta & 0 \\ -\operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Capítulo 3

Transformaciones lineales

En este capítulo se definen y se estudian las propiedades de las transformaciones lineales. Se aprenderá como obtener sus representaciones matriciales y como operar con ellas para obtener la imágen de un vector dado. Además, se mostrará como cambia la matriz que representa a una transformación lineal, mediante una transformación de coordenadas.

3.1. Definición y propiedades

Definición 3.1.1 (Aplicaciones vectoriales) *Si V y W son espacios vectoriales y T es una regla que asocia con cada vector $\boldsymbol{v} \in V$ uno y solo un vector $\boldsymbol{w} \in W$, entonces se dice que T aplica o mapea V en W y se denota por $T : V \to W$. El único vector $\boldsymbol{w}$ que T asocia con el vector $\boldsymbol{v}$ se conoce como la imagen de $\boldsymbol{v}$ bajo T y se denota por $\boldsymbol{w} = T(\boldsymbol{v})$.*

Ejemplo 3.1.1 *Si $T : \mathbb{R}^2 \to \mathbb{R}^3$ está definida como,*

$$T(x, y) = (2x, 2y, x + y).$$

Hallar la imagen de $\boldsymbol{v} = (1, 1)$ bajo T.

Solución:.

$$\mathbf{w} = T(\mathbf{v}) = T(1, 1) = (2(1), 2(1), 1 + 1) = (2, 2, 2).$$

Definición 3.1.2 (Transformaciones lineales) *Si $T : V \to W$ es una aplicación que mapea el espacio vectorial V al espacio vectorial W, entonces T será una transformación lineal si se cumplen las siguientes condiciones:*

a) $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$.

b) $T(k\boldsymbol{v}) = kT(\boldsymbol{v})$.

Ejemplo 3.1.2 *Si* $T : \mathbb{R}^2 \to \mathbb{R}^3$ *está definida como,*

$$T(x, y) = (2x, 2y, x + y),$$

pruebe que T *es una transformación lineal.*

Solución: Sea k un escalar cualquiera y sean $\mathbf{u} = (u_1, u_2)$ y $\mathbf{v} = (v_1, v_2)$ vectores de $\mathbb{R}^2$, se verificará que se cumplen las dos condiciones de la definición.

a)

$$\begin{aligned}
T(\mathbf{u} + \mathbf{v}) &= T((u_1, u_2) + (v_1, v_2)) = T(u_1 + v_1, u_2 + v_2),\\
&= (2(u_1 + v_1), 2(u_2 + v_2), (u_1 + v_1) + (u_1 + v_2)),\\
&= (2u_1 + 2v_1, 2u_2 + 2v_2, u_1 + v_1 + u_2 + v_2),\\
&= (2u_1 + 2v_1, 2u_2 + 2v_2, (u_1 + u_2) + (v_1 + v_2)),\\
&= (2u_1, 2u_2, (u_1 + u_2)) + (2v_1, 2v_2, (v_1 + v_2)),\\
&= T(\mathbf{u}) + T(\mathbf{v}).
\end{aligned}$$

b)

$$\begin{aligned}
T(k(\mathbf{u})) &= T(k(u_1, u_2)) = T(ku_1, ku_2),\\
&= (2ku_1, 2ku_2, ku_1 + ku_2) = k(2u_1, 2u_2, u_1 + u_2),\\
&= kT(\mathbf{u}).
\end{aligned}$$

Debido a que las dos condiciones de la definición se satisfacen, T es una transformación lineal.

Ejemplo 3.1.3 *Sea* A *una matriz de mxn y considérese la notación matricial para los vectores en* $\mathbb{R}^n$ *y* $\mathbb{R}^m$*. Pruebe que* $T : \mathbb{R}^n \to \mathbb{R}^m$ *definida como* $T(\boldsymbol{x}) = A\boldsymbol{x}$ *es una transformación lineal.*

Solución: Sean $\mathbf{x}$ y $\mathbf{y}$ vectores de $\mathbb{R}^n$ y k un escalar cualquiera, donde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad y \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

además,

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

Se verificará que se satisfacen las dos condiciones de la definición.

a)

$$\begin{aligned}
T(\mathbf{x}+\mathbf{y}) &= T\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}\right), \\
&= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}\right), \\
&= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \\
&= T(\mathbf{x}) + T(\mathbf{y}).
\end{aligned}$$

b)

$$\begin{aligned}
T(k(\mathbf{x})) &= T\left(k\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right), \\
&= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} kx_1 \\ kx_2 \\ \vdots \\ kx_n \end{bmatrix}, \\
&= k\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \\
&= kT(\mathbf{x}).
\end{aligned}$$

Se ha verificado que se satisfacen las dos condiciones de la definición, por lo tanto, T es una transformación lineal.

Ejemplo 3.1.4 *Sea V un espacio vectorial de dimensión n y $S = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ una base para V. Si $T : V \to \mathbb{R}^n$ se define como, $T(\boldsymbol{v}) = (\boldsymbol{v})_S$. Pruebe que T es una transformación lineal.*

Solución: Sea k un escalar y sean $\mathbf{u}$ y $\mathbf{v}$ vectores de $\mathbb{R}^n$, como S es una base para V, entonces,

$$\mathbf{u} = \sum_{i=1}^{n} u_i \mathbf{w}_i,$$

$$\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{w}_i,$$

$$\mathbf{u} + \mathbf{v} = \sum_{i=1}^{n} (u_i + v_i)\mathbf{w}_i,$$

y

$$k\mathbf{u} = \sum_{i=1}^{n} (ku_i)\mathbf{w}_i,$$

de manera que,

$$(\mathbf{u})_S = (u_1, u_2, \ldots, u_n),$$

$$(\mathbf{v})_S = (v_1, v_2, \ldots, v_n),$$

$$(\mathbf{u} + \mathbf{v})_S = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n),$$

y

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

Verifíquese que se satisfacen las dos propiedades de la definición.

a)

$$\begin{aligned}
T(\mathbf{u} + \mathbf{v}) &= (\mathbf{u} + \mathbf{v})_S = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n), \\
&= (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n), \\
&= (\mathbf{u})_S + (\mathbf{v})_S = T(\mathbf{u}) + T(\mathbf{v}).
\end{aligned}$$

b)

$$\begin{aligned} T(k\mathbf{u}) &= (k\mathbf{u})_S = (ku_1, ku_1, \ldots, ku_n), \\ &= k(u_1, u_2, \ldots, u_n) = k(\mathbf{u})_S, \\ &= kT(\mathbf{u}). \end{aligned}$$

Dado que se satisfacen las dos condiciones, T es una transformación lineal.

Ejemplo 3.1.5 (Transformación identidad) *Sea V cualquier espacio vectorial y sea $T : V \to V$ definida como $T(\boldsymbol{v}) = \boldsymbol{v}$. Demuestre que T es una transformación lineal.*

Solución: Sea k un escalar cualquiera y sean **u** y **v** vectores en V, entonces,

a)
$$T(\mathbf{u} + \mathbf{v}) = \mathbf{u} + \mathbf{v} = T(\mathbf{u}) + T(\mathbf{v}).$$

b)
$$T(k\mathbf{u}) = k\mathbf{u} = kT(\mathbf{u}).$$

T es una transformación lineal, ya que satisface las condiciones de la definición.

Ejemplo 3.1.6 (Transformaciones de dilatación y contracción) *Sea V cualquier espacio vectorial y k cualquier escalar, pruebe que $T : V \to V$ definida por $T(\boldsymbol{v}) = k\boldsymbol{v}$ es una transformación lineal.*

Solución: Sea r un escalar y sean **u** y **v** vectores en V, entonces,

a)
$$T(\mathbf{u} + \mathbf{v}) = k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v} = T(\mathbf{u}) + T(\mathbf{v}).$$

b)
$$T(r\mathbf{u}) = k(r\mathbf{u}) = r(k\mathbf{u}) = rT(\mathbf{u}).$$

T es una transformación lineal, ya que satisface las dos condiciones de la definición.

Teorema 3.1.1 *Si $T : V \to W$ es una transformación lineal, entonces*

a) $T(\boldsymbol{0}) = \boldsymbol{0}$.

b) $T(-\boldsymbol{v}) = -T(\boldsymbol{v})$.

c) $T(\boldsymbol{u} - \boldsymbol{v}) = T(\boldsymbol{u}) - T(\boldsymbol{v})$.

Demostración:

a) Se sabe que $0\mathbf{v} = \mathbf{0}$, para todo vector $\mathbf{v} \in V$, de manera que,

$$T(\mathbf{0}) = T(0\mathbf{v}) = 0T(\mathbf{v}) = \mathbf{0}.$$

b)

$$T(-\mathbf{v}) = T((-1)\mathbf{V}) = (-1)T(\mathbf{v}) = -T(\mathbf{v})$$

c)

$$\begin{aligned} T(\mathbf{u} - \mathbf{v}) &= T(\mathbf{u} + (-\mathbf{v})) = T(\mathbf{u}) + T(-\mathbf{v}), \\ &= T(\mathbf{u}) - T(\mathbf{v}). \quad \blacksquare \end{aligned}$$

Teorema 3.1.2 *Sea $T : V \to W$ una transformación lineal, entonces para cualesquiera vectores $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ y escalares $\{c_1, c_2, \ldots, c_k\}$ se tiene que,*

$$T(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_k\boldsymbol{v}_k) = c_1T(\boldsymbol{v}_1) + c_2T(\boldsymbol{v}_2) + \ldots + c_kT(\boldsymbol{v}_k).$$

Demostración: Se recurrirá al método de inducción matemática.

a) Verifíquese que el resultado es válido para algún valor de k. En este caso se prueba que el resultado es válido para $k = 2$. Como V es un espacio vectorial, entonces es cerrado bajo la multiplicación por un escalar, de manera que, $\mathbf{u}_1 = c_1\mathbf{v}_1$ y $\mathbf{u}_2 = c_2\mathbf{v}_2$ están en V. Además, V es cerrado bajo la suma; así que $\mathbf{u}_1 + \mathbf{u}_2$ está en V. Como T es una transformación lineal que aplica V en W, entonces,

$$\begin{aligned} T(\mathbf{u}_1 + \mathbf{u}_2) &= T(\mathbf{u}_1) + T(\mathbf{u}_2), \\ T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) &= T(c_1\mathbf{v}_1) + T(c_2\mathbf{v}_2), \\ T(c_1\mathbf{v}_1 + c_2\mathbf{u}_2) &= c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2). \end{aligned}$$

b) Supóngase que el resultado es válido para $k - 1$; es decir,

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_{k-1}\mathbf{v}_{k-1}) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \ldots + c_{k-1}T(\mathbf{v}_{k-1}).$$

c) Finalmente se probará que el resuldado es válido para k. Sean $\mathbf{u}_1 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_{k-1}\mathbf{v}_{k-1}$ y $\mathbf{u}_2 = c_k\mathbf{v}_k$, entonces por lo argumentado en los incisos a) y b) se tiene,

$$\begin{aligned} T(\mathbf{u}_1 + \mathbf{u}_2) &= T(\mathbf{u}_1) + T(\mathbf{u}_2), \\ T((c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_{k-1}\mathbf{v}_{k-1}) + c_k\mathbf{v}_k) &= T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_{k-1}\mathbf{v}_{k-1}) + T(c_k\mathbf{v}_k), \\ T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k) &= c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \ldots + c_kT(\mathbf{v}_k). \quad \blacksquare \end{aligned}$$

Teorema 3.1.3 *Sea $T : V \to W$ una transformación lineal de un espacio vectorial de dimensión n hacia un espacio vectorial W de dimensión m. Sea además, $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ una base de V. Si $\boldsymbol{v}$ es cualquier vector en V, entonces $T(\boldsymbol{v})$ queda completamente determinada por, $\{T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)\}$.*

Demostración: Como S representa una base para V, entonces existe un conjunto único de escalares $c_1, c_2, \ldots, c_n$ tales que,

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n.$$

Por el teorema anterior es posible desarrollar,

$$T(\mathbf{v}) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \ldots + c_nT(\mathbf{v}_n).$$

Se aprecia que $T(\mathbf{v})$ queda completamente determinada por los elementos del conjunto $\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\}$. $\blacksquare$

Ejemplo 3.1.7 *Considere la base $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ para $\mathbb{R}^3$, donde $\boldsymbol{v}_1 = (1, 1, 1), \boldsymbol{v}_2 = (1, 1, 0), \boldsymbol{v}_3 = (1, 0, 0)$ y sea T el operador lineal tal que,*

$$T(\boldsymbol{v}_1) = (2, -1, 4), \quad T(\boldsymbol{v}_2) = (3, 0, 1) \quad y \quad T(\boldsymbol{v}_3) = (-1, 5, 1).$$

Obtener una fórmula para $T(x_1, x_2, x_3)$ y usarla para calcular $T(2, 4, -1)$.

Solución: Exprésese el vector $\mathbf{v} = (x_1, x_2, x_3)$ como combinación lineal de los vectores de la base; es decir,

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3.$$

Sustituyendo los vectores dados se obtiene,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

La ecuación anterior se expresa en forma matricial como,

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

De lo anterior se tiene un sistema de ecuaciones de la forma $A\mathbf{c} = \mathbf{x}$. La inversa de la matriz de coeficientes es:

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

La solución del sistema es entonces,

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ x_2 - x_3 \\ x_1 - x_2 \end{bmatrix}.$$

Por el teorema anterior se tiene que,

$$\begin{aligned} T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) &= c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + c_3T(\mathbf{v}_3), \\ &= x_3 \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} + (x_2 - x_3) \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + (x_1 - x_2) \begin{bmatrix} -1 \\ 5 \\ 1 \end{bmatrix}, \\ &= \begin{bmatrix} -x_1 + 4x_2 - x_3 \\ 5x_1 - 5x_2 - x_3 \\ x_1 + 3x_3 \end{bmatrix}. \end{aligned}$$

Finalmente,

$$T\left(\begin{bmatrix} 2 \\ 4 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} -2 + 4(4) - (-1) \\ 5(2) - 5(4) - (-1) \\ 2 + 3(-1) \end{bmatrix} = \begin{bmatrix} 15 \\ -9 \\ -1 \end{bmatrix}.$$

3.2. Imagen y kernel de una transformación

Definición 3.2.1 (Kernel de la transformación) *Si $T : V \to W$ es una transformación lineal, entonces el conjunto de vectores en V que T aplica al vector cero($\mathbf{0} \in W$) se conoce como núcleo, kernel o espacio nulo de la transformación y se denota por $ker(T)$.*

Definición 3.2.2 (Recorrido de la transformación) *Si $T : V \to W$ es una transformación lineal, entonces el conjunto de todos los vectores en W que son imágenes bajo T de almenos un vector en V se conoce como recorrido de T y se denota por $R(T)$.*

Ejemplo 3.2.1 (Transformación identidad) *Si $T : V \to V$ es la transformación identidad; es decir, $T(\boldsymbol{v})=\boldsymbol{v}$, para todo $\boldsymbol{v}$ en V, entonces el kernel de la transformación consiste únicamente del vector cero. El recorrido de la transformación es el espacio vectorial V.*

Ejemplo 3.2.2 (Transformación cero) *Si $T : V \to W$ es la transformación lineal cero; es decir, $T(\boldsymbol{v})=\boldsymbol{0}$, para todo $\boldsymbol{v}$ en V, entonces $ker(T)=V$ y $R(T) = \{\boldsymbol{0}\}$.*

Ejemplo 3.2.3 (Multiplicación por una matriz) *Si $T : \mathbb{R}^n \to \mathbb{R}^m$ es la transformación lineal definida como $T(\boldsymbol{x})=A\boldsymbol{x}$, donde A es una matriz de tamaño mxn, entonces*

$$ker(T) = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{0}\};$$

es decir, $ker(T)$ es el espacio nulo de la matriz A; además,

$$R(T) = \{\boldsymbol{w} \in \mathbb{R}^m : \boldsymbol{w} = A\boldsymbol{x},\ donde\ \boldsymbol{x} \in \mathbb{R}^n\} = \mathcal{L}(S_c),$$

es decir, $R(T)$ es el espacio de columnas de la matriz A.

Teorema 3.2.1 *Si $T : V \to W$ es una transformación lineal, entonces:*

a) El núcleo de T es un subespacio de V.

b) El recorrido de T es un subespacio de W.

Demostración:

a) Se sabe que para transformaciones lineales $T(\mathbf{0}) = \mathbf{0}$; de manera que, $ker(T)$ es un conjunto diferente del vacío. supóngase ahora que $\mathbf{v}_1$ y $\mathbf{v}_2$ están en $ker(T)$ y que k es un escalar cualquiera. Se requiere probar que $\mathbf{v}_1 + \mathbf{v}_2$ y $k\mathbf{v}_1$ están en $ker(T)$. Para tal fin considérese lo siguiente:

$$T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

y

$$T(k\mathbf{v}_1) = kT(\mathbf{v}_1) = k\mathbf{0} = \mathbf{0}.$$

Hemos verificado que $ker(T)$ satisface las propiedades de cerradura tanto para la suma como para la multiplicación por un escalar; por lo tanto, $ker(T)$ es un subespacio de V.

b) Se sabe que $\mathbf{0} = T(\mathbf{0}) \in R(T)$, de manera que $R(T)$ es un conjunto distinto del vacío. supóngase ahora que $\mathbf{w}_1$ y $\mathbf{w}_2$ están en $R(T)$; es decir, existen elementos $\mathbf{v}_1$ y $\mathbf{v}_2$ en V tales que, $\mathbf{w}_1 = T(\mathbf{v}_1)$ y $\mathbf{w}_2 = T(\mathbf{v}_2)$. Se quiere probar que $\mathbf{w}_1 + \mathbf{w}_2$ y $k\mathbf{w}_1$ están en W, donde k es un escalar cualquiera. Para tal fin considérese lo siguiente:

$$T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2,$$

$$T(k\mathbf{v}_1) = kT(\mathbf{v}_1) = k\mathbf{w}_1.$$

Lo anterior prueba que $\mathbf{w}_1 + \mathbf{w}_2$ y $k\mathbf{w}_1$ son imagenes de vectores en V; por lo tanto, están en $R(T)$, con lo cual se concluye que $R(T)$ es un subespacio de W. ∎

Definición 3.2.3 (Nulidad y rango de la transformación) *La dimensión del kernel de la transformación se conoce como nulidad de la transformación y se representa por $\nu(T)$. En este mismo sentido, la dimensión del recorrido de la transformación se conoce como rango de la transformación y se representa por $\rho(T)$.*

Teorema 3.2.2 *Si A es una matriz de mxn y $T : \mathbb{R}^n \to \mathbb{R}^m$ está definida por T($\boldsymbol{x}$)=A$\boldsymbol{x}$, entonces:*

a) $\nu(T) = \nu(A)$.

b) $\rho(T) = \rho(A)$.

Demostración: Como se mostró en el ejemplo 3.2.3,

$$ker(T) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\};$$

$$R(T) = \mathcal{L}(S_c);$$

por lo tanto, considerando las dimensiones de estos espacios se concluye lo que se afirma en el teorema. ∎

Teorema 3.2.3 *Si $T : V \to W$ es una transformación lineal de un espacio vectorial V, de dimensión n, a un espacio vectorial W, entonces,*

$$\nu(T) + \rho(T) = n.$$

Demostración: Sea k la dimensión de $ker(T)$ y considere los casos en que $1 \leq k \leq n$. Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ una base para $ker(T)$, entonces por el teorema 2.4.8 se sabe que existe una base para V que contiene a S; es decir, existen $(n-k)$ vectores $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ tales que $S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ es una base para V.
Ahora se requiere probar que el conjunto $S'' = \{T(\mathbf{v}_{k+1}), T(\mathbf{v}_{k+2}), \ldots, T(\mathbf{v}_n)\}$ es una base para $R(T)$. Se probará que S'' genera a V, para tal fin supóngase que $\mathbf{w} \in R(T)$, en este caso, existe un $\mathbf{v} \in V$ tal que, $\mathbf{w} = T(\mathbf{v})$. Como S' es una base para V y $\mathbf{v} \in V$, entonces es posible expresar este vector como combinación lineal de los vectores de la base; es decir,

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 \ldots + a_k\mathbf{v}_k + a_{k+1}\mathbf{v}_{k+1}, \ldots + a_n\mathbf{v}_n.$$

Ahora considere $T(\mathbf{v})$.

$$T(\mathbf{v}) = T(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 \ldots + a_k\mathbf{v}_k + a_{k+1}\mathbf{v}_{k+1} + \ldots + a_n\mathbf{v}_n).$$

Por el teorema 3.1.2

$$T(\mathbf{v}) = a_1T(\mathbf{v}_1) + a_2T(\mathbf{v}_2) + \ldots + a_kT(\mathbf{v}_k) + a_{k+1}T(\mathbf{v}_{k+1}) + \ldots + a_nT(\mathbf{v}_n).$$

Los primeros k términos se anulan, ya que la transformación actúa sobre vectores que pertecenen a $ker(T)$, por lo tanto,

$$T(\mathbf{v}) = a_{k+1}T(\mathbf{v}_{k+1}) + \ldots + a_nT(\mathbf{v}_n) = \mathbf{w}.$$

Con esto hemos probado que S'' genera a $R(T)$.
Ahora se demostrará que S'' es linealmente independiente. Para tal fin considerérese la ecuación vectorial,

$$a_{k+1}T(\mathbf{v}_{k+1}) + \ldots + a_nT(\mathbf{v}_n) = \mathbf{0}.$$

Si probamos que los escales en la ecuación anterior son todos iguales a cero, entonces S'' será linealmente independiente. De acuerdo con el teorema 3.1.2, la ecuación anterior puede reagruparse, de manera que,

$$T(a_{k+1}\mathbf{v}_{k+1} + \ldots + a_n\mathbf{v}_n) = \mathbf{0}.$$

Se observar que el vector $a_{k+1}\mathbf{v}_{k+1} + \ldots + a_n\mathbf{v}_n$ está en el kernel de la transformación, por lo tanto, puede ser expresado como combinación lineal de los vectores en S; es decir,

$$a_{k+1}\mathbf{v}_{k+1} + \ldots + a_n\mathbf{v}_n = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \ldots + b_k\mathbf{v}_k.$$

por lo anterior se tiene que,

$$b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \ldots + b_k\mathbf{v}_k - a_{k+1}\mathbf{v}_{k+1} - \ldots - a_n\mathbf{v}_n = \mathbf{0}$$

Como los vectores en S' son linealmente independientes, entonces todos los escalares en la ecuación anterior son iguales a cero, particularmente, $a_{k+1} = \ldots = a_n = 0$. Concluimos que S'' es linealmente independiente.
Como S'' es una base para $R(T)$, entonces $\nu(T) = dim(R(T)) = n - k$. En conclusión, $\nu(T) + \rho(T) = n$. Los casos en que $k = n$ y $k = 0$ se dejan como ejercicio para el estudiante. ▮

Ejemplo 3.2.4 *Sea $T : \mathbb{R}^5 \to \mathbb{R}^4$ la transformación lineal definidad como,*

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & -1 & 3 & -1 \\ 1 & 0 & 0 & 2 & -1 \\ 2 & 0 & -1 & 5 & -1 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

a) Determine una base para $ker(T)$.

b) Determine una base para $R(T)$.

c) Verifique el teorema 3.2.3

Solución:

a) De acuerdo con el teorema 3.2.2, se debe hallar una base para el espacio nulo de la matriz A. A continuación se resuelve la homogénea, reduciendo la matriz mediante operaciones elementales, para determinar dicha base.

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

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

Las variables principales del sistema son, x_1, x_3 y x_5, mientras que las variables libres son x_2 y x_3. Las euaciones asociadas a la matriz escalonada son:

$$\begin{aligned} x_1 + 2x_4 &= 0 \\ x_3 - x_4 &= 0 \\ x_5 &= 0 \end{aligned}$$

Parametrizando las variables libres y despejando las variables principales en términos de las variables libres se obtiene,

$$\begin{aligned} x_1 &= -2t \\ x_2 &= r \\ x_3 &= t \\ x_4 &= t \\ x_5 &= 0 \end{aligned}$$

En forma matricial,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2t \\ r \\ t \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} + r \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Los vectores

$$\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

son linealmente independientes y generan a $ker(T)$, por lo tanto, forman una base para $ker(T)$. En este caso, $\nu(T) = dim(ker(T)) = 2$.

b) Se sabe que $R(T)$ corresponde al espacio de columnas de la matriz A, de manera que reduciendo la transpuesta de la matriz A obtendremos una base para el espacio de columnas, como se mostró en la sección 2.4.2. Considérese ahora la transpuesta de la matriz A y realice operaciones elementales sobre sus renglones.

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

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

Los vectores columna diferentes de cero en la transpuesta de la última matriz constituyen una base para el espacio de columnas de la matriz A, y por lo tanto, para el recorrido de la transformación. Estos vectores son,

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

En este caso $\rho(T) = dim(R(T)) = 3$

c) De auerdo con el teorema 3.2.3, $n = \nu(T) + \rho(T)$, en este caso, 5=2+3. Pudimos haber aplicado dicho teorema de la manera siguiente: Una vez que se sabe que $n = 5$ y $\nu(T) = 2$, entonces se tiene que $\rho(T) = 3$, por lo tanto, toda base para $R(T)$ debe contener tres vectores. Por otro lado, al eliminar la columna cero de matriz A se obtienen cuatro vectores columnas que de acuerdo con el teorema 2.4.4, seguirán generando a $R(T)$; además, se observa que la cuarta columna es combinación lineal de las columnas que le anteceden ($c_4 = 2c_1 - c_3$), de manera que se puede extraer también este vector y el conjunto restante seguirá generando a $R(T)$. Finalmente, tres vectores que generan un espacio de dimensión 3, forman una base para dicho espacio, de acuerdo con el teorema 2.4.6. Por lo tanto, el conjunto de vectores,

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ -1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \end{bmatrix} \right\}$$

constituye una base para $R(T)$.

3.3. Representación de una transformación lineal

Teorema 3.3.1 *Sea $T : V \to W$ una transformaciónlineal de un espacio vectorial V, de dimensión n, a un espacio vectorial W de dimensión m, y sean $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ y $S' = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ bases para V y W, respectivamente, entonces la matriz A de tamaño mxn cuya j-ésima columna es el vector de coordenadas $[T(v_j)]_{S'}$, tiene la siguiente propiedad: Si $\boldsymbol{x}$ está en V, entonces*

$$[T(\boldsymbol{x})]_{S'} = A[\boldsymbol{x}]_S, \tag{3.1}$$

donde $[\boldsymbol{x}]_S$ y $[T(\boldsymbol{x})]_{S'}$ son las matrices de coordenadas de los vectores $\boldsymbol{x}$ y $T(\boldsymbol{x})$ con respecto a las bases S y S', respectivamente. La matriz A es única y se conoce como matriz matriz de la transformación T respecto a las bases S y S'. [1]

Demostración: Se demostrará el teorema construyendo la matriz A y verificando que satisface la propiedad señalada en la ecuación 3.1.
Se sabe que $T(\mathbf{v}_j) \in W$ y S' es una base para este espacio, entones es posible expresar al vector $T(\mathbf{v}_j)$ como combinación lineal de los vectores en S'; es decir,

$$T(\mathbf{v}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{w}_i, \quad 1 \leq j \leq n. \tag{3.2}$$

La ecuación anterior nos permite obtener las matrices de coordenadas,

$$[T(\mathbf{v}_j)]_{S'} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, \quad 1 \leq j \leq n, \tag{3.3}$$

de manera que, la matriz cuya j-ésima columna es el vector $[T(\mathbf{v}_j)]_{S'}$ queda representada por,

$$\begin{aligned} A &= \begin{bmatrix} [T(\mathbf{v}_1)]_{S'} & [T(\mathbf{v}_2)]_{S'} & \ldots & [T(\mathbf{v}_n)]_{S'} \end{bmatrix}, \\ A &= \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}. \end{aligned} \tag{3.4}$$

[1]Algunos autores utilizan la notación $[T]_{S',S}$ para destacar que la matriz es única respecto a las bases S y S'.

Ahora considere un vector arbitrario $\mathbf{x} \in V$. Como S es una base para V y $\mathbf{x} \in V$, entonces es posible expresar este vector de la manera siguiente,

$$\mathbf{x} = \sum_{i=1}^{n} b_i \mathbf{v}_i, \tag{3.5}$$

de modo que,

$$[\mathbf{x}]_S = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}. \tag{3.6}$$

Ahora aplique la transformación lineal al vector $\mathbf{x}$.

$$\begin{aligned} T(\mathbf{x}) &= T\left(\sum_{i=1}^{n} b_i \mathbf{v}_i\right) \\ &= \sum_{i=1}^{n} b_i T(\mathbf{v}_i), \quad \text{teorema 3.1.2} \end{aligned}$$

Se sabe, además, que la matriz de coordenadas es una transformación lineal(ejemplo 3.1.4); por lo tanto,

$$\begin{aligned} [T(\mathbf{x})]_{S'} &= \sum_{i=1}^{n} b_i [T(\mathbf{v}_i)]_{S'} \\ &= b_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + b_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + b_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}, \\ &= \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}, \\ &= = A[x]_S. \end{aligned}$$

Resta probar que la matriz A es única, para tal fin, supóngase que existe otra matriz B que satisface la ecuación 3.1, es decir,

$$[T(\mathbf{x})]_{S'} = B[\mathbf{x}]_S. \quad \forall \mathbf{x} \in V \tag{3.7}$$

Restando las ecuaciones 3.1 y 3.7 se obtiene,

$$(A - B)[\mathbf{x}]_S = 0 \tag{3.8}$$

Sea C la matriz de tamaño mxn definida como, $C = A - B$, entonces la ecuación 3.8 puede expresarse como,

$$C[\mathbf{x}]_S = 0 \tag{3.9}$$

La ecuación 3.9 es válida para toda $\mathbf{x} \in V$, en particular se satisface para los vectores de la base S. Para estos vectores se tiene que,

$$[\mathbf{v}_1]_S = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad [\mathbf{v}_2]_S = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \qquad \ldots, \qquad [\mathbf{v}_n]_S = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Representamos a la matriz C mediante sus entradas, como se muestra a continuación, con el fin de desarrollar la operación indicada en la ecuación 3.9 con los vectores de la base S.

$$C = \begin{bmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \ldots & c_{mn} \end{bmatrix}.$$

De acuerdo con la ecuación 3.9 se tiene,

$$C[\mathbf{v}_1]_S = \begin{bmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

La ecuación anterior muestra que la primer columna de la matriz C es idénticamente cero. Sustiuyendo el resto de los de los vectores de la base se concluye que C es la matriz cero; por lo tanto, $A = B$. ∎

Matriz estándar de la transformación

Si $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es una transformación lineal y si, S y S' son las bases estándares para $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente, entonces la matriz que representa a la transformación respecto a S y S' se conoce como matriz estándar de la transformación y se representa por A_T.

Ejemplo 3.3.1 (Multiplicación por una matriz) *Si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal definida por $T(\boldsymbol{x}) = A\boldsymbol{x}$ donde A es una matriz de tamaño mxn, determine la matriz estándar de la transformación.*

Solución: Sean $S = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ la base estándar para $\mathbb{R}^n$ y

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

Para comenzar apliquemos la transformación a los vectores de la base S.

$$\begin{aligned} T(\mathbf{e}_1) &= \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \\ T(\mathbf{e}_2) &= \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \\ \vdots &= \vdots \\ T(\mathbf{e}_n) &= \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}. \end{aligned}$$

Debido a que S' es la base canónica para $\mathbb{R}^m$, entonces

$$[T(\mathbf{e}_i)]_{S'} = T(\mathbf{e}_i), \quad 1 \leq i \leq n.$$

De esta manera, la matriz estándar de la transformación es:

$$A_T = \left[\ [T(\mathbf{e})_1]_{S'} \,\middle|\, T(\mathbf{e})_1]_{S'} \,\middle|\, \ldots \,\middle|\, T(\mathbf{e})_1]_{S'}\ \right] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Por lo anterior, $A_T = A$; es decir, la matriz estándar corresponde a la matriz A.

Ejemplo 3.3.2 *Considere la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^2$, definida como,*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} y+z \\ x-y \end{bmatrix}.$$

Determine la matriz estándar de la transformación.

Solución: Las bases en este caso son:

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

y

$$S' = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

Ahora aplíquese la transformación a los vectores de la base S.

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

$$T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Como hemos elegido la base canónica para $\mathbb{R}^2$, entonces $T(\mathbf{v}_i) = [T(\mathbf{v}_i)]_{S'}$ para $i = 1, 2, 3$; por lo tanto, la matriz estándar de la transformación es:

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

Transformaciones en un mismo espacio

Si $T : V \to V$ es una transformación lineal, entonces se acostumbra usar una sola base de V para representar a la transformación.

Ejemplo 3.3.3 (Transformación identidad) *Si V es un espacio vectorial de dimensión n, si $T : V \to V$ es la transformación identidad, y sea $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ una base cualquiera para V, entonces la matriz que representa a la transformación es,*

$$\begin{aligned} A_T &= [[T(\boldsymbol{v}_1)]_S|[T(\boldsymbol{v}_2)]_S|\ldots|[T(\boldsymbol{v}_n)_S], \\ &= [[\boldsymbol{v}_1]_S|[\boldsymbol{v}_2]_S|\ldots|[\boldsymbol{v}_n]_S], \\ &= \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \ldots & 1 \end{bmatrix}, \\ &= I_n. \end{aligned}$$

Casos generales

Ejemplo 3.3.4 *Para la transformación lineal del ejemplo 3.3.2,*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} y+z \\ x-y \end{bmatrix},$$

y considerando las bases

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

y

$$S' = \left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}.$$

a) Determine la matriz de la transformación respecto a las bases S y S'.

b) Verifique el teorema 3.3.1 para el vector $\boldsymbol{x} = (1, 0, -1)$.

Solución:

a) Aplíquese la transformación a los vectores de la base:

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$T\left(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ -1 \end{bmatrix},$$

$$T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Procedamos ahora a determinar las matrices de coordenadas respecto a la base S', para cada uno de los vectores anteriores. Esto significa hallar los escalares que nos permiten expresar estos vectores como combinaciones lineales de los vectores en S'; es decir,

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = a\begin{bmatrix} 1 \\ 2 \end{bmatrix} + b\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix},$$

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix} = c\begin{bmatrix} 1 \\ 2 \end{bmatrix} + d\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} c \\ d \end{bmatrix},$$

$$\begin{bmatrix} 2 \\ 0 \end{bmatrix} = e\begin{bmatrix} 1 \\ 2 \end{bmatrix} + f\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} e \\ f \end{bmatrix}.$$

Se puede observar que los tres sistemas de ecuaciones posen la misma matriz de coeficientes. En estos casos es conveniente resolver los sistemas reduciendo la matriz aumentada mediante operaciones elementales, agregando las columnas necesarias como se muestra a continuación:

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

La matriz escalonada en los renglones reducida para el sistema anterior es:

$$\left[\begin{array}{cc|ccc} 1 & 0 & \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ 0 & 1 & \frac{1}{3} & \frac{5}{3} & \frac{4}{3} \end{array}\right].$$

Por lo anterior, la matriz de la transformación respecto a las bases S y S' es,

$$A = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ 1 & 5 & 4 \end{bmatrix}.$$

Se observar que esta matriz es diferente a la que se obtuvo en el ejemplo 3.3.2.

b) Para comenzar determínese $T(\mathbf{x})$.

$$T(\mathbf{x}) = T\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

En este caso particular, resulta sencillo determinar $[T(\mathbf{x})]_{S'}$.

$$[T(\mathbf{x})]_{S'} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \tag{3.10}$$

Para hallar $[\mathbf{x}]_S$ exprésese el vector $\mathbf{x}$ como combinación lineal de los vectores de la base S; es decir,

$$\mathbf{x} = a_1 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + a_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

La solución del sistema anterior es:

$$\begin{cases} a_1 = -1, \\ a_2 = -2, \\ a_3 = 2, \end{cases}$$

De modo que,

$$[\mathbf{x}]_S = \begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix}.$$

De acuerdo con el teorema 3.3.1,

$$[T(\mathbf{x})]_{S'} = A[\mathbf{x}]_S = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ 1 & 5 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

Este resultado concuerda con la ecuación 3.10 como era de esperarse.

Ejemplo 3.3.5 *Sea $T : P_1 \to P_2$ definida como $T(\boldsymbol{p}) = x\boldsymbol{p}$ y sean S y S' las bases estándares para P_1 y P_2, respectivamente,*

a) determine la matriz de la transformación respecto a las bases S y S'.

b) Verifique el teorema 3.10 para el vector $\mathbf{q}(x)=3\text{-}2x$.

Solución:

a) Las bases estándares para P_1 y P_2 son $S = \{1, x\}$ y $S' = \{1, x, x^2\}$, respectivamente. Para comenzar apliquemos la transformación a los vectores de la base.

$$\begin{aligned} T(1) &= x, \\ T(x) &= x^2. \end{aligned}$$

Los vectores de coordenadas respecto a la base S' son:

$$[T(1)]_{S'} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

$$[T(x)]_{S'} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Con los vectores anteriores construimos la matriz de la transformación.

$$A = \left[\ [T(1)]_{S'} \ \middle|\ T(x)]_{S'}\ \right] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

b) Verifiquemos el teorema 3.3.1 para el vector $\mathbf{q} = 3 - 2x$. En este caso $T(\mathbf{q}) = 3x - 2x^2$, de manera que,

$$[T(q)]_{S'} = \begin{bmatrix} 0 \\ 3 \\ -2 \end{bmatrix}.$$

Por otra parte,

$$[\mathbf{q}]_S = \begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$

Finalmente,

$$[T(\mathbf{q})]_{S'} = A[\mathbf{q}]_S = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ -2 \end{bmatrix}.$$

3.4. Cambio de base en la representación matricial de una transformación

Teorema 3.4.1 *Sea $T : V \to V$ una transformación lineal, donde V es un espacio vectorial de dimensión finita n. Sean $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ y $S' = \{\boldsymbol{v}'_1, \boldsymbol{v}'_2, \ldots, \boldsymbol{v}'_n\}_n$ bases para V, y sea P la matriz de transición de S' a S. Si A es la matriz que representa a T respecto a S, entonces $P^{-1}AP$ es la matriz que representa a T con respecto a S'.*

Demostración: Se sabe, por el teorema 3.3.1, que es posible hallar matrices A y B tales que:

$$[T(\mathbf{x})]_S = A[\mathbf{x}]_S, \tag{3.11}$$

$$[T(\mathbf{x})]_{S'} = B[\mathbf{x}]_{S'}, \tag{3.12}$$

para todo vector $\mathbf{x} \in V$.
Además, por el teorema 2.6.2 se tiene:

$$[\mathbf{x}]_S = P[\mathbf{x}]_{S'}, \tag{3.13}$$

$$[\mathbf{x}]_{S'} = Q[\mathbf{x}]_S, \tag{3.14}$$

para todo vector $\mathbf{x} \in V$.
Como $T(\mathbf{x}) \in V$, entonces la ecuación 3.14 nos conduce a,

$$[T(\mathbf{x})]_{S'} = Q[T(\mathbf{x})]_S. \tag{3.15}$$

Sustituyendo la ecuación 3.11 en la ecuación 3.15 se obtiene,

$$[T(\mathbf{x})]_{S'} = QA[\mathbf{x}]_S. \tag{3.16}$$

Ahora, sustituyendo la ecuación 3.13 en la la ecuación 3.16 se tiene,

$$[T(\mathbf{x})]_{S'} = QAP[\mathbf{x}]_{S'}. \tag{3.17}$$

Además, de acuerdo con el teorema 3.3.1, la matriz B en la ecuación 3.12 es única. Comparando las ecuaciones 3.17 y 3.12 se concluye que,

$$B = QAP. \tag{3.18}$$

Finalmente, de acuerdo con el teorema 2.6.3, $Q = P^{-1}$, por lo tanto,

$$B = P^{-1}AP. \quad \blacksquare \tag{3.19}$$

Las matrices A y B que se relacionan mediante la ecuación 3.19 se conocen como matrices semejantes. Las propiedades de las matrices semejantes se estudiarán en el siguiente capítulo.

Ejemplo 3.4.1 *Sea* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *la transformación lineal definida por,*

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x - 3y \\ x - 2y \end{bmatrix},$$

y sean S *y* S' *dos bases de* $\mathbb{R}^2$*, donde*

$$S = \left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\} \quad y \quad S' = \left\{\begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}.$$

a) Determine las matrices de la transformación respecto a las bases S *y* S'*.*

b) Verifique que se satisface el teorema anterior.

Solución:

a) Apliquemos la transformación a los vectores de la base S para determinar la matriz de la transformación correspondiente a esta base,

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -3 \\ -2 \end{bmatrix},$$

de manera que la matriz de la transformación respecto a la base S es,

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

Ahora determine la matriz de la transformación respecto a la base S', para lo cual es necesario aplicar la transformación a los vectores de la base.

$$T\left(\begin{bmatrix} 3 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

Las matrices de coordenada de los vectores anteriores respecto a la base S' son:

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix}_{S'} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ -1 \end{bmatrix}_{S'} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

La matriz de la transformación respecto a la base S' es,

$$B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{3.20}$$

b) Para verificar que se satisface el teorema anterior, se necesita determinar la matriz P de transición de S' a S y también su inversa. Se comenzará por determinar P.

$$\begin{aligned} P &= \left[\left[\begin{array}{c} 3 \\ 1 \end{array}\right]_S \left[\begin{array}{c} 1 \\ 1 \end{array}\right]_S \right], \\ &= \left[\begin{array}{cc} 3 & 1 \\ 1 & 1 \end{array}\right]. \end{aligned}$$

La inversa de esta matriz es,

$$P^{-1} = \frac{1}{2}\left[\begin{array}{rr} 1 & -1 \\ -1 & 3 \end{array}\right].$$

Finalmente, considere el producto $P^{-1}AP$.

$$\begin{aligned} P^{-1}AP &= \frac{1}{2}\left[\begin{array}{rr} 1 & -1 \\ -1 & 3 \end{array}\right]\left[\begin{array}{rr} 2 & -3 \\ 1 & -2 \end{array}\right]\left[\begin{array}{cc} 3 & 1 \\ 1 & 1 \end{array}\right], \\ &= \frac{1}{2}\left[\begin{array}{rr} 1 & -1 \\ -1 & 3 \end{array}\right]\left[\begin{array}{rr} 3 & -1 \\ 1 & -1 \end{array}\right], \\ &= \left[\begin{array}{rr} 1 & 0 \\ 0 & -1 \end{array}\right], \\ &= B. \end{aligned}$$

3.5. Isomorfismos

En esta sección se retomarán algunos conceptos y resultados obtenidos en la sección A.1.1. Comenzaremos por definir a las tranformaciones uno, transformaciones sobre y finalmente, a las tranformaciones uno a uno y sobre, las cuales permitirán obtener una transformación inversa única.

3.5.1. Transformación inversa

Definición 3.5.1 (Transformaciones uno a uno o inyectivas) *Si $T : V \to W$ es una transformación lineal, entonces se dice que T es uno a uno o inyectiva, si para todo $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ en V, $T(\boldsymbol{v}_1) = T(\boldsymbol{v}_2)$ implica que $\boldsymbol{v}_1 = \boldsymbol{v}_2$. De manera equivalente, T es uno a uno si para todo $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ en V, $T(\boldsymbol{v}_1) \neq T(\boldsymbol{v}_2)$ implica que $\boldsymbol{v}_1 \neq \boldsymbol{v}_2$*

Teorema 3.5.1 *Una transformación lineal $T : V \to W$ es uno a uno si y solo si $ker(T) = \{\boldsymbol{0}\}$.*

Demostración: $\Rightarrow$) supóngase que T es uno a uno y sea $\mathbf{v}$ un vector del kernel de la transformación, entonces, $T(\mathbf{v}) = \mathbf{0}$. Ahora, por el teorema 3.1.1a se tiene, $T(\mathbf{0}) = \mathbf{0}$. Como T es uno a uno, entonces $\mathbf{v}$=$\mathbf{0}$; es decir, el kernel de la transformación consiste únicamente del vector cero.

$\Leftarrow$) supóngase que $ker(T) = \{\mathbf{0}\}$. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en V tales que,

$$T(\mathbf{u}) = T(\mathbf{v}),$$

de manera que,

$$T(\mathbf{u}) - T(\mathbf{v}) = \mathbf{0}.$$

Por el teorema 3.1.1c la ecuación anterior se expresa como,

$$T(\mathbf{u} - \mathbf{v}) = \mathbf{0};$$

es decir, $(\mathbf{u} - \mathbf{v})$ está en ker(T). Como $\mathbf{0}$ es el único elemento de $ker(T)$, entonces,

$$\mathbf{u} - \mathbf{v} = \mathbf{0},$$

luego, $\mathbf{u} = \mathbf{v}$, por lo tanto, T es uno a uno. ∎

Definición 3.5.2 (Transformaciones no singulares) *Si $T : V \to W$ es una transformación lineal, entonces decimos que T es una transformación no singular si $ker(T) = \{\mathbf{0}\}$. Por el teorema anterior, una transformación es no singular si y solo si es inyectiva.*

Ejemplo 3.5.1 *Sea $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por,*

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -y \\ 2y - x \\ x + y \end{bmatrix}.$$

Determine si la transformación lineal es uno a uno.

Solución: Determínese el kernel de la transformación; es decir, halle los vectores $\mathbf{x}$ tales que $T(\mathbf{x}) = \mathbf{0}$, de esta manera

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -y \\ 2y - x \\ x + y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Lo anterior nos conduce al sistema de ecuaciones,

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

Por simple inspección, se tiene $x = y = 0$ es la única solución. De manera que $ker(T) = \{\mathbf{0}\}$. De acuerdo con el teorema anterior T es uno a uno.

Ejemplo 3.5.2 *Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por, $T(\boldsymbol{x}) = A\boldsymbol{x}$, donde A es la siguiente matriz.*

$$\begin{bmatrix} 3 & 2 & -1 \\ 4 & 7 & 0 \\ -6 & -4 & 2 \end{bmatrix}.$$

Determine si la transformación lineal es uno a uno.

Solución: *En este caso el kernel de la transformación es el espacio nulo de la matriz A. Como A es una matriz cuadrada, la homogénea $A\boldsymbol{x} = \boldsymbol{0}$ tendrá únicamente la solución trivial si $|A| \neq 0$.*

$$|A| = \begin{vmatrix} 3 & 2 & -1 \\ 4 & 7 & 0 \\ -6 & -4 & 2 \end{vmatrix} = -2 \begin{vmatrix} 3 & 2 & -1 \\ 4 & 7 & 0 \\ 3 & 2 & -1 \end{vmatrix} = 0.$$

Se concluye que la transformación no es uno a uno.

Teorema 3.5.2 *Si $T : V \to W$ es una transformación lineal, entonces T es no singular si y solo si T mapea cada conjunto linealmente independiente de vectores en V en otro conjunto linealmente independiente de vectores en W.*

Demostración: $\Rightarrow$) supóngase que T es no singular y que,

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\},$$

es un conjunto de vectores en V linealmente independiente. Se demostrará que,

$$S' = \{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_k)\},$$

es conjunto linealmente independiente en W. Para tal efecto considérese la ecuación,

$$a_1T(\mathbf{v}_1) + a_2T(\mathbf{v}_2) + \ldots + a_kT(\mathbf{v}_k) = \mathbf{0}.$$

Como T es una transformación lineal, entonces se pueden reagrupar los términos de la ecuación anterior de la manera siguiente:

$$T(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_k\mathbf{v}_k) = \mathbf{0}.$$

Como T es no singular, entonces,

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_k\mathbf{v}_k = \mathbf{0}$$

Como S es un conjunto linealmente independiente, entonces todos los escalares en la ecuación anterior son iguales a cero; por lo tanto, S' es un conjunto linealmente indepediente en W.
$\Leftarrow$) supóngase que T aplica cada conjunto linealmente independiente de vectores en V en un conjunto linealmente independiente en W. En particular, si $\mathbf{v} \neq \mathbf{0}$ es un vector en V, entonces el conjunto $S = \{\mathbf{v}\}$ es linealmente independiente, por lo tanto, $S' = \{T(\mathbf{v})\}$ será linealmente independiente, lo cual implica que $T(\mathbf{v}) \neq \mathbf{0}$. De lo anterior concluimos que el único vector que T mapea al vector cero de W, es el vector cero en V; por lo tanto, T es no singular. $\blacksquare$

Definición 3.5.3 (Transformaciones sobre) *Si $T : V \to W$ es una transformación lineal y si para cada vector $\boldsymbol{w}$ en W, existe un vector $\boldsymbol{v}$ en V, tal que $\boldsymbol{w} = T(\boldsymbol{v})$, entonces se dice que T es una transformación sobre (o que T es una transformación de V sobre W). De manera equivalente, T es sobre si y solo si $R(T) = W$.*

Ejemplo 3.5.3 *Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal definida por,*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + 2y - 3z \\ 7x + 2y \\ x - 12y + 19z \end{bmatrix}.$$

Determine si la transformación lineal es sobre.

Solución: Verifíquese que dado un vector $\mathbf{w} \in \mathbb{R}^3$, existe un vector $\mathbf{v} \in \mathbb{R}^3$ tal que $T(\mathbf{v}) = \mathbf{w}$. Sean $\mathbf{v}$ y $\mathbf{w}$ los vectores,

$$\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad y \quad \mathbf{w} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

La ecuación vectorial, $T(\mathbf{v}) = \mathbf{w}$, nos conduce al siguiente sistema de ecuaciones:

$$\left.\begin{array}{lcl} 2x + 2y - 3z & = & b_1, \\ 7x + 2y & = & b_2, \\ x - 12y + 19z & = & b_3. \end{array}\right\}$$

El sistema anterior se expresa en forma matricial como:

$$\begin{bmatrix} 1 & 2 & -3 \\ 7 & 2 & 0 \\ 1 & -12 & 19 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

De acuerdo con el teorema 2.4.15, el sistema tendrá solución única para toda matriz

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Si y solo si $|A| \neq 0$. En este caso,

$$|A| = \begin{vmatrix} 1 & 2 & -3 \\ 7 & 2 & 0 \\ 1 & -12 & 19 \end{vmatrix} = 30 \neq 0,$$

por lo tanto, el sistema anterior es consistente para toda matriz de la forma

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Se concluye que T es sobre.

Definición 3.5.4 (Composición de transformaciones lineales) *Sean V, W y Z espacios vectoriales. Si $T_1 : V \to W$ y $T_2 : W \to Z$ son transformaciones lineales, entonces la componsición de T_2 con T_1, denotada por $T_2 \circ T_1$, es la transformación de V en W definida como:*

$$(T_2 \circ T_1)(\boldsymbol{v}) = T_2(T_1(\boldsymbol{v})), \quad \forall \boldsymbol{v} \in V.$$

Ejemplo 3.5.4 *Considere las transformaciones $T_1 : \mathbb{R}^3 \to \mathbb{R}^2$ y $T_2 : \mathbb{R}^2 \to P_1$ definidas como:*

$$T_1\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}\right) = \begin{bmatrix} a_1 + 2a_2 - 3a_3 \\ 7a_1 + 2a_2 \end{bmatrix} \quad y \quad T_2\left(\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}\right) = 2b_1x + b_2.$$

Determine la composición de $T_2 \circ T_1$.

Solución:

$$\begin{aligned} T_2 \circ T_1\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}\right) &= T_2\left(T_1\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}\right)\right), \\ &= T_2\left(\begin{bmatrix} a_1 + 2a_2 - 3a_3 \\ 7a_1 + 2a_2 \end{bmatrix}\right), \\ &= 2(a_1 + 2a_2 - 3a_3)x + (7a_1 + 2a_2). \end{aligned}$$

Teorema 3.5.3 *Sean V,W y Z espacios vectoriales y sean $T_1 : V \to W$ y $T_2 : W \to Z$ transformaciones lineales, entonces la composición de T_1 y T_2 es también una transformación lineal.*

Demostración: Sean **u** y **v** elementos de V, y k es un escalar, entonces

a)

$$\begin{aligned}(T_2 \circ T_1)(\mathbf{u}+\mathbf{v}) &= T_2(T_1(\mathbf{u}+\mathbf{v})),\\ &= T_2(T_1(\mathbf{u}) + T_1(\mathbf{v})),\ T_1 \text{ es una transformación lineal}\\ &= T_2(T_1(\mathbf{u})) + T_2(T_1(\mathbf{v})),\ T_2 \text{ es una transformación lineal}\\ &= (T_2 \circ T_1)(\mathbf{u}) + (T_2 \circ T_1)(\mathbf{v}).\end{aligned}$$

b)

$$\begin{aligned}(T_2 \circ T_1)(k\mathbf{u}) &= T_2(kT_1(\mathbf{u})),\ T_1 \text{ es una transformación lineal}\\ &= kT_2(T_1(\mathbf{u})),\ T_2 \text{ es una transformación lineal}\\ &= k(T_2 \circ T_1)(\mathbf{u}). \quad \blacksquare\end{aligned}$$

Teorema 3.5.4 *Si V, W y Z son espacios vectoriales de dimensión finita con bases S, S' y S'', respectivamente, y si además, $T_1 : V \to W$ es una transformación lineal cuya representación respecto a las bases S y S' es la matriz A_1, y de la misma manera, $T_2 : W \to Z$ es una transformación lineal cuya representación respecto a las bases S' y S'' es la matriz A_2 , entonces A_2A_1 es la matriz que representa a $T_2 \circ T_1$ respecto a las bases S y S''*

Demostración: De acuerdo con el teorema 3.3.1, las matrices A_1 y A_2 satisfacen las siguientes relaciones,

$$[T_1(\mathbf{x})]_{S'} = A_1[\mathbf{x}]_S, \tag{3.21}$$

$$[T_2(\mathbf{y})]_{S''} = A_2[\mathbf{y}]_{S'}. \tag{3.22}$$

Como $T_2 \circ T_1 : V \to W$, entonces por el teorema 3.3.1, la matriz A_3 que representa a $T_2 \circ T_1$ respecto a las bases S y S'' debe satisfacer:

$$[(T_2 \circ T_1)(\mathbf{x})]_{S''} = A_3[\mathbf{x}]_S.$$

Por lo tanto,

$$\begin{aligned}A_3[\mathbf{x}]_S &= [(T_2 \circ T_1)(\mathbf{x})]_{S''},\\ &= [T_2((T_1)(\mathbf{x}))]_{S''},\\ &= A_2[T_1(\mathbf{x})]_{S'}, \quad \text{Aplicando la ecuación 3.21}\\ &= A_2A_1[\mathbf{x}]_S. \quad \text{Aplicando la ecuación 3.22}\end{aligned}$$

Como A_3 es única, entonces $A_3 = A_2A_1$. $\blacksquare$

Ejemplo 3.5.5 *Para el ejemplo 3.5.4 verifique que el teorema anterior se satisface para las bases canónicas de los espacios involucrados.*

Solución: En este caso las bases S, S' y S'' son:

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \quad S' = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \quad S'' = \{1, x\}.$$

La matriz que reprenta a T_1 respecto a las bases S y S' es:

$$A_1 = \begin{bmatrix} 1 & 2 & -3 \\ 7 & 2 & 0 \end{bmatrix}.$$

La matriz que reprenta a T_2 respecto a las bases S' y S'' es:

$$A_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

La matriz que representa a la composición $T_2 \circ T_1$, respecto a las base S y S'' es:

$$\begin{aligned} A_2 A_1 &= \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -3 \\ 7 & 2 & 0 \end{bmatrix}, \\ &= \begin{bmatrix} 2 & 4 & -6 \\ 7 & 2 & 0 \end{bmatrix}. \end{aligned}$$

Ahora, sea $\mathbf{x}$ un vector cualquiera de $\mathbb{R}^3$, donde,

$$\mathbf{x} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix},$$

como S es la base canónica para $\mathbb{R}^3$, entonces $\mathbf{x} = [\mathbf{x}]_S$, y de acuerdo con el teorema 3.3.1,

$$\begin{aligned} [T(x)]_{S''} &= A_3 [x]_S, \\ &= \begin{bmatrix} 2 & 4 & -6 \\ 7 & 2 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \\ &= \begin{bmatrix} 2a_1 + 4a_2 - 6a_3 \\ 7a_1 + 2a_2 \end{bmatrix}, \end{aligned}$$

de manera que,

$$T\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}\right) = (2a_1 + 4a_2 - 6a_3)x + 7a_1 + 2a_2.$$

Definición 3.5.5 (La inversa de una transformación lineal) *La transformación lineal $T : V \to W$ se conoce como invertible si existe una función U de W en V talque $U \circ T$ es la identidad sobre V y $T \circ U$ es la identidad sobre W. Hemos probado que si T es invertible, entonces la inversa es única y se denota por T^{-1}; además, T es invertible si y solo si T es uno a uno y sobre. (ver sección A.1.1)*

Teorema 3.5.5 *Si $T : V \to W$ es una transformación lineal invertible, entonces T^{-1} es también una transformación lineal.*

Demostración: a) Sean $\mathbf{w}_1$ y $\mathbf{w}_2$ elementos de W y sean $\mathbf{v}_1$ y $\mathbf{v}_2$ dos vectores en V tales que, $T^{-1}(\mathbf{w}_1) = \mathbf{v}_1$ y $T^{-1}(\mathbf{w}_2) = \mathbf{v}_2$, es decir, $\mathbf{v}_1$ y $\mathbf{v}_2$ son los únicos vectores que T aplica a $\mathbf{w}_1$ y $\mathbf{w}_2$, respectivamente. Por lo tanto,

$$T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2.$$

La última igualdad muestra que, $\mathbf{v}_1 + \mathbf{v}_2$, es el único vector que T aplica a $\mathbf{w}_1 + \mathbf{w}_2$, de manera que,

$$T^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = \mathbf{v}_1 + \mathbf{v}_2 = T^{-1}(\mathbf{w}_1) + T^{-1}(\mathbf{w}_2).$$

b) Si k es un escalar, y, $\mathbf{w}_1$ y $\mathbf{v}_1$ son como en el inciso anterior, entonces, $T(k\mathbf{v}_1) = k\mathbf{w}_1$; por lo tanto,

$$T^{-1}(k\mathbf{w}_1) = k\mathbf{v}_1 = kT^{-1}(\mathbf{w}_1). \quad \blacksquare$$

Teorema 3.5.6 *Si $T : V \to W$ es una transformación lineal, donde V y W son espacios vectoriales tales que $dim(V) = dim(W)$, entonces los siguientes enunciados son equivalentes:*

a) T es invertible.

b) T es no singular.

c) T es sobre.

Demostración: $a) \Rightarrow b)$ Si T es invertible, entonces es uno a uno y por el teorema 3.5.1 $ker(T) = \{\mathbf{0}\}$; es decir, T es no singular.
$b) \Rightarrow c)$ Si T es no singular, entonces $\nu(T) = 0$. Por el teorema de la dimensión (teorema 3.2.3) se tiene,

$$\rho(T) + \nu(T) = n.$$

En este caso, $\rho(T) = n$; es decir, $dim(R(T)) = n$. De manera que $R(T)$ tiene una base con n vectores. Como $R(T)$ es un subespacio de W y W tiene dimensión n, entonces $R(T) = W$, lo cual significa que T es sobre.
$c) \Rightarrow a)$ Si T es sobre, entonces $R(T) = W$, por lo tanto, $dim(R(T)) = dim(W) = n$. Por el teorema de la dimensión:

$$\nu(T) = n - \rho(T) = n - n = 0.$$

Por lo tanto, $ker(T) = \{\mathbf{0}\}$, lo cual implica que T es uno a uno. Finalmente, T es uno a uno y sobre, por el teorema A.1.3, T es invertible. ∎

Ejemplo 3.5.6 *Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal definida mediante,*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + 2y - 2z \\ 2x + 4z \\ -x + 2y \end{bmatrix}.$$

Determine si T es invertible, en caso afirmativo halle T^{-1}.

Solución: Por el teorema anterior, T será invertible si y solo si T es no singular. Investíguese si T es o no singular; es decir, determine el kernel de la transformación.

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + 2y - 2z \\ 2x + 4z \\ -x + 2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

La ecuación anterior nos conduce al sistema de ecuaciones,

$$\begin{cases} x + 2y - 2z &= 0 \\ 2x + 4z &= 0 \\ -x + 2y &= 0 \end{cases}$$

El sistema matricial equivalente correspondiente es:

$$\begin{bmatrix} 1 & 2 & -2 \\ 2 & 0 & 4 \\ -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

De acuerdo con el teorema **??**, este sistema tendrá únicamente la solución trivial si y solo si, $|A| \neq 0$. En este caso $|A| = 12$; por lo tanto, el sistema tiene solamente la solución trivial, en consecuencia, T es no singular y finalmente, T es invertible.
Para determinar la inversa de T recordemos que para todo $\mathbf{b} = (b_1, b_2, b_3)$, y $\mathbf{x} = (x, y, z)$, se debe cumplir que

$$T(\mathbf{x}) = \mathbf{b} \quad y \quad T^{-1}(\mathbf{b}) = \mathbf{x}.$$

Por lo anterior,

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + 2y - 2z \\ 2x + 4z \\ -x + 2y \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

La ecuación anterior nos conduce al sistema de ecuaciones,

$$\begin{cases} x + 2y - 2z &= b_1 \\ 2x + 4z &= b_2 \\ -x + 2y &= b_3 \end{cases}$$

El sistema matricial equivalente al sistema anterior es,

$$\begin{bmatrix} 1 & 2 & -2 \\ 2 & 4 & 0 \\ -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Para hallar la solución del sistema, como A es invertible, determine su inversa. Para este caso,

$$A^{-1} = \frac{1}{12}\begin{bmatrix} 4 & 2 & -4 \\ 2 & 1 & 4 \\ -2 & 2 & 2 \end{bmatrix},$$

por lo tanto,

$$\begin{aligned} \begin{bmatrix} x \\ y \\ z \end{bmatrix} &= \frac{1}{12}\begin{bmatrix} 4 & 2 & -4 \\ 2 & 1 & 4 \\ -2 & 2 & 2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \\ &= \frac{1}{12}\begin{bmatrix} 4b_1 + 2b_2 - 4b_3 \\ 2b_1 + b_2 + 4b_3 \\ -2b_1 + 2b_2 + 2b_3 \end{bmatrix}. \end{aligned}$$

De manera que,

$$T^{-1}\left(\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 4b_1 + 2b_2 - 4b_3 \\ 2b_1 + b_2 + 4b_3 \\ -2b_1 + 2b_2 + 2b_3 \end{bmatrix}.$$

Teorema 3.5.7 *Sean V, W y Z espacios vectoriales. Si $T_1 : V \to W$ y $T_2 : W \to Z$ son transformaciones lineales invertibles, entonces $T_2 \circ T_1$ es invertible y su inversa es:*

$$(T_2 \circ T_1)^{-1} = T_1^{-1} \circ T_2^{-1}.$$

Demostración: Se demostró en la sección A.1.1.

Teorema 3.5.8 *Si $T : V \to V$ es una transformación lineal invertible, y sea A la matriz que representa a T respecto a una base S para V, entonces A^{-1} es la matriz que representa a T^{-1} respecto a la base S.*

Demostración: Como T es invertible, entonces existe la transformación lineal T^{-1} tal que, $T^{-1} \circ T = T_I =$ donde T_I es el operador identidad. Sean A y B las matrices que representan a T y a T^{-1} respecto a la base S, respectivamente, entonces de acuerdo con el teorema 3.22, la matriz que reprenta a $T^{-1} \circ T$ es por un lado, el producto BA. Por otra pate, la matriz que representa a la transformación identidad es la matriz identidad, como se vio en el ejemplo 3.3.3. Como la matriz que representa a una transformación lineal respecto a una base es única, entonces, $BA = I$. Como A es una matriz cuadrada, entonces A es invertible y B es su inversa (teorema 1.3.7); es decir, $B = A^{-1}$. ▮

Ejemplo 3.5.7 *Considere la transformación del ejemplo 3.5.6*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + 2y - 2z \\ 2x + 4z \\ -x + 2y \end{bmatrix}.$$

Determine la transformación inversa acorde con el teorema anterior.

Solución: Sea $S = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ la base canónica para $\mathbb{R}^3$, entonces la matriz que representa a T respecto a esta base es,

$$\begin{aligned} A_T &= [T(\mathbf{e}_1), T(\mathbf{e}_2), T(\mathbf{e}_3)], \\ &= \begin{bmatrix} 1 & 2 & -2 \\ 2 & 0 & 4 \\ -1 & 2 & 0 \end{bmatrix}. \end{aligned}$$

La inversa de esta matriz es,

$$(A_T)^{-1} = \frac{1}{12}\begin{bmatrix} 4 & 2 & -4 \\ 2 & 1 & 4 \\ -2 & 2 & 2 \end{bmatrix}.$$

De acuerdo con el teorema anterior, la representación de la inversa de T respecto a la base estándar será justamente $(A_T)^{-1}$. Por otro lado, de acuerdo con el teorema 3.3.1,

$$(A_T)^{-1}[\mathbf{y}]_S = [T(\mathbf{y})]_S.$$

Como S es la base estándar, entonces, $[\mathbf{y}]_S = \mathbf{y}$; por lo tanto,

$$(A_T)^{-1}\mathbf{y} = T^{-1}(\mathbf{y}).$$

Si $\mathbf{y} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, entonces,

$$\begin{aligned} T^{-1}\left(\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}\right) &= (A_T)^{-1}\mathbf{y}, \\ &= \frac{1}{12}\begin{bmatrix} 4 & 2 & -4 \\ 2 & 1 & 4 \\ -2 & 2 & 2 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \\ &= \frac{1}{12}\begin{bmatrix} 4b_1 + 2b_2 - 4b_3 \\ 2b_1 + b_2 + 4b_3 \\ -2b_1 + 2b_2 + 2b_3 \end{bmatrix}. \end{aligned}$$

3.5.2. Definición y ejemplos de espacios isomorfos

Definición 3.5.6 *Si $T : V \to W$ es una transformación uno a uno y sobre, entonces se dice que T es un isomorfismo de V sobre W. Además, se dice que V es isomorfo a W.*

Ejemplo 3.5.8 *Si V es cualquier espacio vectorial, entonces V es isomorfo a V, pues la transformación identidad es un isomorfismo de V sobre V.*

Teorema 3.5.9 *Si V es isomorfo a W, entonces W es isomorfo a V.*

Demostración: Como V es isomorfo a W, entonces existe un isomorfismo $T : V \to W$. Como T es invertible, entonces $T^{-1} : W \to V$ es un isomorfismo de W sobre V; por lo tanto, W es isomorfo a V. ∎

Por el teorema anterior, simplemente diremos que V y W son isomorfos cuando exista un isomorfismo T de V sobre W.

Teorema 3.5.10 *Sean V, W y Z espacios vectoriales tales que, V y W son isomorfos y por otra parte, W y Z son isomorfos, entonces V y Z son isomorfos*

Demostración: Como V y W son isomorfos, entonces existe un isomorfismo $T_1 : V \to W$, de la misma manera, como W y Z son isomorfos, entonces existe un isomorfismo $T_2 : W \to Z$, de manera que: $T_2 \circ T_1 : V \to Z$ es un isomorfismo de V sobre W; por lo tanto, V y W son isomorfos. ∎

Teorema 3.5.11 *Si V es un espacio vectorial de dimensión n, sobre el campo de los reales, entonces V es isomorfo con* $\mathbb{R}^n$*.*

Demostración: Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base para el espacio V y sea T la transformación $T : V \to \mathbb{R}^n$ definida como $T(\mathbf{v}) = (\mathbf{v})_S$; es decir, T asigna a cada vector $\mathbf{v}$ en V su vector de cordenadas respecto a la base S. Como se vio en el ejemplo 3.1.4, T es una transformación lineal; además, por el teorema 2.4.9, T es uno a uno , y finalmente, por el teorema 3.5.1 T es sobre; de manera que, T es un isomorfismo de V sobre $\mathbb{R}^n$; por lo tanto, V y $\mathbb{R}^n$ son isomorfos. ∎

Ejemplo 3.5.9 *El teorema anterior nos permite establecer formalmente el isomorfismo entre* $\mathbb{M}_{nx1}$ *y* $\mathbb{R}^n$*. Durante el curso hemos hecho uso de este isomorfismo, cambiando un espacio por otro cuando ha resultado conveniente.*

Teorema 3.5.12 *Si V y W son espacios vectoriales de dimensión finita, entonces V y W son isomorfos si y solo si* $dim(V) = dim(W)$*.*

Demostración: $\Rightarrow$) Supóngase que V y W son isomorfos y que $dim(V) = n$. Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base para V y sea $T : V \to W$ un isomorfismo de V sobre W. Se probará que $S' = \{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\}$ es una base para $R(T)$. En efecto, como T es uno a uno, de acuerdo con el teorema 3.5.2 S' es linealmente independiente; además, por el teorema de la dimensión se tiene

$$n = dim(ker(T)) + dim(R(T)).$$

Como T es uno a uno, entonces $dim(ker(T)) = 0$; por lo tanto, $n = dim(R(T))$. Ahora, como S' contiene n vectores que generan un espacio vectorial de dimensión n, entonces de acuerdo con el teorema 2.4.6, S' es una base para $R(T)$. Por último, como T es sobre, entonces $R(T) = W$, luego, $dim(V) = dim(W)$.

$\Leftarrow$) Si $dim(V) = dim(W) = n$, entonces, de acuerdo con el teorema 3.5.11, tanto W como V son isoformos a $\mathbb{R}^n$. Finalmente, el teorema 3.5.10 nos permite concluir que W y V son isomorfos. ∎

Ejemplo 3.5.10 *Determine si* P_2 *y* $\mathbb{R}^3$ *son isomorfos, en caso afirmativo determine un isomorfismo* $T : P_2 \to \mathbb{R}^3$*.*

Solución: Como P_2 y $\mathbb{R}^3$ tienen la misma dimensión, entonces el teorema anterior afirmar que son espacios isomorfos. Para determinar un isomorfismo $T : P_2 \to \mathbb{R}^3$, considérese la base canónica $S = \{1, x, x^2\}$ para P_2, de manera que, $T : P_2 \to \mathbb{R}^3$ definida como $T(\mathbf{p}(x)) = (\mathbf{p}(x))_S$ define un isomorfismo de P_2 sobre $\mathbb{R}^3$. Demanera explícita se tiene:

$$T(ax^2 + bx + c) = (a, b, c).$$

Ejemplo 3.5.11 *Determine si* $\mathbb{M}_{2X2}$ *y* P_3 *son espacios isomorfos.*
Como los dos espacios son de dimensión 4, de acuerdo con el teorema anterior, estos espacios son isomorfos.

Capítulo 4

Aplicaciones

En el capítulo anterior vimos como representar a una transformación lineal mediante una matriz y como ésta se modifica cuando cambiamos de coordenadas. En este caítulo trataremos de hallar un sistema de coordenadas en el cual, la representación matricial de la transformación lineal sea lo más simple, esto nos conduce al problema de la diagonalización de matrices que a su vez requiere del desarrollo del problema de los eigenvalores, lo cual trataremos en la primer sección. Este capítulo se concluye con dos aplicaciones directas de la teoría deasarrollada durante el curso y finalmente, se bosquejará lo que se conoce como forma canónica de Jordan.

4.1. Valores y vectores característico

Definición 4.1.1 (Eigenvalores y eigenvectores) *Si A es una matriz de nxn, entonces un vector $\boldsymbol{x} \neq \boldsymbol{0}$ en $\mathbb{R}^n$ se denomina eigenvector de A si,*

$$A\boldsymbol{x} = \lambda \boldsymbol{x}, \tag{4.1}$$

para algún escalar λ. Este escalar se conoce como eigenvalor, valor característico o valor propio de A correspondiente al eigenvector $\boldsymbol{x}$.

La ecuación 4.1 nos plantea el problema de determinar los eigenvalores y los eigenvectores correspondientes para una matriz A. Para tal fin se procede de la siguiente manera. La ecuación 4.1 puede expresarse como,

$$A\mathbf{x} = \lambda I \mathbf{x}. \tag{4.2}$$

De la ecuación 4.2 se obtiene,

$$(\lambda I - A)\mathbf{x} = 0. \tag{4.3}$$

Si λ es un eigenvalor de la matriz A, entonces la ecuación 4.3 debe admitir soluciones no triviales. De acuerdo con el teorema 2.4.15 la ecuación 4.3 tiene una solución no trivial si y solo si,

$$det(\lambda I - A) = 0. \tag{4.4}$$

La ecuación 4.4 se conoce como ecuación característica de la matriz A.

4.1.1. Definición y polinomio característico

Para comprender la estructura de la ecuación 4.4, representemos a la matriz A en términos de sus entradas; es decir,

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

Sustituyendo la matriz anterior en la ecuación 4.4 obtenemos,

$$det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \lambda - a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & \lambda - a_{nn} \end{vmatrix}.$$

Todos los productos elementales del determinante anterior contienen al menos un factor de la diagonal principal, por lo tanto, cada producto elemental del determinante es un polinomio en λ. El producto elemental que proviene de la diagonal principal es, $(\lambda - a_{11})(\lambda - a_{22})...(\lambda - a_{22})$. Este término es un polinomio de grado n, donde el coeficiente de λ^n es 1. Por lo anterior, la ecuación característica representa un polinomio de grado n.

Definición 4.1.2 (Polinomio característico) *Dada una matriz de tamaño nxn, se conoce como polinomio característico de la matriz A, y se representa por $P(\lambda)$, al polinomio de grado n que resulta del desarrollo de la ecuación característica de la matriz A; es decir,*

$$P(\lambda) = det(\lambda I - A) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} \dots + c_n. \tag{4.5}$$

De acuerdo con la ecuación característica se tiene,

$$det(\lambda I - A) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} \dots + c_n = 0 \tag{4.6}$$

Por lo tanto, los eigenvalores de la matriz A corresponden a las raíces de su polinomio característico. Por el teorema fundametal del álgebra, sabemos que un polinomio de grado n tiene a lo más n raíces diferentes. En general, determinar las raíces de un polinomio no es una tarea sencilla; sin embargo, cuando los coeficientes en todos los términos del polinomio característico son enteros, pueden ser útiles los siguientes resultados.

a) Si $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ son las raíces del polinomio característico, entonces

$$P(\lambda) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} \ldots + c_n = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n).$$

b) Las raíces enteras del polinomio, si éstas existen, deben ser divisores del término constante c_n.

En general, las raíces del polinomio característico pueden ser irracionales o incluso números complejos. El tratamiento de los espacios vectoriales dentro del campo de los números complejos queda fuera del alcance de este curso.

Ejemplo 4.1.1 *Determine los valores propios de la matriz A, donde*

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

Solución: De acuerdo con la ecuación 4.5 tenemos,

$$\begin{aligned} P(\lambda) &= det(\lambda I - A) = \begin{vmatrix} \lambda + 4 & 2 \\ -3 & \lambda - 3 \end{vmatrix}, \\ &= (\lambda + 4)(\lambda - 3) + 6 = \lambda^2 + \lambda - 6 = (\lambda + 3)(\lambda - 2) = 0. \end{aligned}$$

Por lo anterior, los eigenvalores de la matriz A son:

$$\begin{cases} \lambda_1 = -3, \\ \lambda_2 = 2. \end{cases}$$

Ejemplo 4.1.2 *Determine los valores propios de la matriz A, donde*

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

Solución: En este caso la ecuación 4.5 nos conduce a,

$$
\begin{aligned}
P(\lambda) &= det(\lambda I - A) = \begin{vmatrix} \lambda-2 & -2 & -3 \\ -1 & \lambda-2 & -1 \\ -2 & 2 & \lambda-1 \end{vmatrix}, \\
&= \lambda^3 - 5\lambda^2 + 2\lambda + 8 = 0.
\end{aligned}
$$

Como los coeficientes del polinomio característico son números enteros, tratemos de determinar si éste tiene raíces enteras. El procedimiento consiste en verificar si los divirores del término constante son raíces del polinomio. En este caso los divisores de 8 son: $\pm1, \pm2, \pm4, \pm8$. Sustituyendo $\lambda = -1$ en el polinomio caracterítico obtenemos,

$$P(-1) = (-1)^3 - 5(-1)^2 + 2(-1) + 8 = 0,$$

por lo tanto, -1, es una raíz del polinomio.

Por otro lado sabemos que,

$$P(\lambda) = (\lambda+1)(\lambda-\lambda_2)(\lambda-\lambda_3) = (\lambda+1)P_2(\lambda)$$

donde

$$P_2(\lambda) = (\lambda-\lambda_2)(\lambda-\lambda_3)$$

Por lo anterior, $P_2(\lambda)$ puede determinarse dividiendo el polinomio característico entre $(\lambda+1)$. El resultado de esta división es:

$$P_2(\lambda) = \lambda^2 - 6\lambda + 8 = (\lambda-2)(\lambda-4)$$

En este caso P_2 resultó factorizable por simple inspección, en otros casos es necesario utilizar la fórmula general para la ecuación de segundo grado. Finalmente, los eigenvalores para la matriz A son:

$$\begin{cases} \lambda_1 = -1, \\ \lambda_2 = 2, \\ \lambda_3 = 4. \end{cases}$$

4.1.2. Cálculo de vectores característicos

Una vez que sabemos como determinar los eigenvalores de una matriz dada, nuestro siguiente objetivo es determinar los eigenvectores correspondientes con dicho eigenvalor. Estos eignvectores deben satisfacer la ecuación 4.3. El espacio de soluciones de esta ecuación se conoce como **eigenespacio** de A correspondiente al eigenvalor λ, denotaremos este eigenespacio mediante V_λ.

Ejemplo 4.1.3 *Determine los eigenvectores correspondientes a los eigenvalores de la matriz A del ejemplo 4.1.1.*

Solución: La matriz del ejemplo 4.1.1 es:

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

Usando la matriz anterior, la ecuación 4.3 se expresa como:

$$\begin{bmatrix} \lambda + 4 & 2 \\ -3 & \lambda - 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{4.7}$$

Los eigenvalores determinados para la matriz A fueron: $\lambda_1 = -3$ y $\lambda_2 = 2$.

a) Sustituyendo $\lambda_1 = -3$, en la ecuación 4.7 obtenemos:

$$\begin{bmatrix} 1 & 2 \\ -3 & -6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

La solución del sistema anterior es el eigenespacio:

$$V_{\lambda_1} = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : \begin{bmatrix} x \\ y \end{bmatrix} = t \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \text{ para cualquier escalar t} \right\}.$$

De manera que para $t = 1$, $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda = -3$.

b) De la misma manera que en el inciso anterior, sustituyendo $\lambda_2 = 2$, en la ecuación 4.7 obtenemos:

$$\begin{bmatrix} 6 & 2 \\ -3 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

La solución del sistema anterior es el eigenespacio:

$$V_{\lambda_2} = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : \begin{bmatrix} x \\ y \end{bmatrix} = t \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \text{ para cualquier escalar t} \right\}.$$

Para $t = 1$, $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda = 2$.

Ejemplo 4.1.4 *Determine los eigenvectores correspondientes a los eigenvalores de la matriz A del ejemplo 4.1.2.*

Solución: La matriz del ejemplo 4.1.2 es:

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

Para esta matriz, la ecuación 4.3 se expresa como:

$$\begin{bmatrix} \lambda - 2 & -2 & -3 \\ -1 & \lambda - 2 & -1 \\ -2 & 2 & \lambda - 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{4.8}$$

Además, los eigenvalores de la matriz A son: $\lambda_1 = -1$, $\lambda_2 = 2$, $\lambda_3 = 4$.

a) Sustituyendo el eigenvalor $\lambda_1 = -1$ en la ecuación 4.8 obtenemos,

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

La solución del sistema anterior es el eigenespacio,

$$V_{\lambda_1} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

Para $t = 1$, $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda_1 = -1$.

b) Sustituyendo el eigenvalor $\lambda_2 = 2$ en la ecuación 4.8 obtenemos,

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

La solución del sistema anterior es el eigenespacio,

$$V_{\lambda_2} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ -\frac{3}{2} \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

Para $t = 2$, $\begin{bmatrix} -2 \\ -3 \\ 2 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda_2 = 2$.

c) Sustituyendo el eigenvalor $\lambda_3 = 4$ en la ecuación 4.8 obtenemos,

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

La solución del sistema anterior es el eigenespacio,

$$V_{\lambda_3} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 4 \\ \frac{5}{2} \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

Para $t = 2$, $\begin{bmatrix} 8 \\ 5 \\ 2 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda_3 = 4$.

Teorema 4.1.1 *Una matriz cuadrada A es no singular, si y solo si $\lambda = 0$ no es un eigenvalor de A.*

Demostración: Consideremos el polinomio expresado por la ecuación 4.5.

$$P(\lambda) = det(\lambda I - A) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} \ldots + c_n. \tag{4.9}$$

Sustituyendo $\lambda = 0$ obtenemos,

$$\begin{aligned} det(-A) &= c_n \\ (-1)^n det(A) &= c_n \quad \text{teorema 1.4.6a} \end{aligned}$$

Ahora, $\lambda = 0$ es un eigenvalor de A si y solo si $c_n = (-1)^n det(A) = 0$. Luego, $c_n = 0$ si y solo si $det(A) = 0$; es decir, $\lambda = 0$ es un eigenvalor de A si y solo si A es una matriz singular. ∎

Este resultado lo incorporaremos a nuestro teorema resumen.

Teorema 4.1.2 *Si A es una matriz de tamaño nxn, entonces los siguientes enunciados son equivalentes.*

a) A es invertible.

b) Ax=0 solo tiene la solución trivial.

c) A es equivalente respecto a los renglones a I_n.

d) Ax=b tiene solución única para toda matriz b de tamaño nx1.

e) $\det A \neq 0$.

f) *A tiene rango n.*

g) *Los vectores renglón de A son linealmente independientes.*

h) *Los vectores columna de A son linealmente independientes.*

i) *0 no es un eigenvalor de A.*

4.2. Semejanza

Definición 4.2.1 (Matrices semejantes) *Si A y B son matrices cuadradas y si existe una matriz no singular P tal que,*

$$B = P^{-1}AP, \tag{4.10}$$

entonces se dice que B es semejante a A.

Ejemplo 4.2.1 *Para cualquier matriz cuadrada A, A es semejante a A. En efecto, si A es una matriz de tamaño nxn, entonces,* $P = I_n$*, permite que A satisfaga la ecuación 4.10.*

Ejemplo 4.2.2 *De acuerdo con el teorema 3.4.1, la matrices que representan a una transformación lineal* $T : V \to V$ *respecto a bases diferentes son semejantes.*

Ejemplo 4.2.3 *Verifique que la ecuación 4.10 se satisface para las matrices A , B y P, donde,*

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

Para verificar que la ecuación 4.10 se satisface se obtendrá la inversa de la matriz P por cualquiera de los métodos estudiados en el capítulo 1. En este caso,

$$P^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Finalmente,

$$\begin{aligned}
P^{-1}AP &= \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}, \\
&= \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 3 \\ 1 & 0 & -3 \\ 0 & 2 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \\
&= B.
\end{aligned}$$

Teorema 4.2.1 *Si B es semejante a A, entonces A es semejante a B.*

Demostración: Como B es semejante A, entonces existe una matriz invertible P tal que,

$$\begin{aligned}
P^{-1}AP &= B, \\
AP &= PB, && \text{multiplicando por } P \text{ por la izquierda.} \\
A &= PBP^{-1}, && \text{multiplicando por } P^{-1} \text{ por la derecha.}
\end{aligned}$$

Finalmente, haciendo $Q = P^{-1}$ en la última ecuación obtenemos,

$$A = Q^{-1}BQ.$$

De acuerdo con la definición, A es semejante a B. De aquí en adelante simplemente diremos que A y B son semejantes. ∎

Teorema 4.2.2 *Si A y B son matrices semejantes y, B y C también lo son, entonces A y C son semejantes.*

Demostración: Como A y B son semejantes, entonces existe una matriz invertible P tal que,

$$A = P^{-1}BP. \tag{4.11}$$

De la misma manera, como B y C son semejantes, entonces existe una matriz invertible Q tal que,

$$B = Q^{-1}CQ. \tag{4.12}$$

Sea $R = QP$, entonces por el teorema 1.3.2 la matriz R es invertible y su inversa es, $R^{-1} = P^{-1}Q^{-1}$. Sustituyendo la ecuación 4.12 en 4.11 obtenemos,

$$\begin{aligned}
A &= P^{-1}(Q^{-1}CQ)P, \\
&= (QP)^{-1}C(QP), \\
&= R^{-1}CR.
\end{aligned}$$

Con lo cual se concluye que A es semejante a C.

Teorema 4.2.3 *Si A y B son matrices semejantes, entonces* $det(B) = det(A)$.

Demostración: Como A y B son semejantes, entonces existe una matriz invertible P tal que,

$$B = P^{-1}AP,$$

por lo tanto,

$$det(B) = det(P^{-1}AP)$$

por el teorema 1.4.11, la ecuación anterior puede expresarse como,

$$det(B) = det(P^{-1})det(A)det(P.)$$

Finalmente, por el corolario 1.4.1 tenemos,

$$det(P^{-1})det(P) = 1,$$

con lo cual se concluye que, $det(B) = det(A)$. ∎

Teorema 4.2.4 *Las matrices semejantes tienen los mismos eigenvalores.*

Demostración: Probaremos que las matrices semejantes tienen el mismo polinomio característico, con lo cual se concluirá el enunciado del teorema. Consideremos que A y B son matrices semejantes con polinomios característicos $P_A(\lambda)$ y $P_B(\lambda)$, respectivamente. En este caso,

$$P_A(\lambda) = |\lambda I - A| \tag{4.13}$$

y

$$P_B(\lambda) = |\lambda I - B| \tag{4.14}$$

Además, como A y B son semejantes, entonces existe una matriz invertible P tal que,

$$B = P^{-1}AP,$$

Sustituyendo la ecuación anterior en la ecuación 4.14 obtenemos:

$$\begin{aligned}
P_B(\lambda) &= |\lambda I - P^{-1}AP|, \\
&= |\lambda P^{-1}P - P^{-1}AP|, \\
&= |P^{-1}(\lambda P - AP)|, \\
&= |P^{-1}(\lambda IP - AP)|, \\
&= |P^{-1}(\lambda I - A)P)|, \\
&= |P^{-1}||(\lambda I - A)||P|, \qquad \text{teorema 1.4.11} \\
&= |(\lambda I - A)|, \qquad \text{corolario 1.4.1} \\
&= P_A(\lambda). \quad \blacksquare
\end{aligned}$$

4.3. Diagonalización de matrices

Definición 4.3.1 (Matrices Diagonalizables) *Se dice que una matriz cuadrada A es una matriz diagonalizable, si es semejante a una matriz diagonal.*

Ejemplo 4.3.1 *La matriz A del ejemplo 4.2.3 es una matriz diagonalizable.*

Teorema 4.3.1 *Una matriz A de tamaño nxn es diagonalizable si y solo si tiene n eigenvectores linealmente independientes.*

Demostración:

a) Supongamos que A es diagonalizable. En este caso existe una matriz no singular P y una matriz diagonal D tal que,

$$D = P^{-1}AP. \tag{4.15}$$

La ecuación anterior puede expresarse como:

$$AP = PD \tag{4.16}$$

Representemos a las matrices A, P y D en términos de sus entradas,

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

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

Si definimos las columnas de la matriz P como los vectores $\mathbf{x}_j$; es decir,

$$x_j = \begin{bmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{nj} \end{bmatrix}, \quad 1 \leq j \leq n.$$

El lado izquierdo de la ecuación 4.16 puede expresarse en términos de los vectores $\mathbf{x}_j$ como,

$$AP = [A\mathbf{x}_1|A\mathbf{x}_2|\dots|A\mathbf{x}_n]$$

De manera análoga, el lado derecho de la ecuación 4.16 puede expresarse en términos de los vectores $\mathbf{x}_j$ como,

$$PD = [\lambda_1\mathbf{x}_1|\lambda_2\mathbf{x}_2|\ldots|\lambda_n\mathbf{x}_n]$$

De manera que:

$$[A\mathbf{x}_1|A\mathbf{x}_2|\ldots|A\mathbf{x}_n] = [\lambda_1\mathbf{x}_1|\lambda_2\mathbf{x}_2|\ldots|\lambda_n\mathbf{x}_n]$$

Con lo cual se concluye que,

$$A\mathbf{x}_j = \lambda_j\mathbf{x}_j \quad 1 \leq j \leq n.$$

Como la matriz P es no singular, entonces por el teorema 4.1.2 podemos concluir que los vectores $\mathbf{x}_j$ son linealmente independientes.

b) Supongamos ahora que A tiene n eigenvectores linealmente independientes. Si λ_j es el eigenvalor correspondiente al eigenvector $\mathbf{x}_j$, entonces,

$$A\mathbf{x}_j = \lambda_j\mathbf{x}_j \quad 1 \leq j \leq n.$$

Podemos representa a los eigenvectores en términos de sus entradas como:

$$x_j = \begin{bmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{nj} \end{bmatrix}, \quad 1 \leq j \leq n.$$

Con estos eigenvectores construimos la matriz P cuya columna j-ésima es el vector $\mathbf{x}_j$. De manera que,

$$P = [\mathbf{x}_1|\mathbf{x}_2|\ldots|\mathbf{x}_n] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

Como los eigenvectores $\mathbf{x}_j$ son linealmente indepedientes, entonces el teorema 4.1.2, nos permite concluir que P es una matriz no singular.

Ahora consideremos el producto matricial AP.

$$
\begin{aligned}
AP &= [A\mathbf{x}_1|A\mathbf{x}_2|\ldots|A\mathbf{x}_n], \\
&= [\lambda_1\mathbf{x}_1|\lambda_2\mathbf{x}_2|\ldots|\lambda_n\mathbf{x}_n], \\
&= \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \ldots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \ldots & \lambda_n p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \ldots & \lambda_n p_{nn} \end{bmatrix}, \\
&= \begin{bmatrix} p_{11} & p_{12} & \ldots & p_{1n} \\ p_{21} & p_{22} & \ldots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \ldots & p_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}, \\
&= PD.
\end{aligned}
$$

Lo anterior nos permite concluir que $D = P^{-1}AP$; es decir, A es semejante a la matriz diagonal D, donde D es la matriz diagonal cuyas entradas son los eigenvalores de la matriz A. ∎

Teorema 4.3.2 *Si $S = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k\}$ son eigenvectores de una matriz A correspondientes a eigenvalores diferentes $\lambda_1, \lambda_2, \ldots, \lambda_k$, entonces S es un conjunto linealmente independiente.*

Demostración: Supongamos que S es linealmente dependiente. Por el teorema 2.4.3 sabemos que para algún i, $1 < i \leq k$, $\mathbf{x}_i$ es combinación lineal de los vectores que le anteceden. Por lo tanto, podemos expresar $\mathbf{x}_i$ como:

$$\mathbf{x}_i = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \ldots a_{i-1}\mathbf{x}_{i-1}, \tag{4.17}$$

donde el conjunto $S' = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{i-1}\}$ es linealmente independiente. Multiplicando la ecuación 4.17 por A tenemos,

$$A\mathbf{x}_i = a_1A\mathbf{x}_1 + a_2A\mathbf{x}_2 + \ldots a_{i-1}A\mathbf{x}_{i-1}, \tag{4.18}$$

Como los $\mathbf{x}_i$ son eigenvectores de A, entonces

$$A\mathbf{x}_j = \lambda_j\mathbf{x}_j, \quad 1 \leq j \leq k. \tag{4.19}$$

Sustituyendo las ecuacións 4.3 en la ecuación 4.18 obtenemos,

$$\lambda_i\mathbf{x}_i = a_1\lambda_1\mathbf{x}_1 + a_2\lambda_2\mathbf{x}_2 + \ldots a_{i-1}\lambda_{i-1}\mathbf{x}_{i-1}. \tag{4.20}$$

Ahora multiplicando la ecuación 4.17 por λ_i obtenemos:

$$\lambda_i \mathbf{x}_i = a_1\lambda_i\mathbf{x}_1 + a_2\lambda_i\mathbf{x}_2 + \ldots a_{i-1}\lambda_i\mathbf{x}_{i-1}, \tag{4.21}$$

Restando las ecuaciones 4.20 y 4.21 obtenemos,

$$\mathbf{0} = a_1(\lambda_1 - \lambda_i)\mathbf{x}_1 + a_2(\lambda_2 - \lambda_i)\mathbf{x}_2 + \ldots a_{i-1}(\lambda_{i-1} - \lambda_i)\mathbf{x}_{i-1}. \tag{4.22}$$

Como S' es un conjunto linealmente independiente, entonces todos los escalares en la ecuación anterior valen cero, por lo tanto,

$$a_j(\lambda_j - \lambda_i) = 0 \quad 1 \leq j \leq i-1.$$

Como todos los eigenvalores son diferentes, entonces $(\lambda_j - \lambda_i) \neq 0$ con lo cual concluimos que, $a_1 = a_2 = \cdots = a_{i-1} = 0$, lo que a su vez indica que $\mathbf{x}_i = \mathbf{0}$. Este resultado constituye una contradicción, ya que $\mathbf{x}_i$ es un eigenvector de A. La contradicción surge de suponer que S es linealmente dependiente, por lo tanto, concluimos que S es linealmente independiente. ∎

Teorema 4.3.3 *Si una matriz A de tamaño nxn tiene n eigenvalores diferentes, entonces A es diagonalizable.*

Demostración: Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los eigenvalores de la matriz A, entonces

$$A\mathbf{x}_j = \lambda_j\mathbf{x}_j, \quad 1 \leq j \leq n.$$

Por el teorema anterior, sabemos que el conjunto de eigenvectores

$$\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$$

es linealmente independiente y por el teorema 4.3.1 A es diagonalizable. ∎

Definición 4.3.2 (Multiplicidad) *Si A es una matriz de tamaño nxn, entonces su polinomio caracterítico puede expresarse como:*

$$P(\lambda) = (\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2}...(\lambda - \lambda_r)^{k_r}$$

donde $k_1, k_2, \ldots, k_r$ son números enteros tales que,

$$k_1 + k_2 + \ldots + k_r = n$$

El exponente k_i del factor $(\lambda - \lambda_i)^{k_i}$ en el polinomio característico se conoce como multiplicidad de la raíz λ_i.

Si la multiplicidad de cada raíz del polinomio caracterísico es 1, entonces, éste tendrá n raíces diferentes y por el teorema 4.3.3, la matriz A será diagonalizable. Sin embargo, podría suceder que una o más raíces del polinomio aparezcan con una multiplicidad distinta de la unidad, en estos casos la matriz podría ser o no diagonalizable, esto dependerá de la dimensión de los eigenespacios. De hecho, si la suma de la dimesiones de los eigenespacios es igual con n, entonces A será diagonalizable.

Ejemplo 4.3.2 *Considere la matriz*

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

Determine esta la matriz es diagonalizable, en caso afirmativo determine la matriz P que la diagonaliza y verifique que se satisface la ecuación,

$$P^{-1}AP = D.$$

Solución: Esta matriz corresponde a la matriz del ejemplo 4.1.1. Sabemos que la matriz tiene 2 eigenvalores diferentes; por el teorema 4.3.3, A es diagonalizable. Los eigenvalores de A son:

$$\begin{cases} \lambda_1 = -3 \\ \lambda_2 = 2 \end{cases}$$

En el ejemplo 4.1.3 hallamos los eigenespacios correspondientes y adicionalmente, un eigenvector en cada eigenespacio.

$$\text{Para } \lambda_1 = -3\text{:} \quad \begin{bmatrix} -2 \\ 1 \end{bmatrix},$$

$$\text{Para } \lambda_2 = 2\text{:} \quad \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

De acuerdo con el teorema 4.3.2, los eigenvectores correspondientes a eigenvalores diferentes son linealmente independientes; por lo tanto, podemos formar la matriz P con los eigenvectores anteriores, de manera que,

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

La inversa de la matriz P es:

$$P^{-1} = -\frac{1}{5}\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}.$$

Finalmente, verifiquemos que A es diagonalizada por P.

$$
\begin{aligned}
P^{-1}AP &= -\frac{1}{5}\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} -4 & -2 \\ 3 & 3 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ 1 & -3 \end{bmatrix}, \\
&= -\frac{1}{5}\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 6 & 2 \\ -3 & -6 \end{bmatrix}, \\
&= -\frac{1}{5}\begin{bmatrix} 15 & 0 \\ 0 & -10 \end{bmatrix}, \\
&= \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix} \\
&= D.
\end{aligned}
$$

Ejemplo 4.3.3 *Considere la matriz A del ejemplo 4.1.2. Determine una matriz P que la diagonalice.*

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

Solución: Los eigenvalores de la matriz A fueron determinados en el ejemplo 4.2. Los eigenvalores obtenidos son:

$$
\begin{cases} \lambda_1 = -1, \\ \lambda_2 = 2, \\ \lambda_3 = 4. \end{cases}
$$

En ejemplo 4.1.4 determinamos los eigenespacios correspondientes; además, determinamos un vector en cada eigenespacio.

$$
\text{Para } \lambda_1 = -1\text{:} \quad \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},
$$

$$
\text{Para } \lambda_2 = 2\text{:} \quad \begin{bmatrix} -2 \\ -3 \\ 2 \end{bmatrix}
$$

$$
\text{Para } \lambda_3 = 4\text{:} \quad \begin{bmatrix} 8 \\ 5 \\ 2 \end{bmatrix},
$$

Estos tres vectores son linealmente independientes; por lo tanto, con ellos podemos construir la matriz P,

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

Esta matriz P diagonaliza a la matriz A.

4.4. Matrices simétricas y diagonalización ortogonal

Definición 4.4.1 (Matrices simétricas) *Dada una matriz A de tamaño nxn, se dice que A es simétrica si ésta es igual a su transpuesta; es decir,* $A = A^T$.

Definición 4.4.2 (Matriz ortogonal) *Una matriz es ortogonal si ésta es invertible y si su inversa es igual a su transpuesta; es decir,* $A^{-1} = A^T$

Definición 4.4.3 (Diagonalización ortogonal) *Se dice que una matriz A de tamaño nxn es diagonalizable ortogonalmente, si existe una matriz ortogonal P tal que* $P^{-1}AP$ *es una matriz diagonal.*

Teorema 4.4.1 *Si A es una matriz de nxn, entonces,*

a) A es diagonalizable ortogonalmente.

b) A tiene un conjunto ortonormal de n eigenvectores.

c) A es simétrica

Demostración:

a) $\Rightarrow$ *b)* Si A es diagonalizable ortogonalmente, entonces exite una matriz ortogonal P tal que, $P^{-1}AP$ es diagonal. En la demostración del teorema 4.15 se mostró que las columnas de P son los eigenvalores de la matriz A; además, como P es ortogonal, entonces de acuerdo con el teorema 2.6.5 las columnas de P son vectores ortonormales. De esta manera se concluye que A tiene un conjunto ortonormal de n eigenvectores.

b) $\Rightarrow$ *a)* Si A tiene un conjunto ortonormal de n eigenvectores, entonces con estos vectores formamos las columnas de la matriz P. Por el teorema 4.15 sabemos que P diagonaliza a A; además, de acuerdo con el teorema 2.6.5 la matriz P es ortogonal. Se concluye que A es diagonalizable ortogonalmente.

a) $\Rightarrow$ *c)* Si A es diagonalizable ortogonalmente, entonces existe una matriz ortogonal P tal que,

$$P^{-1}AP = D,$$

donde D es una matriz diagonal. De la ecuación anterior obtenemos:

$$A = PDP^{-1}.$$

Como P es una matriz ortogonal, entonces $P^T = P^{-1}$, por lo tanto,

$$A = PDP^T.$$

Ahora determinemos la transpuesta de la matriz A.

$$A^T = (PDP^T)^T = P(PD)^T = PD^TP^T.$$

Como D es una matriz diagonal, entonces $D = D^T$, con lo cual concluimos que:

$$A^T = PDP^T = A;$$

Es decir, A es simétrica.

c) $\Rightarrow$ *a)* La demostración se omite, ya que queda fuera del alcance de este curso.

Teorema 4.4.2 *Si A es una matriz simétrica, entonces:*

a) Los eigenvectores correspondientes a eigenvalores distintos son ortogonales.

b) Todos los eigenvalores de una matriz simétrica son números reales.

Demostración: a) En el elemplo **??** se demostró la siguiente relación para el producto euclideano interior: Si $\mathbf{x}$ y $\mathbf{y}$ son vectores en $\mathbb{R}^n$ y A es una matriz de nxn, entonces

$$\langle A\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, A^T\mathbf{y}\rangle$$

Si $\mathbf{x}_i$ y $\mathbf{x}_j$ son eigenvectores de A correspondientes a eigenvalores diferentes λ_i, y λ_j, respectivamente. Entonces,

$$\begin{aligned}
\langle A\mathbf{x}_i, \mathbf{x}_j\rangle &= \langle \mathbf{x}_i, A^T\mathbf{x}_j\rangle, \\
&= \langle \mathbf{x}_i, A\mathbf{x}_j\rangle, && \text{A es simétrica} \\
\langle \lambda_i\mathbf{x}_i, \mathbf{x}_j\rangle &= \langle \mathbf{x}_i, \lambda_j\mathbf{x}_j\rangle, && \mathbf{x}_j \text{ y } \mathbf{x}_j \text{ son eigenvectores de A} \\
\lambda_i\langle \mathbf{x}_i, \mathbf{x}_j\rangle &= \lambda_j\langle \mathbf{x}_i, \mathbf{x}_j\rangle. && \text{Simetría y homogeneidad del producto interior}
\end{aligned}$$

La última ecuación se puede expresar como,

$$(\lambda_i - \lambda_j)\langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0.$$

Como los eigenvalores son diferentes, entonces $(\lambda_i - \lambda_j) \neq 0$ y, por lo tanto,

$$\langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0;$$

es decir, $\mathbf{x}_j$ y $\mathbf{x}_j$ son ortogonales.

b) La demostración se omite ya que queda fuera del alcance de este curso.

Ejemplo 4.4.1 *Considere la matriz simétrica,*

$$A = \begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix},$$

a) determine la matriz ortogonal P que la diagonaliza,

b) verifique la ecuación $P^T AP = D$

Solución:

a) Determinemos los eigenvalores de la matriz A:

$$P(\lambda) = \begin{vmatrix} \lambda - 4 & 2 \\ 2 & \lambda - 4 \end{vmatrix} = (\lambda-4)(\lambda-4)-4 = \lambda^2-8\lambda+12 = (\lambda-2)(\lambda-6).$$

Las raíces del polinomio característico son: $\lambda_1 = 2$ y $\lambda_2 = 6$. Los eigenespacios que se obtiene al resolver la homogénea,

$$\left[\begin{array}{cc|c} \lambda - 4 & 2 & 0 \\ 2 & \lambda - 4 & 0 \end{array}\right]$$

son:

Para $\lambda_1 = 2$,

$$V_{\lambda_1} = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : \begin{bmatrix} x \\ y \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

Para $t = 1$, $\mathbf{p}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda_1 = 1$.

Para $\lambda_2 = 6$,

$$V_{\lambda_2} = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : \begin{bmatrix} x \\ y \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

Para $t = 1$, $\mathbf{p}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda_2 = 1$. Para finalizar, se normalizan los eigenvectores obtenidos. La matriz P resulta ser,

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

b) Verifiquemos las ecuación $P^TAP = D$.

$$\begin{aligned} P^TAP &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \\ &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \\ &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -6 \\ 2 & 6 \end{bmatrix}, \\ &= \frac{1}{2} \begin{bmatrix} 4 & 0 \\ 0 & 12 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}, \\ &= D. \end{aligned}$$

Ejemplo 4.4.2 *Consideremos la matriz simétrica*

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

determine una matriz ortogonal P que la diagonalice.

Solución: Determinemos sus eigenvalores,

$$\begin{aligned} P(\lambda) &= \begin{vmatrix} \lambda-2 & 1 & 0 \\ 1 & \lambda-2 & 0 \\ 0 & 0 & \lambda-1 \end{vmatrix}, \\ &= (\lambda-1)[(\lambda-2)^2-1], \\ &= \lambda^3-5\lambda^2+7\lambda-3, \\ &= (\lambda-1)^2(\lambda-3). \end{aligned}$$

Por lo anterior, los eigenvalores de la matriz A son: $\lambda_1 = 1$ y $\lambda_2 = 3$. Para determinar los eigenespacios, debemos resolver los sistemas homogéneos $[\lambda_i I - A] = 0$. Los eigenespacios correspondientes a los eigenvalores obtenidos son:

Para $\lambda_1 = 1$,

$$V_{\lambda_1} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = r \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{para cualquier pareja de escalares r,t} \right\}.$$

Para $r = 0$ y $t = 1$, $\mathbf{p}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda_1 = 1$.

Para $r = 1$ y $t = 0$, $\mathbf{p}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Podemos apreciar que estos vectores son ortogonales.

Para $\lambda_2 = 3$,

$$V_{\lambda_2} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

Para $t = 1$, $\mathbf{p}_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda_2 = 3$.

Finalmente, se normalizan los eigenvectores anteriores y se construye la matriz P. Además, recordemos que el orden en que se colocan los eigenvectores para formar la matriz P, solo tiene efecto sobre el orden en que aparecen los eigenvalores en la matriz diagonal D. Por lo tanto, si formamos la matriz de la siguiente manera,

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

entonces

$$D = P^T AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

4.5. Formas cuadráticas y seciones cónicas

Definición 4.5.1 (Formas cuadráticas) *Si A es una matriz simétrica de tamaño nxn, entonces la función $g : \mathbb{R}^n \to \mathbb{R}$ definida como:*

$$g(\boldsymbol{x}) = \boldsymbol{x}^T A\boldsymbol{x}, \tag{4.23}$$

donde,

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

se conoce como forma cuadrática real en las variables $x_1, x_2, \ldots, x_n$. La matriz A se conoce como la matriz de la forma cuadrática g.

Ejemplo 4.5.1 *Escriba de manera explícita la forma cuadrática para las variables x y y.*
Para desarrollar la forma cuadrática debemos considerar una matriz simétrica general de tamaño 2x2. En este caso,

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix},$$

de modo que,

$$\begin{aligned} g(x,y) &= \boldsymbol{x}^T A\boldsymbol{x}, \\ &= [x,y]\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}, \\ &= [x,y]\begin{bmatrix} ax+by \\ bx+cy \end{bmatrix}, \\ &= x(ax+by)+y(bx+cy), \\ &= ax^2+2bxy+cy^2. \end{aligned}$$

Estructura de la forma cuadrática general

Si consideramos que $\boldsymbol{x}$ es una matriz de tamaño nx1 y que $\boldsymbol{x}^T$ es una matriz de tamaño 1xn, entonces podemos representarlas como:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix},$$

$$\boldsymbol{x}^T = [x_1^T, x_2^T, \ldots, x_n^T] = [x_{11}^T, x_{21}^T, \ldots, x_{n1}^T].$$

Además, el producto $A\boldsymbol{x}$ es una matriz de tamaño nx1. Denotaremos a sus entradas como $[A\boldsymbol{x}]_{k1}$. Por otra parte, podemos observar que $\boldsymbol{x}^T(A\boldsymbol{x})$ corresponde al producto interior definido en el ejemplo **??**. *Por lo tanto,*

$$\begin{aligned} g(\boldsymbol{x}) &= \boldsymbol{x}^T(A\boldsymbol{x}), \\ &= \sum_j [\boldsymbol{x}^T]_{1j}[A\boldsymbol{x}]_{j1}, \\ &= \sum_j [\boldsymbol{x}^T]_{1j}\sum_i [A]_{ji}[\boldsymbol{x}]_{i1}, \\ &= \sum_j\sum_i [\boldsymbol{x}]_{j1}[A]_{ji}[\boldsymbol{x}]_{i1}, \\ &= \sum_j\sum_i a_{ij}x_jx_i, \\ &= \sum_i a_{ii}x_i^2 + \sum_{i\neq j} a_{ij}x_ix_j. \end{aligned}$$

Los términos $a_{ij}x_ix_j$ en la segunda sumatoria de la última ecuación se conocen como términos de producto cruzado en la forma cuadrática.

Teorema 4.5.1 *Sea $\boldsymbol{x}^T A\boldsymbol{x}$ una forma cuadrática en los valores $x_1, x_2, \ldots, x_n$, donde A es una matriz simétrica. Si P diagonaliza ortogonalmente a A y si las nuevas variables $y_1, y_2, \ldots y_n$ están definidas por la ecuación $\boldsymbol{x} = P\boldsymbol{y}$, entonces al sustituir esta ecuación en $\boldsymbol{x}^T A\boldsymbol{x}$ se obtiene,*

$$\boldsymbol{x}^T A\boldsymbol{x} = \boldsymbol{y}^T D\boldsymbol{y} = \sum_{i=1}^{n} \lambda_i y_i^2,$$

donde, $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los eigenvalores de la matriz A y

$$D = P^T AP = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$

Demostración: Como A es una matriz simétrica, entonces por el teorema 4.4.1, existe una matriz ortogonal P que diagonaliza a A. Además, las columnas de la matriz P se forman con los eigenvectores normalizados de V, de manera que,

$$D = P^T AP = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix},$$

donde, $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los eigenvalores de la matriz A.
Si las nuevas variables se definen como

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

es decir, $\mathbf{y} = P\mathbf{x}$, entonces

$$\begin{aligned} g(\mathbf{x}) &= \mathbf{x}^T A\mathbf{x}, \\ &= (P\mathbf{y})^T A(P\mathbf{y}), \\ &= \mathbf{y}^T (P^T AP)\mathbf{y}, \\ &= \mathbf{y}^T D\mathbf{y}, \\ &= \sum_{i=1}^{n} \lambda_i y_i^2. \quad \blacksquare \end{aligned}$$

Secciones cónica

Uno de los problemas que se pueden abordar con la teoría desarrollada en este curso, es el estudio de las cónicas. A continuación se desarrollará un enfoque matricial para la interpretación de la ecuación de una cónica.

Definición 4.5.2 (Ecuación cuadrática) *Se conoce como ecuación cuadrática en las variables x y y a la ecuación,*

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0, \tag{4.24}$$

donde, a, b, c, d, e, f son todos números reales y por lo menos uno de los números a, b, c es diferente de cero.

Definición 4.5.3 *Dada una ecuación cuadrática, se conoce como forma cuadrática asociada al término,*

$$ax^2 + 2bxy + cy^2. \tag{4.25}$$

Las gráficas de la ecuación 4.24, se conocen como secciones cónicas o simplemente como cónicas. El nombre hace referencia a que dichas gráficas se obtienen al intersecar un cono circular recto de dos hojas con un plano. Las secciones cónicas pueden ser esferas, elipses, parábolas, hipérbolas e incluso un punto, una recta, dos rectas, el conjunto vacío. Se conocen como casos no degenerados a las esferas, elipses, parábolas, hipérbolas, mientras que el resto de las gráficas se conocen como casos degenerados.
Se dice que una cónica no degenerada se encuentra en posición estándar, normal o canónica, si su ecuación se pueden expresar en una de las formas que se muestran en la tabla 4.1. Las gráficas estas ecuaciones se muestran en la figura 4.1.
Podemos observar que ninguna cónica en su posición normal contiene el término de producto cruzado xy. Podemos notar, además, que las ecuaciones no contienen simultáneamente los términos x^2 y x ni tampoco y^2 y y. Si la ecuación de la cónica contine el término cruzado xy, este hecho implica que la cónica está rotada y posiblemente transladada respecto a su posición estándar. Si la ecuación de la cónica contiene el término cuadrático y lineal en al menos una de las variables, de manera simultánea, entonces la cónica se encuentra transladada y posiblemente rotada respecto a su posición estándar. En la figura 4.2 se muestran algunas de las posibles gráficas.

Tabla 4.1: Ecuaciones de las cónicas en posición estándar

Cónica	Ecuación
a) Circunferencia	$\frac{x^2}{a^2} + \frac{y^2}{a^2} = 1$
b) Elipse	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1; \quad a > b > 0$
c) Elipse	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1; \quad b > a > 0$
d) Parábola	$x^2 = ay; \quad a > 0$
e) Parábola	$x^2 = ay; \quad a < 0$
f) Parábola	$y^2 = ax; \quad a > 0$
g) Parábola	$y^2 = ax; \quad a < 0$
h) Hipérbola	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1; \quad a > 0, b > 0$
i) Hipérbola	$\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1; \quad a > 0, b > 0$

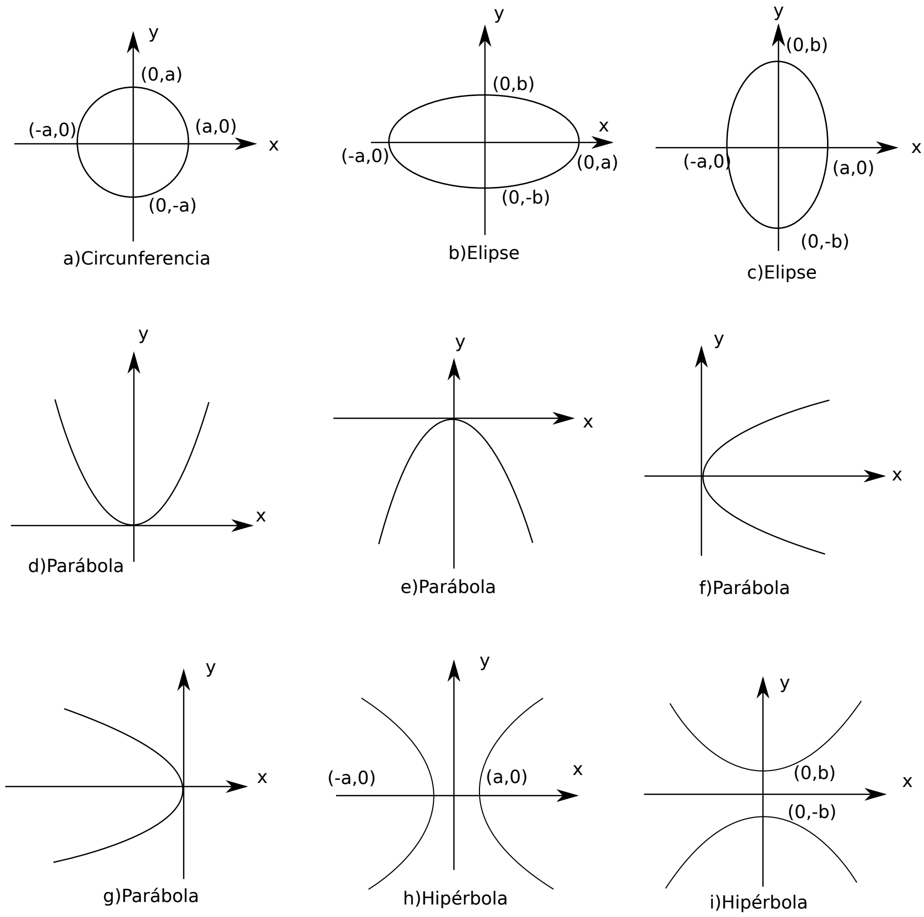

Figura 4.1: Secciones cónicas en posiciones canónicas

Teorema 4.5.2 (Teorema de los ejes principales para $\mathbb{R}^2$) *Sea*

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0,$$

la ecuación de una cónica C, sea

$$\boldsymbol{x}^T A\boldsymbol{x} = ax^2 + 2bxy + cy^2,$$

la forma cuadrática asociada, entonces los ejes de coordenadas se pueden girar de modo que la ecuación de la cónica en el nuevo sistema de coordenadas $x'y'$ sea de la forma,

$$\lambda_1 x'^2 + \lambda_2 y'^2 + d'x' + e'y' + f = 0,$$

donde λ_1 y λ_2 son los eigenvalores de A. La rotación se puede efectuar mediante la sustitución $\boldsymbol{x} = P\boldsymbol{x}'$, donde P diagonaliza ortogonalmente a $\boldsymbol{x}^T A\boldsymbol{x}$ y $det(P) = 1$.

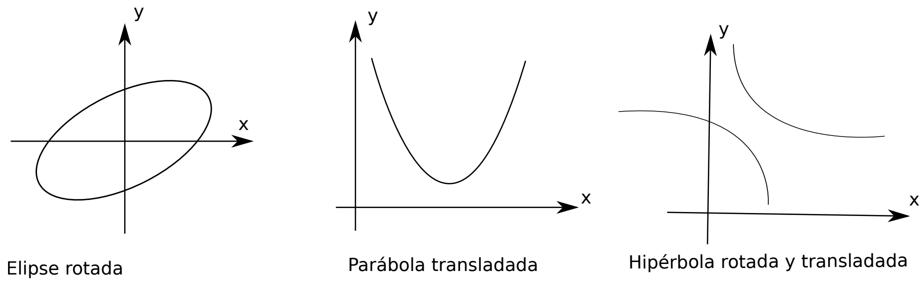

Figura 4.2: Secciones cónicas rotadas o transladadas

Demostración: La ecuación 4.24, puede representar como:

$$\mathbf{x}^T A\mathbf{x} + B\mathbf{x} + f = 0, \tag{4.26}$$

donde

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad y \quad B = [d, e].$$

Como A es una matriz simétrica, entonces de acuerdo con el teorema 4.5.1, existe una matriz ortogonal P tal que, P^TAP es una matriz diagonal; de manera más precisa,

$$P^TAP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

donde λ_1 y λ_2 son los eigenvalores de la matriz A.

Si hacemos $\mathbf{x} = P\mathbf{x}'$, y sustituimos en la ecuación 4.26 obtenemos,

$$\begin{aligned} \mathbf{x}^T A\mathbf{x} + B\mathbf{x} + f &= 0, \\ \mathbf{x}'^T P^T AP\mathbf{x}' + BP\mathbf{x}' + f &= 0, \\ \mathbf{x}'^T DP\mathbf{x}' + BP\mathbf{x}' + f &= 0, \end{aligned} \tag{4.27}$$

haciendo

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix}, \quad y \quad P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix},$$

y sustituyendo en la ecuación 4.27 obtenemos,

$$\begin{aligned} [x', y'] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [d, e] \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + f &= 0, \\ \lambda_1 x'^2 + \lambda_2 y'^2 + (dp_{11} + ep_{21})x' + (dp_{12} + ep_{22})y' + f &= 0. \end{aligned} \tag{4.28}$$

Finalmente, sean

$$d' = dp_{11} + ep_{21}$$

y

$$e' = dp_{12} + ep_{22}$$

entonces la ecuación 4.28 se expresa como:

$$\lambda_1 x'^2 + \lambda_2 y'^2 + d'x' + e'y' + f = 0.$$

Como vimos en el ejemplo 2.6.5, para una rotación positiva; es decir, en el sentido contrario de las manecillas del reloj, los ejes x' y' están a lo largo de los vectores propios de la matriz A y además, $detP = 1$. Para garantizar que tengamos una rotación positiva, podemos intercambiar los eigenvectores en la matriz P cuando sea necesario. Sabemos también que si $\mathbf{x}$ es un eigenvector de A, entonces $-\mathbf{x}$ también será un eigenvector de la matriz , lo cual equivale a multiplicar una columna de P por (-1), todo con el fin de que $detP = 1$. ▮

Ejemplo 4.5.2 *Describa y grafique la cónica representada por la ecuación cuadrática,*

$$5x^2 + 5y^2 - 8xy - 36 = 0.$$

Solución: La ecuación cuadrática anterior puede expresarse como:

$$\begin{aligned} 5x^2 + 5y^2 - 8xy - 36 &= 0, \\ \mathbf{x}^T A\mathbf{x} - 36 &= 0, \\ \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - 36 &= 0. \end{aligned} \tag{4.29}$$

Determinemos los eigenvalores de la matriz A:

$$\begin{aligned} P(\lambda) &= \begin{vmatrix} \lambda - 5 & 4 \\ 4 & \lambda - 5 \end{vmatrix} \\ &= (\lambda - 5)^2 - 16. \end{aligned}$$

Los eigenvalores de la matriz A son: $\lambda_1 = 1$ y $\lambda_2 = 9$. Los eigenespacios correspondientes pueden expresarse como:

Para $\lambda = 1$

$$\left\{ \begin{bmatrix} x \\ y \end{bmatrix} : \begin{bmatrix} x \\ y \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

De manera que para $t = 1$, $\mathbf{e}_1' = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, es un eigevector de A.

Para $\lambda = 9$

$$\left\{ \begin{bmatrix} x \\ y \end{bmatrix} : \begin{bmatrix} x \\ y \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

De manera que para $t = 1$, $\mathbf{e}_2' = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, es un eigenvector de A.

Finalmente, con los eigenvectores normalizados construiremos la matriz P.

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

La matriz P diagonaliza ortogonalmente a la matriz A, además $detA = 1$, de manera que no hay necesidad de intercambiar las columnas de P. La matriz P representa una rotación de 45°, como se estudió en el ejemplo 2.6.5. La matriz P satisface la ecuación:

$$P^T A P = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}.$$

Haciendo la sustitución $\mathbf{x} = P\mathbf{x}'$ en la ecuación 4.29 obtenemos:

$$x'^2 + 9y'^2 - 36 = 0,$$

de manera que,

$$\frac{x'^2}{6^2} + \frac{y'^2}{2^2} = 1.$$

Esta ecuación representa una elipse con su eje principal sobre el eje x'. Pueden también consultarse la tabla 4.1 y la figura 4.1 para identificar dicha cónica. La elipse se encuentra rotada respecto a su posición estándar en un ángulo de 45°, como se muestra en la figura 4.3.

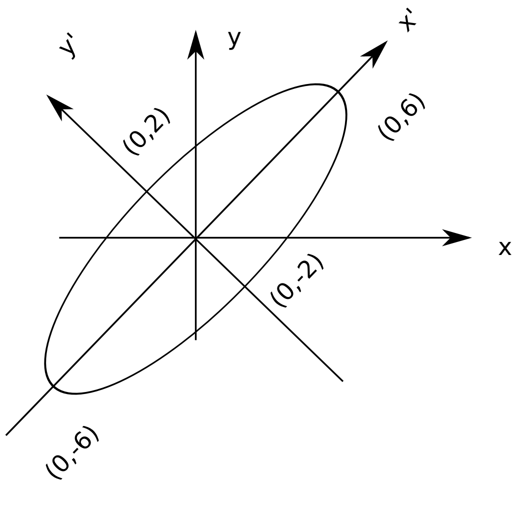

Figura 4.3: Elipse rotada

Ejemplo 4.5.3 *Describa y grafique la cónica representada por la ecuación cuadrática,*

$$5x^2 + 5y^2 - 8xy + 75\sqrt{2}x - 105\sqrt{2}y + 1089 = 0. \tag{4.30}$$

Solución: La ecuación cuadrática anterior puede expresarse como:

$$5x^2 + 5y^2 - 8xy + 75\sqrt{2}x - 105\sqrt{2}y + 1089 = 0,$$

$$\begin{aligned} \mathbf{x}^T A\mathbf{x} + B\mathbf{x} + f &= 0, \\ \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [75\sqrt{2}, -105\sqrt{2}] \begin{bmatrix} x \\ y \end{bmatrix} + 1089 &= 0. \end{aligned} \tag{4.31}$$

Procediendo como en el ejemplo anterior, se debe determinar la matriz de rotación P que diagonaliza a la matriz A. Retomando los resultados del ejemplo anterior tenemos,

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Haciendo la sustitución $\mathbf{x} = P\mathbf{x}'$ en la ecuación 4.29 obtenemos,

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}}x' - \frac{1}{\sqrt{2}}y' \\ \frac{1}{\sqrt{2}}x' + \frac{1}{\sqrt{2}}y' \end{bmatrix} \tag{4.32}$$

Sustituyendo la ecuación anterior en la ecuación 4.31 obtenemos,

$$\begin{aligned}
x'^2 + 9y'^2 + [75\sqrt{2}, -105\sqrt{2}] \begin{bmatrix} \frac{1}{\sqrt{2}}x' - \frac{1}{\sqrt{2}}y' \\ \frac{1}{\sqrt{2}}x' + \frac{1}{\sqrt{2}}y' \end{bmatrix} + 1089 &= 0, \\
x'^2 + 9y'^2 + [75, -105] \begin{bmatrix} x' - y' \\ x' + y' \end{bmatrix} + 1125 &= 0, \\
x'^2 + 9y'^2 - 30x' - 180y' + 1089 &= 0, \\
(x'^2 - 2(15)x') + 9(y'^2 - 2(10)y') + 1089 &= 0. \quad (4.33)
\end{aligned}$$

Finalmente, completando los trinomios cuadrados perfectos se obtiene,

$$\begin{aligned}
(x'^2 - 2(15)x') + 9(y'^2 - 2(10)y') + 1089 &= 0, \\
(x'^2 - 2(15)x' + 15^2) + 9(y'^2 - 2(10)y' +^(10)^2) - 36 &= 0, \\
(x' - 15)^2 + 9(y' - 10)^2 - 36 &= 0. \quad (4.34)
\end{aligned}$$

De la ecuación 4.34 obtenemos,

$$\frac{(x' - 15)^2}{6^2} + \frac{(y' - 10)^{2\,2}}{2} = 1.$$

Haciendo el cambio de variables

$$\mathbf{x}'' = \mathbf{x}' - \mathbf{x}'_0 = \begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix} - \begin{bmatrix} 15 \\ 10 \end{bmatrix} = \begin{bmatrix} x' - 15 \\ y' - 10 \end{bmatrix}. \quad (4.35)$$

obtenemos:

$$\frac{x''^2}{6^2} + \frac{y''^2}{2^2} = 1.$$

Desde el sistema $x''y''$, tenemos una elipse en su forma canónica, con si eje principal sobre el eje x". El cambio de coordenadas representado por la ecuación 4.32 representa una rotación positiva de 45°, mientras que el cambio de coordenadas representado por 4.35 representa una translación. Respecto al sistema $x'y'$, la elipse se encuente transladada a la posición $(15, 10)$. Por lo tanto, desde el sistema xy, la ecuación 4.30 representa una elipse rotada 45° respecto a su eje principal y translada a la posición $(\frac{5}{\sqrt{2}}, \frac{25}{\sqrt{2}})$ como se muestra en la figura 4.4.

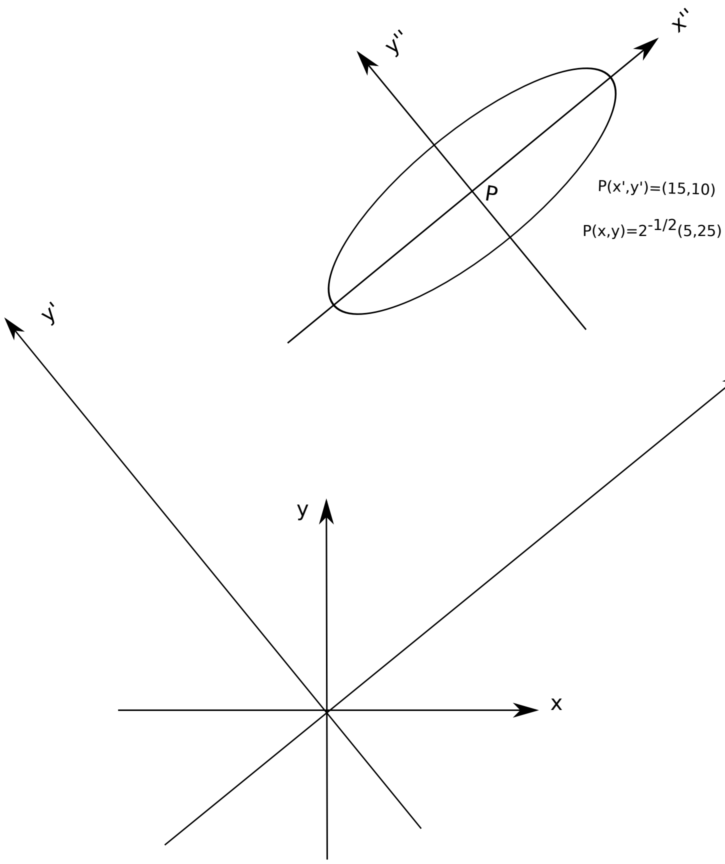

Figura 4.4: Elipse rotada y transladada

4.6. Aplicaciones a las ecuaciones diferenciales matriciales

Las ecuaciones diferenciales, son igualdades que relacionan una función desconocida y sus derivadas. Un ejemplo simple de una ecuación diferencial es,

$$y' = ay. \tag{4.36}$$

En la ecuación 4.36, nuestro interés se centra en determinar la función o familia de funciones, $y = f(x)$, definidas en los reales, que satisfagan dicha ecuación. Podemos verificar fácilmente que,

$$y = be^{ax}, \tag{4.37}$$

donde b es cualquier constante, satisface la ecuación 4.36, ya que

$$y' = abe^a x = ay.$$

La ecuación 4.37 se conoce como solución general de la ecuación $y' = ay$. Para determinar una solución particular de la ecuación diferencial 4.36, debemos conocer o imponer, lo que se conoce como condición inicial. Por ejemplo, si se impone la condición inicial $y(0) = 1$, entonces, debemos evaluar la ecuación 4.37 y obtener el valor de los parámetros, en este caso el parámetro es la constante b, que permiten satisfacer la condición inicial. Para el caso que hemos planteado,

$$y(0) = be^{a0} = b = 1,$$

de manera que la solución particular que satisface la condición inicial, $y(0) = 1$, es

$$y(x) = e^{ax}.$$

Definición 4.6.1 (Sistemas lineales homogéneos) *Se conoce como sistema lineal homogéneo de ecuaciones diferenciales o simplemente como sistema lineal homogéneo, al sistema lineal de ecuaciones diferenciales representado mediante,*

$$\begin{array}{rcl} y_1'(x) & = & a_{11}y_1(x) + a_{12}y_2(x) + \ldots + a_{1n}y_n(x), \\ y_2'(x) & = & a_{21}y_1(x) + a_{22}y_2(x) + \ldots + a_{2n}y_n(x), \\ \vdots & = & \vdots \\ y_n'(x) & = & a_{n1}y_1(x) + a_{n2}y_2(x) + \ldots + a_{nn}y_n(x), \end{array} \tag{4.38}$$

donde los elementos a_{ij} son constantes y $y_1 = f_1(x)$, $y_2 = f_2(x)$, $\ldots$, $y_n = f_n(x)$ son funciones de valor real.

La ecuación 4.38 puede representarse de forma matricial como,

$$\begin{bmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \tag{4.39}$$

o equivalentemente,

$$\mathbf{y}' = A\mathbf{y}. \tag{4.40}$$

La función

$$\mathbf{y}(x) = \begin{bmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{bmatrix}. \tag{4.41}$$

que satisface la ecuación 4.38 se conoce como solución del sistema dado.
Si $S = \{\mathbf{y}^1(x), \mathbf{y}^2(x), \ldots, \mathbf{y}^n(x)\}$ es un conjunto de soluciones de la ecuación 4.38, entonces, cualquier combinación lineal de estas soluciones, será también una solución del sistema. En efecto, sea

$$\mathbf{y}(x) = \sum_{i=1}^{n} b_i \mathbf{y}^i(x), \tag{4.42}$$

una combinación lineal de los vectores en S, entonces

$$\mathbf{y}'(x) = \sum_{i=1}^{n} b_i \mathbf{y}'^i(x).$$

Por otra parte,

$$\begin{aligned} A\mathbf{y}(x) &= A\left(\sum_{i=1}^{k} b_i \mathbf{y}^i(x)\right), \\ &= \sum_{i=1}^{k} b_i A\mathbf{y}^i(x), \\ &= \sum_{i=1}^{k} b_i \mathbf{y}'^i(x), \\ &= \mathbf{y}'(x). \end{aligned}$$

El conjunto S se conoce como **conjunto fundamental**, si cualquier solución de la ecuación 4.38 puede expresarse como combinación lineal de sus elementos. En este caso, la ecuación 4.42, se conoce como **solución general** del sistema representado por la eccuación 4.38.
Para obtener una **solución particular** de la solución general, debemos conocer o imponer una condición inicial. Por ejemplo, conocer el valor de la función en un punto $x = 0$, de manera que,

$$\mathbf{y}(0) = \mathbf{y}_0. \tag{4.43}$$

Esta ecuación en realidad representa un sistema de ecuaciones, ya que, de acuerdo con la ecuación 4.42 tenemos,

$$\mathbf{y}(0) = b_1\mathbf{y}^1(0) + b_2\mathbf{y}^2(0) + \ldots + b_n\mathbf{y}^n(0). \tag{4.44}$$

Las incógnitas son las contantes $b_1, b_2, \ldots, b_n$, de modo que la ecuación 4.43 puede representarse como,

$$C\mathbf{b} = \mathbf{y}_0, \tag{4.45}$$

donde ,

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}, \tag{4.46}$$

y C es la matriz cuyas columnas son $\mathbf{y}^1(0), \mathbf{y}^2(0), \ldots, \mathbf{y}^n(0)$, respectivamente. Si el conjunto S es un cojunto fundamental para el sistema 4.38, entonces C es una matriz no singular, de manera que la ecuación 4.45, siempre tiene solución.

Si la matriz A en la ecuación 4.47 es una matriz diagonal, el sistema puede resolver con relativa facilidad, pues en este caso,

$$\begin{bmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & \ldots & \\ 0 & a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}. \tag{4.47}$$

El sistema anterior reprenta n ecuaciones de la forma mostrada en la ecuación 4.36, en este caso las soluciones son,

$$\begin{cases} y_1 = b_1 e^{a_{11}x}, \\ y_2 = b_2 e^{a_{22}x}, \\ \vdots, \\ y_n = b_n e^{a_{nn}x}. \end{cases}$$

La solución general en forma vectorial se expresa como,

$$\begin{aligned} \mathbf{y}(x) &= \begin{bmatrix} b_1 e^{a_{11}x}, \\ b_2 e^{a_{22}x}, \\ \vdots, \\ b_n e^{a_{nn}x}. \end{bmatrix} \\ &= b_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} e^{a_{11}x} + b_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} e^{a_{22}x} + \vdots + b_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} e^{a_{nn}x}. \end{aligned}$$

Teorema 4.6.1 *Si una matriz A de nxn tiene n vectores propios linealmente independientes $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$ asociados con los valores propios $\lambda_1, \lambda_2 \ldots, \lambda_n$,* *respectivamente, entonces la solución general del sistema lineal homogéneo de ecuaciones diferenciales*

$$\boldsymbol{y}' = A\boldsymbol{y}, \tag{4.48}$$

está dado por,

$$\boldsymbol{y}(x) = b_1\boldsymbol{p}_1 e^{\lambda_1 x} + b_2\boldsymbol{p}_2 e^{\lambda_2 x} + \ldots + b_n\boldsymbol{p}_n e^{\lambda_n x}. \tag{4.49}$$

Demostración: Como A tiene n eigenvectores linealmente independientes, entonces de acuerdo con el teorema 4.3.1, A es diagonalizable. La matriz no singular P que diagonaliza a la matriz A, tiene por columnas los eigenvectores de ésta, es decir,

$$P = [\mathbf{p}_1|\mathbf{p}_2|\ldots|\mathbf{p}_n] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix},$$

además,

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$

Ahora consideremos el cambio de variables mediante,

$$\mathbf{y}(x) = P\mathbf{u}(x), \tag{4.50}$$

como P es una matriz constante, entonces

$$\mathbf{y}'(x) = P\mathbf{u}'(x). \tag{4.51}$$

Sustituyendo las ecuaciones 4.50 y 4.51 en la ecuación 4.48 obtenemos,

$$P\mathbf{u}' = AP\mathbf{u} \tag{4.52}$$

La ecuación anterior puede expresarse como,

$$\begin{aligned} \mathbf{u}' &= P^{-1}AP\mathbf{u}, \\ &= D\mathbf{u}, \end{aligned}$$

De manera explícita tenemos,

$$\begin{bmatrix} u_1' \\ u_2' \\ \vdots \\ u_n' \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}. \tag{4.53}$$

La solución del sistema anterior es,

$$\begin{cases} u_1 = b_1 e^{\lambda_1 x}, \\ u_2 = b_1 e^{\lambda_2 x}, \\ \vdots = \vdots \\ u_n = b_1 e^{\lambda_n x}. \end{cases}$$

De manera vectorial podemos representar la solución anterior mediante:

$$\begin{aligned} \mathbf{u}(x) &= \begin{bmatrix} b_1 e^{\lambda_1 x} \\ b_2 e^{\lambda_2 x} \\ \vdots \\ b_n e^{\lambda_n x} \end{bmatrix} \\ &= b_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} e^{\lambda_1 x} + b_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} e^{\lambda_2 x} + \cdots + b_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} e^{\lambda_n x}. \end{aligned}$$

Finalmente, multiplicando por la matriz P obtenemos,

$$\begin{aligned} \mathbf{y}(x) &= P\mathbf{u}(x) \\ &= b_1 P \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} e^{\lambda_1 x} + b_2 P \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} e^{\lambda_2 x} + \ldots + b_n P \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} e^{\lambda_n x}, \\ &= b_1 \begin{bmatrix} p_{11} \\ p_{21} \\ \vdots \\ p_{n1} \end{bmatrix} e^{\lambda_1 x} + b_2 \begin{bmatrix} p_{12} \\ p_{22} \\ \vdots \\ p_{n2} \end{bmatrix} e^{\lambda_2 x} + \ldots + b_n \begin{bmatrix} p_{1n} \\ p_{2n} \\ \vdots \\ p_{nn} \end{bmatrix} e^{\lambda_n x}, \\ &= b_1 \mathbf{p}_1 e^{\lambda_1 x} + b_2 \mathbf{p}_2 e^{\lambda_2 x} + \ldots + b_n \mathbf{p}_n e^{\lambda_n x}. \quad \blacksquare \end{aligned}$$

Ejemplo 4.6.1 *Considere el sistema homogéneo representado por*

$$\begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

a) Determine la solución general del sistema.

b) Determine la solución particular del sistema si $\boldsymbol{y}(0) = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$.

Solución:

a) Determinemos los eigenvalores de la matriz A.

$$P(\lambda) = \begin{vmatrix} \lambda - 1 & -2 \\ 1 & \lambda - 4 \end{vmatrix} = (\lambda-1)(\lambda-4)+2 = \lambda^2-5\lambda+6 = (\lambda-2)(\lambda-3).$$

de manera que los eigenvalores de la matriz A son: $\lambda_1 = 2$ y $\lambda_2 = 3$. Los eigenespacios correspondientes son:
Para $\lambda_1 = 2$,

$$\left\{ \begin{bmatrix} x \\ y \end{bmatrix} : \begin{bmatrix} x \\ y \end{bmatrix} = t \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

de manera, que para t=1, obtenenemos el eigenvector

$$\mathbf{p}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Para $\lambda_2 = 3$,

$$\left\{ \begin{bmatrix} x \\ y \end{bmatrix} : \begin{bmatrix} x \\ y \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

de manera, que para t=1, obtenemos el eigenvector,

$$\mathbf{p}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Finalmente, de acuerdo con el teorema anterior, la solución general es,

$$\begin{aligned} \mathbf{y}(x) &= b_1\mathbf{p}_1 e^{\lambda_1 x} + b_2\mathbf{p}_2 e^{\lambda_2 x}, \\ &= b_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^{2x} + b_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{3x}. \end{aligned}$$

b) Si $\mathbf{y}(0) = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$ entonces,

$$\mathbf{y}(0) = b_1\mathbf{p}_1 + b_2\mathbf{p}_2 = b_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + b_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

La solución de este sistema de ecuaciones es, $b_1 = -3$, $b_2 = 6$, de modo que la solución particular es:

$$\mathbf{y}(x) = -3 \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^{2x} + 6 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{3x}.$$

Ejemplo 4.6.2 *Determine la solución general para el sistema lineal homogéneo*

$$\begin{bmatrix} y_1' \\ y_2' \\ y_3' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix},$$

Solución: Comencemos por determinar los eigenvalores de la matriz A.

$$P(\lambda) = \begin{vmatrix} \lambda - 1 & 0 & -2 \\ 1 & \lambda - 2 & -2 \\ 0 & 0 & \lambda - 3 \end{vmatrix} = (\lambda - 1)(\lambda - 2)(\lambda - 3) = 0.$$

De acuerdo con lo anterior, los eigevalores de A son: 1, 2, 3. Los eigenespacios correspondientes son:
Para $\lambda_1 = 1$,

$$\left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

Para $t = 1$, $\mathbf{p}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda_1 = 1$.

Para $\lambda_2 = 2$,

$$\left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

Para $t = 1$, $\mathbf{p}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda_2 = 2$.

Para $\lambda_3 = 3$,

$$\left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

Para $t = 1$, $\mathbf{p}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, es un eigenvector para el eigenvalor $\lambda_3 = 3$.

Por último, de acuerdo con el teorema anterior, la solución general es,

$$\begin{aligned} \mathbf{y}(x) &= b_1\mathbf{p}_1 e^{\lambda_1 x} + b_2\mathbf{p}_2 e^{\lambda_2 x} + b_3\mathbf{p}_3 e^{\lambda_3 x}, \\ &= b_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} e^{x} + b_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} e^{2x} + b_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} e^{3x}. \end{aligned}$$

4.7. Forma canónica de Jordan

En esta parte, a partir de la una matriz particular, se hará un pequeño bosquejo de lo que constituye el tema de la Forma canónica de Jordan. Consideremos la matriz

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

Como primer punto, tratemos de indagar si esta matriz es diagonalizable. Para tal fin, determinemos su polinomio característico. En este caso,

$$\begin{aligned} P(\lambda) &= |\lambda I - A| \\ &= \begin{bmatrix} \lambda & -3 & -1 \\ -2 & \lambda+1 & 1 \\ 2 & 1 & \lambda+1 \end{bmatrix} \\ &= (\lambda-2)(\lambda+2)^2 \end{aligned}$$

Los eigenvalores de la matriz A son:

$$\begin{cases} \lambda_1 = 2 \\ \lambda_2 = -2 \end{cases}$$

Para determinar los eigenespacios, debemos resolver la homogénea

$$(\lambda_i I - A)\mathbf{x} = \mathbf{0}.$$

Para $\lambda_1 = 2$ obtenemos,

$$V_{\lambda_1} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

Para $\lambda_2 = -2$ obtenemos,

$$V_{\lambda_2} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \quad \text{para cualquier escalar t} \right\}.$$

En este caso, la multiplicidad del eigenvalor λ_2, es diferente de la dimensión de su subespacio. Lo anterior implica que A no tiene tres eigenvectores linealmente independientes y, de acuerdo con el teorema 4.3.1, A no es una matriz diagonalizable. Ahora asignaremos al eigenvalor $\lambda_2 = -2$ un segundo eigenespacio que denotaremos por $V^*_{\lambda_2}$, donde $V^*_{\lambda_2} = ker(\lambda_2 I - A)^2$. Resolviendo la homogénea obtenemos,

$$V^*_{\lambda_2} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + r \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \quad \text{para cualquier r y t} \right\}.$$

En este caso $V^*_{\lambda_2}$ es un espacio de dimensión 2, además $V_{\lambda_2} \subset V^*_{\lambda_2}$ Deseamos determinar una base para $V^*_{\lambda_2}$ que contenga un vector de V_{λ_2}. Consideremos el vector $\mathbf{v}_3$, donde,

$$\mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

El vector anterior no se encuetra en V_{λ_2}; por lo tanto,

$$(\lambda_2 I - A)\mathbf{v}_3 \neq \mathbf{0}.$$

Definamos el vector $\mathbf{v}_2$ como $\mathbf{v}_2 = (A - \lambda_2 I)\mathbf{v}_3$, es decir,

$$\begin{aligned} \mathbf{v}_2 &= (A - \lambda_2 I)\mathbf{v}_3 \\ &= \begin{bmatrix} 2 & 3 & 1 \\ 2 & 1 & -1 \\ -2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \end{aligned}$$

Sea $S' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, donde

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

este conjunto constituye una base para $\mathbb{R}^3$.
Sea T la transformación lineal definida como $T(\mathbf{x}) = A(\mathbf{x})$, determinemos la representación de T, respecto a la base S'.
Por la forma en que hemos definido los vectores de la base S' tenemos que:

$$\begin{aligned} A\mathbf{v}_1 &= 2\mathbf{v}_1, \\ A\mathbf{v}_2 &= -2\mathbf{v}_2, \\ A\mathbf{v}_3 &= \mathbf{v}_2 + \lambda_2 \mathbf{v}_3, \end{aligned}$$

de manera que la representación de T, respecto a la base S' es:

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

Además, la matriz de transición de la base S' a la base S, donde S es la base estándar, está determinada por:

$$P = \begin{bmatrix} -1 & 1 & 0 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix},$$

Las matrices A y J están relacionadas mediante, $J = P^{-1}AP$. La matriz J no es una matriz diagonal; sin embargo, se encuentra en una forma más "simple"que la matriz A, en el sentido de que es más sencillo realizar operaciones con esta matriz que con la matriz A. Se dice que la matriz J representa la forma canónica de Jordan la matriz A. En general, la forma canónica de Jordan es una generalización del problema de diagonalización; sin embargo, su desarrollo teórico queda fuera del alcance de un primer curso de álgebra lineal. Quien así lo desee, encontrará un desarrollo formal del tema en la referencia [**?**]

Bibliografía

[1] Stanley I. Grossman & José Job Flores Godoy, Álgebra Lineal, 7a. Edición, Editorial Mc Graw-Hill, México (2012).

[2] Howard Anton, Introducción al Álgebra Líneal, 4a. Edición, Editorial Limusa, México (2010).

[3] Fernando Barrera Mora, Álgebra Líneal, 1a. Edición, Editorial Patria, México (2007).

[4] Bernard Kolman & David R. Hill, Álgebra Líneal, 8a. Edición, Editorial Pearson, México (2006).

[5] S. Lipschutz, Álgebra Líneal, 2a. Edición, Editorial Mc Graw-Hill, México (1992).

[6] K. Hoffman & R. Kunze, Álgebra Líneal, 2a. Edición, Editorial Prentice-Hall, México (1973).

[7] Frank Ayres, Jr., Modern Abstract Algebra, Editorial Mc Graw-Hill, USA (1965).

Apéndice

Apéndice A

Permutaciones

A.1. Permutaciones

El estudio de las permutaciones es un tema fundamental para el desarrollo de los determinantes. Para comenzar el estudio de las permutaciones es necesario recurrir al tema de las aplicaciones. En esta parte solo se enuncian las definiciones y se prueban algunos resultados generales, sin adentrarse en ejemplos concretos más allá de las propias permutaciones.

A.1.1. Definiciones básicas

Definición A.1.1 (Mapeo o aplicación) *Sean A y B dos conjuntos cualesquiera, entonces diremos que una aplicación o un mapeo de A en B es una asociación o correspondencia que asigna a cada elemento del conjunto A un único elemento del conjunto B. El único elemento de B que se asocia con cualquier elemento de A se conoce como imagen del elemeto de A en el mapeo.*

Denotaremos un mapeo α del conjunto A en el conjunto B, mediante:

$$\alpha : A \to B.$$

Además, si $b \in B$ es la imagen de $a \in A$ bajo α escribiremos, $b = \alpha(a)$.

Definición A.1.2 (Mapeo uno a uno)

Un mapeo α de un conjunto A en un conjunto B se conoce como mapeo uno a uno de A en B si las imágenes de distintos elementos de A corresponden a distintos elementos de B.

Definición A.1.3 (Mapeo sobre) *Si en un mapeo de un conjunto A en un conjunto B cada elemento del conjunto B es una imagen, diremos que el mapeo es de A sobre B. Si algún elemento de B no es una imagen, entonces diremos que el mapeo es de A en B, pero no de A sobre B.*

Definición A.1.4 (Mapeo uno a uno y sobre) *Si un mapeo α de un cojunto A un conjunto B satisface las dos definiciones anteriores, diremos entonces que α es un mapeo uno a uno de A sobre B.*

Definición A.1.5 (Mapeo identidad) *Dado cualquier conjuto A, se conoce como mapeo identidad ϵ, a la aplicación uno a uno de A sobre A, definida como,*

$$\epsilon(a) = a, \quad \forall a \in A.$$

En general un mapeo de un conjunto A en un conjunto B es un subconjunto de $A \times B$[1] donde cada elemento de A ocurre una y solamente una vez como la primer componente en los elementos del subconjunto.

En cualquier mapeo α de un conjunto A en un conjunto B, el conjunto A se conoce como dominio y el conjunto B como codominio de α. Si el mapeo α es de A sobre B, entonces B se conoce como rango de α, en otro caso, el rango es un subconjunto propio de B que consiste de las imágenes de todos los elementos de A.

Definición A.1.6 (Mapeos equivalentes o iguales) *Dieremos que dos mapeos α y β son iguales o equivalentes si estos mapeos tienen el mismo dominio A y si $\alpha(a) = \beta(a)$ para todo $a \in A$.*

Definición A.1.7 (Producto o composición de aplicaciones) *Sea α un mapeo de A en B y sea β un mapeo de B en C, sabemos que el efecto de α es mapear cada elemento $a \in A$ en $\alpha(a) \in B$ y el efecto de B es mapear cada $\alpha(a) \in B$ en $\beta(\alpha(a)) \in C$. El resultado final de aplicar α seguido de β es un mapeo de A en C definido por,*

$$(\beta \circ \alpha)(a) = \beta(\alpha(a)), \quad \forall a \in A,$$

$\beta \circ \alpha$ se conoce como producto o composición de los mapeos α y β en este orden.

[1]Este conjunto se conoce como conjunto producto y consta de todas las parejas ordenadas (a,b) tales que $a \in A$, $b \in B$.

Teorema A.1.1 *Si α es una aplicación de A en B, entonces*

$$\begin{aligned} \epsilon_B \circ \alpha &= \alpha \\ \alpha \circ \epsilon_A &= \alpha \end{aligned}$$

donde ϵ_A y ϵ_B son las aplicaciones identidad en A y B, respectivamente.

Demostración: a) Como α es una aplicación de A en B y ϵ_B es una aplicación de B en B, entonces, $\epsilon_B \circ \alpha$ es una aplicación de A en B. Además, si $\alpha(a) = b$, entonces

$$\begin{aligned} (\epsilon_B \circ \alpha)(a) &= \epsilon_B(\alpha(a)), \\ &= \epsilon_B(b), \\ &= b, \\ &= \alpha(a). \end{aligned}$$

Por lo tanto, se concluye que, $\epsilon_B \circ \alpha = \alpha$.

b) De manera análoga al inciso anterior, como α es una aplicación de A en B y ϵ_A es una aplicación de A en A, entonces, $\alpha \circ \epsilon_A$ es una aplicación de A en B, de manera que si $\alpha(a) = b$, entonces,

$$\begin{aligned} (\alpha \circ \epsilon_A)(a) &= \alpha(\epsilon_A(a)), \\ &= \alpha(a). \end{aligned}$$

Concluimos que, $\alpha \circ \epsilon_A = \alpha$. ∎

Teorema A.1.2 (Asociatividad de las composiciones) *Consideremos las aplicaciones α, β y γ tales que α mapea A en B, β mapea B en C y γ mapea C en D, entonces*

$$(\gamma \circ \beta) \circ \alpha = \gamma \circ (\beta \circ \alpha). \tag{A.1}$$

Demostración: Verifiquemos que las aplicaciones que aparecen en la ecuación A.1 estén permitidas. Las aplicaciones están definidas de la siguiente manera:

$$\alpha : A \to B,$$

$$\beta : B \to C,$$

$$\gamma : C \to D.$$

Además para las composiciones se tiene,

$$\gamma \circ \beta : B \to D,$$

$$\beta \circ \alpha : A \to C.$$

de manera que,

$$(\gamma \circ \beta) \circ \alpha : A \to D, \tag{A.2}$$

y

$$\gamma \circ (\beta \circ \alpha) : A \to D. \tag{A.3}$$

Las ecuaciones A.2 y A.3 muestran que las composiciones involucradas en la ecuación A.1 tienen el mismo dominio. Verifiquemos ahora que los mapeos de elementos iguales son los mismos.

De la definición de composición tenemos,

$$\forall a \in A,$$

$$\begin{aligned} ((\gamma \circ \beta) \circ \alpha)(a) &= (\gamma \circ \beta)(\alpha(a)), \\ &= \gamma(\beta(\alpha(a))). \end{aligned} \tag{A.4}$$

Procediendo de la misma manera,

$$\begin{aligned} (\gamma \circ (\beta \circ \alpha))(a) &= \gamma((\beta \circ \alpha)(a)), \\ &= \gamma(\beta(\alpha(a))). \end{aligned} \tag{A.5}$$

Comparando las ecuaciones A.4 y A.5 se concluye la demostración. ∎

Definición A.1.8 (Mapeo inverso) *Si α es una aplicación de A en B y si existe una aplicación de β de B en A, tal que*

$$\left\{ \begin{aligned} \beta \circ \alpha &= \epsilon_A, \\ \alpha \circ \beta &= \epsilon_B, \end{aligned} \right. \tag{A.6}$$

donde ϵ_A y ϵ_B son las aplicaciones identidad en A y B, respectivamente, entonces decimos que α es invertible y que β es su inversa.

Teorema A.1.3 *Si α es un mapeo uno a uno de un conjunto A sobre un conjunto B, entonces α tiene una inversa única, e inversamente.*

Demostración: $\Rightarrow$) Como α es uno a uno y sobre, entonces si $a \in A$, existe un único $b \in B$ tal que, $\alpha(a) = b$. Podemos entonces asociar cada elemento $b \in B$ con un único elemento $a \in A$. Es decir, α induce una aplicación uno a uno y sobre que mapea B sobre A. Denotaremos este mapeo como β y verificaremos que satisface la definición anterior. Sea,

$$\beta : B \to A,$$

y

$$\beta(b) = a, \text{ para todo } b \in B.$$

Ahora consideremos las composiciones $\alpha \circ \beta$ y $\beta \circ \alpha$.

$$\alpha \circ \beta : B \to B,$$

$$\beta \circ \alpha : A \to A,$$

Determinemos las composiciones:

$$\begin{aligned}(\alpha \circ \beta)(b) &= \alpha(\beta(b)),\\ &= \alpha(a),\\ &= b.\end{aligned}$$

Por lo anterior, $\alpha \circ \beta = \epsilon_B$.
Ahora consideremos,

$$\begin{aligned}(\beta \circ \alpha)(a) &= \beta(\alpha(a)),\\ &= \beta(b),\\ &= a.\end{aligned}$$

Se observa que, $\beta \circ \alpha = \epsilon_A$.
En conclusión,

$$\begin{cases} \beta \circ \alpha = \epsilon_A \\ \alpha \circ \beta = \epsilon_B \end{cases}$$

Ahora demostraremos que la aplicación inversa es única, para lo cual supondremos que existe otra aplicación γ que satisface la ecuación A.6; es decir,

$$\gamma \circ \alpha = \epsilon_A,$$

y

$$\alpha \circ \gamma = \epsilon_B.$$

Consideremos la segunda ecuación y formemos la composición $(\gamma \circ \alpha) \circ \beta$; es decir,

$$\begin{aligned} (\gamma \circ \alpha) \circ \beta &= \epsilon_A \circ \beta, \\ \gamma \circ (\alpha \circ \beta) &= \beta, \\ \gamma \circ \epsilon_B &= \beta, \\ \gamma &= \beta. \end{aligned}$$

Concluimos que la inversa es única. De aquí en adelante la denotaremos como α^{-1}. Además α^{-1} satisface las ecuaciones:

$$\begin{cases} \alpha^{-1} \circ \alpha = \epsilon_A, \\ \alpha \circ \alpha^{-1} = \epsilon_B. \end{cases}$$

$\Leftarrow$) Supongamos ahora que α tiene una inversa única, nuestro propósito es demostrar que α es uno a uno y sobre.

i) En esta parte probaremos que α es uno o uno. Supongamos para tal efecto que existen a_1 y $a_2 \in A$ tales que, $a_1 \neq a_2$ y además,

$$\alpha(a_1) = \alpha(a_2).$$

En este caso α no sería un mapeo uno a uno de A en B; sin embargo, aplicando α^{-1} a la ecuación anterior obtenemos,

$$\begin{aligned} \alpha(a_1) &= \alpha(a_2), \\ \alpha^{-1}(\alpha(a_1)) &= \alpha^{-1}(\alpha(a_2)), \\ (\alpha^{-1} \circ \alpha)(a_1) &= (\alpha^{-1} \circ \alpha)(a_2), \\ \epsilon_A(a_1) &= \epsilon_A(a_2), \\ a_1 &= a_2. \end{aligned}$$

Por lo tanto, existe una contradicción, la cual surge de suponer que α no era una aplicación uno a uno de A en B. En conclusión α es una aplicación uno a uno de A en B.

ii) Demostraremos en esta parte que α es una aplicación de A sobre B. Consideremos un elemento $b \in B$, luego,

$$\begin{aligned} b &= \epsilon_B b, \\ b &= (\alpha \circ \alpha^{-1})(b), \\ b &= \alpha(\alpha^{-1}(b)). \end{aligned}$$

La última ecuación muestra que existe un $a \in A$ tal que $\alpha(a) = b$; por lo tanto, b es imagen de a bajo α. Como esto es cierto para todo $b \in B$, entonces B es una aplicación de A sobre B. ∎

Teorema A.1.4 *Si α es un mapeo uno a uno de un conjunto A sobre un conjunto B y β es un mapeo uno a uno del conjunto B sobre un conjunto C, entonces $(\beta \circ \alpha)$ es invertible y su inversa es $(\beta \circ \alpha)^{-1} = \alpha^{-1} \circ \beta^{-1}$.*

Demostración: Como α y β son aplicaciones uno a uno y sobre, entonces tienen inversas tales que,

$$\alpha^{-1} : B \to A,$$
$$\beta^{-1} : C \to B,$$

por lo anterior,

$$\alpha^{-1} \circ \beta^{-1} : C \to A,$$

por otro lado,

$$\beta \circ \alpha : A \to C,$$

además,

$$(\alpha^{-1} \circ \beta^{-1}) \circ (\beta \circ \alpha) : A \to A.$$

Consideremos entonces la última composición y utilicemos la propiedad asociativa para desarrollarla.

$$\begin{aligned}
(\alpha^{-1} \circ \beta^{-1}) \circ (\beta \circ \alpha) &= \alpha^{-1} \circ (\beta^{-1} \circ (\beta \circ \alpha)), \\
&= \alpha^{-1} \circ ((\beta^{-1} \circ \beta) \circ \alpha), \\
&= \alpha^{-1} \circ (\epsilon_B \circ \alpha), \\
&= \alpha^{-1} \circ \alpha, \\
&= \epsilon_A.
\end{aligned}$$

De manera análoga, consideremos la composición

$$(\beta \circ \alpha) \circ (\alpha^{-1} \circ \beta^{-1}) : C \to C.$$

$$\begin{aligned}
(\beta \circ \alpha) \circ (\alpha^{-1} \circ \beta^{-1}) &= \beta \circ (\alpha \circ (\alpha^{-1} \circ \beta^{-1})), \\
&= \beta \circ ((\alpha \circ \alpha^{-1}) \circ \beta^{-1})), \\
&= \beta \circ (\epsilon_B \circ \beta^{-1}) \\
&= \beta \circ \beta^{-1} \\
&= \epsilon_C.
\end{aligned}$$

Con esto hemos demostrado, de acuerdo con la definición, que $\beta \circ \alpha$ es invertible y por el teorema A.1.3 sabemos que la inversa es única, de manera que:

$$(\beta \circ \alpha)^{-1} = \alpha^{-1} \circ \beta^{-1}. \quad \blacksquare$$

Definición A.1.9 (Permutaciones) *Una permutación del conjunto de números enteros* $\{1, 2, 3, \ldots, n\}$ *es un mapeo uno a uno y sobre sí mismo. Denotaremos una permutación mediante,*

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix},$$

o simplemente,

$$\sigma = j_1 j_2 \ldots j_n = \sigma(1)\sigma(2) \ldots \sigma(n),$$

donde, $\sigma(i) = j_i$.

Dado que σ es uno a uno y sobre, entonces la secuencia $j_1 j_2 \ldots j_n$ es un arreglo de los números $1, 2, 3, \ldots, n$. Denotaremos al conjunto de todas las permutaciones σ mediante S_n. Ahora, si $\sigma \in S_n$ el mapeo inverso $\sigma^{-1} \in S_n$, y, si σ y $\tau \in S_n$, es decir, son arreglos de los números $\{1, 2, \ldots, n\}$, entonces $\tau \circ \sigma$ es un rearreglo de un arreglo, lo que constituye un arreglo en sí mismo de los números $\{1, 2, \ldots, n\}$, por lo que, $\sigma \circ \tau \in S_n$. En partiular la permutación identidad, $\epsilon = \sigma^{-1} \circ \sigma \in S_n$, donde $\epsilon = 12, \ldots, n$.

Ejemplo A.1.1 *Determine todas las permutaciones* σ *de* S_3.

Solución: Un método sencillo para determinar las permutaciones consiste en emplear un diagrama de árbol, como se muestra a continuación,

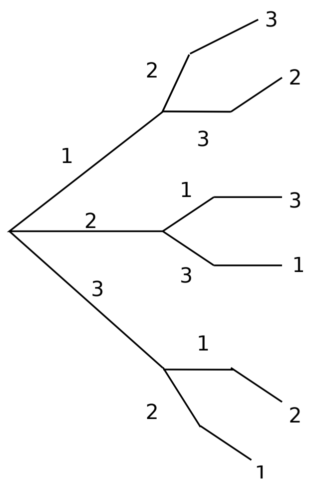

Figura A.1: Diagrama de árbol

Las permutaciones se forman iniciando en el punto común hasta llegar al final o rama del árbol. Con lo anterior obtenemos las siguientes permutaciones, $123, 132, 213, 231, 312, 321$. En este caso el número total de permutaciones es 6.

En general, si se tiene el conjunto $\{1, 2, \ldots, n\}$ y se desea hallar el número de permutaciones de tamaño n (tomados n elementos a la vez), podemos proceder de la siguiente manera. Construya una línea con n cuadros, como se muestra a continuación.

$$\boxed{\text{n}}\ \boxed{\text{n-1}}\boxed{\text{n-2}}\ \cdots\ \boxed{\text{1}}$$

Figura A.2: Línea de cuadros para contar las permutaciones

En el primer cuadro se coloca el número de maneras de escoger el primer elemento de la permutación; es decir, n. En el segundo cuadro se coloca el numero de maneras de escoger el segundo elemento de la pernmutación; es decir, n-1, ya que el elemento usado para llenar el primer cuadro ya no puede volver a usarse en la permutación. El tercer cuadro puede llenarse de n-2 manera y así sucesivamente el n-ésimo cuadro puede llenarse de una sola manera. Por el principio fundamental de conteo[2] se tienen,

$$_nP_n = n(n-1)(n-2)\ldots 1,$$

permutaciones diferentes de tamaño n.

El resultado anterior se expresa de manera cotidiana en términos de la función factorial,

$$_nP_n = n!$$

Ejemplo A.1.2 *Sean $\sigma = 132$ y $\tau = 321$ permutaciones de S_3 hallar $\sigma \circ \tau$ y $\tau \circ \sigma$*

Solución: De la definición de permutación tenemos que,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

y

$$\tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

De lo anterior se tiene que,

$$\sigma(1) = 1,\ \sigma(2) = 3,\ \sigma(3) = 2,$$

[2]Basado en el diagrama de árbol: Si un proceso puede realizarse de n_1 maneras diferentes, un segundo proceso puede realizarse de n_2 maneras diferentes y así sucesivamente un k-ésimo proceso puede realizarse de n_k maneras diferentes, entonces el proceso compuesto en la secuencua especificada puede efectuarse de $n_1 x n_2 x \ldots x n_k$ maneras diferentes.

y

$$\tau(1) = 3,\ \tau(2) = 2,\ \tau(3) = 1.$$

Determinemos la permutación, $\sigma \circ \tau$.

$$\begin{array}{cccc} & 1 & 2 & 3 \\ \sigma & \downarrow & \downarrow & \downarrow \\ & 1 & 3 & 2 \\ \tau & \downarrow & \downarrow & \downarrow \\ & 3 & 1 & 2 \end{array}$$

De lo cual concluimos que $\sigma \circ \tau = 312$.

Ahora consideremos la permutación $\tau \circ \sigma$.

$$\begin{array}{cccc} & 1 & 2 & 3 \\ \tau & \downarrow & \downarrow & \downarrow \\ & 3 & 2 & 1 \\ \sigma & \downarrow & \downarrow & \downarrow \\ & 2 & 3 & 1 \end{array}$$

Se concluye que $\tau \circ \sigma = 231$.

Ejemplo A.1.3 *Sean $\sigma = 231$ y $\tau = 312$ permutaciones de S_3, halle las permutaciones inversas correspondientes.*

Solución: Sabemos que, $\sigma \circ \sigma^{-1} = \epsilon$ y de la misma manera $\tau \circ \tau^{-1} = \epsilon$, donde $\epsilon = 123$. Procediendo como en el ejemplo anterior se tiene,

$$\begin{array}{cccc} & 1 & 2 & 3 \\ \sigma & \downarrow & \downarrow & \downarrow \\ & 2 & 3 & 1 \\ \sigma^{-1} & \downarrow & \downarrow & \downarrow \\ & 1 & 2 & 3 \end{array}$$

se concluye que,

$$\sigma^{-1}(1) = 3,\ \sigma^{-1}(2) = 1,\ \sigma^{-1}(3) = 2,$$

por lo tanto,

$$\sigma^{-1} = 312.$$

Consideremos ahora la permutación τ,

$$\begin{array}{cccc} & 1 & 2 & 3 \\ \tau & \downarrow & \downarrow & \downarrow \\ & 3 & 1 & 2 \\ \tau^{-1} & \downarrow & \downarrow & \downarrow \\ & 1 & 2 & 3 \end{array}$$

$$\tau^{-1}(1) = 2,\ \tau^{-1}(2) = 3,\ \tau^{-1}(3) = 1,$$

por lo tanto,

$$\tau^{-1} = 231.$$

Definición A.1.10 (Inversión) *Decimos que ocurre una inversión en una permutación* $(j_1, j_2, \ldots, j_n)$*, siempre que un entero mayor precede a uno menor.*

Para determinar el número total de inversiones en una permutación, se procede de la siguiente manera:

- Determine el número de enteros menores a j_1 y que siguen a j_1 en la permutación.
- Determine el número de enteros menores a j_2 y que siguen a j_2 en la permutación.
- Se continúa con al proceso para $j_3, j_4, \ldots, j_{(n-1)}$.
- La suma de estos números será el número total de inversiones en la permutación.

Ejemplo A.1.4 *Determine el número de inversiones en la permutación* $\sigma = 53241$.

Solución:
Se tienen 4 números menores a 5 y que le siguen en la permutación.
Se tienen 2 números menores que 3 y que le siguen en la permutación.
Se tiene 1 número menor que 2 y que le siguen en la permutación.
Se tiene 1 número menor que 4 y que le siguen en la permutación.
Finalmente, el total de inversiones para la permutación es: 4+2+1+1+1=8.

Definición A.1.11 (Paridad de una permutación) *Se dice que una permutación es par si el número total de inversiones es un entero par, en caso contrario se dice que la permutación es impar.*

Ejemplo A.1.5 *Clasifique las diversas permutaciones de* $\{1, 2, 3\}$ *en pares e impares.*

Solución: En la siguiente tabla se muestran las permutaciones y sus propiedades.

Tabla A.1: Paridad de las permutaciones del conjunto {1,2,3}

Permutación	No. de inversiones	Paridad
123	0	par
132	1	impar
213	1	impar
231	2	par
312	2	par
321	3	impar

Definición A.1.12 (Signo de una permutación) *Dada una permutación $\sigma = j_1 j_2 \ldots j_n$ de los enteros $\{1, 2, \ldots, n\}$, definimos el signo de la permutación, denotado por $sgn(\sigma)$ como,*

$$sgn(\sigma) = \begin{cases} +1, & si\ \sigma\ es\ par. \\ -1, & si\ \sigma\ es\ impar. \end{cases}$$

Definición A.1.13 (Transposición) *Una permutación que solo intercambia un par de índices, manteniendo el resto sin cambios, se conoce como transposición*

A.1.2. Propiedades de las permutaciones

Teorema A.1.5 *Toda transposición τ es una permutación impar; es decir, $sgn(\tau) = -1$.*

Demostración: Supongamos que $i > j$ y sea τ la permutación que intercambia los índices i y j mientras que deja sin cambio el resto; es decir,

$$\tau(i) = j, \qquad \tau(j) = i, \qquad \tau(k) = k, \quad \forall k \neq i, j.$$

Por lo anterior,

$$\tau = 12 \ldots (i-1)j(i+1) \ldots (j-1)i(j+1) \ldots n.$$

Contemos las inversiones que ocurren en la permutación.

- Desde 1 hasta $(i-1)$ no existen inversiones.
- Los enteros $\{i+1, i+2, \ldots, j-1\}$ son menores a j y le siguen en la permutación; por lo tanto, para j tenemos $(j-i)$ inversiones.

- Para cada entero desde $(i{+}1)$ hasta $(j{-}1)$ tenemos un entero menor que le sigue en la permutación(este entero es i); es decir, tenemos $j-i+1$ inversiones de este tipo. Los enteros mayores que i que le anteceden en la permutación son $\{i+1, i+2, \ldots, j-1\}$.
- Para i no tenemos inversiones ya que todos los enteros que le siguen en la permutación son mayores.
- Desde $(j+1)$ hasta n no existen inversiones.

Por lo anterior tenemos en total $2(j-i)+1$ inversiones, lo que representa un número impar de permutaciones. ∎

Teorema A.1.6 *Sea τ una permutación de S_n, entonces*

$$S_n = \{\tau \circ \sigma : \sigma \in S_n.\}$$

Esto significa que si σ corre sobre S_n, entonces $\tau \circ \sigma$ también corre sobre S_n

Demostración: Sea $S^* = \{\tau \circ \sigma \ : \sigma \in S_n\}$ probaremos que $S^* = S_n$.

$S^* \subseteq S_n$) Si $\tau \in S_n$ y σ_i representa cualquiera de las $n!$ permutaciones de S_n, entonces $\tau \circ \sigma_i \in S_n$. Por lo que, $S^* \subset S_n$.

$S_n \subseteq S^*$) Si σ_i representa cualquiera de las $n!$ permutaciones de S_n y $\tau \in S_n$, entonces $\tau^{-1} \in S_n$, además, $\sigma_k = \tau^{-1} \circ \sigma_i \in S_n$. Por otro lado, $\sigma_i = \epsilon \circ \sigma_i$, además usando el teorema A.1.2 tenemos,

$$\begin{aligned}
\sigma_i &= \epsilon \circ \sigma_i, \\
&= (\tau \circ \tau^{-1}) \circ \sigma_i, \\
&= \tau \circ (\tau^{-1} \circ \sigma_i), \\
&= \tau \circ \sigma_k.
\end{aligned}$$

Por lo anterior, $\sigma_i \in S^*$, por lo tanto, $S_n \subset S^*$

En conclusión, como $S^* \subseteq S_n$ y $S_n \subseteq S^*$, entonces $S_n = S^*$. ∎

Parejas de inversión en una permutación

Consideremos la siguiente permutación de S_5.

$$\sigma = 53241.$$

Observemos las parejas (i, k), tales que $i > k$, con la condición de que i precede a k en la permutación. Las parejas que satisfacen la condición anterior son: $(5,3)(5,2)(5,4)(5,1)(3,2)(3,1)(2,1)(4,1)$. Es claro que el número de parejas que cumplen la condición anterior determina la paridad de la permutación. Consideremos la pareja (5,3), por ejemplo, para esta pareja se tiene que $\sigma(1) = 5$ y $\sigma(2) = 3$; es decir, la pareja (1,2) satisface la condición de que $1 < 2$ y $\sigma(1) > \sigma(2)$. En general esto se cumple para cada pareja y para cada permutación. El resultado general se establece de la siguiente manera:

Teorema A.1.7 *Si $\sigma = \sigma(1)\sigma(2)\ldots\sigma(n)$ es cualquier permutación de S_n, entonces para cada pareja (i, k) tal que $i > k$, donde i precede a k en la permutación, existe una pareja (i', k') tal que,*

$$i' < k' \quad y \quad \sigma(i') > \sigma(k') \tag{A.7}$$

y también viceversa. De esta manera σ es par o impar dependiendo del número de parejas que satisfacen la ecuación A.7

Demostración: Sean i' y k' tales que $\sigma(i') = i$ y $\sigma(k') = k$, entonces $i > k$ si y solo si $\sigma(i') > \sigma(k')$, además i precede a k en la permutacion si y solo si $i' < k'$. ∎

Teorema A.1.8 *Sea $\sigma \in S_n$ con la propiedad de que $\sigma(n) = n$ y sea $\sigma^* \in S_{n-1}$ definida como $\sigma^*(x) = \sigma(x)$, entonces*

i) $sgn(\sigma^*) = sgn(\sigma)$

ii) Si σ corre sobre de todas las permutaciones de S_n donde $\sigma(n) = n$, entonces σ^* corre sobre de todas las permutaciones de S_{n-1}.

Demostración: i) Consideremos una permutación de S_n para la cual $\sigma(n) = j_n = n$; es decir,

$$\sigma = j_1 j_2 \ldots j_{n-1} n,$$

en este caso σ^* es

$$\sigma^* = j_1 j_2 \ldots j_{n-1}.$$

Para hallar las inversiones de σ^* debemos contar el número de parejas (i, k) con $i < k$ para las cuales $j_i > j_k$. Sean l_i con $i = 1, 2, \ldots (n-1)$ los números de parejas que satisfacen la condición para cada j_i respectivamente, entonces el número total de inversiones para σ^* es

$$l_1 + l_2 + \ldots + l_{n-1}$$

Para hallar las inversiones de σ debemos repetir el proceso de contar las parejas de inversión para $j_1, j_2, \ldots, n$; sin embargo, observamos que

$j_n = n$ no forma pareja de inversión con ningún j_i. Por lo anterior, el número de inversiones de σ es

$$l_1 + l_2 + \ldots + l_{n-1}.$$

Con esto concluimos que σ y σ^* tienen la misma paridad, por lo tanto,

$$sgn(\sigma) = sgn(\sigma^*)$$

ii) Consideremos ahora las permutaciones de S_n: $\sigma = j_1 j_2 \ldots j_{n-1} j_n$. El número total de permutaciones es n! Ahora si imponemos la condición de que $j_n = n$, entonces el número total de permutaciones de este tipo es (n-1)!, ya que el último elemento permanece fijo y solo se pueden permutar los (n-1) elementos restantes, pero estos elementos que pueden permutarse corresponde a los enteros $\{1, 2, \ldots (n-1)\}$, como bien sabemos, las permutaciones de estos números corresponden a S_{n-1}; por lo tanto, cuando $\sigma \in S_n$ corre sobre todas las permutaciones de S_n con $j_n = n$, entonces σ^* corre sobre todas las permutaciones de S_{n-1}. ∎

Definición A.1.14 (La función polinomial g) *Consideremos la función polinomial g definida de la siguiente manera*

$$g = g(x_1, x_2, \ldots, x_n) = \prod_{i<j} (x_i - x_j). \tag{A.8}$$

Donde el símbolo Π indica la multiplicación de los factores permitidos; es decir, aquellos que satisfagan la condición $i < j$. Desarrollemos algunos casos particulares de la función g.

1. Si n=3, entonces

$$g = g(x_1, x_2, x_3) = \prod_{i<j} (x_i - x_j) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3).$$

2. Si n=4, entonces

$$\begin{aligned} g &= g(x_1, x_2, x_3, x_4) \\ &= \prod_{i<j} (x_i - x_j) \\ &= (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4) \end{aligned}$$

Ahora, sea $\sigma = \sigma(1)\sigma(2)\ldots\sigma(n)$ una permutación de S_n, definimos la función $\sigma(g)$ de la siguiente manera,

$$\sigma(g) = \prod_{i<j} (x_{\sigma(i)} - x_{\sigma(j)}). \tag{A.9}$$

Desarrollemos algunos ejemplos de la función $\sigma(g)$ para algunos casos particulares,

1. Consideremos la permutación $\sigma = \sigma(1)\sigma(2)\sigma(3) = 321$

$$\begin{aligned}
\sigma(g) &= \prod_{i<j}(x_{\sigma(i)} - x_{\sigma(j)}) \\
&= (x_{\sigma(1)} - x_{\sigma(2)})(x_{\sigma(1)} - x_{\sigma(3)})(x_{\sigma(2)} - x_{\sigma(3)}) \\
&= (x_3 - x_2)(x_3 - x_1)(x_2 - x_1) \\
&= (-1)^3(x_1 - x_2)(x_1 - x_3)(x_2 - x_3) \\
&= -g(x_1, x_2, x_3)
\end{aligned}$$

2. Ahora consideremos la permutación $\sigma = \sigma(1)\sigma(2)\sigma(3) = 231$

$$\begin{aligned}
\sigma(g) &= \prod_{i<j}(x_{\sigma(i)} - x_{\sigma(j)}) \\
&= (x_{\sigma(1)} - x_{\sigma(2)})(x_{\sigma(1)} - x_{\sigma(3)})(x_{\sigma(2)} - x_{\sigma(3)}) \\
&= (x_2 - x_3)(x_2 - x_1)(x_3 - x_1) \\
&= (-1)^2(x_1 - x_2)(x_1 - x_3)(x_2 - x_3) \\
&= g(x_1, x_2, x_3)
\end{aligned}$$

En esta forma es fácil apreciar que $\sigma(g) = g$ o $\sigma(g) = -g$, de acuerdo a si tenemos un número par o impar de términos $x_i - x_j$ donde, $i > j$. Lo anterior se enuncia de manera general en el siguiente teorema.

Teorema A.1.9 *Si σ una permutación de S_n, entonces $\sigma(g) = sgn(\sigma)g$.*

Demostración: En general, debido a que σ es uno a uno y sobre, entonces

$$\sigma(g) = \prod_{i<j}(x_{\sigma(i)} - x_{\sigma(j)}) = \prod_{i<j \ o \ i>j}(x_i - x_j),$$

entonces $\sigma(g) = g$ o $\sigma(g) = -g$ dependiendo de si el número de factores de la forma $(x_i - x_j)$ con $i > j$ es par o impar respectivamente. De acuerdo con el teorema A.1.7, se tiene que,

$$\text{para cada } i < j \text{ tal que } \sigma(i) > \sigma(j), \tag{A.10}$$

habrá un factor $(x_{\sigma(i)} - x_{\sigma(i)})$ en $\sigma(g)$ para el cual $\sigma(i) > \sigma(j)$ y por lo tanto, $\sigma(g) = g$ si y solo si tenemos un número par de parejas que que satisfagan

A.10, es decir, $\sigma(g) = g$ si solo si, σ es una parmutación par. De la misma manera, $\sigma(g) = -g$ si y solo si tenemos un número impar de parejas que que satisfagan A.10, lo que se cumple si y solo si, σ es una permutación impar. En conclusión, $\sigma(g) = sgn(\sigma)g$. ∎

Teorema A.1.10 *Sean σ y τ permutaciones de S_n, entonces $sgn(\sigma \circ \tau) = sgn(\sigma)sgn(\tau)$*

Demostración: Por el teorema anterior se tiene ,

$$sgn(\sigma \circ \tau)g = (\sigma \circ \tau)(g) = \sigma(\tau(g)) = \sigma(sgn(\tau)g) = sgn(\sigma)sgn(\tau)g.$$

De lo anterior se concluye que,

$$sgn(\sigma \circ \tau) = sgn(\sigma)sgn(\tau). \quad \blacksquare$$

Teorema A.1.11 *Si σ es una permutación de S_n y σ^{-1} es la permutación inversa de σ, entonces*

i) $sgn(\sigma) = sgn(\sigma^{-1})$.

ii) Para escalares a_{ij} se cumple que, $a_{1j_1}a_{2j_2}\dots a_{nj_n} = a_{k_1 1}a_{k_1 2}\dots a_{k_n n}$, donde $\sigma^{-1} = k_1k_2\dots k_n$.

Demostración: i) Sabemos que $\sigma \circ \sigma^{-1} = \epsilon$, donde ϵ es la permutación identidad. Además ϵ es una permutación par. Por el teorema anterior,

$$\begin{aligned} sgn(\sigma)sgn(\sigma^{-1}) &= sgn(\sigma \circ \sigma^{-1}), \\ &= sgn(\epsilon), \\ &= 1. \end{aligned}$$

Concluimos que $sgn(\sigma) = sgn(\tau)$. De otro modo el producto de los signos sería negativo.

ii) Como $\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix} = \sigma(1)\sigma(2)\dots\sigma(n)$, entonces existen escalares $k_1k_2\dots k_n$ tales que,

$$\sigma(k_1) = 1, \sigma(k_2) = 2, \dots, \sigma(k_n) = n,$$

De tal manera que,

$$a_{1j_1}a_{2j_2}\dots a_{nj_n} = a_{k_1 1}a_{k_1 2}\dots a_{k_n n}.$$

Ahora sea τ la permutación defida como, $\tau = \begin{pmatrix} 1 & 2 & \dots & n \\ k_1 & k_2 & \dots & k_n \end{pmatrix}$, entonces

$$(\sigma \circ \tau)(i) = \sigma(\tau(i)) = \sigma(k_i) = i,$$

es decir, $\sigma \circ \tau = \epsilon$, por lo tanto, $\tau = \sigma^{-1}$ ∎.