Álgebra Lineal

Pablo V. Negrón Marrero

Universidad de Puerto Rico
Departamento de Matemáticas
Humacao, PR 00791

Tabla de Contenido

Prefacio

El álgebra lineal es la rama de las matemáticas que estudia conceptos tales como vectores, matrices, sistemas de ecuaciones lineales y en un enfoque más formal, espacios vectoriales, y transformaciones lineales. Es un área de la matemática bien activa y que tiene conexiones con muchas áreas dentro y fuera de las matemáticas como el análisis funcional, las ecuaciones diferenciales, investigación operativa, los gráficos por computadora, la ingeniería, etc.

El estudio de la caracterización de las soluciones de un sistema de ecuaciones lineales es central a la disciplina del álgebra lineal. Al plantearnos la pregunta de si un sistema tiene solución o no, lo podemos hacer tomando uno de tres enfoques:

i) visualizando cada ecuación como la de un "hiperplano" en el espacio, y buscamos o calculamos la intersección de estos hiperplanos. Este punto de visto lo llamamos *geométrico*.

ii) El sistema se puede visualizar como una ecuación matricial y podemos utilizar toda la maquinaria de matrices para estudiar la existencia y multiplidad de soluciones. Podemos tambien, utilizando las llamadas *operaciones elementales de fila*, calcular las soluciones del sistema. Este es el enfoque esencialmente de los Capitulo 1–2 de este libro.

iii) Las soluciones de un sistema de ecuaciones lineales se pueden interpretar como los coeficientes de una *combinación lineal* de las columnas de la matriz de coeficientes del sistema, que de igual al lado derecho del sistema. Aquí entran las nociones de espacios vectoriales, independencia lineal, bases, dimensión, y ortogonalidad las cuales se desarrollan en los Capítulos 3–5.

Los aspectos computacionales del álgebra lineal son parte fundamental de la mayoria de los procesos numéricos que se efectúan en cálculos científicos hoy en día. En particular el computo de las soluciones de un sistema, los autovalores de matrices, factorizaciones de matrices, por mencionar solo algunos. En este libro no vamos a entrar en los aspectos numéricos del álgebra lineal pero si estudiaremos los fundamentos para la mayoría de estos métodos.

Las versiones preliminares de este libro fueron utilizadas por varios años en el curso de álgebra lineal de la Universidad de Puerto Rico en Humacao. Le agradezco a todos los estudiantes que con sus sugerencias y comentarios ayudaron a mejorar el texto.

Pablo V. Negrón Marrero
Universidad de Puerto Rico
Departamento de Matemáticas
Humacao, PR 00791

1 Matrices y sistemas de ecuaciones

1.1 Sistemas lineales y soluciones

Una *ecuación lineal* en n variables es una ecuación de la forma

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b, \tag{1.1}$$

donde los números b, a_i, $i = 1,\ldots,n$ son dados o conocidos, y $x_1, x_2, \ldots, x_n$ son las variables o desconocidas[1] Un sistema lineal consiste de un número finito de ecuaciones de la forma (1.1). Esto es, si $m \geq 1$ es un entero dado, un *sistema lineal de m ecuaciones en n desconocidas*, o lo mismo, un *sistema* $m \times n$, es uno especificado por:

$$\left\{ \begin{array}{ccccccccc} a_{11} x_1 & + & a_{12} x_2 & + & \cdots & + & a_{1n} x_n & = & b_1, \\ a_{21} x_1 & + & a_{22} x_2 & + & \cdots & + & a_{2n} x_n & = & b_2, \\ \vdots & & \vdots & & \ddots & & \vdots & \vdots & \vdots \\ a_{m1} x_1 & + & a_{m2} x_2 & + & \cdots & + & a_{mn} x_n & = & b_m, \end{array} \right. \tag{1.2}$$

donde los números b_i, a_{ij}, $i = 1,\ldots,m$, $j = 1,\ldots,n$ son dados o conocidos.

Una *solución del sistema* (1.2) consiste de n números $\alpha_1, \alpha_2, \ldots, \alpha_n$ que al sustituirse en el sistema (1.2) por las variables $x_1, x_2, \ldots, x_n$ respectivamente, cada ecuación se cumple, esto es, ambos lados de cada ecuación son iguales. En tal caso decimos que $x_1 = \alpha_1$, $x_2 = \alpha_2$,..., $x_n = \alpha_n$ o lo mismo que $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ es solución del sistema.

Ejemplo 1.1. Considere el sistema 2×2:

$$\left\{ \begin{array}{rcl} x + y & = & 1, \\ x - 2y & = & 0. \end{array} \right.$$

Note que $x = 2/3$, $y = 1/3$ es solución del sistema. Podemos escribir o decir también que $(2/3, 1/3)$ es solución del sistema. $\square$

El conjunto de todas las soluciones de un sistema se llama el *conjunto solución del sistema*. Si un sistema no tiene solución, o lo mismo que el

[1]En el caso $n = 2$ escribimos usualmente x, y en lugar de x_1, x_2. Cuando $n = 3$, la costumbre es escribir x, y, z en lugar de x_1, x_2, x_3.

Tipo de sistema	número de soluciones
inconsistente	0
consistente–independiente	1
consistente–dependiente	∞

Tabla 1.1: Tipo y número de soluciones de un sistema lineal

conjunto solución es (el conjunto) vacío, decimos que el sistema es *inconsistente*. Si el sistema tiene (al menos una) solución, el sistema se dice que es *consistente*. Si el sistema tiene solo una solución (el conjunto solución consiste de un solo n–tuplo), entonces el sistema se llama *independiente*. Un sistema consistente con más de una solución se llama *dependiente*. Como consecuencia de los resultados de este capitulo, veremos que un sistema consistente–dependiente siempre tiene un número infinito de soluciones. La Tabla (1.1) incluye este resultado y resume las definiciones que acabamos de discutir.

Ejemplo 1.2. Más adelante veremos que $(2/3, 1/3)$ es la única solución del sistema del Ejemplo (1.1). De modo que ese sistema es consistente–independiente. El sistema

$$\begin{cases} x+y &= 1, \\ x+y &= 0. \end{cases}$$

claramente es inconsistente. Por otra parte, el sistema

$$\begin{cases} x+y &= 1, \\ 2x+2y &= 2. \end{cases}$$

es consistente–dependiente ya que ambas ecuaciones son esencialmente iguales (una es un múltiplo de la otra). El conjunto solución en este caso consiste de todos los puntos en la recta con ecuación $x+y=1$. □

1.2 Sistemas equivalentes

El procedimiento más común para resolver sistemas lineales consiste en transformar el sistema original a uno cuya solución sea fácil de calcular.

Para que este proceso sea efectivo es importante que al transformar el sistema original, el conjunto solución de éste no cambie. Las *transformaciones elementales* que veremos ahora tienen ésta propiedad. También es necesario que definamos lo que es un sistema *fácil* de resolver. Estos sistemas son los llamados *sistemas triangulares*.

Dos sistemas de ecuaciones con el mismo conjunto solución se llaman *sistemas equivalentes*.

Ejemplo 1.3. Los sistemas

$$\left\{\begin{array}{rcr} x_1 + x_2 & = & 1, \\ x_1 - x_2 & = & -1, \end{array}\right. \qquad \left\{\begin{array}{lcr} 2x_1 + 2x_2 & = & 2, \\ 2x_1 & = & 0, \end{array}\right.$$

tienen la solución única $x_1 = 0, x_2 = 1$ por lo que son equivalentes. □

Las operaciones que utilizamos para transformar un sistema se llaman *operaciones elementales de ecuaciones (o filas)* y están das por:

i) Intercambiar dos ecuaciones de un sistema.

ii) Multiplicar una ecuación de un sistema por un número distinto de cero.

iii) Sumarle a una ecuación un múltiplo de otra ecuación.[2]

Claramente las operaciones elementales (i), (ii) preservan el conjunto solución del sistema. Se puede verificar que (iii) también preserva el conjunto solución.

Por el momento limitamos la discusión a sistemas[3] $n \times n$. Un sistema $n \times n$ se llama *triangular* si en la k–esima ecuación del sistema, los coeficientes de las primeras $k-1$ variables son cero y el de la variable x_k es distinto de cero, esto para $k = 1, 2, \ldots, n$.

Ejemplo 1.4. El sistema

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

[2] En la notación que veremos en esta sección, la operación elemental de fila (iii), siempre es de la forma $e_i = e_i + \alpha e_j$. ¡La operación $e_i = e_j + \alpha e_i$ no es una operación elemental!

[3] Los sistemas $n \times n$ se llaman *sistemas cuadrados*.

es uno triangular. Note que éste sistema es fácil de resolver: de la ultima o tercera ecuación del sistema, tenemos que $x_3 = 2$. Sustituyendo este valor en la segunda ecuación y despejando para x_2, tenemos que $x_2 = 1$. Finalmente, sustituyendo los valores de x_2, x_3 en la primera ecuación y despejando para x_1, tenemos que $x_1 = -3$. La solución del sistema es $(x_1, x_2, x_3) = (-3, 1, 2)$. Note de paso que la solución es única. □

El método de solución para resolver un sistema triangular descrito en el ejemplo anterior se llama *sustitución para atrás*. Si un sistema no está en forma triangular, podemos intentar transformarlo a uno triangular utilizando operaciones elementales de ecuaciones.

Ejemplo 1.5. Considere el sistema:

$$\left\{\begin{array}{rcl} x_1 + 2x_2 + x_3 & = & 1, \\ 2x_1 - x_2 + 3x_3 & = & 3, \\ x_1 - x_2 + 2x_3 & = & -1. \end{array}\right. \tag{1.3}$$

Para describir una operación elemental de ecuación, representamos la ecuación k con e_k. Por ejemplo, si le sumamos a la segunda ecuación -4 veces la tercera, escribimos $e_2 = e_2 - 4e_3$. Tenemos ahora que el sistema dado lo podemos transformar a uno triangular con las siguientes operaciones:

$$\begin{array}{c} e_2 = e_2 - 2e_1 \\ \longrightarrow \\ e_3 = e_3 - e_1 \end{array} \left\{\begin{array}{rcr} x_1 + 2x_2 + x_3 & = & 1, \\ -5x_2 + x_3 & = & 1, \\ -3x_2 + x_3 & = & -2. \end{array}\right.$$

$$\begin{array}{c} e_3 = e_3 - \frac{3}{5}e_2 \\ \longrightarrow \end{array} \left\{\begin{array}{rcr} x_1 + 2x_2 + x_3 & = & 1, \\ -5x_2 + x_3 & = & 1, \\ \frac{2}{5}x_3 & = & -\frac{13}{5}. \end{array}\right.$$

Usando sustitución para atrás obtenemos que $x_1 = 21/2$, $x_2 = -3/2$, $x_3 = -13/2$.

Este proceso se puede llevar acabo también con la llamada *matriz aumentada* del sistema. Para el sistema (1.3) la matriz aumentada es:

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

El proceso para triangularizar el sistema es igual al de antes pero usamos f_k en lugar de e_k para describir la fila k de la matriz aumentada. Veamos:

$$\begin{array}{c} f_2 = f_2 - f_1 \\ \longrightarrow \\ f_3 = f_3 - f_1 \end{array} \left[\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 0 & -5 & 1 & 1 \\ 0 & -3 & 1 & -2 \end{array}\right]$$

$$\begin{array}{c} f_3 = f_3 - \frac{3}{5} f_2 \\ \longrightarrow \end{array} \left[\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 0 & -5 & 1 & 1 \\ 0 & 0 & \frac{2}{5} & -\frac{13}{5} \end{array}\right]$$

El proceso de sustitución para atrás ahora es igual que antes, identificando cada columna con la variable correspondiente. □

La *matriz de coeficientes* del sistema (1.2) es el arreglo $m \times n$ de números dado por:

$$\left[\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{array}\right],$$

y la correspondiente *matriz aumentada*:

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

El proceso de transformar un sistema a uno triangular por medio de operaciones elementales de filas y luego resolver el sistema resultante por sustitución para atrás se llama el *método de eliminación Gaussiana*. Pero hay todavía un detalle importante que abordar: no todo sistema puede ser transformado a uno triangular. Para empezar, si el sistema no es cuadrado, entonces no va a poder ser transformado a uno triangular. Pero aun cuando el sistema sea cuadrado, no necesariamente puede ser transformado a uno triangular.

Ejemplo 1.6. Considere la siguiente matriz de coeficientes de un cierto sistema:

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

Al eliminar la entrada de la fila dos, columna uno con la operación $f_2 = f_2 + f_1$, obtenemos la matriz:

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

que no puede ser transformada a una triangular. No obstante, por medio de las operaciones $(1/2)f_2$, $f_3 = f_3 - f_2$, podemos transformar esta matriz a la siguiente:

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

Decimos que esta matriz está en forma *echelon por filas*. □

Una matriz $m \times n$ está en forma *echelon por filas* si se cumplen las siguientes:

i) si las entradas de una fila no son todas cero, entonces la primera entrada distinta de cero en la fila es 1.

ii) Si la fila k no consiste completamente de ceros, entonces el número de ceros al principio de la fila $k+1$ es mayor que el correspondiente número para la fila k.

iii) Todas las filas cuyas entradas son todas igual a cero, están en la parte de abajo de la matriz, esto es, debajo de las filas con entradas distintas de cero.

Nota 1.7. Toda matriz puede ser transformada, por medio de operaciones elementales de fila, a una que esté en forma echelon por filas. No obstante, la forma echelon de una matriz no es única. □

Ejemplo 1.8.

$$\begin{bmatrix} 1 & 4 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \text{ está en forma echelon por filas.}$$

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \text{ no está en forma echelon por filas.}$$

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{no está en forma echelon por filas.}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{no está en forma echelon por filas.}$$

□

El proceso de eliminación gaussiana lo extendemos ahora a sistemas $m \times n$ de la forma siguiente. Primero transformamos el sistema original a uno cuya matriz aumentada esté en forma echelon por filas. Si la forma echelon de la matriz aumentada tiene una fila de la forma:

$$\left(\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & 1 \end{array} \right).$$

entonces el sistema es inconsistente. De lo contrario, el sistema es consistente e independiente si las filas distintas de cero corresponden a un sistema triangular. De lo contrario el sistema es dependiente. En este último caso las variables correspondientes a las primeras entradas distintas de cero en cada fila se llaman *variables lideres* y las restantes variables son las *variables libres*.

Ejemplo 1.9. Considere el sistema

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

La matriz aumentada del sistema es:

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

Transformamos esta matriz a forma echelon por filas:

$$\begin{array}{c} f_2 = f_2 - 2f_1 \\ \longrightarrow \\ f_3 = f_3 - f_1 \\ f_4 = f_4 - 2f_1 \end{array} \left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & -3 & -3 & 2 \\ 0 & -2 & 1 & 1 & 3 \\ 0 & 3 & -4 & -4 & 4 \end{array} \right]$$

$$\begin{array}{c} f_3 = f_3 + 2f_2 \\ \longrightarrow \\ f_4 = f_4 - 3f_2 \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & -3 & -3 & 2 \\ 0 & 0 & -5 & -5 & 7 \\ 0 & 0 & 5 & 5 & -2 \end{array}\right]$$

$$\begin{array}{c} f_4 = f_4 + f_3 \\ \longrightarrow \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & -3 & -3 & 2 \\ 0 & 0 & -5 & -5 & 7 \\ 0 & 0 & 0 & 0 & 5 \end{array}\right]$$

$$\begin{array}{c} -(1/5)f_3 \\ \longrightarrow \\ (1/5)f_4 \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & -3 & -3 & 2 \\ 0 & 0 & 1 & 1 & -7/5 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

Tenemos entonces que este sistema es inconsistente.

Consideramos ahora el mismo sistema pero con el lado derecho de la ultima ecuación igual a -1. La eliminación procede de la siguiente forma:

$$\begin{array}{c} f_2 = f_2 - 2f_1 \\ \longrightarrow \\ f_3 = f_3 - f_1 \\ f_4 = f_4 - 2f_1 \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & -3 & -3 & 2 \\ 0 & -2 & 1 & 1 & 3 \\ 0 & 3 & -4 & -4 & -1 \end{array}\right]$$

$$\begin{array}{c} f_3 = f_3 + 2f_2 \\ \longrightarrow \\ f_4 = f_4 - 3f_2 \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & -3 & -3 & 2 \\ 0 & 0 & -5 & -5 & 7 \\ 0 & 0 & 5 & 5 & -7 \end{array}\right]$$

$$\begin{array}{c} f_4 = f_4 + f_3 \\ \longrightarrow \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & -3 & -3 & 2 \\ 0 & 0 & -5 & -5 & 7 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\begin{array}{c} -(1/5)f_3 \\ \longrightarrow \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & -3 & -3 & 2 \\ 0 & 0 & 1 & 1 & -7/5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

El sistema ahora es consistente–dependiente. Las variables lideres son x_1, x_2, x_3 mientras que x_4 es la variable libre. El sistema se puede escribir ahora, moviendo los términos de x_4 al lado derecho, como:

$$\begin{cases} x_1 + x_2 + x_3 &= -x_4, \\ x_2 - 3x_3 &= 2 + 3x_4, \\ x_3 &= -7/5 - x_4. \end{cases}$$

Este sistema triangular en x_1, x_2, x_3 se resuelve ahora para cualquier valor de x_4, esto es, como si supiéramos el valor de x_4, usando sustitución para atrás. El resultado es que $x_1 = 18/5$, $x_2 = -11/5$, $x_3 = -7/5 - x_4$. La solución del sistema es:

$$\left(\frac{18}{5}, -\frac{11}{5}, -\frac{7}{5} - x_4, x_4\right), \quad x_4 \in \mathbb{R}.$$

Note que para cualquier valor de x_4, este 4–tuplo es solución, i.e., tenemos un número infinito de soluciones. La solución también podría escribirse como:

$$\left(\frac{18}{5}, -\frac{11}{5}, -\frac{7}{5} - \alpha, \alpha\right), \quad \alpha \in \mathbb{R}.$$

$\square$

Un sistema de ecuaciones lineales de m ecuaciones en n desconocidas se llama *sobre–determinado* si $m > n$, i.e., hay más ecuaciones que desconocidas. Los sistemas sobre–determinados por lo general son inconsistentes. Si $m < n$, es decir, menos ecuaciones que desconocidas, decimos que el sistema es *sub–determinado*. Los sistemas sub–determinados son por lo general consistentes. Si un sistema sub–determinado es consistente, entonces en la forma echelon por fila de la matriz aumentada siempre habrán variables libres. Tenemos entonces el siguiente resultado:

Teorema 1.10. *Si un sistema sub–determinado es consistente, entonces es dependiente.*

Un sistema lineal se llama *homogéneo* si el lado derecho de cada una de las ecuaciones del sistema es cero. Todo sistema homogéneo es consistente ya que poniendo todas las variables igual a cero, todas las ecuaciones se cumplen. Esta solución se llama la *solución trivial*. Si el sistema homogéneo es además sub–determinado, entonces es dependiente y por consiguiente tiene una solución distinta a la solución trivial.

Teorema 1.11. *Todo sistema homogéneo sub–determinado tiene soluciones no–triviales.*

1.3 Álgebra de Matrices

Una *matriz* es un arreglo rectangular de números. En particular una matriz $m \times n$ tiene la forma:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \tag{1.4}$$

Utilizaremos letras mayúsculas para denotar matrices. La *entrada* de la matriz A en la fila i y columna j se denota por a_{ij}. En forma compacta, escribimos la matriz (1.4) como $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$, o de forma aún más compacta como $A = (a_{ij})$.

Definición 1.12. Dos matrices A, B son iguales si $a_{ij} = b_{ij}$ para toda i, j.

1.3.1 Multiplicación por escalar

Si $A = (a_{ij})$ es una matriz y $\alpha \in \mathbb{R}$, entonces αA es la matriz que se define por $\alpha A = (\alpha a_{ij})$.

Ejemplo 1.13. Si

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

entonces

$$-4A = \begin{bmatrix} 8 & -16 \\ -12 & 4 \end{bmatrix}.$$

□

1.3.2 Suma de matrices

Si $A = (a_{ij})$, $B = (b_{ij})$ son matrices $m \times n$ ambas, entonces definimos:

$$A \pm B = (a_{ij} \pm b_{ij}).$$

Ejemplo 1.14.

$$\begin{bmatrix} 8 & -16 & 4 \\ -12 & 4 & -1 \end{bmatrix} - \begin{bmatrix} -2 & 4 & 7 \\ -9 & -7 & 3 \end{bmatrix} = \begin{bmatrix} 10 & -20 & -3 \\ -3 & 11 & -4 \end{bmatrix}.$$

□

Si O es la matriz $m \times n$ cuyas entradas son todas igual a cero, entonces es fácil ver que:

$$A + O = A = O + A, \quad A + (-1)A = O = (-1)A + A.$$

De la segunda de éstas identidades tenemos que el inverso aditivo de A es $(-1)A$, i.e.,

$$-A = (-1)A.$$

1.3.3 Multiplicación de matrices

La multiplicación de matrices es la operación más complicada con matrices. Examinamos este proceso en etapas.

Matriz fila por matriz columna

Sea A una matriz $1 \times n$ y B otra matriz $n \times 1$. Si escribimos $A = (a_{1j})$ y $B = (b_{j1})$, entonces

$$AB = \sum_{j=1}^{n} a_{1j} b_{j1}.$$

A esta operación se le llama de forma genérica, un *producto interior*.

Ejemplo 1.15.

$$\begin{bmatrix} -1 & 4 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 10 \end{bmatrix} = (-1)(2) + (4)(-3) + (6)(10) = 46.$$

□

Matriz por matriz columna

Si $A = (a_{ij})$ es una matriz $m \times n$ y $B = (b_{j1})$ es una matriz *columna* $n \times 1$, entonces AB es la matriz columna de tamaño $m \times 1$ cuyas entradas se obtienen multiplicando las filas de A por B, esto es:

$$AB = \left(\sum_{j=1}^{n} a_{ij} b_{j1} \right),$$

donde $i = 1, \ldots, m$.

Ejemplo 1.16.

$$\begin{bmatrix} -2 & 1 & 3 \\ 4 & 3 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -4 \end{bmatrix} = \begin{bmatrix} (-2)(3)+(1)(2)+(3)(-4) \\ (4)(3)+(3)(2)+(6)(-4) \end{bmatrix} = \begin{bmatrix} -16 \\ -6 \end{bmatrix}$$

□

Matriz por matriz (caso general)

Si $A = (a_{ij})$ es $m \times n$ y $B = (b_{ij})$ es $n \times p$, entonces AB es la matriz de tamaño $m \times p$ cuyas entradas se obtienen multiplicando las filas de A por las columnas de B, esto es:

$$AB = \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right),$$

donde $i = 1, \ldots, m$, $k = 1, \ldots, p$.

Ejemplo 1.17. Para

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

tenemos que

$$\begin{aligned} AB &= \begin{bmatrix} (3)(2)+(-1)(0)+(1)(1) & (3)(1)+(-1)(1)+(1)(1) \\ (0)(2)+(2)(0)+(4)(1) & (0)(1)+(2)(1)+(4)(1) \end{bmatrix}, \\ &= \begin{bmatrix} 7 & 3 \\ 4 & 6 \end{bmatrix}. \end{aligned}$$

□

Note que el producto de matrices no es conmutativo, esto es $AB \neq BA$ en general. De hecho es posible que AB se pueda calcular pero BA no. Aún cuando ambos productos se puedan calcular, que es posible solo si ambas matrices son $n \times n$, los productos no tienen porque ser iguales.

Ejemplo 1.18. Para las matrices:

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

tenemos que

$$AB = \begin{bmatrix} -13 & 0 \\ 23 & 4 \end{bmatrix}, \quad BA = \begin{bmatrix} 1 & 6 \\ 7 & -10 \end{bmatrix},$$

i.e., $AB \neq BA$. $\square$

Las operaciones de suma, resta, multiplicación por escalar, y multiplicación de matrices, tienen las siguientes propiedades algebraicas.

Teorema 1.19. *Sean α, β números reales y A, B, C matrices de las dimensiones apropiadas para que cada una de las operaciones indicadas estén definidas. Entonces:*

i) $A + B = B + A$

ii) $(A + B) + C = A + (B + C)$

iii) $(AB)C = A(BC)$

iv) $A(B + C) = AB + AC$

v) $(A + B)C = AC + BC$

vi) $(\alpha\beta)A = \alpha(\beta A)$

vii) $\alpha(AB) = (\alpha A)B = A(\alpha B)$

viii) $(\alpha + \beta)A = \alpha A + \beta A$

ix) $\alpha(A + B) = \alpha A + \alpha B$

Demostración: Verificamos las propiedades (i) y (vii) únicamente. Sean $A = (a_{ij})$, $B = (b_{ij})$ matrices $m \times n$. Entonces

$$A + B = (a_{ij} + b_{ij}) = (b_{ij} + a_{ij}) = B + A.$$

Para la propiedad (vii), tomamos $A = (a_{ij})$ de tamaño $m \times n$ y $B = (b_{ij})$ de tamaño $n \times p$. Ponemos $C = AB$ y $D = \alpha A$. Queremos ver que $\alpha C = DB$. Como

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \quad d_{ij} = \alpha a_{ij},$$

tenemos que

$$\begin{aligned} \alpha C &= (\alpha c_{ij}) = \left(\alpha \sum_{k=1}^{n} a_{ik} b_{kj} \right), \\ &= \left(\sum_{k=1}^{n} (\alpha a_{ik}) b_{kj} \right) = \left(\sum_{k=1}^{n} d_{ik} b_{kj} \right) = DB. \end{aligned}$$

De igual forma se demuestra que $(\alpha A)B = A(\alpha B)$. □

Otra operación importante que realizamos con matrices es la de *transposición*.

Definición 1.20. Si $A = (a_{ij})$ es una matriz $m \times n$, entonces la *matriz transpuesta* de A se denota por A^t y es la matriz $n \times m$ dada por:

$$A^t = (a_{ji}).$$

Note que las filas de A^t son las columnas de A pero colocadas de forma horizontal.

Ejemplo 1.21.

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 6 & 1 & 2 \end{bmatrix}, \quad A^t = \begin{bmatrix} 1 & 6 \\ 2 & 1 \\ 4 & 2 \end{bmatrix},$$

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

□

Definición 1.22. Una matriz A de tamaño $n \times n$ se dice que es *simétrica* si $A^t = A$.

Note que A es simétrica si y solo si $a_{ij} = a_{ji}$ para toda i, j.

Ejemplo 1.23. Las siguientes matrices son simétricas

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & -1 & 2 \\ 4 & 2 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} -7 & 4 \\ 4 & 2 \end{bmatrix}.$$

□

La operación de transposición tiene las siguientes propiedades:

i) $(A^t)^t = A$

ii) $(\alpha A)^t = \alpha A^t$

iii) $(A+B)^t = A^t + B^t$, donde A, B son matrices de las mismas dimensiones.

iv) $(AB)^t = B^t A^t$, donde A es $m \times n$ y B es $n \times p$.

El *espacio euclidiano* de dimensión n se denota por $\mathbb{R}^n$ y se define por:

$$\mathbb{R}^n = \Big\{(a_1, a_2, \ldots, a_n)^t \,:\, a_i \in \mathbb{R}, \quad 1 \leq i \leq n\Big\}. \tag{1.5}$$

Los elementos de $\mathbb{R}^n$, ésto es, las matrices columnas, también se les llama *vectores*. En lugar de usar letras mayúsculas, representaremos los elementos de $\mathbb{R}^n$ con letras minúsculas con una flecha arriba, ésto es $\vec{x}, \vec{y}$, etc. Usando la definición de multiplicación de matrices de ésta sección, es fácil ver ahora que el sistema (1.2) puede escribirse de la forma:

$$A\vec{x} = \vec{b}, \tag{1.6}$$

donde A es la matriz de coeficientes del sistema y $\vec{x}, \vec{b}$ son matrices columnas dadas por:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Usando el transpuesto podemos escribir $\vec{x} = (x_1, x_2, \ldots, x_n)^t$, lo mismo con $\vec{b}$. Tenemos pues que en (1.6), $\vec{x} \in \mathbb{R}^n$, $\vec{b} \in \mathbb{R}^m$.

1.4 Matrices especiales

Vamos ahora a ver como el proceso de eliminación gaussiana puede escribirse en forma de una ecuación matricial. Usando las matrices especiales llamadas *matrices elementales*, veremos que el proceso de eliminación gaussiana se puede realizar multiplicando en ambos lados la ecuación (1.6) por una secuencia apropiada de matrices elementales.

Matrices triangulares

Una matriz $A = (a_{ij})$ de tamaño $n \times n$ es *triangular superior* si $a_{ij} = 0$ para $i > j$. Se dice que A es *triangular inferior* si $a_{ij} = 0$ para $i < j$. Una matriz es *triangular* si es triangular superior o triangular inferior. Una matriz es *diagonal* si es triangular superior y triangular inferior, ambas.

Ejemplo 1.24.

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \text{es triangular superior,}$$

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \text{es triangular inferior,}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{son diagonales.}$$

□

1.4.1 Matrices nosingulares

Para una matriz $A = (a_{ij})$ de tamaño $n \times n$, el vector $(a_{11}, a_{22}, \dots, a_{nn})$ se llama la *diagonal principal* de A. La matriz diagonal con solo unos en la diagonal principal se llama la *matriz identidad* y se denota con la letra I. Esto es $I = (\delta_{ij})$ donde

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

Para cualquier matriz A de tamaño $n \times n$, se cumple que:

$$AI = A = IA.$$

Definición 1.25. Una matriz A de tamaño $n \times n$ se llama *nosingular* o *invertible* si existe otra matriz que se denota por A^{-1}, con la propiedad de que:

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

A^{-1} se llama el *inverso multiplicativo* de A.

Ejemplo 1.26. Se puede verificar (multiplicando A por A^{-1}) que si

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

entonces

$$A^{-1} = \begin{bmatrix} -\frac{1}{10} & \frac{2}{5} \\ \frac{3}{10} & -\frac{1}{5} \end{bmatrix}.$$

□

Los inversos de matrices cuando existen son únicos. Para ver esto, suponga que B es otro inverso de la matriz A, esto es, $AB = BA = I$. Entonces

$$B = BI = B(AA^{-1}) = (BA)A^{-1} = IA^{-1} = A^{-1}.$$

Teorema 1.27. *Sean A, B matrices nosingulares. Entonces AB es nosingular y $(AB)^{-1} = B^{-1}A^{-1}$.*

Demostración: Verificamos esto multiplicando:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$$

De igual forma se verifica que $(B^{-1}A^{-1})(AB) = I$ de modo que $B^{-1}A^{-1}$ es un inverso para AB. Como los inversos son únicos, podemos concluir que $(AB)^{-1} = B^{-1}A^{-1}$. □

Sistemas equivalentes

Considere el sistema $m \times n$ dado por la ecuación (1.6). Sea M una matriz $m \times m$ nosingular. Si $\vec{y}$ es tal que $A\vec{y} = \vec{b}$, entonces multiplicando en ambos lados por la matriz M tenemos que $(MA)\vec{y} = M\vec{b}$. Por otro lado si $\vec{z}$ es tal que $(MA)\vec{z} = M\vec{b}$, entonces multiplicando por M^{-1} en ambos lados tenemos que:

$$\begin{aligned} M^{-1}(MA)\vec{z} &= M^{-1}(M\vec{b}), \\ (M^{-1}M)A\vec{z} &= (M^{-1}M)\vec{b}, \\ IA\vec{z} &= I\vec{b}, \end{aligned}$$

$$A\vec{z} = b.$$

Esto demuestra que el conjunto solución del sistema (1.6) es igual al del sistema $(MA)\vec{x} = M\vec{b}$, i.e., los sistemas son equivalentes. Vemos entonces que si multiplicamos por la izquierda con una matriz nosingular, ambos lados de la ecuación matricial de un sistema, obtenemos un sistema equivalente al original. Ahora veremos que cada una de las operaciones elementales de fila que definimos en la Sección (1.2), se puede representar con una multiplicación matricial en ambos lados del sistema por una matriz nosingular. Tenemos entonces por la discusión anterior, que las operaciones elementales de fila preservan el conjunto solución del sistema al cual se apliquen.

1.4.2 Matrices elementales

Una *matriz elemental* es una que se obtiene a partir de la matriz identidad I por medio de una sola operación elemental de fila. Hay tres tipos de matrices elementales dependiendo del tipo de operación elemental que se utilizó.

a) (Tipo I) Esta matriz elemental es la que se obtiene intercambiando dos filas de la matriz identidad. Si E es una matriz elemental del tipo I, y A es una matriz cualquiera $n \times n$, entonces EA es igual a la matriz A con el mismo intercambio de filas que se le hizo a la matriz identidad para obtener a E.

   Ejemplo 1.28. Considere las matrices:

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

   Note que E se obtiene intercambiando las filas dos y tres de la matriz identidad. Ahora

$$EA = \begin{bmatrix} 3 & 4 & 2 \\ 2 & 1 & 3 \\ 1 & 6 & -1 \end{bmatrix},$$

   i.e., la matriz A con las filas dos y tres intercambiadas. □

b) (Tipo II) Esta matriz elemental es la que se obtiene multiplicando una fila de la matriz identidad por un escalar distinto del cero. El efecto de pre–multiplicar una matriz A por una matriz elemental del tipo II es que la correspondiente fila de A se multiplica por el mismo escalar que se uso para generar la matriz elemental.

Ejemplo 1.29. Considere las matrices:

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

Note que E se obtiene a partir de la matriz identidad multiplicando la fila dos de ésta por el número -3. Tenemos ahora que:

$$EA = \begin{bmatrix} 3 & 4 & 2 \\ -3 & -18 & 3 \\ 2 & 1 & 3 \end{bmatrix},$$

i.e., la matriz A con la fila dos multiplicada por -3. □

c) (Tipo III) Esta matriz elemental se obtiene a partir de la matriz identidad luego de realizar sobre ésta una sola operación elemental de fila de tipo III. El efecto de pre–multiplicar una matriz A por una matriz elemental del tipo III es el de hacerle a ésta la misma operación elemental de tipo III que se usó para generar la matriz elemental.

Ejemplo 1.30. Considere las matrices:

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

Note que E se obtiene a partir de la matriz identidad reemplazando la fila dos con la fila dos menos cuatro veces la tercera fila. Tenemos ahora que:

$$EA = \begin{bmatrix} 3 & 4 & 2 \\ -7 & 2 & -13 \\ 2 & 1 & 3 \end{bmatrix},$$

i.e., se remplazó la fila dos de A con la fila dos menos cuatro veces la tercera fila. □

Vemos entonces que las operaciones elementales de fila se pueden realizar pre–multiplicando por matrices elementales. Otra propiedad importante de las matrices elementales es la siguiente:

Teorema 1.31. *Si E es una matriz elemental, entonces E es invertible y su inversa E^{-1} es otra matriz elemental del mismo tipo.*

Los inversos de las matrices elementales se pueden calcular de forma bien fácil:

i) Si E es del tipo I, entonces $E^{-1} = E$.

ii) Si E es del tipo II, entonces E^{-1} se obtiene a partir de E tomando los recíprocos de las entradas diagonales.

iii) Si E es del tipo III, entonces E^{-1} se obtiene a partir de E tomando el opuesto de las entradas fuera de la diagonal.

Ejemplo 1.32.

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad E_1^{-1} = E_1,$$

$$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}.$$

□

Nota 1.33. Es importante entender que aúnque una operación de filas como $f_i = f_j + \alpha f_i$ es válida, ésta no corresponde a una operación elemental de fila. ¡De hecho son dos operaciones elementales de fila! Por ejemplo, en el caso 3×3, si aplicamos la operación $f_2 = f_1 + 2f_2$ a la matriz identidad, tenemos:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{f_2 = f_1 + 2f_2} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Esta última matriz no es una matriz elemental ya que se obtiene a partir de la matriz identidad aplicando las operaciones elementales de fila $2f_2$ y $f_2 = f_2 + f_1$ (en este orden). La operación elemental de fila de tipo III, siempre es de la forma $f_i = f_i + \alpha f_j$. □

El método de eliminación gaussiana se puede representar ahora en forma de una ecuación matricial utilizando las matrices elementales.

Ejemplo 1.34. Considere la matriz:

$$A = \begin{bmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{bmatrix}.$$

Vamos a transformarla a su forma echelon, sin el requisito de que la primera entrada en cada fila distinta de cero, sea un uno:

$$\begin{bmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{bmatrix} \underset{f_3 = f_3 - 2f_1}{\overset{f_2 = f_2 - \frac{1}{2}f_1}{\longrightarrow}} \begin{bmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & -9 & 5 \end{bmatrix} \overset{f_3 = f_3 + 3f_2}{\longrightarrow} \begin{bmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 8 \end{bmatrix}$$

Las operaciones de fila que hicimos se pueden representar con las matrices elementales:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}.$$

El proceso de arriba se puede escribir ahora como:

$$U = \begin{bmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 8 \end{bmatrix} = E_3 E_2 E_1 A.$$

De aquí que:

$$A = E_1^{-1} E_2^{-1} E_3^{-1} U = LU,$$

donde

$$L = E_1^{-1} E_2^{-1} E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 2 & -3 & 1 \end{bmatrix}.$$

□

La ecuación $A = LU$ en el Ejemplo (1.34) se llama la *factorización LU* de la matriz A.

Definición 1.35. Sea A una matriz $n \times n$. Decimos que $A = LU$ es la factorización LU de la matriz A si:

i) U es triangular superior;

ii) L es triangular inferior con las entradas en la diagonal principal todas igual a uno.

Se puede verificar que toda matriz nosingular tiene una factorización LU, y que ésta es única. Del Ejemplo (1.34) vemos que la factorización LU de A se calcula reduciendo la matriz A a forma escalonada, usando solo operaciones del tipo III. Luego las entradas de L debajo de la diagonal principal, se obtienen a partir de las operaciones de fila realizadas, siempre visualizando éstas de la forma $f_i = f_i - \alpha f_j$.

Ejemplo 1.36. Para la matriz

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

tenemos que:

$$\begin{bmatrix} 1 & 3 & -2 \\ 2 & 5 & 1 \\ -3 & 4 & 2 \end{bmatrix} \xrightarrow[f_3 = f_3 - (-3)f_1]{f_2 = f_2 - 2f_1} \begin{bmatrix} 1 & 3 & -2 \\ 0 & -1 & 5 \\ 0 & 13 & -4 \end{bmatrix}$$

$$\xrightarrow{f_3 = f_3 - (-13)f_2} \begin{bmatrix} 1 & 3 & -2 \\ 0 & -1 & 5 \\ 0 & 0 & 61 \end{bmatrix} = U.$$

Con

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & -13 & 1 \end{bmatrix},$$

tenemos que $A = LU$. □

1.4.3 Matrices equivalentes

Decimos que una matriz B es *equivalente por filas* a otra matriz A, escribimos $B \sim A$, si existen matrices elementales $E_1, E_2, \dots, E_k$ tal que

$$B = E_k E_{k-1} \cdots E_2 E_1 A.$$

Esto es lo mismo que decir que B se obtiene a partir de A por medio de un número finito de operaciones elementales de fila. Usando el Teorema (1.31) es fácil ver que:

i) Si $B \sim A$, entonces $A \sim B$.

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

Como claramente $A \sim A$, éstas propiedades demuestran que "$\sim$" es una *relación de equivalencia* en el conjunto de matrices $n \times n$. Usando los Teoremas (1.27) y (1.31) se puede demostrar ahora el siguiente importante resultado.

Teorema 1.37. *Sea A una matriz $n \times n$. Las siguientes aseveraciones son todas equivalentes:*

i) A es nosingular.

ii) El sistema $A\vec{x} = \vec{0}$ tiene como única solución $\vec{x} = \vec{0}$.

iii) $A \sim I$.

iv) El sistema $A\vec{x} = \vec{b}$ tiene solución que es única para todo $\vec{b} \in \mathbb{R}^n$.

Demostración: Suponga que A es nosingular y que $A\vec{x} = \vec{0}$. Entonces:

$$\begin{aligned} A\vec{x} &= \vec{0}, \\ A^{-1}(A\vec{x}) &= A^{-1}\vec{0}, \\ (A^{-1}A)\vec{x} &= \vec{0}, \\ I\vec{x} &= \vec{0}, \\ \vec{x} &= \vec{0}. \end{aligned}$$

Tenemos entonces que i) implica ii).

Suponga ahora que el sistema $A\vec{x} = \vec{0}$ tiene como única solución $\vec{x} = \vec{0}$. Esto quiere decir que la forma echelon de la matriz de coeficientes del sistema homogéneo, tiene la forma:

$$\begin{bmatrix} 1 & * & \cdots & * \\ 0 & 1 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = E_k E_{k-1} \cdots E_1 A,$$

donde los asteriscos indican que los valores en dichas entradas son números cualesquiera, y las matrices E_i's son elementales. Como todos los elementos de la diagonal en la matriz reducida, son distintos de cero, podemos eliminar la parte de arriba de cada columna. Esto es, existen otras matrices elementales $T_1, T_2, \ldots, T_p$ tal que:

$$I = T_p T_{p-1} \cdots T_1 (E_k E_{k-1} \cdots E_1 A),$$

lo que implica que $A \sim I$.

Si $A \sim I$, entonces la forma echelon de la matriz aumentada del sistema $A\vec{x} = \vec{b}$ no tiene filas completas de cero, ni tampoco filas de la forma:

$$\left(\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & 1 \end{array} \right).$$

Esto es, el sistema es consistente e independiente por lo que tiene solución para cualquier $\vec{b} \in \mathbb{R}^n$ y ésta solución es única.

Finalmente, si el sistema $A\vec{x} = \vec{b}$ tiene solución para todo $\vec{b} \in \mathbb{R}^n$, entonces los sistemas:

$$A\vec{x}_i = \vec{e}_i, \quad 1 \le i \le n,$$

tienen todos solución. Aquí $\vec{e}_i$ es el vector con todas las entradas cero excepto en la posición i que tiene un uno. Es fácil verificar ahora que:

$$A^{-1} = [\vec{x}_1, \ldots, \vec{x}_n],$$

es decir, que A es nosingular. $\square$

Cómputo de inversos

El Teorema 1.37 nos dá un método para calcular el inverso de una matriz A cuando ésta es nosingular. Si A es nosingular, entonces $A \sim I$ por lo que existen matrices elementales $E_1, E_2, \ldots, E_k$ tal que:

$$I = E_k E_{k-1} \cdots E_1 A.$$

Multiplicando por la derecha ambos lados de ésta ecuación por A^{-1}, tenemos que:

$$\begin{aligned} IA^{-1} &= (E_k E_{k-1} \cdots E_1 A)A^{-1}, \\ A^{-1} &= E_k E_{k-1} \cdots E_1 (AA^{-1}), \\ A^{-1} &= E_k E_{k-1} \cdots E_1 I. \end{aligned}$$

Esto es, las mismas operaciones de fila que transforman a la matriz A a la matriz identidad, transforman la matriz identidad a la matriz inversa de A. Así que el método para calcular A^{-1} consiste en:

i) Forme la matriz aumentada $[A|I]$.

ii) Use operaciones elementales de fila hasta transformar la matriz del paso anterior a la forma $[I|B]$.

iii) La matriz B es la inversa de A, i.e., $B = A^{-1}$.

De hecho si en este proceso, el paso ii) falla en darnos una matriz reducida de la forma $[I|B]$, entonces la matriz A no tiene inversa.

Ejemplo 1.38. Considere la matriz:

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

Aumentamos la matriz A con la identidad a la derecha y reducimos el lado izquierdo a la identidad:

$$\left[\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & -1 & 0 & 1 & 0 \\ 1 & -1 & 1 & 0 & 0 & 1 \end{array}\right],$$

$$\underset{f_3=f_3-f_1}{\overset{f_2=f_2-2f_1}{\longrightarrow}} \left[\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{array}\right],$$

$$\overset{f_3=f_3+f_2}{\longrightarrow} \left[\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -3 & -3 & 1 & 1 \end{array}\right],$$

$$\underset{f_2=f_2-f_3}{\overset{f_1=f_1+\frac{1}{3}f_3}{\longrightarrow}} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & -3 & -3 & 1 & 1 \end{array}\right],$$

$$\overset{-\frac{1}{3}f_3}{\longrightarrow} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 & -\frac{1}{3} & -\frac{1}{3} \end{array}\right].$$

De aquí tenemos que:

$$A^{-1} = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ 1 & 0 & -1 \\ 1 & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}.$$

□

En la Sección 1.2 describimos el método de eliminación gaussiana para resolver sistemas lineales $m \times n$. En el caso de sistemas cuadrados, podemos describir otro método, siempre y cuando la matriz de coeficientes del sistema sea nosingular. Considere el sistema $A\vec{x} = \vec{b}$ donde A es nosingular. Multiplicando por la izquierda ambos lados de ésta ecuación por A^{-1} obtenemos:

$$\begin{aligned} A^{-1}(A\vec{x}) &= A^{-1}\vec{b}, \\ (A^{-1}A)\vec{x} &= A^{-1}\vec{b}, \\ I\vec{x} &= A^{-1}\vec{b}, \\ \vec{x} &= A^{-1}\vec{b}. \end{aligned}$$

Así que otro método para resolver un sistema $n \times n$ es:

i) Calcular A^{-1}, la inversa de la matriz de coeficientes.

ii) Formar el producto $\vec{x} = A^{-1}\vec{b}$.

Ejemplo 1.39. El sistema:

$$\left\{\begin{array}{rcr} x_1 \qquad + x_3 & = & 1, \\ 2x_1 + x_2 - x_3 & = & -2, \\ x_1 - x_2 + x_3 & = & 4, \end{array}\right.$$

tiene matriz de coeficientes:

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

Del ejemplo anterior tenemos que:

$$A^{-1} = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ 1 & 0 & -1 \\ 1 & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}.$$

Con $\vec{b} = (1, -2, 4)^t$ y $\vec{x} = (x_1, x_2, x_3)^t$, tenemos que:

$$\vec{x} = A^{-1}\vec{b} = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ 1 & 0 & -1 \\ 1 & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ -3 \\ \frac{1}{3} \end{bmatrix}.$$

□

¡Aunque el método de éste ejemplo para resolver un sistema (multiplicando por la inversa de la matriz de coeficientes) parece ser más directo, en realidad no lo es! Un examen más cuidadoso de éste método y el de eliminación gaussiana refleja que éste último es mucho más eficiente que el de calcular la inversa primero. Es por ésto que el método de la inversa para resolver sistemas lineales solo se utiliza en la practica en ocasiones bien especiales donde sea imperativo calcular el inverso. La mayoría de los paquetes de computadoras, tanto comerciales como de dominio público, utilizan el método de eliminación gaussiana para resolver sistemas lineales, siempre y cuando el tamaño de la matriz de coeficientes no sea demasiado grande.

1.5 Matrices particionadas

Una matriz A $m \times n$ se puede visualizar en forma de bloques de muchas formas distintas. Una forma común de verla es en bloques de columnas. Esto es, podemos escribir que

$$A = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n],$$

donde

$$\vec{a}_j = (a_{1j}, a_{2j}, \ldots, a_{mj})^t \in \mathbb{R}^m, \quad 1 \le j \le n,$$

y $A = (a_{ij})$. Podemos también particionar a A por filas como:

$$A = \begin{bmatrix} \vec{b}_1^t \\ \vec{b}_2^t \\ \vdots \\ \vec{b}_m^t \end{bmatrix},$$

donde $\vec{b}_i \in \mathbb{R}^n$ y:

$$\vec{b}_i = (a_{i1}, a_{i2}, \ldots, a_{in})^t, \quad 1 \le i \le m.$$

Ejemplo 1.40.

$$A = \begin{bmatrix} 1 & 2 & 4 \\ -1 & 3 & -2 \end{bmatrix} = [\vec{a}_1, \vec{a}_2, \vec{a}_3] = \begin{bmatrix} \vec{b}_1^t \\ \vec{b}_2^t \end{bmatrix},$$

donde

$$\vec{a}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \vec{a}_3 = \begin{bmatrix} 4 \\ -2 \end{bmatrix},$$

$$\vec{b}_1 = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \quad \vec{b}_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}.$$

□

La matriz identidad I de tamaño $n \times n$ se puede particionar por columnas como:

$$I = [\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n],$$

donde $\vec{e}_i \in \mathbb{R}^n$ es un vector con todas las entradas cero excepto por un uno en la posición i, $1 \le i \le n$. El conjunto de vectores $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ se llama la *base estándar de* $\mathbb{R}^n$.

Ejemplo 1.41. Para $n = 3$ tenemos que:

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [\vec{e}_1, \vec{e}_2, \vec{e}_3],$$

donde

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

□

Sea A una matriz $m \times n$ y B otra matriz $n \times r$. Escribimos a B en su forma particionada por columnas como:

$$B = [\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_r], \quad \vec{b}_j \in \mathbb{R}^n, \quad 1 \le j \le r.$$

Ahora podemos escribir el producto AB como:

$$AB = A[\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_r] = [A\vec{b}_1, A\vec{b}_2, \ldots, A\vec{b}_r].$$

Esto es, la j–esima columna de AB está dada por $A\vec{b}_j$, o lo mismo, la matriz A por la j–esima columna de B.

Ejemplo 1.42. Para

$$A = \begin{bmatrix} 1 & 2 & 4 \\ -1 & 3 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 7 & 4 & 6 \\ -2 & 2 & 1 & 2 \\ 6 & 1 & -1 & 3 \end{bmatrix},$$

tenemos que la tercera columna del producto AB es:

$$A\vec{b}_3 = \begin{bmatrix} 1 & 2 & 4 \\ -1 & 3 & -2 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

□

1.6 Ejercicios

Ejercicio 1.1. Para las matrices:

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

calcule AC, $-3B^t$, y $B + C$.

Ejercicio 1.2. Calcule $3A^t - BC$ dado que:

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

Ejercicio 1.3. Encuentre todas las soluciones del sistema (homogéneo):

$$\begin{cases} -2x + 6y &= 0, \\ x - 3y &= 0, \\ -7x + 21y &= 0. \end{cases}$$

Ejercicio 1.4. Considere el sistema de ecuaciones:

$$\begin{cases} 2x_1 - 3x_2 + 2x_3 - 3x_4 &= 2, \\ 4x_1 + 5x_2 - 5x_3 + x_4 &= -3. \end{cases}$$

a) Escriba la matriz aumentada del sistema.

b) Transforme la matriz de la parte (a) a una de forma echelon.

c) Clasifique el sistema como consistente o inconsistente, dependiente o independiente.

d) De ser consistente el sistema, halle todas las soluciones.

Ejercicio 1.5. Considere el sistema de ecuaciones:

$$\begin{cases} 2x_1 + 2x_2 - 4x_3 &= -3, \\ -4x_1 - x_2 + 10x_3 &= 2, \\ 6x_1 + 12x_2 - 10x_3 &= 1. \end{cases}$$

a) Escriba la matriz aumentada del sistema.

b) Transforme la matriz de la parte (a) a una de forma escalonada.

c) Usado la matriz de la parte (b), halle la solución del sistema de ecuaciones.

Ejercicio 1.6. Para el siguiente sistema de ecuaciones

$$\left\{\begin{array}{rcl} 2x-y-Kz &=& 0, \\ x-y-2z &=& 1, \\ -x+2y &=& K, \end{array}\right. ,$$

determine los valores de K para el cual:

a) el sistema no tiene solución;

b) el sistema tiene un numero infinito de soluciones;

c) el sistema tiene exactamente una solución.

Ejercicio 1.7. Suponga que $\vec{x}_1, \vec{x}_2$ son soluciones del sistema $A\vec{x} = \vec{b}$, esto es:

$$A\vec{x}_1 = \vec{b}, \quad A\vec{x}_2 = \vec{b}.$$

a) Verifique que $A(\alpha\vec{x}_1 + \beta\vec{x}_2) = (\alpha + \beta)\vec{b}$ para cualesquiera $\alpha, \beta \in \mathbb{R}$. (5 puntos)

b) Concluya que $\alpha\vec{x}_1 + \beta\vec{x}_2$ es solución del sistema $A\vec{x} = \vec{b}$ para todo α, β tal que $\alpha + \beta = 1$. (5 puntos)

c) Si $\vec{x}_1 \neq \vec{x}_2$, ¿cúantas soluciones tiene el sistema $A\vec{x} = \vec{b}$? (3 puntos)

Ejercicio 1.8. Para la matriz

$$A = \begin{bmatrix} -2 & 4 & 5 \\ 2 & -1 & -6 \\ -6 & 0 & 26 \end{bmatrix},$$

halle la factorización LU de A.

Ejercicio 1.9. Usando el método de eliminación, halle la inversa de la matriz:

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

Use A^{-1} para resolver el sistema $A\vec{x} = \vec{b}$, donde $\vec{b} = (-2, 2)^t$.

Ejercicio 1.10. Considere la matriz

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

a) Halle A^{-1}.

b) Use A^{-1} para resolver el sistema $A\vec{x} = \vec{b}$, donde $\vec{b} = (1,1,1)^t$.

Ejercicio 1.11. Determine cuales de las siguientes aseveraciones son ciertas o falsas. En cada caso explique su contestación, con una verificación o con un contra ejemplo, según sea el caso.

a) Todo sistema 2×2 tiene a lo más dos soluciones.

b) El sistema $A\vec{x} = \vec{0}$ tiene siempre como única solución $\vec{x} = \vec{0}$.

c) Dos matrices que sean equivalentes por filas, tienen la misma forma echelon.

d) Si A y B son matrices $n \times n$ diagonales, entonces $C = AB$ es también diagonal.

e) Si A y B son matrices $n \times n$ no singulares, entonces la matriz $C = A + B$ es también no singular.

f) Si A y B son matrices $n \times n$ elementales, entonces $C = AB$ es también matriz elemental.

g) Si A y B son matrices $n \times n$ simétricas, entonces $C = AB$ es también simétrica.

Ejercicio 1.12. Indique si la siguiente identidad

$$A^2 + 3AB + 2B^2 = (A + 2B)(A + B),$$

siempre se cumple o no. Si cree que no siempre se cumple diga una condición sobre las matrices A ó B para que la igualdad sea cierta.

Ejercicio 1.13. Determine si las siguientes matrices son elementales o no. Cuando lo sean, calcule su inversa:

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

Ejercicio 1.14. Encuentre dos matrices elementales E_1 y E_2 tales que $E_2E_1A = B$, donde (8 puntos)

$$A = \begin{bmatrix} 2 & 1 & 3 \\ -2 & 4 & 5 \\ 3 & 1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 & 3 \\ 3 & 1 & 4 \\ -4 & 8 & 10 \end{bmatrix}.$$

Ejercicio 1.15. Dado que $U = E_3E_2E_1A$ donde:

$$\begin{aligned} E_1 &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -5 & 0 & 1 \end{bmatrix}, \\ E_3 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{3} & 1 \end{bmatrix}, \quad U = \begin{bmatrix} -4 & 2 & 9 \\ 0 & 5 & -3 \\ 0 & 0 & 7 \end{bmatrix}, \end{aligned}$$

halle la factorización LU de la matriz A.

Ejercicio 1.16. Usando el método de eliminación, halle la inversa de la matriz:

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

Ejercicio 1.17. Verifique que la matriz

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

es su propia inversa si y solo si $A = \pm I$ ó $a_{11} = -a_{22}$ y $a_{21}a_{12} = 1 - a_{11}^2$.

Ejercicio 1.18. Sean A, B matrices $n \times m$ y $m \times n$ respectivamente de modo que AB es de tamaño $n \times n$. Verifique que si $n > m$, entonces AB es singular, i.e., no tiene inversa. **Ayuda:** Use el Teorema 1.11 para argumentar que el sistema $B\vec{x} = \vec{0}$ tiene solución no trivial y continúe a partir de esto.

Ejercicio 1.19. Verifique que una matriz diagonal A tiene inverso si y solo si $a_{ii} \neq 0$ para toda i. En tal caso verifique que A^{-1} es una matriz diagonal con entradas en la diagonal $b_{ii} = \frac{1}{a_{ii}}$ para toda i, donde $B = A^{-1}$. Usando ésto calcule la inversa de la matriz

$$A = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & \pi \end{bmatrix}.$$

Ejercicio 1.20. Verifique que la matriz

$$A = \begin{bmatrix} \sin\theta & \cos\theta & 0 \\ \cos\theta & -\sin\theta & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

es no singular para cualquier $\theta \in \mathbb{R}$ y calcule A^{-1}.

Ejercicio 1.21. Sean A y B matrices simétricas $n \times n$.

a) Verifique que $A + B$ es simétrica.

b) Verifique que $(AB)^t = BA$.

c) Verifique que la matriz $\frac{1}{2}(A + A^t)$ es simétrica.

d) Verifique que la matriz $\frac{1}{2}(A - A^t)$ es anti–simétrica, es decir que $A^t = -A$.

Ejercicio 1.22. Para cualquier matriz A de tamaño $n \times n$, verifique que los productos AA^t y A^tA se pueden calcular ambos y que dan como resultado matrices simétricas.

2 Determinantes

Los *determinantes* son una herramienta del álgebra lineal con una historia bastante extensa, pero con aplicabilidad mayormente teórica. La definición del determinante utiliza el concepto de permutaciones y su paridad o imparidad. En este capitulo no utilizaremos esta definición. Tomaremos como la definición del determinante una caracterización o consecuencia de la definición que usa permutaciones, la llamada expansión por filas. Esta definición o caracterización es un tanto más simple para propósitos de calcular los determinantes. Es importante tener en mente que los determinantes son utilizados en raras ocasiones en cálculos numéricos. No obstante, éstos tienen una valía teórica muy importante, en particular para obtener el resultado sobre la llamada *regla de Cramer*.

2.1 Definiciones y conceptos básicos

Vamos a definir el determinante de una matriz en forma "recursiva". Sea A una matriz 2×2, ésto es:

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

En este caso, definimos el *determinante* de A por:

$$\det A = |A| = a_{11}a_{22} - a_{12}a_{21}. \tag{2.1}$$

Supongamos ahora, que A es una matriz $n \times n$, donde $n > 2$. Para cada par de índices i, j, definimos una matriz M_{ij} de tamaño $(n-1) \times (n-1)$ que se obtiene a partir de la matriz A eliminando la fila i y la columna j. El determinante $\left|M_{ij}\right|$ se llama el *menor* del elemento a_{ij} de A , y el número

$$A_{ij} = (-1)^{i+j} \left|M_{ij}\right|,$$

se llama el *cofactor* de a_{ij}. El determinante de la matriz A se define ahora como:

$$\det A = |A| = \sum_{j=1}^{n} a_{ij}A_{ij}, \quad \text{para cualquier } i. \tag{2.2}$$

Nota 2.1. Se puede verificar que esta definición es independiente del índice i. $\square$

La sumatoria (2.2) se llama una *expansión de* $\det A$ *por cofactores usando la fila* i. Se puede verificar también que

$$\det A = |A| = \sum_{i=1}^{n} a_{ij} A_{ij}, \quad \text{para cualquier } j, \tag{2.3}$$

lo cual se llama una *expansión de cofactores usando la columna* j.

Nota 2.2. Tanto (2.2) como (2.3) expresan el determinante de una matriz $n \times n$ en términos de determinantes de matrices $(n-1)\times(n-1)$. Esto junto con la definición (2.1) completa la definición recursiva del determinante. $\square$

Ejemplo 2.3. Para

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

tenemos, usando (2.1), que

$$\det A = (-2)(2) - (3)(1) = -7.$$

Note, en particular, que el determinante de una matriz no tiene que ser positivo. $\square$

Ejemplo 2.4. Para la matriz

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

tenemos mediante una expansión de cofactores usando la primera fila, que:

$$\begin{aligned} \det A &= (1)\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} - (2)\begin{vmatrix} -1 & 2 \\ 4 & 1 \end{vmatrix} + (4)\begin{vmatrix} -1 & 3 \\ 4 & 2 \end{vmatrix}, \\ &= -1 + 18 - 56 = -39. \end{aligned}$$

$\square$

Ejemplo 2.5. Considere la matriz

$$A = \begin{bmatrix} -1 & 0 & 32 \\ -6 & 0 & 17 \\ 3 & 1 & 9 \end{bmatrix}.$$

Para calcular $\det A$, podemos aprovechar la estructura de A. Si hacemos una expansión de cofactores usando la segunda columna, tenemos que:

$$\det A = -(1) \begin{vmatrix} -1 & 32 \\ -6 & 17 \end{vmatrix} = -175.$$

$\square$

Las siguientes propiedades del determinante se pueden verificar a partir de la definición.

Proposición 2.6. *Sea A una matriz $n \times n$. Entonces:*

i) $\det A^t = \det A$.

ii) Si A es una matriz triangular, entonces

$$\det A = a_{11} a_{22} \cdots a_{nn}.$$

iii) Si A tiene una fila o una columna con todas las entradas cero, entonces $\det A = 0$.

iv) Si A tiene dos filas o dos columnas idénticas, entonces $\det A = 0$.

Ejemplo 2.7. Para la matriz

$$A = \begin{bmatrix} -2 & 6 & 7 & 72 \\ 0 & 3 & 101 & \pi \\ 0 & 0 & -1 & 10^6 \\ 0 & 0 & 0 & 4 \end{bmatrix},$$

tenemos que

$$\det A = (-2)(3)(-1)(4) = 24.$$

$\square$

2.2 Propiedades adicionales de los determinantes

En esta sección vamos a estudiar algunas propiedades adicionales de los determinantes. Comenzamos con un resultado que nos dice como cambia el determinante de una matriz, luego de aplicarle a ésta una operación elemental de fila.

Teorema 2.8. *Sea A una matriz $n \times n$ y E una matriz elemental. Entonces*

$$\det(EA) = (\det E)(\det A) = \det(AE).$$

Además:

$$\det E = \begin{cases} -1 & , \ \textit{si } E \textit{ es del tipo I,} \\ \alpha \neq 0 & , \ \textit{si } E \textit{ es del tipo II,} \\ 1 & , \ \textit{si } E \textit{ es del tipo III.} \end{cases}$$

Demostración: Primero vamos a verificar que:

$$\det(EA) = \begin{cases} -\det A & , \ \text{si } E \text{ es del tipo I,} \\ \alpha \det A & , \ \text{si } E \text{ es del tipo II,} \\ \det A & , \ \text{si } E \text{ es del tipo III.} \end{cases} \tag{2.4}$$

Este resultado se puede verificar directamente para el caso $n = 2$. Suponemos, para argumentar por inducción, que el resultado es cierto para matrices de tamaño $k \geq 2$. Ahora tomamos $n = k + 1$.

1. Si E es del tipo I, denotamos por i_1, i_2 las filas de A que E intercambia. Hacemos una expansión para $\det EA$ en cofactores usando la fila i, para cualquier índice i distinto de ambos i_1, i_2. En este caso, la fila i de EA y la de A son iguales, y las submatrices M_{ij} de EA que se obtienen eliminando la fila i y cualquier columna j de EA, corresponden a las submatrices M_{ij} de A pero con dos filas intercambiadas. Usando la hipótesis de inducción, tenemos entonces que el cofactor $(EA)_{ij}$ es el negativo de A_{ij}. Así que

$$\det(EA) = \sum_{j=1}^{n} a_{ij}(EA)_{ij} = \sum_{j=1}^{n} a_{ij}(-A_{ij}) = -\det A.$$

2. Suponga que E es del tipo II y que multiplica la fila i de A por $\alpha \neq 0$. Note que las submatrices M_{ij} de EA que se obtienen eliminando la

fila i y cualquier columna j de EA, son idénticas a las que se obtienen a partir de A. Así que $(EA)_{ij} = A_{ij}$ para todo j. Entonces haciendo una expansión para $\det EA$ en cofactores usando la fila i, tenemos que:

$$\det(EA) = \sum_{j=1}^{n} \alpha a_{ij} A_{ij} = \alpha \det A.$$

3. Suponga que E es del tipo III y que le suma α veces la fila k a la fila i. Entonces las filas de EA son todas iguales a las de A excepto por la i, cuyas entradas son $a_{ij} + \alpha a_{kj}$, $j = 1, \ldots, n$. Esto en particular implica que las submatrices M_{ij} de EA que se obtienen eliminando la fila i y cualquier columna j de EA, son idénticas a las que se obtienen a partir de A, por lo que $(EA)_{ij} = A_{ij}$ para todo j. Así que, haciendo una expansión para $\det(EA)$ en cofactores usando la fila i, tenemos que:

$$\begin{aligned} \det(EA) &= \sum_{j=1}^{n} (a_{ij} + \alpha a_{kj}) A_{ij} = \sum_{j=1}^{n} a_{ij} A_{ij} + \alpha \sum_{j=1}^{n} a_{kj} A_{ij}, \\ &= \det A + \alpha \sum_{j=1}^{n} a_{kj} A_{ij}. \end{aligned}$$

Pero $\sum_{j=1}^{n} a_{kj} A_{ij}$ corresponde a la expansión en cofactores usando la fila i, del determinante de una matriz igual a A pero con la fila i idéntica a la k. Por consiguiente $\sum_{j=1}^{n} a_{kj} A_{ij} = 0$, y tenemos que $\det EA = \det A$.

La formula para $\det E$ se obtiene ahora tomando $A = I$ en (2.4) y usando que $\det I = 1$. Finalmente $\det(EA) = \det(AE)$ se obtiene a partir de la parte (i) de la Proposición 2.6 y que la transpuesta de una matriz elemental, es también elemental del mismo tipo. □

Usando este resultado, podemos ahora verificar el siguiente:

Teorema 2.9. *Una matriz A de tamaño $n \times n$, es singular si y solo si $detA = 0$.*

Demostración: Existen matrices elementales $E_1, E_2, \ldots, E_k$ tal que

$$U = E_k E_{k-1} \cdots E_1 A,$$

donde U está en forma echelon. Por el Teorema 2.8 tenemos que

$$\det U = \det E_k \det E_{k-1} \cdots \det E_1 \det A.$$

También, por el Teorema 2.8, el producto $\det E_k \det E_{k-1} \cdots \det E_1$ es distinto de cero, por lo que $\det U = 0$ si y solo si $\det A = 0$. Si A es singular, entonces por el Teorema 1.37, Parte (iii), U tiene una fila con todas las entradas cero por lo que $\det U = 0$. Si A es nosingular, entonces U es triangular superior con todos 1's en la diagonal por lo que $\det U = 1$, y por consiguiente $\det A \neq 0$. □

El resultado del Teorema 2.8 sugiere un método alterno para calcular el determinante. En este método, primero reducimos la matriz A usando operaciones elementales de fila de tipos I y III, a una matriz en forma echelon "reducida", donde la primera entrada de cada fila no es necesariamente uno. Luego multiplicamos la diagonal de la matriz resultante, y ajustamos el signo dependiendo del número de operaciones del tipo I que se hicieron.

Ejemplo 2.10. Ilustramos el método descrito antes, calculando el determinante de la siguiente matriz:

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

Tenemos ahora que

$$\begin{aligned} \det A &= \begin{vmatrix} 1 & 2 & 4 \\ 0 & 5 & 6 \\ 0 & -6 & -15 \end{vmatrix} \begin{matrix} f_2 = f_2 + f_1 \\ f_3 = f_3 - 4f_1 \end{matrix} \\ &= \begin{vmatrix} 1 & 2 & 4 \\ 0 & 5 & 6 \\ 0 & 0 & -\frac{39}{5} \end{vmatrix} \; f_3 = f_3 + \tfrac{6}{5} f_2 \\ &= (1)(5)\left(-\frac{39}{5}\right) = -39. \end{aligned}$$

□

Hay tres métodos para calcular el determinante de una matriz:

1. usando la definición basada en permutaciones;
2. usando expansiones en cofactores;
3. el método de eliminación o reducción.

De estos tres métodos resulta que el último es el más eficiente para calcular determinantes, especialmente para matrices de tamaño grande, ya que utiliza la menor cantidad de operaciones aritméticas.

Nota 2.11. Aunque $\det A = 0$ si y solo si A es singular, el tamaño de $\det A$ no es un buen indicador de singularidad. Por ejemplo, la matriz

$$A = \begin{bmatrix} 10^{-16} & 0 \\ 0 & 10^{-16} \end{bmatrix},$$

tiene determinante 10^{-32}, pero es una matriz perfectamente nosingular desde el punto de vista numérico. Por otro lado, la matriz

$$B = \begin{bmatrix} 10^{-32} & 1 \\ 0 & 1 \end{bmatrix},$$

tiene determinante 10^{-32} también, pero desde el punto de vista numérico, es lo que se conoce como una matriz *mal acondicionada*, ya que está bien "cerca" de ser singular (las filas son casi idénticas). Un método más eficiente y efectivo para detectar cuando una matriz esta cerca de ser singular, es el que utiliza la *descomposición de valores singulares* de la matriz. □

Cerramos esta sección con la siguiente propiedad multiplicativa del determinante.

Teorema 2.12. *Sean A, B matrices $n \times n$. Entonces*

$$\det(AB) = \det A \det B. \tag{2.5}$$

Esto es, el determinante de un producto de matrices, es el producto de los determinantes de las matrices.

Demostración: Si A ó B son singulares, entonces el lado derecho de (2.5) es cero. En el caso en que B es singular, entonces existe un vector $\vec{x} \neq \vec{0}$ tal que $B\vec{x} = \vec{0}$. Pero entonces $AB\vec{x} = \vec{0}$, por lo que AB es singular. De igual forma, usando transpuestos, si A es singular, entonces AB es singular. En cualquier caso tendríamos que $\det(AB) = 0$ por lo que (2.5) se cumple.

Suponga ahora que ambas A y B son nosingulares. Entonces por el Teorema 1.37, Parte iii), tenemos que existen matrices elementales $E_1, E_2, \ldots, E_k$ y $U_1, U_2, \ldots, U_m$ tal que

$$A = E_1 E_2 \cdots E_k I, \quad B = U_1 U_2 \cdots U_m I.$$

Ahora, usando el Teorema 2.8, obtenemos que:

$$\det A = \det E_1 \, \det E_2 \cdots \det E_k, \quad \det B = \det U_1 \, \det U_2 \cdots \det U_m,$$

donde usamos que $\det I = 1$. Usando nuevamente el Teorema 2.8, tenemos ahora que:

$$\begin{aligned} \det(AB) &= \det\left(E_1 E_2 \cdots E_k U_1 U_2 \cdots U_m\right), \\ &= \det E_1 \, \det E_2 \cdots \det E_k \, \det U_1 \, \det U_2 \cdots \det U_m, \\ &= \det A \, \det B. \end{aligned}$$

□

2.3 La regla de Cramer

Recuerde que para una matriz A cualquiera de tamaño $n \times n$, A_{ij} denota el cofactor asociado al elemento a_{ij} de A, donde $i, j = 1, \ldots, n$. Definimos el *adjunto* de A como la matriz transpuesta de los cofactores de A, ésto es:

$$\operatorname{adj} A = (A_{ij})^t = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}^t. \tag{2.6}$$

Ejemplo 2.13. Para la matriz

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

tenemos que

$$A_{11} = \begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} = -1, \quad A_{12} = -\begin{vmatrix} -1 & 2 \\ 4 & 1 \end{vmatrix} = 9, \quad A_{13} = \begin{vmatrix} -1 & 3 \\ 4 & 2 \end{vmatrix} = -14,$$

$$A_{21} = -\begin{vmatrix} 2 & 4 \\ 2 & 1 \end{vmatrix} = 6, \quad A_{22} = \begin{vmatrix} 1 & 4 \\ 4 & 1 \end{vmatrix} = -15, \quad A_{23} = -\begin{vmatrix} 1 & 2 \\ 4 & 2 \end{vmatrix} = 6,$$

$$A_{31} = \begin{vmatrix} 2 & 4 \\ 3 & 2 \end{vmatrix} = -8, \quad A_{32} = -\begin{vmatrix} 1 & 4 \\ -1 & 2 \end{vmatrix} = -6, \quad A_{33} = \begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix} = 5.$$

De modo que

$$\operatorname{adj} A = \begin{bmatrix} -1 & 9 & -14 \\ 6 & -15 & 6 \\ -8 & -6 & 5 \end{bmatrix}^t = \begin{bmatrix} -1 & 6 & -8 \\ 9 & -15 & -6 \\ -14 & 6 & 5 \end{bmatrix}.$$

□

Se puede verificar ahora que:

Proposición 2.14. *Sea A una matriz $n \times n$. Entonces*

$$A(\operatorname{adj} A) = (\operatorname{adj} A)A = (\det A)I.$$

En particular, si A es nosingular, entonces

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A.$$

Ejemplo 2.15. Para la matriz

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

sabemos que

$$\det A = -39, \quad \operatorname{adj} A = \begin{bmatrix} -1 & 6 & -8 \\ 9 & -15 & -6 \\ -14 & 6 & 5 \end{bmatrix}.$$

Así que A es no singular y por el resultado en la Proposición 2.14, tenemos que:

$$A^{-1} = \frac{1}{\det A}\operatorname{adj} A = \begin{bmatrix} \frac{1}{39} & -\frac{6}{39} & \frac{8}{39} \\ -\frac{9}{39} & \frac{15}{39} & \frac{6}{39} \\ \frac{14}{39} & -\frac{6}{39} & -\frac{5}{39} \end{bmatrix}.$$

□

Usando la representación de A^{-1} dada en la Proposición 2.14, podemos verificar el siguiente resultado.

Teorema 2.16 (Regla de Cramer)**.** *Sea A una matriz nosingular $n \times n$ y $\vec{b} \in \mathbb{R}^n$. Para cualquier índice i denotamos por A_i la matriz que se obtiene a partir de A remplazando la columna i con el vector $\vec{b}$. Entonces la solución $\vec{x} = (x_1, x_2, \ldots, x_n)^t$ del sistema $A\vec{x} = \vec{b}$, está dada por:*

$$x_i = \frac{\det A_i}{\det A}, \quad 1 \leq i \leq n. \tag{2.7}$$

Demostración: Sabemos que $\vec{x} = A^{-1}\vec{b}$, y que $A^{-1} = \frac{1}{\det A}\operatorname{adj} A$. Note que

$$\begin{aligned} (\operatorname{adj} A)\vec{b} &= \left[\sum_{j=1}^{n} (\operatorname{adj} A)_{ij} b_j\right], \\ &= \left[\sum_{j=1}^{n} A_{ji} b_j\right]. \end{aligned}$$

Pero $\sum_{j=1}^{n} A_{ji} b_j = \sum_{j=1}^{n} b_j A_{ji}$ es igual a la expansión en cofactores de $\det A_i$ usando la columna i de A_i. Así que

$$\begin{aligned} \vec{x} &= \frac{1}{\det A}(\operatorname{adj} A)\vec{b}, \\ &= \frac{1}{\det A}\left[\sum_{j=1}^{n} A_{ji} b_j\right], \\ &= \left[\frac{\det A_i}{\det A}\right]. \end{aligned}$$

□

Ejemplo 2.17. Vamos a resolver el siguiente sistema usando la regla de Cramer:

$$\begin{cases} x_1 + 2x_2 + 4x_3 &= 3, \\ -x_1 + 3x_2 + 2x_3 &= -6, \\ 4x_1 + 2x_2 + x_3 &= 1. \end{cases}$$

Este sistema se puede escribir como $A\vec{x} = \vec{b}$ donde:

$$A = \begin{bmatrix} 1 & 2 & 4 \\ -1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 3 \\ -6 \\ 1 \end{bmatrix}.$$

Sabemos que $\det A = -39$. Calculando, tenemos que:

$$\begin{aligned} \det A_1 &= \begin{vmatrix} 3 & 2 & 4 \\ -6 & 3 & 2 \\ 1 & 2 & 1 \end{vmatrix} = -47, \\ \det A_2 &= \begin{vmatrix} 1 & 3 & 4 \\ -1 & -6 & 2 \\ 4 & 1 & 1 \end{vmatrix} = 111, \\ \det A_3 &= \begin{vmatrix} 1 & 2 & 3 \\ -1 & 3 & -6 \\ 4 & 2 & 1 \end{vmatrix} = -73. \end{aligned}$$

Usando la regla de Cramer (2.7), tenemos que la solución del sistema es:

$$x_1 = \frac{47}{39}, \quad x_2 = -\frac{111}{39}, \quad x_3 = \frac{73}{39}.$$

□

2.4 Ejercicios

Ejercicio 2.1. Calcule el determinante de las siguientes matrices usando el método de cofactores.

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}, \quad \begin{bmatrix} -1 & 1 & 0 \\ 2 & 1 & 4 \\ 1 & 5 & 6 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 1 \\ -2 & -1 & 4 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 6 & 2 & 0 \\ 1 & 1 & -2 & 3 \end{bmatrix}.$$

Ejercicio 2.2. Calcule $\det(A)$ dado que:

$$A = \begin{bmatrix} 2 & 0 & 0 \\ -\pi & -4 & 0 \\ \sqrt{2} & \sqrt{3} & 3 \end{bmatrix} \begin{bmatrix} 3 & 0 & 4 \\ -10 & 2 & 100 \\ -2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1}.$$

Ejercicio 2.3. Usando el mtodo de eliminacin, calcule

$$\begin{vmatrix} 3 & 2 & -1 \\ 1 & 6 & 3 \\ 2 & -4 & 0 \end{vmatrix}.$$

Ejercicio 2.4. Para la matriz

$$A = \begin{bmatrix} -2 & 4 & 5 \\ 2 & -1 & -6 \\ -6 & 0 & 26 \end{bmatrix},$$

halle:

a) La factorización LU de A.

b) Usando la parte (a), encuentre $\det A$.

c) Dado que

$$\operatorname{adj} A = \begin{bmatrix} -26 & -104 & -19 \\ -16 & -22 & -2 \\ -6 & -24 & -6 \end{bmatrix},$$

halle A^{-1}.

Ejercicio 2.5. Sean A, B, C matrices 4×4 tal que $\det A = -4$, $\det(AB) = 10$, y $\det(BC) = -15$. Halle el valor de $\det(AC) + \det(2B)$.

Ejercicio 2.6. Suponga que A y B son matrices 4×4 tales que $\det(A) = -3$ y $\det(B) = 4$.

a) Calcule los siguientes: $\det(AB)$, $\det(3A)$, $\det(2AB)$ y $\det(A^{-1}B)$.

b) ¿Será posible hallar $\det(A + B)$ con la información dada? (Justifique su respuesta.)

Ejercicio 2.7. Resuelva el sistema usando la regla de Cramer:

$$\begin{cases} -5x+7y &= 3, \\ 4x-11y &= -2. \end{cases}$$

Ejercicio 2.8. Sean F, A matrices $n \times n$, donde F es nosingular. Verifique que

$$\det(F^{-1}AF) = \det A.$$

Ejercicio 2.9. Resuelve el siguiente sistema de ecuaciones usando la regla de Cramer:

$$\begin{cases} x_1 + 2x_3 &= 6, \\ -3x_1 + 4x_2 + 6x_3 &= 30, \\ -x_1 - 2x_2 + 3x_3 &= 8. \end{cases}$$

Puede utilizar la siguiente información:

$$\det\begin{bmatrix} 1 & 0 & 6 \\ -3 & 4 & 30 \\ -1 & -2 & 8 \end{bmatrix} = 152, \quad \det\begin{bmatrix} 1 & 6 & 2 \\ -3 & 30 & 6 \\ -1 & 8 & 3 \end{bmatrix} = 72,$$

$$\det\begin{bmatrix} 1 & 0 & 2 \\ -3 & 4 & 6 \\ -1 & -2 & 3 \end{bmatrix} = 44, \quad \det\begin{bmatrix} 6 & 0 & 2 \\ 30 & 4 & 6 \\ 8 & -2 & 3 \end{bmatrix} = -40.$$

Ejercicio 2.10. Para cualquier matriz A de tamaño 2×2 y para cualquier número real α, verifique que:

$$\det(\alpha A) = \alpha^2 \det(A).$$

¿Qué puede decir en el caso en que A es $n \times n$?

Ejercicio 2.11. Verifique que si A es una matriz 2×2 cuyas entradas son números enteros y con $\det(A) = 1$, entonces su inversa tiene también entradas enteras y determinante uno. ¿Qué puede decir en el caso $n \times n$?

Ejercicio 2.12. Si $\det(A) = \det(B)$, ¿podemos concluir que $A = B$, donde A y B son matrices $n \times n$? Explique su contestación.

Ejercicio 2.13. Si A y B son matrices $n \times n$ equivalentes por fila, ¿podemos concluir que $\det(A) = \det(B)$? Explique su contestación.

Ejercicio 2.14. Use el resultado de la Proposición 2.14 para verificar las siguientes propiedades del operador adj (todas las matrices son $n \times n$):

a) $\operatorname{adj}(I) = I$ donde I es la matriz identidad.

b) $\operatorname{adj}(\alpha A) = \alpha^{n-1} \operatorname{adj}(A)$ para cualquier numero real α.

c) Si A, B son ambas nosingulares, entonces $\operatorname{adj}(AB) = \operatorname{adj}(B)\operatorname{adj}(A)$. En particular, concluya que $\operatorname{adj}(A^{-1}) = [\operatorname{adj}(A)]^{-1}$.

d) $\operatorname{adj}(A^m) = [\operatorname{adj}(A)]^m$ para cualquier entero positivo m.

e) $\operatorname{adj}(A^t) = [\operatorname{adj}(A)]^t$.

f) $\det(\operatorname{adj}(A)) = [\det(A)]^{n-1}$.

g) Si A es nosingular, entonces $\operatorname{adj}(\operatorname{adj}(A)) = [\det(A)]^{n-2} A$.

Ejercicio 2.15. Sea A una matriz triangular superior.

a) Verifique que A es nosingular si y solo si $a_{ii} \neq 0$ para toda i.

b) Si A es nosingular, verifique que A^{-1} es también triangular superior. **Ayuda:** Verifique que los cofactores A_{ij} de A son todos cero siempre que $i < j$ y use ésto en la Proposición 2.14.

Ejercicio 2.16. El *producto cruz* de dos vectores $\vec{x}, \vec{y} \in \mathbb{R}^3$ se define por el vector $\vec{x} \times \vec{y}$ con componentes:

$$\vec{x} \times \vec{y} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix}.$$

Verifique que si A es nosingular, entonces

$$(A\vec{x}) \times (A\vec{y}) = (\det A) A^{-t} (\vec{x} \times \vec{y}) = [\operatorname{adj}(A)]^t (\vec{x} \times \vec{y}).$$

Ejercicio 2.17. Verifique para cualesquiera números reales x_1, x_2, x_3, tenemos que

$$\det \begin{bmatrix} 1 + x_1 & x_2 & x_3 \\ x_1 & 1 + x_2 & x_3 \\ x_1 & x_2 & 1 + x_3 \end{bmatrix} = 1 + x_1 + x_2 + x_3.$$

Ejercicio 2.18. Verifique para cualesquiera números reales x_1, x_2, x_3, tenemos que

$$\det \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{bmatrix} = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$

Este determinante se llama el *determinante de Vandermonde* 3×3.

Ejercicio 2.19. Calcule $\det A$ dado que $A^2 = A$.

3 Espacios vectoriales

3.1 El espacio tri–dimensional $\mathbb{R}^3$

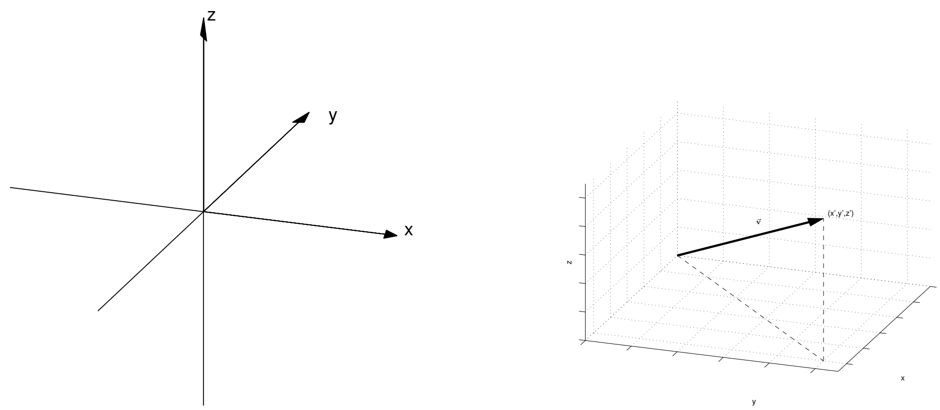
El sistema de *coordenadas cartesianas* en $\mathbb{R}^3$ consiste de tres rectas perpendiculares entre si llamadas los *ejes de coordenadas*. Normalmente estos ejes se denotan por las letras x, y, z. Suponemos que los ejes están orientados de acuerdo a la *regla de la mano derecha*. (Vea la Figura 3.1a.) El punto de intersección de los tres ejes de coordenadas se llama el *origen*. Cada par de éstas rectas determina un plano los cuales se llaman los *planos de coordenadas* Por ejemplo, el *plano de coordenadas xy* consiste de la intersección del eje de x con el de y, con definiciones similares para los planos de coordenadas xz, y yz. Los planos de coordenadas dividen a $\mathbb{R}^3$ en ocho regiones llamadas *octantes*.

Todo punto en $\mathbb{R}^3$ se puede representar con un triple (a, b, c), donde a representa la distancia (con signo) al plano yz; b es la distancia al plano xz; y c es la distancia al plano xy. Tenemos entonces que el origen tiene coordenadas $(0, 0, 0)$, y el plano xy consiste de todos los puntos (a, b, c) tal que $c = 0$, etc. El *primer octante* es el conjunto de todos los puntos (a, b, c) donde $a, b, c \geq 0$. Otros octantes se pueden definir de forma similar.

3.1.1 Vectores flechas

Un *vector flecha* $\vec{\mathbf{v}}$ en $\mathbb{R}^3$ consiste de un segmento dirigido que comienza en el origen según se muestra en la Figura 3.1b. Al conjunto de todas las flechas en $\mathbb{R}^3$ lo denotamos por $\mathbb{E}^3$. Note que un vector tiene *magnitud* y *dirección*. Ejemplos de vectores lo serían la fuerza, la velocidad, aceleración, etc.. Asociamos al vector flecha $\vec{\mathbf{v}}$ su punto terminal $(x', y', z') \in \mathbb{R}^3$, pero ambos son objetos distintos. Dos vectores flecha $\vec{\mathbf{v}}, \vec{\mathbf{w}}$ son *iguales* si y solo si tienen el mismo punto terminal.

Nota 3.1. Las coordenadas del punto terminal de un vector flecha $\vec{\mathbf{v}}$ son respecto al sistema de coordenadas cartesianas de $\mathbb{R}^3$. Los resultados que vamos a discutir sobre la relación entre multiplicación escalar y suma de flechas, no son necesariamente validos con respecto a otros sistemas de coordenadas. $\square$



(a) Ejes orientados según regla de la mano derecha.

(b) Vector flecha $\vec{\mathbf{v}}$ con punto terminal (x', y', z').

Figura 3.1: Ejes de coordenadas y vector flecha.

La *suma de las flechas* $\vec{\mathbf{v}}$ *y* $\vec{\mathbf{w}}$ se define geométricamente de acuerdo al diagrama de la Figura (3.2). Esto es, el origen y los puntos terminales de las flechas definen un plano y un paralelogramo en dicho plano. La flecha que representa a $\vec{\mathbf{v}} + \vec{\mathbf{w}}$ está dada por la diagonal de dicho paralelogramo que emana del origen. Usando argumentos de triángulos similares se puede demostrar ahora que:

Teorema 3.2. *Sean* $\vec{\mathbf{v}}$ *y* $\vec{\mathbf{w}}$ *elementos de* $\mathbb{E}^3$ *con puntos terminales* (x', y', z') *y* (x^*, y^*, z^*) *respectivamente. Entonces* $\vec{\mathbf{v}} + \vec{\mathbf{w}}$ *tiene punto terminal* $(x' + x^*, y' + y^*, z' + z^*)$.

Dado un vector $\vec{\mathbf{v}}$ y un escalar $\alpha \in \mathbb{R}$, entonces *la multiplicación escalar* $\alpha\vec{\mathbf{v}}$ representa el segmento dirigido de largo $|\alpha|$ veces el largo de $\vec{\mathbf{v}}$, en la misma dirección de $\vec{\mathbf{v}}$ si $\alpha \geq 0$, ó en la dirección opuesta de $\vec{\mathbf{v}}$ si $\alpha < 0$. Nuevamente usando un argumento geométrico se obtiene que:

Teorema 3.3. *Si* $\vec{\mathbf{v}}$ *tiene punto terminal* (x', y', z') *y* $\alpha \in \mathbb{R}$, *entonces* $\alpha\vec{\mathbf{v}}$ *tiene punto terminal* $(\alpha x', \alpha y', \alpha z')$.

Nota 3.4. Estos resultados dicen que el conjunto de flechas $\mathbb{E}^3$, con las operaciones que definimos de suma de flechas y multiplicación de una flecha por un escalar, es *isomorfo* a $\mathbb{R}^3$ con la suma de matrices y multiplicación escalar de $\mathbb{R}^3$. Es por esta razón que en las discusiones en este

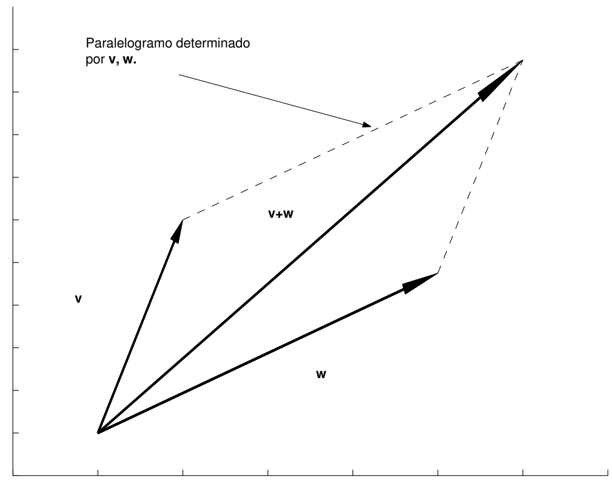


Figura 3.2: Suma de las flechas $\vec{\mathbf{v}}$ y $\vec{\mathbf{w}}$.

libro sobre flechas en $\mathbb{R}^3$, no haremos distinción entre la flecha como objeto geométrico (o sea, como elemento de $\mathbb{E}^3$), y su punto terminal. $\square$

Ejemplo 3.5. Sea $\vec{\mathbf{v}}$ la flecha con punto terminal $(-1,2,4)$ y $\vec{\mathbf{w}}$ la flecha con punto terminal $(5,-3,10)$. Entonces $\vec{\mathbf{v}}+\vec{\mathbf{w}}$ tiene punto terminal:

$$(-1,2,4)+(5,-3,10)=(-1+5,2-3,4+10)=(4,-1,14).$$

El punto terminal de $-2\vec{\mathbf{v}}$ es:

$$-2(-1,2,4)=(2,-4,-8).$$

$\square$

3.2 Definición de un espacio vectorial

Las operaciones de suma y multiplicación escalar en $\mathbb{R}^n$ y el conjunto de las matrices $m\times n$, tienen unas propiedades algebraicas bien especiales las cuales se recogen en el Teorema 1.19. Otros conjuntos, con otras operaciones de suma y multiplicación escalar, podrían tener propiedades similares. En general, estos conjuntos con sus operaciones de suma y multiplicación escalar se llaman *espacios vectoriales*.

Definición 3.6. Sea X un conjunto y suponga que existen dos funciones suma : $X \times X \longrightarrow X$ y msclr : $\mathbb{R} \times X \longrightarrow X$. Escribimos:

$$x +^* y = \text{suma}(x, y), \quad \alpha \cdot x = \text{msclr}(\alpha, x).$$

El triple $(X, +^*, \cdot)$ se llama un *espacio vectorial* si las siguientes aseveraciones se cumplen:

A1: $x +^* y = y +^* x$ para todo $x, y \in X$.

A2: $(x +^* y) +^* z = x +^* (y +^* z)$ para todo $x, y, z \in X$.

A3: Existe un elemento $\theta \in X$ tal que $x +^* \theta = x$ para todo $x \in X$.

A4: Correspondiente a cualquier $x \in X$, existe un elemento $-x \in X$ tal que $x +^* (-x) = \theta$.

A5: $\alpha \cdot (x +^* y) = \alpha \cdot x +^* \alpha \cdot y$, para todo $\alpha \in \mathbb{R}$, $x, y \in X$.

A6: $(\alpha + \beta) \cdot x = \alpha \cdot x +^* \beta \cdot x$, para todo $\alpha, \beta \in \mathbb{R}$, $x \in X$.

A7: $(\alpha\beta) \cdot x = \alpha \cdot (\beta \cdot x)$, para todo $\alpha, \beta \in \mathbb{R}$, $x \in X$.

A8: $1 \cdot x = x$ para todo $x \in X$.

Los elementos de X se llaman *vectores*.

Notación: Usualmente y cuando no halla confusión, escribiremos αx en lugar $\alpha \cdot x$.

Partiendo de la definición de un espacio vectorial se pueden derivar una serie de propiedades, las cuales son ciertas en todo espacio vectorial. Veamos algunas de estas propiedades.

Teorema 3.7. *Sea $(X, +^*, \cdot)$ un espacio vectorial. Entonces para cualquier $x \in X$, tenemos que:*

i) $0 \cdot x = \theta$

ii) *Si $x +^* y = \theta$, entonces $y = -x$. (¡Unicidad de los inversos aditivos!)*

iii) $(-1) \cdot x = -x$

Nota 3.8. Observe que $-x$ es el inverso aditivo de x, mientras que $(-1) \cdot x$ representa la operación de multiplicar el escalar -1 por el vector x. □

Demostración:

i)

$$\begin{aligned} x &= 1\cdot x, \quad \text{axioma A8,} \\ &= (1+0)\cdot x, \quad \text{axioma de } \mathbb{R}, \\ &= 1\cdot x +^* 0\cdot x, \quad \text{axioma A6,} \\ &= x +^* 0\cdot x, \quad \text{axioma A8.} \end{aligned}$$

Sumando $-x$ en ambos lados tenemos que:

$$\begin{aligned} -x +^* x &= -x +^* (x +^* 0\cdot x), \\ \theta &= (-x +^* x) +^* 0\cdot x, \quad \text{axiomas A2 y A4,} \\ \theta &= \theta +^* 0\cdot x, \quad \text{axioma A4,} \\ \theta &= 0\cdot x, \quad \text{axioma A3.} \end{aligned}$$

ii) Partiendo de $x +^* y = \theta$, tenemos que

$$\begin{aligned} -x +^* (x +^* y) &= -x +^* \theta, \\ (-x +^* x) +^* y &= -x, \quad \text{axiomas A2 y A3,} \\ \theta +^* y &= -x, \quad \text{axioma A4,} \\ y &= -x, \quad \text{axioma A3.} \end{aligned}$$

iii) De la parte (i) sabemos que $0\cdot x = \theta$. De modo que $(-1+1)\cdot x = 0\cdot x = \theta$. Así que:

$$\begin{aligned} (-1)\cdot x +^* 1\cdot x &= \theta, \quad \text{axioma A6,} \\ (-1)\cdot x +^* x &= \theta, \quad \text{axioma A8,} \\ x +^* (-1)\cdot x &= \theta, \quad \text{axioma A1,} \end{aligned}$$

Usando el resultado de la parte (ii), podemos concluir ahora que

$$(-1)\cdot x = -x$$

.

□

Ejemplo 3.9. Tomamos $X = \mathbb{R}^+$ como el conjunto de los números reales positivos. Definimos la suma y multiplicación escalar en $\mathbb{R}^+$ por:

$$x +^* y = xy, \quad \alpha \cdot x = x^\alpha, \quad x, y \in \mathbb{R}^+, \alpha \in \mathbb{R}.$$

(La suma de x y y en $\mathbb{R}^+$, se define como el producto en $\mathbb{R}$ de x y y.) Note que tanto $x +^* y$, como $\alpha \cdot x$ producen resultados en $\mathbb{R}^+$. Veamos que $(\mathbb{R}^+, +^*, \cdot)$ es un espacio vectorial. Para ver esto, verificamos los axiomas A1–A8 en la Definición 3.6. Este procedimiento utiliza varias de las propiedades de la multiplicación y la exponenciación en $\mathbb{R}$.

A1: $x +^* y = xy = yx = y +^* x$ para todo $x, y \in \mathbb{R}^+$.

A2: $(x +^* y) +^* z = (xy)z = x(yz) = x +^* (y +^* z)$ para todo $x, y, z \in \mathbb{R}^+$.

A3: Tomamos $\theta = 1$. Entonces $x +^* \theta = x +^* 1 = x(1) = x$. Por lo tanto, $\theta = 1$ es la identidad aditiva de $\mathbb{R}^+$.

A4: Para cualquier $x \in \mathbb{R}^+$, tenemos que $\frac{1}{x} \in \mathbb{R}^+$. Tenemos ahora que $x +^* \frac{1}{x} = x\left(\frac{1}{x}\right) = 1 = \theta$. Así que $-x = \frac{1}{x}$ es el inverso aditivo de x.

A5: $\alpha \cdot (x +^* y) = (xy)^\alpha = x^\alpha y^\alpha = (\alpha \cdot x)(\alpha \cdot y) = \alpha \cdot x +^* \alpha \cdot y$, para todo $\alpha \in \mathbb{R}$, $x, y \in \mathbb{R}^+$.

A6: $(\alpha + \beta) \cdot x = x^{\alpha+\beta} = x^\alpha x^\beta = (\alpha \cdot x)(\beta \cdot x) = \alpha \cdot x +^* \beta \cdot x$, para todo $\alpha, \beta \in \mathbb{R}$, $x \in \mathbb{R}^+$.

A7: $(\alpha\beta) \cdot x = x^{\alpha\beta} = \left(x^\beta\right)^\alpha = (\beta \cdot x)^\alpha = \alpha \cdot (\beta \cdot x)$, para todo $\alpha, \beta \in \mathbb{R}$, $x \in \mathbb{R}^+$.

A8: $1 \cdot x = x^1 = x$ para todo $x \in X$.

Podemos concluir ahora que $(\mathbb{R}^+, +^*, \cdot)$ es un espacio vectorial. □

Nota 3.10. De aquí en adelante, y siempre que no haya posibilidad de confusión, utilizaremos el símbolo "+" en lugar de "$+^*$" para representar la suma del espacio vectorial. Del contexto de la discusión se podrá inferir cuando este símbolo se refiere a la suma del espacio vectorial o a la suma de escalares en $\mathbb{R}$. □

3.3 Ejemplos importantes de espacios vectoriales

En esta sección vamos a describir varios espacios vectoriales que utilizaremos en el transcurso del libro como ejemplos para ilustrar la teoría y otros resultados.

No vamos a verificar los axiomas A1–A8 en cada caso, pero al indicar que dichos conjuntos con las correspondientes operaciones son espacios vectoriales, damos por entendido que estas propiedades se pueden verificar. El lector más curioso debe trabajar estas verificaciones por su cuenta.

Ejemplo 3.11. Para n, m dados, denotamos el conjunto de matrices $m \times n$ por:

$$M_{m\times n} = \{A | A \text{ es matriz } m \times n\}.$$

Recordamos que para $\alpha \in \mathbb{R}$ y $A, B \in M_{m\times n}$, la suma y multiplicación escalar se definen por:

$$A + B = (a_{ij} + b_{ij}), \quad \alpha \cdot A = (\alpha a_{ij}).$$

Veamos que $(M_{m\times n}, +, \cdot)$ es un espacio vectorial verificando los axiomas A1–A8 en la Definición 3.6. Este procedimiento utiliza varias de las propiedades de la suma y multiplicación en $\mathbb{R}$.

A1: $A + B = (a_{ij} + b_{ij}) = (b_{ij} + a_{ij}) = B + A$.

A2: La entrada ij de $(A + B) + C$ es $(a_{ij} + b_{ij}) + c_{ij}$. Pero

$$(a_{ij} + b_{ij}) + c_{ij} = a_{ij} + (b_{ij} + c_{ij}),$$

el cual es el elemento ij de $A+(B+C)$. Como ij es arbitrario tenemos que $(A + B) + C = A + (B + C)$.

A3: Tomando $O = (o_{ij})$ donde $o_{ij} = 0$ para toda i, j. tenemos que $A + O = (a_{ij} + o_{ij}) = (a_{ij} + 0) = (a_{ij}) = A$.

A4: Para $A = (a_{ij})$ definimos $-A = (-a_{ij})$. La entrada ij de $A + (-A)$ es $a_{ij} + (-a_{ij}) = a_{ij} - a_{ij} = 0 = o_{ij}$. Como ij es arbitrario tenemos que $A + (-A) = O$.

A5: La entrada ij de $\alpha(A+B)$ es $\alpha(a_{ij}+b_{ij}) = \alpha a_{ij} + \alpha b_{ij}$ el cual es el elemento ij de $\alpha A+\alpha B$. Como ij es arbitrario tenemos que $\alpha(A+B) = \alpha A + \alpha B$.

A6: $(\alpha+\beta)A = (\alpha+\beta)(a_{ij}) = ((\alpha+\beta)a_{ij}) = (\alpha a_{ij} + \beta a_{ij}) = (\alpha a_{ij}) + (\beta a_{ij}) = \alpha(a_{ij}) + \beta(a_{ij}) = \alpha A + \beta A$.

A7: El elemento ij de $(\alpha\beta)A$ es $(\alpha\beta)a_{ij} = \alpha(\beta a_{ij})$ el cual es el elemento ij de $\alpha(\beta A)$. Como ij es arbitrario tenemos que $(\alpha\beta)A = \alpha(\beta A)$.

A8: $(1)A = (1)(a_{ij}) = ((1)a_{ij}) = (a_{ij}) = A$.

Esto confirma de forma definitiva que $(M_{m\times n}, +, \cdot)$ es un espacio vectorial. El caso especial $n = 1$ lo denotamos por $\mathbb{R}^m$ y lo llamamos el *espacio euclidiano de dimensión m*. $\square$

Ejemplo 3.12. Para cualquier entero $n \geq 1$, definimos

$$P_n = \{p : \mathbb{R} \longrightarrow \mathbb{R} \,|\, p \text{ es polinomio de grado} < n\}.$$

Note que P_n consiste de todos los polinomios con coeficientes reales de grado menor que n. De aquí que un polinomio cualquiera $p \in P_n$, se puede escribir como $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots a_{n-1} x^{n-1}$, donde $a_0, a_1, \ldots, a_{n-1} \in \mathbb{R}$. Para $\alpha \in \mathbb{R}$ y $p, q \in P_n$, definimos $p+q$ y $\alpha \cdot p$ por:

$$(p+q)(x) = p(x) + q(x), \quad (\alpha \cdot p)(x) = \alpha p(x), \quad x \in \mathbb{R}.$$

Note que $p+q$ y $\alpha \cdot p$ pertenecen a P_n. (¿Por qué?) Verificamos ahora que $(P_n, +, \cdot)$ es un espacio vectorial. Suponemos en la discusión que $p, q, g \in P_n$ y que $\alpha, \beta \in \mathbb{R}$.

A1: Como $(p+q)(x) = p(x) + q(x) = q(x) + p(x) = (q+p)(x)$ para toda $x \in \mathbb{R}$, tenemos que $p+q = q+p$.

A2: Dado que

$$\begin{aligned} ((p+q)+g)(x) &= (p+q)(x) + g(x) = (p(x)+q(x)) + g(x), \\ &= p(x) + (q(x)+g(x)) = p(x) + (q+g)(x) \\ &= (p+(q+g))(x), \end{aligned}$$

para toda $x \in \mathbb{R}$, tenemos que $(p+q)+g = p+(q+g)$.

A3: Definimos θ por $\theta(x) = 0$ para toda x. Es claro que $\theta \in P_n$. Como $(p+\theta)(x) = p(x)+\theta(x) = p(x)+0 = p(x)$ para toda x, tenemos que $p+\theta = p$.

A4: Para $p \in P_n$, definimos $-p$ por $(-p)(x) = -p(x)$ para toda x. Tenemos que $-p \in P_n$ y que $(p+(-p))(x) = p(x)+(-p)(x) = p(x)+(-p(x)) = 0 = \theta(x)$ para toda x. Tenemos entonces que $p+(-p) = \theta$.

A5: Como

$$\begin{aligned}(\alpha\cdot(p+q))(x) &= \alpha(p+q)(x) = \alpha(p(x)+q(x)) \\ &= \alpha p(x)+\alpha q(x) = (\alpha\cdot p)(x)+(\alpha\cdot q)(x) \\ &= (\alpha\cdot p+\alpha\cdot q)(x),\end{aligned}$$

para toda x, tenemos que $\alpha\cdot(p+q) = \alpha\cdot p+\alpha\cdot q$.

A6: Como

$$\begin{aligned}((\alpha+\beta)\cdot p)(x) &= (\alpha+\beta)p(x) = \alpha p(x)+\beta p(x) \\ &= (\alpha\cdot p)(x)+(\beta\cdot p)(x) = (\alpha\cdot p+\beta\cdot p)(x),\end{aligned}$$

para toda x, podemos concluir que $(\alpha+\beta)\cdot p = \alpha\cdot p+\beta\cdot p$.

A7: Note que

$$\begin{aligned}((\alpha\beta)\cdot p)(x) &= (\alpha\beta)p(x) = \alpha(\beta p(x)) \\ &= \alpha(\beta\cdot p)(x) = (\alpha\cdot(\beta\cdot p))(x),\end{aligned}$$

para toda x, de donde obtenemos que $(\alpha\beta)\cdot p = \alpha\cdot(\beta\cdot p)$.

A8: Finalmente $(1\cdot p)(x) = (1)p(x) = p(x)$ para toda x, por lo que tenemos que $1\cdot p = p$.

$\square$

Un concepto bien importante asociado a un espacio vectorial es el de *dimensión*, el cual discutiremos más adelante en este capitulo. Por el momento solo comentamos que los ejemplos de $\mathbb{R}^m$, $M_{m\times n}$, y P_n, son de espacios vectoriales de dimensión finita. Los espacios en el siguiente ejemplo, son todos de *dimención infinita*.

Ejemplo 3.13. Sea $I \subset \mathbb{R}$ un intervalo. Para cualquier entero $n \geq 0$ definimos[1]

$$C^n(I) = \left\{ f : I \longrightarrow \mathbb{R} \mid f^{(j)} \text{ es continua en } I,\ 0 \leq j \leq n \right\}.$$

Para $\alpha \in \mathbb{R}$ y $f, g \in C^n(I)$, definimos $f + g$ y $\alpha \cdot f$ por:

$$(f+g)(x) = f(x) + g(x), \quad (\alpha \cdot f)(x) = \alpha f(x), \quad x \in I.$$

Usando que la propiedad de continuidad se preserva bajo suma de funciones y multiplicación escalar, tenemos que $f + g$ y $\alpha \cdot f$ pertenecen a $C^n(I)$. Tomando θ como la función cero sobre I, esto es

$$\theta(x) = 0, \quad x \in I,$$

y definiendo el inverso aditivo $-f$ como,

$$(-f)(x) = -f(x), \quad x \in I,$$

se puede verificar que $(C^n(I), +, \cdot)$ es un espacio vectorial. (La verificación de ésto es similar a la del Ejemplo 3.12.) El caso especial cuando $n = 0$ se denota por $C(I)$ y consiste de todas las funciones continuas sobre I. □

3.4 Subespacios

En muchas ocasiones tendremos que considerar algún subconjunto de alguno de nuestros espacios vectoriales básicos $\mathbb{R}^m$, $M_{m\times n}$, P_n, y $C^n(I)$, En tal situación puede ser de interés, asegurar o determinar, si dicho subconjunto tiene propiedades algebraicas que de alguna forma "hereda" del espacio original. En tal caso decimos que el subconjunto es un *subespacio* del espacio vectorial original.

Definición 3.14. Sea $(V, +, \cdot)$ un espacio vectorial y $S \subset V$. Decimos que S es un *subespacio* de V si:

i) $\alpha \cdot x \in S$ para todo $\alpha \in \mathbb{R}$, $x \in S$.

ii) $x + y \in S$ para todo $x, y \in S$.

[1]Cuando $I = [a, b]$, escribimos $C^n[a, b]$ en lugar de $C^n([a, b])$.

Esto es, S es un subespacio de V, si S es *cerrado* bajo las operaciones de suma y multiplicación escalar de V.

Nota 3.15. Se desprende de la definición que, si S es un subespacio de V, entonces $(S,+,\cdot)$ es un espacio vectorial por si solo, con las operaciones de suma y multiplicación escalar de V. □

Nota 3.16. Las condiciones (i) y (ii) en la definición de subespacio, se pueden verificar de forma conjunta de la forma siguiente:

$$\alpha \cdot x + \beta \cdot y \in S \text{ para todo } \alpha, \beta \in \mathbb{R}, \quad x, y \in S.$$

□

Ejemplo 3.17. Sea S el subconjunto de $\mathbb{R}^2$ dado por

$$S = \left\{ (x_1, x_2)^t \, : \, x_1, x_2 \geq 0 \right\}.$$

Note que si $\vec{x}, \vec{y} \in S$, entonces

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad x_1, x_2 \geq 0,$$

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad y_1, y_2 \geq 0.$$

Tenemos ahora que

$$\vec{x} + \vec{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix} \in S,$$

ya que $x_1 + y_1, x_2 + y_2 \geq 0$. De modo que S tiene clausura bajo la suma de $\mathbb{R}^2$. No obstante, no tiene clausura bajo multiplicación escalar ya que si tomamos $\alpha = -1$, $\vec{x} = (1,2)^t$, entonces

$$\alpha \vec{x} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} \notin S.$$

Así que S no es un subespacio de $\mathbb{R}^2$. □

Ejemplo 3.18. Considere el siguiente subconjunto de $M_{2\times 2}$:

$$D_{2\times 2} = \left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} : a, b \in \mathbb{R} \right\}.$$

Si $\alpha \in \mathbb{R}$ y $A \in D_{2\times 2}$, entonces

$$\alpha A = \alpha \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = \begin{bmatrix} \alpha a & \alpha 0 \\ \alpha 0 & \alpha b \end{bmatrix} = \begin{bmatrix} \alpha a & 0 \\ 0 & \alpha b \end{bmatrix} \in D_{2\times 2}.$$

Sean $A_1, A_2 \in D_{2\times 2}$ donde

$$A_1 = \begin{bmatrix} a_1 & 0 \\ 0 & b_1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} a_2 & 0 \\ 0 & b_2 \end{bmatrix}.$$

Entonces

$$A_1 + A_2 = \begin{bmatrix} a_1 + a_2 & 0 + 0 \\ 0 + 0 & b_1 + b_2 \end{bmatrix} = \begin{bmatrix} a_1 + a_2 & 0 \\ 0 & b_1 + b_2 \end{bmatrix} \in D_{2\times 2}.$$

Podemos entonces concluir que $D_{2\times 2}$ es un subespacio de $M_{2\times 2}$. □

Ejemplo 3.19. Sea

$$S = \left\{ f \in C[a,b] : \int_a^b f(x)\,\mathrm{d}x = 0 \right\}.$$

Note que $S \subset C[a,b]$. Sean $\alpha, \beta \in \mathbb{R}$ y $f, g \in S$. Entonces, por propiedades de funciones continuas, $\alpha f + \beta g$ es continua, es decir $\alpha f + \beta g \in C[a,b]$. También, como $f, g \in S$, tenemos que

$$\int_a^b f(x)\,\mathrm{d}x = 0, \quad \int_a^b g(x)\,\mathrm{d}x = 0.$$

De aquí que:

$$\begin{aligned} \int_a^b (\alpha f + \beta g)(x)\,\mathrm{d}x &= \int_a^b (\alpha f(x) + \beta g(x))\,\mathrm{d}x, \\ &= \alpha \int_a^b f(x)\,\mathrm{d}x + \beta \int_a^b g(x)\,\mathrm{d}x, \\ &= \alpha(0) + \beta(0) = 0, \end{aligned}$$

es decir $\alpha f + \beta g \in S$. Podemos ahora concluir que S es un subespacio de $C[a,b]$. □

Ejemplo 3.20. Sea

$$S = \{f \in C[-1,1] : f \text{ es impar}\}.$$

Note que $S \subset C[-1,1]$. Recuerde que $f \in C[-1,1]$ es impar si $f(-x) = -f(x)$ para toda $x \in [-1,1]$. Sean $\alpha, \beta \in \mathbb{R}$ y $f, g \in S$. Entonces, f, g son impares y por propiedades de funciones continuas, $\alpha f + \beta g \in C[-1,1]$. Veamos que $\alpha f + \beta g$ es impar:

$$\begin{aligned} (\alpha f + \beta g)(-x) &= \alpha f(-x) + \beta g(-x), \\ &= \alpha(-f(x)) + \beta(-g(x)), \\ &= -(\alpha f(x) + \beta g(x)), \\ &= -(\alpha f + \beta g)(x). \end{aligned}$$

Podemos entonces concluir que S es un subespacio de $C[-1,1]$. □

Más adelante veremos que toda matriz A tiene asociada a ésta, cuatro espacios fundamentales. Uno de estos espacios fundamentales es el *espacio nulo* o *núcleo* de la matriz.

Definición 3.21. Sea A una matriz $m \times n$. El *espacio nulo* de A se define por:

$$N(A) = \left\{\vec{\mathbf{x}} \in \mathbb{R}^n : A\vec{\mathbf{x}} = \vec{\mathbf{0}}\right\}.$$

Ejemplo 3.22. ¿Cómo verificamos si un $\vec{\mathbf{x}}$ pertenece o no al espacio nulo de una matriz A? Pues multiplicamos a $\vec{\mathbf{x}}$ por A, y verificamos si el resultado es el vector cero. Suponga que

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

Si $\vec{\mathbf{x}} = (1, -3, 2)^t$, entonces como

$$A\vec{\mathbf{x}} = \begin{bmatrix} 1 & 4 & 5 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

podemos concluir que $\vec{\mathbf{x}} \notin N(A)$. Para $\vec{\mathbf{y}} = (1, 1, -1)^t$ tenemos

$$A\vec{\mathbf{y}} = \begin{bmatrix} 1 & 4 & 5 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

por lo que $\vec{\mathbf{y}} \in N(A)$. □

Ejemplo 3.23. Vamos a calcular el espacio nulo de la matriz

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

Para lograr ésto, tenemos que buscar todas las soluciones del sistema homogéneo $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$. Para esto reducimos la matriz aumentada a su forma echelon:

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 0 \\ 2 & 1 & 3 & 0 \end{array}\right] \underset{\frac{1}{3}f_3}{\xrightarrow{f_2 = f_2 - 2f_1}} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 0 \\ 0 & 1 & -\frac{1}{3} & 0 \end{array}\right].$$

De aquí que $x_1 = -\frac{5}{3}x_3$, $x_2 = \frac{1}{3}x_3$. Poniendo $x_3 = \alpha$, tenemos que

$$N(A) = \left\{ \alpha \begin{bmatrix} -5/3 \\ 1/3 \\ 1 \end{bmatrix} : \alpha \in \mathbb{R} \right\}.$$

□

Proposición 3.24. *Sea A una matriz $m \times n$. Entonces $N(A)$ es un subespacio de $\mathbb{R}^n$.*

Demostración: Note que $N(A) \subset \mathbb{R}^n$. Sean $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in N(A)$. Entonces

$$A\vec{\mathbf{x}} = \vec{\mathbf{0}}, \quad A\vec{\mathbf{y}} = \vec{\mathbf{0}}.$$

Ahora si $\alpha, \beta \in \mathbb{R}$, tenemos que

$$A(\alpha\vec{\mathbf{x}} + \beta\vec{\mathbf{y}}) = \alpha A\vec{\mathbf{x}} + \beta A\vec{\mathbf{y}} = \alpha\vec{\mathbf{0}} + \beta\vec{\mathbf{0}} = \vec{\mathbf{0}},$$

es decir que $\alpha\vec{\mathbf{x}} + \beta\vec{\mathbf{y}} \in N(A)$. Podemos concluir entonces que $N(A)$ es un subespacio de $\mathbb{R}^n$. □

3.5 Conjuntos generadores

Un concepto central en el álgebra lineal es el de *combinación lineal*. Considere el sistema lineal $A\vec{x} = \vec{b}$, donde

$$A = \begin{bmatrix} -2 & 4 & 1 \\ 3 & -6 & 2 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}.$$

Escribiendo $\vec{x} = (x, y, z)^t$, el sistema es equivalente a la ecuación:

$$x\begin{bmatrix} -2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 4 \\ -6 \end{bmatrix} + z\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}.$$

El lado izquierdo de esta ecuación es una *combinación lineal* de las columnas de A. De modo que la pregunta de si el sistema tiene solución o no, es equivalente a preguntarse si es posible escribir el lado derecho del sistema como una combinación lineal de las columnas de A.

Definición 3.25. Sea $(V, +, \cdot)$ un espacio vectorial y $v_1, v_2, \ldots, v_n$ vectores en V. Para cualesquiera escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$, la expresión

$$\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \cdots + \alpha_n \cdot v_n,$$

de llama una *combinación lineal!definición* de los vectores $v_1, v_2, \ldots, v_n$. El conjunto de todas las combinaciones lineales de los vectores $v_1, v_2, \ldots, v_n$, se llama el *conjunto o subespacio generado* por los vectores $v_1, v_2, \ldots, v_n$ y se denota[2] por $\operatorname{span}\{v_1, v_2, \ldots, v_n\}$. En notación de conjuntos*combinación lineal!conjunto generado*,

$$\operatorname{span}\{v_1, v_2, \ldots, v_n\} = \{\alpha_1 \cdot v_1 + \cdots + \alpha_n \cdot v_n \,|\, \alpha_i \in \mathbb{R}, 1 \le i \le n\}.$$

Ejemplo 3.26. Con los vectores

$$\vec{v}_1 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 4 \\ -6 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

podemos formar la siguiente combinación lineal:

$$2\vec{v}_1 - 3\vec{v}_2 + 2\vec{v}_3 = 2\begin{bmatrix} -2 \\ 3 \end{bmatrix} - 3\begin{bmatrix} 4 \\ -6 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -14 \\ 28 \end{bmatrix}.$$

Poniendo $A = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$ y $\vec{b} = (4, 2)^t$, tenemos que el sistema $A\vec{x} = \vec{b}$ tiene solución si y solo si $\vec{b}$ es combinación lineal de $\vec{v}_1, \vec{v}_2, \vec{v}_3$, es decir, si $\vec{b} \in \operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. □

Ejemplo 3.27. En el Ejemplo 3.23, podemos escribir que $N(A) = \operatorname{span}\{\vec{v}\}$, donde $\vec{v} = (-5/3, 1/3, 1)^t$. □

[2]En ocasiones, poniendo que $S = \{v_1, v_2, \ldots, v_n\}$, escribiremos $\operatorname{span} S$ en lugar de $\operatorname{span}\{v_1, v_2, \ldots, v_n\}$.

Ejemplo 3.28. El conjunto

$$S = \left\{a + bx^2 \,:\, a, b \in \mathbb{R}\right\},$$

que es un subespacio de P_3, se puede escribir como $S = \operatorname{span}\left\{1, x^2\right\}$ ya que para cualquier $p \in S$,

$$p(x) = a + bx^2 = a \cdot 1 + b \cdot x^2.$$

□

Ejemplo 3.29. Como

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

tenemos que el subespacio $D_{2\times 2}$ del Ejemplo 3.18, se puede escribir como

$$D_{2\times 2} = \operatorname{span}\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}.$$

□

Teorema 3.30. *Sea $(V, +, \cdot)$ un espacio vectorial y $v_1, v_2, \ldots, v_n$ vectores en V. Entonces* $\operatorname{span}\{v_1, v_2, \ldots, v_n\}$ *es un subespacio de V.*

Demostración: Sea $S = \operatorname{span}\{v_1, v_2, \ldots, v_n\}$. Claramente $S \subset V$. Si $x, y \in S$, entonces existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ y $\beta_1, \beta_2, \ldots, \beta_n$ tal que

$$x = \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \cdots + \alpha_n \cdot v_n, \quad y = \beta_1 \cdot v_1 + \beta_2 \cdot v_2 + \cdots + \beta_n \cdot v_n.$$

Tenemos ahora que si $\alpha, \beta \in \mathbb{R}$, entonces

$$\begin{aligned}
\alpha \cdot x + \beta \cdot y &= \alpha \cdot (\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \cdots + \alpha_n \cdot v_n) \\
&\quad + \beta \cdot (\beta_1 \cdot v_1 + \beta_2 \cdot v_2 + \cdots + \beta_n \cdot v_n), \\
&= ((\alpha\alpha_1) \cdot v_1 + (\alpha\alpha_2) \cdot v_2 + \cdots + (\alpha\alpha_n) \cdot v_n) \\
&\quad + ((\beta\beta_1) \cdot v_1 + (\beta\beta_2) \cdot v_2 + \cdots + (\beta\beta_n) \cdot v_n), \\
&= (\alpha\alpha_1 + \beta\beta_1) \cdot v_1 + (\alpha\alpha_2 + \beta\beta_2) \cdot v_2 \\
&\quad + \cdots + (\alpha\alpha_n + \beta\beta_n) \cdot v_n \in S.
\end{aligned}$$

Tenemos entonces que S es un subespacio de V. $\square$

Este teorema nos dice que $\operatorname{span}\{v_1, v_2, \ldots, v_n\}$ siempre es un subespacio de V. Nos interesa ahora la pregunta en la otra dirección, esto es, dado un espacio vectorial V, ¿es posible encontrar un conjunto de vectores $\{v_1, v_2, \ldots, v_n\}$ en V de modo que $V = \operatorname{span}\{v_1, v_2, \ldots, v_n\}$? La siguiente definición recoge ésta idea.

Definición 3.31. Los vectores $\{v_1, v_2, \ldots, v_n\}$ del espacio vectorial V son un *conjunto generador* para V si

$$V = \operatorname{span}\{v_1, v_2, \ldots, v_n\}.$$

Esto es, todo $v \in V$ se puede escribir como una combinación lineal de los vectores $v_1, v_2, \ldots, v_n$.

Nota 3.32. $V = \operatorname{span}\{v_1, v_2, \ldots, v_n\}$ si para todo vector $v \in V$ existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ tal que

$$v = \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \cdots + \alpha_n \cdot v_n.$$

$\square$

Ejemplo 3.33. El conjunto de $\mathbb{R}^3$:

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\},$$

es un conjunto generador de $\mathbb{R}^3$. Para ver ésto, tomamos $\vec{b} = (a, b, c)^t \in \mathbb{R}^3$ arbitrario. Verificamos que existen escalares c_1, c_2, c_3 tal que:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}.$$

Esta ecuación es equivalente al sistema:

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Como el determinante de la matriz de coeficientes de este sistema es 2, que es distinto de cero, tenemos que este sistema tiene solución para cualquier $(a,b,c)^t \in \mathbb{R}^3$, por lo que el conjunto dado de vectores es un conjunto generador para $\mathbb{R}^3$. $\square$

Ejemplo 3.34. Veamos que $\left\{x-2, x+1, x^2+1\right\}$ es un conjunto generador para P_3. Si $a+bx+cx^2 \in P_3$, entonces buscamos escalares c_1, c_2, c_3 tal que:

$$a+bx+cx^2 = c_1(x-2)+c_2(x+1)+c_3(x^2+1).$$

Esto es equivalente a:

$$a+bx+cx^2 = -2c_1+c_2+c_3+(c_1+c_2)x+c_3x^2.$$

Igualando potencias de x en ambos lados obtenemos que:

$$\left\{\begin{array}{rcl} -2c_1+c_2+c_3 &=& a, \\ c_1+c_2 \quad &=& b, \\ c_3 &=& c. \end{array}\right.$$

La matriz de coeficientes de este sistema es nosingular por lo que el sistema tiene solución para cualesquiera (a,b,c). Como $a+bx+cx^2 \in P_3$ es arbitrario, tenemos que $\operatorname{span}\left\{x-2, x+1, x^2+1\right\} = P_3$. $\square$

Ejemplo 3.35. Veamos que

$$\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ -1 \end{bmatrix}\right\},$$

es generador para $\mathbb{R}^2$. Sea $(a,b)^t \in \mathbb{R}^2$. Buscamos c_1, c_2, c_3 tal que:

$$\begin{bmatrix} a \\ b \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} 2 \\ -1 \end{bmatrix} + c_3\begin{bmatrix} 5 \\ -1 \end{bmatrix}.$$

Esto es equivalente al sistema:

$$\begin{bmatrix} 1 & 2 & 5 \\ 1 & -1 & -1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

La forma echelon de la matriz aumentada de este sistema es:

$$\left[\begin{array}{ccc|c} 1 & 2 & 5 & a \\ 0 & 1 & 2 & \frac{a-b}{3} \end{array}\right].$$

Así que el sistema es consistente–dependiente por lo que siempre es posible encontrar una solución (c_1, c_2, c_3) del sistema. (De hecho hay un número infinito de soluciones.) Como $(a,b)^t \in \mathbb{R}^2$ es arbitrario, tenemos que:

$$\operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ -1 \end{bmatrix}\right\} = \mathbb{R}^2.$$

□

3.6 Independencia lineal

En el Ejemplo 3.35, también se puede verificar que:

$$\operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}\right\} = \mathbb{R}^2.$$

¿Qúe pasa con el vector $(5,-1)^t$? Note que:

$$(1)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (2)\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}. \tag{3.1}$$

Esto implica que cualquier combinación lineal de:

$$\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ -1 \end{bmatrix}\right\},$$

se puede escribir como una combinación lineal de:

$$\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}\right\}.$$

Veamos:

$$c_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} 2 \\ -1 \end{bmatrix} + c_3\begin{bmatrix} 5 \\ -1 \end{bmatrix} \quad = \quad c_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$
\begin{aligned}
& +c_3\left((1)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (2)\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right), \\
= \quad & (c_1+c_3)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (c_2+2c_3)\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \\
\in \quad & \operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}\right\}.
\end{aligned}
$$

Esto es:

$$\operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}\right\} = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ -1 \end{bmatrix}\right\}.$$

Note que (3.1) se puede escribir como:

$$(1)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (2)\begin{bmatrix} 2 \\ -1 \end{bmatrix} + (-1)\begin{bmatrix} 5 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

esto es, los vectores $\{(1,1)^t, (2,-1)^t, (5,-1)^t\}$ son *linealmente dependientes*.

Definición 3.36. Los vectores $\{v_1, v_2, \ldots, v_n\}$ del espacio vectorial V son *linealmente dependientes* si podemos encontrar escalares $c_1, c_2, \ldots, c_n$, no todos cero, tal que:

$$c_1 \cdot v_1 + c_2 \cdot v_2 + \cdots + c_n \cdot v_n = \theta,$$

donde θ es el vector cero de V. De lo contrario, si esto no es posible, decimos que los vectores $\{v_1, v_2, \ldots, v_n\}$ son *linealmente independientes*.

Nota 3.37. Si $\{v_1, v_2, \ldots, v_n\}$ son linealmente dependientes, entonces al menos uno de estos vectores es combinación lineal de los restantes vectores. □

Nota 3.38. Si $\{v_1, v_2, \ldots, v_n\}$ son linealmente independientes, entonces la única forma en que:

$$c_1 \cdot v_1 + c_2 \cdot v_2 + \cdots + c_n \cdot v_n = \theta,$$

es con $c_1 = c_2 = \cdots = c_n = 0$. Así que para verificar si un conjunto de vectores es linealmente independiente, hay que verificar que la única combinación lineal de éstos vectores que es igual al vector cero, es la que tiene todos los coeficientes igual a cero. □

Ejemplo 3.39. Veamos que:

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},$$

son linealmente independientes en $\mathbb{R}^3$. Para ver esto suponemos que hay c_1, c_2 tal que:

$$c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}.$$

Pero ésta ecuación es equivalente al sistema:

$$c_1 + c_2 = 0, \quad c_1 = 0, \quad c_2 = 0,$$

que evidentemente tiene como única solución $c_1 = c_2 = 0$. Así que los vectores $\vec{v}_1, \vec{v}_2$ son linealmente independientes en $\mathbb{R}^3$. $\square$

Ejemplo 3.40. Los polinomios $\left\{1, 1+x, x^2-1\right\}$ son linealmente independientes en P_3. Para ver esto, supongamos que existen escalares c_1, c_2, c_3 tal que:

$$c_1(1) + c_2(1+x) + c_3(x^2-1) = 0.$$

Esta ecuación se puede escribir como:

$$(c_1 + c_2 - c_3) + c_2x + c_3x^2 = 0.$$

Igualando coeficientes de x en ambos lados, tenemos que:

$$c_1 + c_2 - c_3 = 0, \quad c_2 = 0, \quad c_3 = 0,$$

es decir, $c_1 = c_2 = c_3 = 0$ por lo que los polinomios dados son linealmente independientes. $\square$

Ejemplo 3.41. Las matrices:

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

son linealmente independientes en $M_{2\times 2}$. Para ver esto, note que si

$$c_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c_2\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_3\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

entonces

$$\begin{bmatrix} c_1 & c_3 \\ c_2+c_3 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

De aquí que

$$c_1 = 0, \quad c_3 = 0, \quad c_2 + c_3 = 0,$$

lo que implica que $c_1 = c_2 = c_3 = 0$. □

Ejemplo 3.42. Los vectores:

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \right\},$$

son linealmente dependientes en $\mathbb{R}^3$. Para ver esto, note que:

$$c_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

es equivalente al sistema:

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Como la matriz de coeficientes de éste sistema es singular, este sistema tiene soluciones no triviales, ésto es, con c_1, c_2, c_3 no todos cero. Por lo tanto los vectores dados son linealmente dependientes. □

Este ejemplo es un caso especial del siguiente teorema.

Teorema 3.43. *Sean $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$ vectores de $\mathbb{R}^n$ y defina la matriz A $n \times n$ por:*

$$A = [\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n].$$

Entonces los vectores $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$ son linealmente independientes si y solo si la matriz A es nosingular.

Hay un resultado parecido al anterior pero para funciones. Para presentar éste otro resultado, necesitamos discutir primero el concepto del wronskiano.

Definición 3.44. Sean $\{f_1, f_2, \dots, f_n\}$ funciones en $C^{(n-1)}(I)$. El *wronskiano* de éstas funciones está dado por:

$$W(x) = \det\begin{bmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{bmatrix}, \quad x \in I.$$

Nota 3.45. Observe que W es una función de x. Para mantener la notación simple, no escribimos explícitamente la dependencia de W en las funciones $\{f_1, f_2, \dots, f_n\}$. □

Ejemplo 3.46.

a) Para $\left\{x^2+1, x^3-2\right\}$ tenemos con $I = \mathbb{R}$, que:

$$W(x) = \det\begin{bmatrix} x^2+1 & x^3-2 \\ 2x & 3x^2 \end{bmatrix} = 3x^2(x^2+1) - 2x(x^3-2) = x^4 + 3x^2 + 4x.$$

b) Para $\left\{e^{2x}, e^{3x}\right\}$ tenemos con $I = \mathbb{R}$, que:

$$W(x) = \det\begin{bmatrix} e^{2x} & e^{3x} \\ 2e^{2x} & 3e^{3x} \end{bmatrix} = 3e^{5x} - 2e^{5x} = e^{5x}.$$

c) Para $\left\{x^2, \operatorname{sen} x, \cos x\right\}$ tenemos con $I = \mathbb{R}$, que:

$$W(x) = \det\begin{bmatrix} x^2 & \operatorname{sen} x & \cos x \\ 2x & \cos x & -\operatorname{sen} x \\ 2 & -\operatorname{sen} x & -\cos x \end{bmatrix} = -x^2 - 2.$$

d) Para $\left\{x^{-1}, \ln x\right\}$ tenemos con $I = (0, \infty)$, que:

$$W(x) = \det\begin{bmatrix} x^{-1} & \ln x \\ -x^{-2} & x^{-1} \end{bmatrix} = x^{-2}(1 + \ln x).$$

□

Tenemos ahora:

Teorema 3.47. *Sean $f_1, f_2, \dots, f_n \in C^{(n-1)}(I)$. Entonces:*

i) Si $\{f_1, f_2, \dots, f_n\}$ son linealmente dependientes, entonces $W(x) = 0$ para todo $x \in I$.

ii) Si $W(x_0) \neq 0$ para algún $x_0 \in I$, entonces $\{f_1, f_2, \dots, f_n\}$ son linealmente independientes.

Demostración:

i) Si $\{f_1, f_2, \dots, f_n\}$ son linealmente dependientes, existen $c_1, c_2, \dots, c_n$ no todos cero tal que:

$$c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) = 0, \quad \forall\, x \in I.$$

Si diferenciamos ésta ecuación $n-1$ veces, y montando las ecuaciones resultantes en un sistema, obtenemos que:

$$\begin{bmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \forall\, x \in I.$$

De aquí que este sistema tiene solución no trivial para cada $x \in I$. Por lo tanto la matriz de coeficientes de este sistema, es singular para cada $x \in I$. Pero el determinante de ésta matriz es $W(x)$. Así que $W(x) = 0$ para toda $x \in I$.

ii) Si $\{f_1, f_2, \dots, f_n\}$ fueran linealmente dependientes, por la parte (i) tendríamos que $W(x) = 0$ para toda $x \in I$. Pero $W(x_0) \neq 0$, así que $\{f_1, f_2, \dots, f_n\}$ son linealmente independientes.

□

Ejemplo 3.48. Usando este teorema y en referencia al Ejemplo 3.46, tenemos que:

a) $\left\{x^2 + 1, x^3 - 2\right\}$ son linealmente independientes, digamos en P_4 ó $C(\mathbb{R})$.

b) $\left\{e^{2x}, e^{3x}\right\}$ son linealmente independientes en $C(\mathbb{R})$.

c) $\left\{x^2, \operatorname{sen} x, \cos x\right\}$ son linealmente independientes en $C(\mathbb{R})$.

d) $\left\{x^{-1}, \ln x\right\}$ son linealmente independientes en $C((0,\infty))$.

$\square$

Ejemplo 3.49. Si $f_1, f_2, \ldots, f_n$ son linealmente dependientes, entonces $W(x) = 0$ para toda $x \in I$. Pero $W(x) = 0$ para toda $x \in I$ no implica necesariamente que $f_1, f_2, \ldots, f_n$ son linealmente dependientes. Consideremos las funciones $f_1(x) = x^2$, $f_2(x) = x|x|$ sobre $I = [-1, 1]$. Entonces es fácil ver que:

$$W(x) = \det\begin{bmatrix} x^2 & x|x| \\ 2x & 2|x| \end{bmatrix} = 2x^2|x| - 2x^2|x| = 0,$$

para todo $x \in I$. Pero $\{f_1, f_2\}$ son linealmente independientes ya que si $c_1 f_1(x) + c_2 f_2(x) = 0$ para todo $x \in I$, entonces poniendo $x = -1, 1$ en ésta ecuación, obtenemos el sistema:

$$c_1 + c_2 = 0, \quad c_1 - c_2 = 0,$$

que tiene solución $c_1 = c_2 = 0$. $\square$

3.7 Bases y dimensión

En el Ejemplo (3.35) y la discusión al principio de la Sección 3.6, vimos que si un conjunto de vectores es linealmente dependiente, entonces es posible eliminar o sacar un vector del conjunto, sin afectar el conjunto generado por éstos. Ese ejemplo es un caso especial del siguiente resultado.

Teorema 3.50. *Sean $\{v_1, v_2, \ldots, v_n\}$ vectores linealmente dependientes en el espacio vectorial V. Entonces para algún índice i, $1 \le i \le n$, tenemos que*

$$\operatorname{span}\{v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\} = \operatorname{span}\{v_1, v_2, \ldots, v_n\}.$$

Demostración: Como $\{v_1, v_2, \ldots, v_n\}$ son linealmente dependientes, existen $c_1, c_2, \ldots, c_n$ no todos cero, tal que:

$$c_1 \cdot v_1 + c_2 \cdot v_2 + \cdots + c_n \cdot v_n = \theta.$$

Sin perdida de generalidad, podemos suponer que $c_1 \neq 0$. Entonces podemos reescribir la ecuación anterior como:

$$v_1 = -\left(\frac{c_2}{c_1}\right) \cdot v_2 - \cdots - \left(\frac{c_n}{c_1} \cdot\right) v_n.$$

Tenemos ahora que

$$\begin{aligned}
\sum_{k=1}^{n} \alpha_k \cdot v_k &= \alpha_1 \cdot \left[-\left(\frac{c_2}{c_1}\cdot\right) v_2 - \cdots - \left(\frac{c_n}{c_1}\cdot\right) v_n \right] + \sum_{k=2}^{n} \alpha_k \cdot v_k, \\
&= \sum_{k=2}^{n} \left[\alpha_k - \alpha_1 \left(\frac{c_k}{c_1}\right) \right] \cdot v_k,
\end{aligned}$$

lo que implica que

$$\operatorname{span}\{v_1, v_2, \dots, v_n\} \subseteq \operatorname{span}\{v_2, \dots, v_n\}.$$

Como siempre

$$\operatorname{span}\{v_2, \dots, v_n\} \subseteq \operatorname{span}\{v_1, v_2, \dots, v_n\},$$

tenemos que

$$\operatorname{span}\{v_1, v_2, \dots, v_n\} = \operatorname{span}\{v_2, \dots, v_n\}.$$

□

De modo que en un conjunto de vectores linealmente dependientes, siempre podemos eliminar al menos uno de los vectores sin afectar el conjunto generado por los vectores. Por el contrario, si los vectores son linealmente independientes, no es posible eliminar un vector sin alterar el conjunto generado por los vectores. Esto motiva la siguiente definición.

Definición 3.51. Los vectores $\{v_1, v_2, \dots, v_n\}$ forman una *base* del espacio vectorial V si

i) $\{v_1, v_2, \dots, v_n\}$ son linealmente independientes;

ii) $V = \operatorname{span}\{v_1, v_2, \dots, v_n\}$.

Nota 3.52. Como veremos más adelante, todo espacio vectorial tiene un número infinito de bases, todas con el mismo número de elementos. □

Ejemplo 3.53. Veamos que $\left\{1, 1+x, x^2-3\right\}$ es una base de P_3. Primero verificamos que son linealmente independientes. Si c_1, c_2, c_3 son tal que

$$c_1(1) + c_2(1+x) + c_3(x^2-3) = 0,$$

entonces

$$c_1 + c_2 - 3c_3 + c_2 x + c_3 x^2 = 0.$$

Igualando potencias de x en ambos lados de esta ecuación, obtenemos que $c_1 = c_2 = c_3 = 0$, es decir, que $\left\{1, 1+x, x^2-3\right\}$ son linealmente independientes.

Veamos ahora que $P_3 = \operatorname{span}\left\{1, 1+x, x^2-3\right\}$. Para ver ésto, tomamos un $p \in P_3$ arbitrario, y verificamos que existen escalares c_1, c_2, c_3 tal que

$$c_1(1) + c_2(1+x) + c_3(x^2-3) = p(x).$$

Pero $p \in P_3$ implica que $p(x) = a_1 + a_2 x + a_3 x^2$ para algunos escalares (conocidos) a_1, a_2, a_3. Sustituyendo en la ecuación anterior, y agrupando las potencias similares de x, tenemos que

$$c_1 + c_2 - 3c_3 + c_2 x + c_3 x^2 = a_1 + a_2 x + a_3 x^2.$$

Igualando las potencias de x en ambos lados de esta ecuación, obtenemos el sistema (para c_1, c_2, c_3):

$$\left\{ \begin{array}{rcl} c_1 + c_2 - 3c_3 & = & a_1, \\ c_2 & = & a_2, \\ c_3 & = & a_3. \end{array} \right.$$

Es fácil ver que este sistema tiene la solución:

$$c_1 = a_1 - a_2 + 3a_3, \quad c_2 = a_2, \quad c_3 = a_3.$$

Como el $p \in P_3$ que usamos es arbitrario, esto verifica que

$$P_3 = \operatorname{span}\left\{1, 1+x, x^2-3\right\},$$

que combinado con la independencia lineal de $\left\{1, 1+x, x^2-3\right\}$ implica que $\left\{1, 1+x, x^2-3\right\}$ es una base de P_3. $\square$

Ejemplo 3.54. Sea $S_{2\times 2}$ el conjunto de matrices simétricas 2×2. Esto es

$$S_{2\times 2} = \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} : a, b, c \in \mathbb{R} \right\}.$$

$S_{2\times 2}$ es un subespacio de $M_{2\times 2}$. Veamos que $\{A_1, A_2, A_3\}$ es una base de $S_{2\times 2}$, donde

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Note que si $c_1A_1 + c_2A_2 + c_3A_3 = O$, donde O es la matriz cero, entonces

$$\begin{bmatrix} c_1 & c_2 \\ c_2 & c_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

lo que implica que $c_1 = c_2 = c_3 = 0$. De modo que $\{A_1, A_2, A_3\}$ son linealmente independientes.

Para $B \in S_{2\times 2}$ arbitrario, veamos ahora que existen c_1, c_2, c_3 tal que $c_1A_1 + c_2A_2 + c_3A_3 = B$. Pero $B \in S_{2\times 2}$ implica que

$$B = \begin{bmatrix} a & b \\ b & c \end{bmatrix},$$

para algunos a, b, c (conocidos). La ecuación $c_1A_1 + c_2A_2 + c_3A_3 = B$ es ahora equivalente a

$$\begin{bmatrix} c_1 & c_2 \\ c_2 & c_3 \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix},$$

lo que implica que $c_1 = a$, $c_2 = b$, $c_3 = c$. Esto es $\operatorname{span}\{A_1, A_2, A_3\} = S_{2\times 2}$, es decir que $\{A_1, A_2, A_3\}$ es una base para $S_{2\times 2}$. □

Ejemplo 3.55. Para cualquier i, $1 \leq i \leq n$, definimos el vector $\vec{e}_i \in \mathbb{R}^n$, como el vector con un uno en la posición i, y ceros en las otras entradas. Es fácil ver ahora que $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ es una base para $\mathbb{R}^n$.

Para cualesquiera índices i, j con $1 \leq i \leq n$ y $1 \leq j \leq m$, definimos la matriz $E_{ij} \in M_{n\times m}$ como la matriz con todas las entradas cero excepto por la posición (i, j) donde tiene un uno. Es fácil ver ahora que $\{E_{11}, E_{12}, \ldots, E_{nm}\}$ es una base para $M_{n\times m}$.

De forma similar, se puede verificar que $\left\{1, x, x^2, \ldots, x^{n-1}\right\}$ es una base para P_n.

Estas bases de $\mathbb{R}^n$, $M_{n\times m}$, y P_n se llaman las *bases canónicas o estándar* de estos espacios vectoriales. □

El siguiente resultado simplifica el proceso de verificar que un conjunto dado de vectores es una base de un espacio vectorial.

Proposición 3.56. *Suponga que* $V = \text{span}\{v_1, v_2, \ldots, v_n\}$ *y que para cualquier vector* $v \in V$*, la representación de* v *en términos de* $\{v_1, v_2, \ldots, v_n\}$ *es única. Entonces* $\{v_1, v_2, \ldots, v_n\}$ *es una base para* V*.*

Demostración: Claramente los vectores generan a V. Por la unicidad de representación, tenemos que la única forma en que

$$c_1 \cdot v_1 + c_2 \cdot v_2 + \cdots + c_n \cdot v_n = \theta,$$

es con $c_1 = c_2 = \cdots = c_n = 0$, por lo que los vectores son linealmente independientes, y por consiguiente, son una base de V. □

Para $\vec{a}, \vec{b} \in \mathbb{R}^n$, el *producto tensorial* de $\vec{a}$ y $\vec{b}$ de denota por $\vec{a} \otimes \vec{b}$ y se define por:

$$\vec{a} \otimes \vec{b} = (a_i b_j).$$

Note que el producto tensorial de dos vectores es una matriz $n \times n$. Recordando que los vectores en $\mathbb{R}^n$ son matrices $n \times 1$, es fácil ver que $\vec{a} \otimes \vec{b}$ es el producto matricial de $\vec{a}$ con $\vec{b}^t$, esto es $\vec{a} \otimes \vec{b} = \vec{a}\,\vec{b}^t$. De aquí se obtiene que para cualquier $\vec{x} \in \mathbb{R}^n$, tenemos que

$$(\vec{a} \otimes \vec{b})\vec{x} = (\vec{b}^t \vec{x})\vec{a}.$$

Ejemplo 3.57. Dado que $\vec{a} = [-4, 5]^t$ y que $\vec{b} = [3, 2]^t$, entonces

$$\vec{a} \otimes \vec{b} = \vec{a}\,\vec{b}^t = \begin{bmatrix} -4 \\ 5 \end{bmatrix} [3, 2] = \begin{bmatrix} -12 & -8 \\ 15 & 10 \end{bmatrix}.$$

□

El siguiente ejemplo utiliza estos productos tensoriales para construir diferentes bases para el conjunto de las matrices cuadradas.

Ejemplo 3.58. Sean $\vec{a}, \vec{b} \in \mathbb{R}^2$ vectores linealmente independientes. Entonces la matriz $[\vec{a}, \vec{b}]$ es nosingular (Teorema 3.43) por lo que $\left\{\vec{a}, \vec{b}\right\}$ genera a $\mathbb{R}^2$, y por consiguiente es una base para $\mathbb{R}^2$. Consideremos ahora el conjunto de matrices 2×2 dado por:

$$S = \left\{\vec{a} \otimes \vec{a}, \vec{a} \otimes \vec{b}, \vec{b} \otimes \vec{a}, \vec{b} \otimes \vec{b}\right\}.$$

Veamos que este conjunto es una base de $M_{2\times 2}$. Usando el resultado de la Proposición 3.56, vemos que para hacer esto, es suficiente verificar que dada $A \in M_{2\times 2}$, existen escalares c_1, c_2, c_3, c_4 únicos tal que

$$c_1 \vec{a} \otimes \vec{a} + c_2 \vec{a} \otimes \vec{b} + c_3 \vec{b} \otimes \vec{a} + c_4 \vec{b} \otimes \vec{b} = A. \tag{3.2}$$

Como $\left\{\vec{a}, \vec{b}\right\}$ es una base para $\mathbb{R}^2$, la matriz A queda completamente determinada si conocemos los resultados o los valores de $A\vec{a}$ y $A\vec{b}$. Multiplicando ahora la ecuación de arriba por la derecha, primero por $\vec{a}$ y luego con $\vec{b}$, obtenemos que:

$$\begin{aligned} \left[c_1(\vec{a}^t\vec{a}) + c_2(\vec{b}^t\vec{a})\right]\vec{a} + \left[c_3(\vec{a}^t\vec{a}) + c_4(\vec{b}^t\vec{a})\right]\vec{b} &= A\vec{a}, &\quad (3.3\text{a})\\ \left[c_1(\vec{a}^t\vec{b}) + c_2(\vec{b}^t\vec{b})\right]\vec{a} + \left[c_3(\vec{a}^t\vec{b}) + c_4(\vec{b}^t\vec{b})\right]\vec{b} &= A\vec{b}. &\quad (3.3\text{b}) \end{aligned}$$

Como $\left\{\vec{a}, \vec{b}\right\}$ es una base de $\mathbb{R}^2$, podemos encontrar escalares únicos $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ tal que

$$\alpha_1\vec{a} + \alpha_2\vec{b} = A\vec{a}, \quad \alpha_3\vec{a} + \alpha_4\vec{b} = A\vec{b}.$$

Por la unicidad de los α's, tenemos que las ecuaciones (3.3) implican ahora que

$$\begin{cases} c_1(\vec{a}^t\vec{a}) + c_2(\vec{b}^t\vec{a}) &= \alpha_1,\\ c_3(\vec{a}^t\vec{a}) + c_4(\vec{b}^t\vec{a}) &= \alpha_2,\\ c_1(\vec{a}^t\vec{b}) + c_2(\vec{b}^t\vec{b}) &= \alpha_3,\\ c_3(\vec{a}^t\vec{b}) + c_4(\vec{b}^t\vec{b}) &= \alpha_4. \end{cases}$$

Estas ecuaciones se pueden reescribir como

$$\begin{bmatrix} \vec{a}^t\vec{a} & \vec{b}^t\vec{a} \\ \vec{a}^t\vec{b} & \vec{b}^t\vec{b} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_3 \end{bmatrix}, \quad \begin{bmatrix} \vec{a}^t\vec{a} & \vec{b}^t\vec{a} \\ \vec{a}^t\vec{b} & \vec{b}^t\vec{b} \end{bmatrix} \begin{bmatrix} c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} \alpha_2 \\ \alpha_4 \end{bmatrix}.$$

Note que la matriz de coeficientes de ambos sistemas es igual, y es nosingular ya que

$$\begin{bmatrix} \vec{a}^t\vec{a} & \vec{b}^t\vec{a} \\ \vec{a}^t\vec{b} & \vec{b}^t\vec{b} \end{bmatrix} = [\vec{a}, \vec{b}]^t[\vec{a}, \vec{b}].$$

Tenemos entonces que estos dos sistemas tienen soluciones que son únicas. Así que para cualquier $A \in M_{2\times 2}$, existen escalares c_1, c_2, c_3, c_4 únicos tal que (3.2) se cumple, por lo que S es una base para $M_{2\times 2}$. $\square$

El ejemplo anterior se puede generalizar a $\mathbb{R}^n$ de la forma siguiente: sean $\vec{u}_1$, $\vec{u}_2$,...,$\vec{u}_n$ vectores linealmente independientes. Entonces

$$\left\{\vec{u}_i \otimes \vec{u}_j \,:\, i,j = 1,\ldots,n\right\},$$

es una base para $M_{n\times n}$.

En el próximo ejemplo utilizamos conceptos de ecuaciones diferenciales ordinarias, en particular, el teorema de existencia y unicidad para problemas de valor inicial. El lector que no este familiarizado con este tema, puede omitir la lectura del ejemplo sin ningún problema de perder el hilo de la discusión restante del capitulo.

Ejemplo 3.59. Sea

$$V = \left\{f \in C^2[-1,1] \,:\, f'' + f = 0\right\}.$$

V es un espacio vectorial. Verificamos ahora que $\{\operatorname{sen} x, \cos x\}$ es una base para V. Primero note que $\operatorname{sen} x \in V$:

$$(\operatorname{sen} x)'' + \operatorname{sen} x = -\operatorname{sen} x + \operatorname{sen} x = 0.$$

De igual forma se verifica que $\cos x \in V$. Como

$$W(x) = \det\begin{bmatrix} \operatorname{sen} x & \cos x \\ \cos x & -\operatorname{sen} x \end{bmatrix} = -1, \quad x \in [-1,1],$$

entonces $\{\operatorname{sen} x, \cos x\}$ son linealmente independientes. Veamos ahora que $V = \operatorname{span}\{\operatorname{sen} x, \cos x\}$.

Sea $f \in V$. Buscamos $c_1, c_2 \in \mathbb{R}$ tal que

$$f(x) = c_1 \operatorname{sen} x + c_2 \cos x, \quad x \in [-1,1]. \tag{3.4}$$

Fijamos por el momento un $x_0 \in [-1,1]$, y buscamos c_1, c_2 tal que

$$\begin{aligned} f(x_0) &= c_1 \operatorname{sen} x_0 + c_2 \cos x_0, \\ f'(x_0) &= c_1 \cos x_0 - c_2 \operatorname{sen} x_0. \end{aligned}$$

Estas ecuaciones son equivalentes al sistema lineal

$$\begin{bmatrix} \operatorname{sen} x_0 & \cos x_0 \\ \cos x_0 & -\operatorname{sen} x_0 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f'(x_0) \end{bmatrix}.$$

La matriz de coeficientes de este sistema, tiene determinante $W(x_0) = -1$, esto es, es una matriz nosingular. Asi que el sistema tiene solución única c_1, c_2. Tenemos entonces que las funciones $f(x)$ y $c_1 \operatorname{sen} x + c_2 \cos x$ son soluciones ambas de la misma ecuación diferencial (de orden dos) y satisfacen las mismas condiciones iniciales en x_0. Por el teorema de existencia y unicidad para ecuaciones diferenciales lineales, ambas funciones coinciden en todo $x \in [-1,1]$, es decir, que (3.4) se cumple. Así que $V = \operatorname{span}\{\operatorname{sen} x, \cos x\}$ y en conjunto con la independencia lineal que ya verificamos, muestra que $\{\operatorname{sen} x, \cos x\}$ es una base para V. □

3.8 Dimensión de un espacio vectorial

Anteriormente mencionamos que todo espacio vectorial tiene un número infinito de bases, pero que todas tienen el mismo número de elementos. Este número común a todas las bases de un espacio vectorial se llama la *dimensión* del espacio vectorial. Veamos como se obtiene este importante resultado.

Teorema 3.60. *Sea $\{v_1, v_2, \dots, v_n\}$ una base del espacio vectorial $(V, +, \cdot)$. Si $\{u_1, u_2, \dots, u_m\} \subseteq V$, donde $m > n$, entonces los vectores $\{u_1, u_2, \dots, u_m\}$ son linealmente dependientes.*

Demostración: Como $V = \operatorname{span}\{v_1, v_2, \dots, v_n\}$, para cada i, $1 \leq i \leq m$, existen escalares $a_{i1}, a_{i,2}, \dots, a_{in}$, tal que

$$u_i = \sum_{j=1}^{n} a_{ij} \cdot v_j. \tag{3.5}$$

Queremos verificar que existen escalares $c_1, c_2, \dots, c_m$ no todos cero, tal que

$$\sum_{i=1}^{m} c_i \cdot u_i = \theta. \tag{3.6}$$

Usando (3.5) tenemos que la ecuación anterior es equivalente a

$$\theta = \sum_{i=1}^{m} c_i \cdot \left(\sum_{j=1}^{n} a_{ij} \cdot v_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} c_i \right) \cdot v_j = \sum_{j=1}^{n} (A^t \vec{c})_j \cdot v_j,$$

donde $A = (a_{ij})$ es $m \times n$, y $(A^t\vec{c})_j$ denota el componente j del vector $A^t\vec{c}$. Como $\{v_1, v_2, \ldots, v_n\}$ son linealmente independientes, la ecuación anterior implica que $(A^t\vec{c})_j = 0$ para $1 \leq j \leq n$, es decir que

$$A^t\vec{c} = \vec{0}.$$

Esta es la condición en los escalares $c_1, c_2, \ldots, c_m$ para que (3.6) se cumpla. Pero como $m > n$, el sistema $A^t\vec{c} = \vec{0}$ es homogéneo y subdeterminado, que siempre son dependientes. Así que existe un $\vec{c} \neq \vec{0}$ tal que $A^t\vec{c} = \vec{0}$, es decir (3.6) se cumple con escalares no todos cero, y por consiguiente $\{u_1, u_2, \ldots, u_m\}$ son linealmente dependientes. □

Este resultado tiene como consecuencia inmediata lo siguiente.

Corolario 3.61. *Cualesquiera dos bases de un espacio vectorial tienen el mismo número de elementos.*

Nota 3.62. Aunque las bases de un espacio vectorial tienen el mismo número de elementos, dos bases cualesquiera no tienen que ser iguales. □

El corolario anterior motiva o justifica la siguiente definición.

Definición 3.63. Sea $\{v_1, v_2, \ldots, v_n\}$ una base para el espacio vectorial $(V, +, \cdot)$. Decimos que V tiene *dimensión* n, y escribimos $n = \dim V$. Si V no tiene base finita alguna, entonces decimos que V es de *dimensión infinita*.

Nota 3.64. La dimensión del espacio vectorial $\{\theta\}$ se define o se toma como cero. □

Ejemplo 3.65. En referencia al Ejemplo 3.55, tenemos que $\dim \mathbb{R}^n = n$, $\dim M_{n\times m} = nm$, y $\dim P_n = n$. □

Ejemplo 3.66 (Subespacios de $\mathbb{R}^3$). Sea S_1 un subespacio de $\mathbb{R}^3$ de dimensión uno. Entonces

$$S_1 = \operatorname{span}\{\vec{x}\}, \quad \vec{x} \neq \vec{0}.$$

Pensando en $\vec{x}$ como una flecha que emana del origen y en lo que significa multiplicar la flecha por un escalar, tenemos que geométricamente, S_1 consiste de todos los puntos en la recta que contiene el origen y el punto terminal de $\vec{x}$.

Sea S_2 un subespacio cualquiera de $\mathbb{R}^3$ de dimensión dos. Entonces

$$S_2 = \operatorname{span}\{\vec{x}, \vec{y}\}, \quad \vec{x}, \vec{y} \neq \vec{0}.$$

Nuevamente, visualizando a $\vec{x}$ y $\vec{y}$ como flechas que emanan del origen y pensando en lo que significa sumar flechas y multiplicar éstas por un escalar, tenemos que geométricamente S_2 consiste de todos los puntos en el plano que contiene el origen y los puntos terminales de $\vec{x}$ y $\vec{y}$.

Como los únicos subespacios de $\mathbb{R}^3$ de dimensiones cero y tres, son $\left\{\vec{0}\right\}$ y $\mathbb{R}^3$ respectivamente, podemos resumir la discusión anterior diciendo que geométricamente, los subespacios de $\mathbb{R}^3$ son:

subespacio	**dimensión**
el origen	cero
rectas que contienen el origen	uno
planos que contienen el origen	dos
el espacio $\mathbb{R}^3$	tres

$\square$

Veamos ahora algunas consecuencias del concepto de base.

Teorema 3.67. *Sea* $(V, +, \cdot)$ *espacio vectorial con* $\dim V = n$. *Entonces:*

i) Cualquier conjunto de n vectores linealmente independientes en V, es una base para V.

ii) Cualquier conjunto de n vectores que genere a V, tiene que ser linealmente independiente.

iii) Si $\{u_1, u_2, \ldots, u_k\} \subset V$, donde $k < n$, entonces $\operatorname{span}\{u_1, u_2, \ldots, u_k\} \subsetneq V$.

iv) Si $\{u_1, u_2, \ldots, u_k\} \subset V$, donde $k < n$, son linealmente independientes en V, entonces existen $\{u_{k+1}, \ldots, u_n\}$ en V tal que $\{u_1, u_2, \ldots, u_n\}$ son linealmente independientes, y por consiguiente, forman una base para V.

Demostración:

i) Suponga $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente en V, donde $\dim V = n$. Por el Teorema 3.60, tenemos que para cualquier $v \in V$, el conjunto $\{v_1, v_2, \ldots, v_n, v\}$ es linealmente dependiente. Así que existen escalares $c_1, c_2, \ldots, c_{n+1}$, no todos cero, de forma que

$$\sum_{i=1}^{n} c_i \cdot v_i + c_{n+1} \cdot v = \theta.$$

Si $c_{n+1} = 0$, tendríamos una combinación lineal notrivial de $\{v_1, v_2, \ldots, v_n\}$ que es igual a θ (el vector cero de V), es decir $\{v_1, v_2, \ldots, v_n\}$ sería linealmente dependiente, lo que sería una contradicción. Así que $c_{n+1} \neq 0$. En tal caso, podemos despejar la ecuación anterior para v obteniendo así que:

$$v = \sum_{i=1}^{n} \left(-\frac{c_i}{c_{n+1}} \right) \cdot v_i \in \operatorname{span}\{v_1, v_2, \ldots, v_n\}.$$

Como $v \in V$ es arbitrario, esto demuestra que $\operatorname{span}\{v_1, v_2, \ldots, v_n\} = V$, es decir que $v_1, v_2, \ldots, v_n$ es una base para V (ya que son linealmente independientes).

ii) Si $\{v_1, v_2, \ldots, v_n\}$ genera a V pero es linealmente dependiente, entonces por el Teorema 3.50, podemos eliminar un vector de $\{v_1, v_2, \ldots, v_n\}$ y seguir generando a V. Pero esto implica que podemos encontrar una base para V con menos de n vectores. Esto contradice que $n = \dim V$.

iii) Si $\operatorname{span}\{u_1, u_2, \ldots, u_k\} = V$ con $k < n$, tendríamos una base con menos de n vectores, contradiciendo nuevamente que $n = \dim V$.

iv) Si $\{u_1, u_2, \ldots, u_k\} \subset V$, donde $k < n$, entonces por la parte (iii) tenemos que $\operatorname{span}\{u_1, u_2, \ldots, u_k\} \subsetneq V$. Sea $u_{k+1} \in V \setminus \operatorname{span}\{u_1, u_2, \ldots, u_k\}$. Si $\{u_1, u_2, \ldots, u_k, u_{k+1}\}$ fuese linealmente dependiente, entonces existen escalares $c_1, c_2, \ldots, c_{k+1}$, no todos cero, tal que

$$\sum_{i=1}^{k+1} c_i \cdot u_i = \theta.$$

Note que $c_{k+1} \neq 0$ ya que $\{u_1, u_2, \ldots, u_k\}$ son linealmente independientes. Así que despejando en la ecuación anterior para u_{k+1}, tendríamos que

$$u_{k+1} = \sum_{i=1}^{k} \left(-\frac{c_i}{c_{k+1}} \right) \cdot u_i \in \operatorname{span}\{u_1, u_2, \ldots, u_k\}.$$

Pero esto es una contradicción ya que $u_{k+1} \notin \operatorname{span}\{u_1, u_2, \ldots, u_k\}$. De modo que $\{u_1, u_2, \ldots, u_k, u_{k+1}\}$ tiene que ser linealmente independiente. Si $k+1 \neq n$, repetimos el proceso, hasta llegar a los n vectores linealmente independientes.

□

Ejemplo 3.68. Veamos que

$$\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\},$$

es una base para $\mathbb{R}^2$. Como

$$\det \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} = 3 \neq 0,$$

por el Teorema 3.43 tenemos que estos vectores son linealmente independientes. Como $\dim \mathbb{R}^2 = 2$, tenemos por la parte (i) del Teorema 3.67, que los vectores son una base para $\mathbb{R}^2$. De igual forma se puede verificar que los vectores

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

forman una base de $\mathbb{R}^4$. Esta base es un caso especial de las llamadas *bases de Fourier* las cuales se utilizan en los métodos de compresión de imágenes.

□

La demostración de la parte (iv) del Teorema 3.67 describe un método para expandir un conjunto de vectores linealmente independiente hasta obtener una base para el espacio vectorial. El siguiente ejemplo ilustra este procedimiento.

Ejemplo 3.69. Los vectores

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix},$$

son linealmente independientes ya que no son proporcionales. Queremos encontrar un vector $\vec{u}_3$ de modo que $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ sean linealmente independientes y por consiguiente, una base para $\mathbb{R}^3$. De acuerdo a la demostración de la parte (iii) del Teorema 3.67, es suficiente conseguir un vector $\vec{u}_3 \notin \operatorname{span}\{\vec{u}_1, \vec{u}_2\}$. Note que $(a,b,c)^t \in \operatorname{span}\{\vec{u}_1, \vec{u}_2\}$ si y solo si

$$\begin{bmatrix} 1 & 3 \\ 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix},$$

para algunos escalares $\alpha, \beta \in \mathbb{R}$. Como

$$\left[\begin{array}{rr|c} 1 & 3 & a \\ 1 & 0 & b \\ -1 & 1 & c \end{array}\right] \begin{array}{c} f_2 = f_2 - f_1 \\ \longrightarrow \\ f_3 = f_3 + f_1 \end{array} \left[\begin{array}{rr|c} 1 & 3 & a \\ 0 & -3 & b-a \\ 0 & 4 & a+c \end{array}\right]$$

$$\begin{array}{c} f_3 = f_3 + \frac{4}{3} f_2 \\ \longrightarrow \end{array} \left[\begin{array}{rr|c} 1 & 3 & a \\ 0 & -3 & b-a \\ 0 & 0 & a+c+\frac{4}{3}(b-a) \end{array}\right],$$

Tenemos que $(a,b,c)^t \in \operatorname{span}\{\vec{u}_1, \vec{u}_2\}$ si y solo si $a + c + \frac{4}{3}(b-a) = 0$. Así que tomando $a = 1, b = 1, c = 2$, tenemos que $\vec{u}_3 = (1,1,2)^t \notin \operatorname{span}\{\vec{u}_1, \vec{u}_2\}$, y por consiguiente que $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ son linealmente independientes, es decir, una base de $\mathbb{R}^3$. □

3.9 Espacios fila y columna de una matriz

Recuerde que una matriz A de tamaño $m \times n$ se puede particionar por columnas como:

$$A = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n],$$

donde $\vec{a}_j \in \mathbb{R}^m$, $1 \leq j \leq n$, son los *vectores columna* de A. El *espacio columna* de A se denota por $R(A)$ y se define por:

$$R(A) = \operatorname{span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}.$$

También podemos particionar a A por filas como:

$$A = \begin{bmatrix} \vec{b}_1^t \\ \vec{b}_2^t \\ \vdots \\ \vec{b}_m^t \end{bmatrix},$$

donde $\vec{b}_i \in \mathbb{R}^n$ son los *vectores fila* de A. El *espacio fila* de A se define como $\operatorname{span}\left\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m\right\}$. Como el espacio fila de A es igual al espacio columna de A^t, tenemos que el espacio fila de A se puede representar por $R(A^t)$, es decir:

$$R(A^t) = \operatorname{span}\left\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m\right\}.$$

Ejemplo 3.70. Para

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

tenemos que el espacio columna es:

$$R(A) = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}\right\}.$$

El espacio fila de A es:

$$R(A^t) = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 2 \\ 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}\right\}.$$

□

Luego de aplicar varias operaciones elementales de fila a una matriz cualquiera A, las filas de la matriz B resultante, son combinaciones lineales de las filas de la matriz A. De modo que el espacio fila de B está contenido en el espacio fila de A, es decir $R(B^t) \subseteq R(A^t)$.

Ejemplo 3.71. Sean

$$A = \begin{bmatrix} 4 & 1 \\ 1 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -2 \\ 0 & 9 \end{bmatrix}.$$

Note que B se obtiene a partir de A, aplicando a A las operaciones elementales de fila $f_1 \leftrightarrow f_2$, $f_2 = f_2 - 4f_1$. Si

$$A = \begin{bmatrix} \vec{b}_1^t \\ \vec{b}_2^t \end{bmatrix}, \quad B = \begin{bmatrix} \vec{c}_1^t \\ \vec{c}_2^t \end{bmatrix},$$

entonces

$$\vec{c}_1 = \vec{b}_2, \quad \vec{c}_2 = \vec{b}_1 - 4\vec{b}_2.$$

De aquí que $\operatorname{span}\{\vec{c}_1, \vec{c}_2\} \subseteq \operatorname{span}\left\{\vec{b}_1, \vec{b}_2\right\}$, es decir

$$R(B^t) \subseteq R(A^t).$$

□

Como consecuencia de esta observación, tenemos el siguiente resultado.

Teorema 3.72. *Sean A, B matrices equivalentes por fila, es decir $A \sim B$. Entonces A y B tienen el mismo espacio fila, es decir que:*

$$R(A^t) = R(B^t).$$

Demostración: Como B se obtiene a partir de A por un número finito de operaciones elementales de fila, entonces $R(B^t) \subseteq R(A^t)$. Como "$\sim$" es una relación de equivalencia, también A se obtiene a partir de B por medio de un número finito de operaciones elementales de filas, por lo que $R(A^t) \subseteq R(B^t)$. Tenemos pues que $R(A^t) = R(B^t)$. □

Nota 3.73. Aunque las operaciones elementales de fila no cambian el espacio fila de una matriz, éstas si cambian el espacio columna. Veamos:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix} \overset{f_2 = f_2 + f_1}{\longrightarrow} \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} = B.$$

Note que

$$R(A) = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}, \quad R(B) = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\},$$

por lo que $R(A) \neq R(B)$, ya que el segundo componente de los elementos de $R(A)$ siempre es cero. No obstante $\dim R(A) = \dim R(B)$, lo cual es siempre cierto, ésto es, las operaciones elementales de fila preservan la dimensión del espacio columna (Corolario 3.77). □

Tenemos entonces que para estudiar el espacio fila de una matriz, primero reducimos ésta, por medio de operaciones elementales de fila, a su forma echelon. Tenemos luego que las filas no cero en la forma echelon forman una base del espacio fila. Antes de ver un ejemplo de este procedimiento, vamos a definir el concepto de *rango* de una matriz.

Definición 3.74. Sea A una matriz $m \times n$. El *rango* de A, escrito $\operatorname{rank} A$, es la dimensión del espacio fila de A. Esto es

$$\operatorname{rank} A = \dim R(A^t).$$

Ejemplo 3.75. Sea

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

Reducimos A a su forma echelon:

$$\begin{bmatrix} 1 & 2 & -3 & 1 \\ 4 & 2 & 0 & 1 \\ 2 & -2 & 6 & -1 \end{bmatrix} \begin{array}{c} f_2 = f_2 - 4f_1 \\ \longrightarrow \\ f_3 = f_3 - 2f_1 \end{array} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & -6 & 12 & -3 \\ 0 & -6 & 12 & -3 \end{bmatrix},$$

$$\begin{array}{c} f_3 = f_3 - f_2 \\ \longrightarrow \\ -\frac{1}{6}f_2 \end{array} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & -2 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

De aquí tenemos que $\operatorname{rank} A = 2$, que $\left\{(1,2,-3,1)^t, (0,1,-2,\frac{1}{2})^t\right\}$ es una base del espacio fila, y que:

$$R(A^t) = \operatorname{span}\left\{(1,2,-3,1)^t, (0,1,-2,\frac{1}{2})^t\right\}.$$

□

Hay una relación bien importante entre el espacio fila y el espacio columna de una matriz. Aunque estos espacios no son iguales, siempre tienen la misma dimensión.

Teorema 3.76. *Sea A una matriz $m \times n$. Entonces la dimensión del espacio fila de A es igual a la dimensión del espacio columna de A, ésto es*

$$\dim R(A^t) = \dim R(A).$$

En particular, $\operatorname{rank} A = \operatorname{rank} A^t$.

Demostración: Sea U la forma echelon de la matriz A. Entonces U tiene $\operatorname{rank} A$ filas distintas de cero. Poniendo $r = \operatorname{rank} A$, sea $\hat{U}$ la matriz $m \times r$ que se obtiene a partir de U eliminando las columnas de U que corresponden a las variables libres. Note que las columnas de $\hat{U}$ son linealmente independientes. Sea $\hat{A}$ la matriz $m \times r$ que se obtiene a partir de A, eliminando las columnas correspondientes a los índices de las que se eliminaron de U. Note que $\hat{U}$ y $\hat{A}$ siguen siendo equivalentes por fila. De hecho, $\hat{U}$ es la forma echelon de $\hat{A}$. Como las columnas de $\hat{U}$ son linealmente independientes, también lo son las columnas de $\hat{A}$. Así que

$$\dim R(A^t) = \operatorname{rank} A = r = \dim R(\hat{A}) \leq \dim R(A),$$

esto es, $\dim R(A^t) \leq \dim R(A)$. Intercambiando los roles de A y A^t, tenemos que $\dim R(A) \leq \dim R(A^t)$, esto es, $\dim R(A^t) = \dim R(A)$.

Usando que $\operatorname{rank} A = \dim R(A^t)$ y que $\operatorname{rank} A^t = \dim R(A)$, tenemos que $\operatorname{rank} A = \operatorname{rank} A^t$. □

Combinando los Teoremas 3.72 y 3.76, tenemos el resultado de que las operaciones elementales de fila preservan la dimensión del espacio columna.

Corolario 3.77. *Sean A y B matrices equivalentes por fila. Entonces*

$$\dim R(A) = \dim R(B).$$

Demostración: Como A y B son equivalentes por fila, tenemos usando el Teorema 3.72 que $R(A^t) = R(B^t)$. Usando ésto podemos ahora escribir que:

$$\begin{aligned} \dim R(A) &= \dim R(A^t) \quad \text{(Teorema 3.76)}, \\ &= \dim R(B^t) \quad \text{(ya que } R(A^t) = R(B^t)\text{)}, \\ &= \dim R(B) \quad \text{(Teorema 3.76)}. \end{aligned}$$

□

Hay también una relación bien importante entre las dimensiones del espacio fila de una matriz y la del espacio nulo.

Teorema 3.78. *Sea A una matriz $m \times n$. Entonces*

$$\operatorname{rank} A + \dim N(A) = n.$$

Demostración: Sea U la forma echelon de A. Entonces el sistema $A\vec{x} = \vec{0}$ es equivalente al sistema $U\vec{x} = \vec{0}$. Este último sistema tiene $\operatorname{rank} A$ variables lideres y $n - \operatorname{rank} A$ variables libres. El número de variables libres es la dimensión del espacio nulo, es decir:

$$\dim N(A) = n - \operatorname{rank} A,$$

de donde se obtiene el resultado. $\square$

Ejemplo 3.79. Para la matriz

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

buscamos bases para $R(A^t)$, $N(A)$, y $R(A)$. La forma echelon de A es:

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

Tenemos entonces que $\operatorname{rank} A = 2$ y que

$$R(A^t) = \operatorname{span}\left\{(1,-2,-4,-3)^t, (0,1,2,-1)^t\right\}.$$

Usando la matriz U, tenemos que $N(A)$ consiste de todas las soluciones del sistema:

$$\begin{cases} x_1 - 2x_2 - 4x_3 - 3x_4 &= 0, \\ \qquad\quad x_2 + 2x_3 - x_4 &= 0. \end{cases}$$

Las soluciones de este sistema son $x_1 = 5x_4$, $x_2 = -2x_3 + x_4$. De modo que $A\vec{x} = \vec{0}$, donde $\vec{x} = (x_1, x_2, x_3, x_4)^t$, si

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

Tenemos entonces que

$$N(A) = \operatorname{span}\left\{(0,-2,1,0)^t, (5,1,0,1)^t\right\},$$

por lo que $\dim N(A) = 2$. Vea que $\operatorname{rank} A + \dim N(A) = 2 + 2 = 4$, que es igual al número de columnas de A.

Para hallar una base para $R(A)$, usamos el procedimiento descrito en la demostración del Teorema 3.76. Las columnas de U que corresponden a los 1's lideres en las filas no cero de U, son la uno y la dos. Por lo tanto las correspondientes columnas de A forman una base para $R(A)$, ésto es

$$R(A) = \operatorname{span}\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

□

Concluimos este capitulo con un resultado relacionado a nuestra discusión a principios de la Sección 3.5. Sea A una matriz $m \times n$. Si particionamos a A por columnas como:

$$A = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n],$$

donde $\vec{a}_j \in \mathbb{R}^m$, $1 \leq j \leq n$, entonces es fácil ver que

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

donde $\vec{x} = (x_1, x_2, \ldots, x_n)^t$. La ecuación $A\vec{x} = \vec{b}$ es ahora equivalente a:

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b}.$$

Usando esta observación podemos ahora demostrar el siguiente teorema.

Teorema 3.80. *Sea A una matriz $m \times n$. Entonces tenemos lo siguiente:*

i) El sistema $A\vec{x} = \vec{b}$ tiene solución si y solo si $\vec{b} \in R(A)$.

ii) El sistema $A\vec{x} = \vec{b}$ tiene solución para todo vector $\vec{b} \in \mathbb{R}^m$ si y solo si $R(A) = \mathbb{R}^m$.

iii) La solución del sistema $A\vec{x} = \vec{b}$ es única si y solo si los vectores columnas de A son linealmente independientes.

Demostración:

i) Note que $A\vec{x} = \vec{b}$ si y solo si

$$\vec{b} = A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n \in \operatorname{span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\} = R(A).$$

ii) Por la parte (i), $A\vec{x} = \vec{b}$ tiene solución si y solo si $\vec{b} \in R(A)$. Para que $A\vec{x} = \vec{b}$ tenga solución para todo lado derecho $\vec{b}$, entonces es necesario y suficiente que $R(A) = \mathbb{R}^m$.

iii) Si $A\vec{x} = \vec{b}$, y $A\vec{y} = \vec{b}$, entonces $A(\vec{x} - \vec{y}) = \vec{0}$. Poniendo $\vec{z} = \vec{x} - \vec{y}$, con $\vec{z} = (z_1, z_2, \ldots, z_n)^t$, tenemos que $A\vec{z} = \vec{0}$, lo que es equivalente a:

$$z_1\vec{a}_1 + z_2\vec{a}_2 + \cdots + z_n\vec{a}_n = \vec{0}.$$

Si $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ es linealmente independiente, entonces $\vec{z} = \vec{0}$, es decir que $\vec{x} = \vec{y}$.

Supongamos ahora que la solución de $A\vec{x} = \vec{b}$ es única y que

$$z_1\vec{a}_1 + z_2\vec{a}_2 + \cdots + z_n\vec{a}_n = \vec{0}. \tag{3.7}$$

Suponga $\vec{z} = (z_1, z_2, \ldots, z_n)^t \neq \vec{0}$. Note que (3.7) es equivalente a $A\vec{z} = \vec{0}$. Sea $\vec{u}$ una solución de $A\vec{x} = \vec{b}$. Entonces $\vec{y} = \vec{u} + \vec{z}$ es también solución de $A\vec{x} = \vec{b}$:

$$A\vec{y} = A(\vec{u} + \vec{z}) = A\vec{u} + A\vec{z} = \vec{b} + \vec{0} = \vec{b}.$$

Pero esto contradice la unicidad de soluciones de $A\vec{x} = \vec{b}$. Así que (3.7) implica que $\vec{z} = \vec{0}$, es decir, que $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ son linealmente independiente.

□

3.10 Ejercicios

Ejercicio 3.1. Determine si las siguientes aseveraciones son ciertas o falsas. En cada caso justifique su contestación citando algún teorema o resultado del libro de texto o con un contra ejemplo de ser necesario.

a) Un conjunto de cinco vectores en $\mathbb{R}^4$ podría ser linealmente independiente.

b) Tres vectores en $\mathbb{R}^3$ no necesariamente forman una base de $\mathbb{R}^3$.

c) El conjunto $\{\vec{\mathbf{x}}, 2\vec{\mathbf{x}}, 3\vec{\mathbf{x}}\}$ es linealmente dependiente en $\mathbb{R}^3$.

d) Cualquier subconjunto de un conjunto de vectores linealmente independiente, es linealmente independiente.

e) Todas las bases de un espacio vectorial son iguales pues tienen el mismo numero de elementos.

f) El conjunto de matrices singulares $n \times n$ forma un subespacio de $M_{n\times n}$.

g) Si $\vec{v}_1 \neq \vec{w}_1$ y $\vec{v}_2 \neq \vec{w}_2$, entonces $\operatorname{span}\{\vec{v}_1, \vec{v}_2\} \neq \operatorname{span}\{\vec{w}_1, \vec{w}_2\}$.

h) Si A es una matriz 5×3, entonces $\operatorname{rank}(A) \leq 3$.

Ejercicio 3.2. Determine si el vector $\vec{b} = (1, 1, -1, 2)^t$ pertenece o no al espacio generado por $\vec{v}_1 = (2, 0, -1, 1)^t$, $\vec{v}_2 = (3, 1, 1, 1)^t$, $\vec{v}_3 = (-3, 1, 2, 1)^t$.

Ejercicio 3.3. Determine si las siguientes matrices son linealmente independientes o no en $M_{2\times 2}$:

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

Ejercicio 3.4. Verifique que las siguientes funciones $\mathrm{e}^x, \mathrm{e}^{-x}$ son linealmente independientes en el espacio vectorial indicado.

a) $\left\{4, x^2 + 1, 2x - 1\right\}$ en P_3.

b) $\{\mathrm{e}^x, x\mathrm{e}^x\}$ en $C^1(\mathbb{R})$.

c) $\{\cos(\ln x), \operatorname{sen}(\ln x)\}$ en $C^1(0, \infty)$.

Ejercicio 3.5. Verifique que el conjunto de matrices simétricas 2×2, esto es

$$S_{2\times 2} = \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} : a, b, c \in \mathbb{R} \right\},$$

es un subespacio de $M_{2\times 2}$.

Ejercicio 3.6. Considere el siguiente subconjunto de $M_{2\times 2}$:

$$S = \left\{ \begin{pmatrix} a & a+b \\ a+b & b \end{pmatrix} : a, b \in \mathbb{R} \right\}.$$

Verifique que S es un subespacio de $M_{2\times 2}$.

Ejercicio 3.7. Considere el subconjunto de P_4 dado por:

$$S = \{p \in P_4 : p(0) = 2p(1)\}.$$

Determine si S es o no un subespacio de P_4.

Ejercicio 3.8. Sea $S = \{p(x) \in P_3 \,|\, p(1) = 0\}$.

a) Verifique que S es un subespacio de P_3.

b) Verifique que $\left\{x-1, x^2-1\right\}$ es una base para S.

Ejercicio 3.9. Sea

$$S = \left\{ \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} \middle| \, a \in \mathbb{R} \right\}.$$

Verifique que S es un subespacio de $M_{2\times 2}$ y encuentre una base para S.

Ejercicio 3.10. Sea $S = \left\{(x, y, z) \in \mathbb{R}^3 \,|\, 2x + 3y - z = 0\right\}$.

a) Verifique que S es un subespacio de $\mathbb{R}^3$.

b) Halle una base para S.

c) Halle $\dim S$.

Ejercicio 3.11. Halle una base para el subespacio U de $\mathbb{R}^4$ dado por:

$$U = \left\{ \begin{bmatrix} x-2y \\ 2y-x \\ x-2y+z \\ z \end{bmatrix} : x, y, z \in \mathbb{R} \right\}.$$

¿Cúal es la dimensión de U?

Ejercicio 3.12. Halle una base para el subespacio S de P_3 dado por:

$$S = \left\{(a+2b+4c)+(a+2c)x+(b+c)x^2 \,|\, a,b,c \in \mathbb{R}\right\}.$$

¿Cúal es la dimensión de S?

Ejercicio 3.13. Considere los siguientes vectores:

$$\vec{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

a) Verifique que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es una base de $\mathbb{R}^3$.

b) Escriba el vector $\vec{u} = (2,-2,4)^t$ como una combinación lineal de $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

Ejercicio 3.14. Determine si el conjunto $\left\{x, 1+x^2, x+x^2\right\}$ es o no una base de P_3.

Ejercicio 3.15. Sean $\vec{v}_1 = (1,0,0)^t$, $\vec{v}_2 = (0,1,1)^t$, $\vec{v}_3 = (1,0,1)^t$, $\vec{v}_4 = (1,2,3)^t$.

a) Determine si los $\vec{v}_i$'s son linealmente dependientes o independientes.

b) Halle la dimensión de $\text{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$.

Ejercicio 3.16. Halle una base para el espacio generado por los vectores:

$$\vec{w}_1 = \begin{bmatrix} 1 \\ -3 \\ 1 \\ -1 \end{bmatrix}, \quad \vec{w}_2 = \begin{bmatrix} -2 \\ 6 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{w}_3 = \begin{bmatrix} -4 \\ 12 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{w}_4 = \begin{bmatrix} 3 \\ -9 \\ 3 \\ 2 \end{bmatrix}.$$

Ejercicio 3.17. La matriz A tiene a la matriz M como su forma echelon reducida donde:

$$A = \begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & 1 \\ 3 & 1 & 2 & -1 \end{pmatrix}, \quad M = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

a) Halle una base para $R(A)$ (el espacio columna de A).

b) Halle una base para el espacio fila de A.

c) Halle una base para $N(A)$ (el espacio nulo de A).

d) Halle $\dim N(A)$ y $\operatorname{rank} A$.

Ejercicio 3.18. La matriz A tiene a la matriz M como su forma echelon reducida donde:

$$A = \begin{bmatrix} -8 & 16 & -15 & 21 & 3 \\ 18 & -36 & 20 & -6 & -2 \\ -3 & 6 & -5 & 6 & 1 \\ 2 & -4 & -8 & 30 & 2 \\ -11 & 22 & -22 & 33 & 5 \end{bmatrix}, \quad M = \begin{bmatrix} 1 & -2 & 0 & 3 & 0 \\ 0 & 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

a) Halle una base para $R(A)$ (el espacio columna de A).

b) Halle una base para el espacio fila de A.

c) Halle una base para $N(A)$ (el espacio nulo de A).

d) Halle $\dim N(A)$ y $\operatorname{rank} A$.

e) ¿Será la matriz A nosingular? Explique su contestación.

Ejercicio 3.19. Sea A una matriz 5×7 con $\operatorname{rank} A = 4$. Complete los siguientes blancos:

a) Las columnas de A son linealmente __________________. (independientes o dependientes)

b) Las filas de A son linealmente __________________. (independientes o dependientes)

c) $\dim N(A) =$ __________________.

d) $\operatorname{rank} A^t =$ __________________.

e) $\dim N(A^t) =$ __________________.

Ejercicio 3.20. Sea M una matriz 6×4 con $\operatorname{rank} M = 4$. Complete los siguientes blancos:

a) Las columnas de M son linealmente __________________. (independientes o dependientes)

b) Las filas de M son linealmente ____________________. (independientes o dependientes)

c) $\dim N(M) =$____________________.

d) $\operatorname{rank} M^t =$____________________.

e) $\dim N(M^t) =$____________________.

Ejercicio 3.21. Sean U, V subespacios del espacio vectorial $(X, +, \cdot)$.

a) Verifique que $U \cap V$ es un subespacio de X.

b) Verifique que $U \oplus V$ es un subespacio de X, donde

$$U \oplus V = \{u + v \,:\, u \in U, \quad v \in V\}.$$

c) Determina si $U \cup V$ es o no subespacio de X.

Ejercicio 3.22. Sean A y B matrices $m \times n$ y $n \times p$ respectivamente. Si $C = AB$, verifique que $N(B) \subset N(C)$ y que $\operatorname{rank} C \leq \operatorname{rank} B$.

4 Transformaciones lineales

Utilizando cualquier matriz A de dimensiones $m \times n$, podemos asociar una función L mediante la siguiente regla:

$$L(\vec{x}) = A\vec{x}, \quad \vec{x} \in \mathbb{R}^n.$$

Note que L es una función de $\mathbb{R}^n$ a $\mathbb{R}^m$, es decir $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Como consecuencia del Teorema 1.19 tenemos que L tiene las siguientes propiedades:

$$\begin{aligned} L(\alpha\vec{x}) &= A(\alpha\vec{x}) = \alpha A\vec{x} = \alpha L(\vec{x}), \quad \forall\, \alpha \in \mathbb{R}, \quad \vec{x} \in \mathbb{R}^n, \\ L(\vec{x} + \vec{y}) &= A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y} = L(\vec{x}) + L(\vec{y}), \quad \forall\, \vec{x}, \vec{y} \in \mathbb{R}^n. \end{aligned}$$

Es por esta razón que decimos que $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es una *transformación lineal*. En este capitulo vamos a estudiar transformaciones lineales mas generales que L, entre espacios vectoriales mas generales que $\mathbb{R}^n$ y $\mathbb{R}^m$. No obstante, veremos que al momento de especificar bases de los espacios vectoriales en cuestión, cualquier transformación lineal se puede representar como producto de una matriz (llamada la representación matricial de la transformación) por un vector en algún $\mathbb{R}^n$.

4.1 Definiciones y algunos ejemplos

Al momento de leer la siguiente definición, conviene pensar en el ejemplo anterior con $V = \mathbb{R}^n$ y $W = \mathbb{R}^m$.

Definición 4.1. Sean V, W espacios vectoriales y L una función de V a W, es decir $L : V \longrightarrow W$. Decimos que L es una *transformación o operador lineal* de V a W si

$$\begin{aligned} L(\alpha \cdot v) &= \alpha \cdot L(v), \quad \forall\, \alpha \in \mathbb{R}, \quad v \in V, \\ L(v + w) &= L(v) + L(w), \quad \forall\, v, w \in V. \end{aligned}$$

Nota 4.2. Observe que $L : V \longrightarrow W$ es lineal si y solo si

$$L(\alpha \cdot v + \beta \cdot w) = \alpha \cdot L(v) + \beta \cdot L(w), \quad \forall\, \alpha, \beta \in \mathbb{R}, \quad v, w \in V.$$

□

Ejemplo 4.3. Para α un escalar fijo, defina $L_\alpha : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ por:

$$L_\alpha(\vec{x}) = \alpha\vec{x},$$

Entonces, como

$$\begin{aligned} L_\alpha(\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2) &= \alpha(\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2) \\ &= \alpha(\alpha_1\vec{x}_1) + \alpha(\alpha_2\vec{x}_2) = \alpha_1(\alpha\vec{x}_1) + \alpha_2(\alpha\vec{x}_2) \\ &= \alpha_1 L_\alpha(\vec{x}_1) + \alpha_2 L_\alpha(\vec{x}_2), \quad \forall\, \alpha_1, \alpha_2 \in \mathbb{R}, \quad \vec{x}_1, \vec{x}_2 \in \mathbb{R}^n, \end{aligned}$$

tenemos que L_α es lineal. □

Ejemplo 4.4. Sea $\{\vec{e}_1, \ldots, \vec{e}_n\}$ la base canónica de $\mathbb{R}^n$ (cf. el Ejemplo 3.55). Fijamos un índice i entre 1 y n, y definimos $L_i : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ por

$$L_i(\vec{x}) = (\vec{x}^t\vec{e}_i)\vec{e}_i.$$

Tenemos ahora que

$$\begin{aligned} L_i(\alpha\vec{x}_1 + \beta\vec{x}_2) &= ((\alpha\vec{x}_1 + \beta\vec{x}_2)^t\vec{e}_i)\vec{e}_i, \\ &= (\alpha(\vec{x}_1^t\vec{e}_i) + \beta(\vec{x}_2^t\vec{e}_i))\vec{e}_i, \\ &= \alpha(\vec{x}_1^t\vec{e}_i)\vec{e}_i + \beta(\vec{x}_2^t\vec{e}_i)\vec{e}_i, \\ &= \alpha L_i(\vec{x}_1) + \beta L_i(\vec{x}_2), \quad \forall\, \alpha, \beta \in \mathbb{R}, \quad \vec{x}_1, \vec{x}_2 \in \mathbb{R}^n. \end{aligned}$$

Tenemos entonces que L_i es un operador lineal. $L_i(\vec{x})$ es la proyección vectorial del vector $\vec{x}$ al i–esimo eje de coordenadas en $\mathbb{R}^n$. □

Ejemplo 4.5. Defina $L : \mathbb{R}^2 \longrightarrow \mathbb{R}$ por

$$L(\vec{x}) = L(x_1, x_2) = x_1x_2.$$

L no es un operador lineal ya que

$$L(\alpha\vec{x}) = L(\alpha x_1, \alpha x_2) = (\alpha x_1)(\alpha x_2) = \alpha^2 x_1 x_2,$$

y esto es igual a $\alpha L(\vec{x})$ únicamente cuando $\alpha = 0, 1$. □

Ejemplo 4.6. Defina la función $L : C[0,1] \longrightarrow C^1[0,1]$ por:

$$L(f)(s) = \int_0^s f(t)\,\mathrm{d}t.$$

Entonces como

$$\begin{aligned}
L(\alpha f+\beta g)(s) &= \int_0^s (\alpha f+\beta g)(t)\,\mathrm{d}t,\\
&= \int_0^s (\alpha f(t)+\beta g(t))\,\mathrm{d}t,\\
&= \alpha\int_0^s f(t)\,\mathrm{d}t+\beta\int_0^s g(t)\,\mathrm{d}t,\\
&= \alpha L(f)(s)+\beta L(g)(s)=(\alpha L(f)+\beta L(g))(s),
\end{aligned}$$

para toda $s \in [0,1]$. Esto implica que

$$L(\alpha f+\beta g)=\alpha L(f)+\beta L(g), \quad \forall\, \alpha,\beta\in\mathbb{R} \quad f,g\in C[0,1],$$

es decir, que L es lineal. □

El siguiente resultado aplica a cualquier operador lineal entre cualesquiera espacios vectoriales V y W. En el enunciado del mismo utilizamos la notación θ_V, θ_W para representar los vectores cero de V y W respectivamente.

Proposición 4.7. *Sean V, W espacios vectoriales y $L : V \longrightarrow W$ un operador lineal. Entonces*

i) $L(\theta_V) = \theta_W$.

ii) Para cualquier entero $n \geq 1$, tenemos que

$$L(\alpha_1\cdot v_1+\cdots\alpha_n\cdot v_n)=\alpha_1\cdot L(v_1)+\cdots+\alpha_n\cdot L(v_n).$$

Demostración:

i) Usando la parte (i) del Teorema 3.7 y la linealidad de L, tenemos que:

$$L(\theta_V)=L(0\cdot\theta_V)=0\cdot L(\theta_V)=\theta_W.$$

ii) Esto se demuestra usando inducción matemática en n.

□

Como una aplicación de este resultado tenemos lo siguiente: suponga que $\{v_1, v_2, \ldots, v_n\}$ es una base del espacio vectorial V y que $L : V \longrightarrow W$ es un operador lineal. Entonces L queda completamente determinado por los valores de $L(v_1), L(v_2), \ldots, L(v_n)$. Esto es así ya que para cualquier $v \in V$, existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ tal que $v = \sum_{i=1}^{n} \alpha_i v_i$, y por la proposición anterior

$$L(v) = L\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i L(v_i).$$

Ejemplo 4.8. Suponga que $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ es una transformación lineal tal que

$$L(\vec{e}_1) = [-2, 4]^t, \quad L(\vec{e}_2) = [3, 3]^t, \quad L(\vec{e}_3) = [5, -1]^t,$$

donde $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ es la base estándar de $\mathbb{R}^3$. Entonces podemos calcular $L(\vec{x})$ para cualquier vector $\vec{x} \in \mathbb{R}^3$. En el caso particular en que

$$\vec{x} = [5, -2, 6]^t = 5\vec{e}_1 - 2\vec{e}_2 + 6\vec{e}_3,$$

tenemos que

$$\begin{aligned} L(\vec{x}) &= 5L(\vec{e}_1) - 2L(\vec{e}_2) + 6L(\vec{e}_3), \\ &= 5[-2, 4]^t - 2[3, 3]^t + 6[5, -1]^t = [14, 8]^t. \end{aligned}$$

□

4.2 Vectores de Coordenadas

Sea $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ una base para $\mathbb{R}^n$. Si $\vec{v} \in \mathbb{R}^n$, sabemos que existen escalares (únicos) $x_1, x_2, \ldots, x_n$ tal que

$$\vec{v} = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n.$$

El vector $(x_1, x_2, \ldots, x_n)^t$ se llama el *vector de coordenadas de $\vec{v}$ con respecto a la base S*, y escribimos

$$(\vec{v})_S = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Nota 4.9. En este libro, si $\vec{v}$ es un vector cualquiera en $\mathbb{R}^n$, y escribimos que $\vec{v} = [x_1, \ldots, x_n]^t$, se sobre entenderá que $\vec{v} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$ donde $\{\vec{e}_1, \ldots, \vec{e}_n\}$ es la base canónica de $\mathbb{R}^n$. $\square$

Ejemplo 4.10. Los vectores

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

son una base de $\mathbb{R}^2$. Considere el vector $\vec{v} = (3,1)^t \in \mathbb{R}^2$. Buscamos x_1, x_2 tal que

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

Esto es equivalente al sistema lineal:

$$\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Este sistema tiene solución $x_1 = 5/3, x_2 = -2/3$ por lo que

$$(\vec{v})_S = \begin{bmatrix} \dfrac{5}{3} \\ \\ -\dfrac{2}{3} \end{bmatrix},$$

es el vector de coordenadas de $\vec{v}$ con respecto a la base S, esto es

$$\vec{v} = \frac{5}{3}\vec{v}_1 - \frac{2}{3}\vec{v}_2.$$

En la Figura muestra los vectores $\{\vec{v}_1, \vec{v}_2\}$ y como los componentes $\frac{5}{3}\vec{v}_1$ y $-\frac{2}{3}\vec{v}_2$ suman al vector $\vec{v}$. $\square$

Podemos generalizar el concepto de vectores de coordenadas a cualquier espacio vectorial. Sea $S = \{v_1, v_2, \ldots, v_n\}$ una base para el espacio vectorial V. Si $v \in V$, entonces existen escalares (únicos) $x_1, x_2, \ldots, x_n$ tal que

$$v = x_1 \cdot v_1 + x_2 \cdot v_2 + \cdots + x_n \cdot v_n.$$

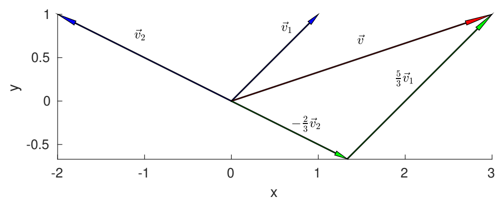


Figura 4.1: Componentes del vector $\vec{v}$ con respecto a la base $\{\vec{v}_1, \vec{v}_2\}$ de $\mathbb{R}^2$ del Ejemplo 4.10.

El vector $(x_1, x_2, \ldots, x_n)^t$ se llama el *vector de coordenadas de v con respecto a la base S*, y escribimos

$$(v)_S = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Ejemplo 4.11. El conjunto $S = \operatorname{span}\left\{x+1, x-1, x^2+1\right\}$ es una base de P_3. Vamos a calcular $(p)_S$ para $p(x) = -3 + 2x - 4x^2$. Escribiendo $(p)_S = (a, b, c)^t$, tenemos que

$$a(x+1) + b(x-1) + c(x^2+1) = -3 + 2x - 4x^2.$$

Agrupando las potencias de x similares, tenemos que

$$a - b + c + (a+b)x + cx^2 = -3 + 2x - 4x^2.$$

Tenemos pues que $a = 3/2$, $b = 1/2$, $c = -4$, por lo que

$$(p)_S = \begin{bmatrix} 3/2 \\ 1/2 \\ -4 \end{bmatrix}.$$

□

Usando la definición de vector de coordenadas en un espacio vectorial arbitrario, en particular la unicidad de éste, se puede verificar el siguiente resultado. El mismo tiene como consecuencia que todo espacio vectorial de dimensión (finita) n es *isomorfo*[1] a $\mathbb{R}^n$.

Teorema 4.12. *Sea V un espacio vectorial con base $S = \{v_1, v_2, \ldots, v_n\}$. Sean $v, w \in V$ y $\alpha, \beta \in \mathbb{R}$. Entonces*

$$(\alpha \cdot v + \beta \cdot w)_S = \alpha(v)_S + \beta(w)_S.$$

En particular, la transformación $L : V \longrightarrow \mathbb{R}^n$ definida por $L(v) = (v)_S$ es lineal, 1–1, y sobre.

4.3 Representaciones Matriciales

La especificación de una transformación lineal $L : V \longrightarrow W$ entre los espacios vectoriales V y W de dimensión finita, no depende en general de las bases de estos espacios. No obstante, si especificamos bases S_1, S_2 de V y W respectivamente, veremos que es posible asociar con L una matriz $m \times n$ que en cierto sentido representa completamente a la transformación L. Antes de discutir el caso general, examinamos el caso en que $V = \mathbb{R}^n$ y $W = \mathbb{R}^m$.

4.3.1 Transformaciones de $\mathbb{R}^n$ a $\mathbb{R}^m$

Sean

$$\begin{aligned} S_1 &= \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} \text{ una base para } \mathbb{R}^n, \\ S_2 &= \{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m\} \text{ una base para } \mathbb{R}^m. \end{aligned}$$

Nota 4.13. El orden de los vectores de las bases se mantiene según se estipula en las definiciones de S_1, S_2. Es por esto que decimos que las bases son *ordenadas*. □

[1] Dos espacios vectoriales V, W son isomorfos si existe una transformación lineal $L : V \longrightarrow W$ que es 1–1 y sobre.

Suponga que $L:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ es una transformación lineal. Sabemos que para cada vector $\vec{v}_j$ de la base S_1 existen escalares $(a_{1j}, a_{2j}, \ldots, a_{mj})^t$ tal que

$$L(\vec{v}_j) = \sum_{i=1}^{m} a_{ij}\vec{w}_i, \quad (L(\vec{v}_j))_{S_2} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, \quad j = 1, 2, \ldots, n.$$

Definimos la matriz $A \in M_{m\times n}$ por

$$A = (a_{ij}) = \Big[(L(\vec{v}_1))_{S_2}, (L(\vec{v}_2))_{S_2}, \ldots, (L(\vec{v}_n))_{S_2}\Big]. \tag{4.1}$$

Sea $\vec{x} \in \mathbb{R}^n$ un vector cualquiera con:

$$\vec{x} = \sum_{j=1}^{n} x_j\vec{v}_j, \quad (\vec{x})_{S_1} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Tenemos ahora que

$$\begin{aligned} L(\vec{x}) &= L\left(\sum_{j=1}^{n} x_j\vec{v}_j\right) = \sum_{j=1}^{n} x_j L(\vec{v}_j), \\ &= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}\vec{w}_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)\vec{w}_i. \end{aligned}$$

De esta última ecuación podemos concluir que:

$$(L(\vec{x}))_{S_2} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{bmatrix} = A\,(\vec{x})_{S_1}. \tag{4.2}$$

La matriz A se llama la *representación matricial de la transformación lineal L con respecto a las bases S_1 y S_2 de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente.*

Nota 4.14. Recuerde que la base canónica o estándar de $\mathbb{R}^n$ está dada por $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ donde los vectores $\{\vec{e}_i\}$ son las columnas de la matriz identidad $n \times n$. Cuando las bases S_1 y S_2 de arriba son las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente, entonces la matriz A se llama *la representación matricial canónica o estándar de* L . □

Ejemplo 4.15. Considere la función $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ dada por:

$$L(\vec{x}) = \begin{bmatrix} x_1 \\ 2x_2 \\ x_1 + x_2 \end{bmatrix}, \quad \vec{x} = [x_1, x_2]^t.$$

Primeramente verificamos que L es lineal. Para esto note que si $\vec{x} = (x_1, x_2)^t$, $\vec{y} = (y_1, y_2)^t$ son vectores arbitrarios de $\mathbb{R}^2$, y $\alpha, \beta \in \mathbb{R}$, entonces

$$\begin{aligned} L(\alpha\vec{x} + \beta\vec{y}) &= L(\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2), \\ &= \begin{bmatrix} \alpha x_1 + \beta y_1 \\ 2(\alpha x_2 + \beta y_2) \\ (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) \end{bmatrix}, \\ &= \begin{bmatrix} \alpha x_1 \\ 2(\alpha x_2) \\ \alpha(x_1 + x_2) \end{bmatrix} + \begin{bmatrix} \beta y_1 \\ 2(\beta y_2) \\ \beta(y_1 + y_2) \end{bmatrix}, \\ &= \alpha \begin{bmatrix} x_1 \\ 2x_2 \\ x_1 + x_2 \end{bmatrix} + \beta \begin{bmatrix} y_1 \\ 2y_2 \\ y_1 + y_2 \end{bmatrix}, \\ &= \alpha L(\vec{x}) + \beta L(\vec{y}). \end{aligned}$$

Podemos ahora concluir que L es una transformación lineal.

Buscamos ahora la representación matricial canónica de L. Según nuestra convención explicada en la Nota 4.13, $\vec{x} = [x_1, x_2]^t$ quiere decir que $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2$ donde $\{\vec{e}_1, \vec{e}_2\}$ es la base canónica de $\mathbb{R}^2$, y $L(\vec{x}) = [x_1, 2x_2, x_1 + x_2]^t$ quiere decir que $L(\vec{x}) = x_1\vec{u}_1 + 2x_2\vec{u}_2 + (x_1 + x_2)\vec{u}_3$ donde $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es la base canónica de $\mathbb{R}^3$. Así que L se puede escribir también de la forma:

$$L(x_1\vec{e}_1 + x_2\vec{e}_2) = x_1\vec{u}_1 + 2x_2\vec{u}_2 + (x_1 + x_2)\vec{u}_3.$$

Tenemos entonces que:

$$L(\vec{e}_1) \quad = \quad L(1,0) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = (1)\vec{u}_1 + (0)\vec{u}_2 + (1)\vec{u}_3,$$

$$L(\vec{e}_2) = L(0,1) = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = (0)\vec{u}_1 + 2\vec{u}_2 + (1)\vec{u}_3.$$

Podemos concluir que la representación matricial canónica de L es:

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

por lo que $L(\vec{x}) = A\vec{x}$. □

Ejemplo 4.16. Buscamos ahora la representación matricial para el mismo operador del ejemplo anterior pero ahora con respecto a las bases:

$$S_1 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}, \quad \text{para } \mathbb{R}^2,$$

$$S_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}, \quad \text{para } \mathbb{R}^3.$$

Necesitamos, primero que nada, buscar los vectores de coordenadas:

$$(L(1,1))_{S_2}, \quad (L(-2,1))_{S_2}.$$

Para lograr ésto usamos que:

$$L(1,1) = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha_1 + \alpha_3 \\ \alpha_2 + \alpha_3 \\ \alpha_1 + \alpha_2 \end{bmatrix},$$

$$L(-2,1) = \begin{bmatrix} -2 \\ 2 \\ -1 \end{bmatrix} = \beta_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \beta_3 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \beta_1 + \beta_3 \\ \beta_2 + \beta_3 \\ \beta_1 + \beta_2 \end{bmatrix}.$$

Así que tenemos que resolver dos sistemas lineales con la misma matriz de coeficientes:

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

y lados derecho $(1,2,2)^t$ y $(-2,2,-1)^t$ respectivamente. Usando transformaciones de filas llegamos a:

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

De aquí se obtiene que:

$$\alpha_1 = \frac{1}{2}, \quad \alpha_2 = \frac{3}{2}, \quad \alpha_3 = \frac{1}{2}, \quad \beta_1 = -\frac{5}{2}, \quad \beta_2 = \frac{3}{2}, \quad \beta_3 = \frac{1}{2}.$$

La representación matricial del operador L del ejemplo anterior con respecto a las bases S_1, S_2 está dada entonces por:

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

□

Ejemplo 4.17. Para el operador L, con las mismas bases y representación matricial del Ejemplo 4.16, calculamos:

$$(L(3,1))_{S_2}.$$

Del Ejemplo 4.10 sabemos que

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix}_{S_1} = \begin{bmatrix} 5/3 \\ -2/3 \end{bmatrix}.$$

Usando la representación matricial de L del ejemplo anterior y la formula (4.2), tenemos que

$$(L(3,1))_{S_2} = A \begin{bmatrix} 3 \\ 1 \end{bmatrix}_{S_1} = \begin{bmatrix} 1/2 & -5/2 \\ 3/2 & 3/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 5/3 \\ -2/3 \end{bmatrix} = \begin{bmatrix} 5/2 \\ 3/2 \\ 1/2 \end{bmatrix}.$$

Esto quiere decir que

$$L(3,1) = \left(\frac{5}{2}\right) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \left(\frac{3}{2}\right) \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \left(\frac{1}{2}\right) \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

□

Ejemplo 4.18. Considere la función $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ dada por:

$$L(\vec{x}) = L(x_1, x_2, x_3) = x_1 \vec{b}_1 + (x_2 + x_3)\vec{b}_2,$$

donde

$$\vec{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Es fácil ver que L es lineal. Buscamos ahora la representación matricial de L con respecto a la base estándar de $\mathbb{R}^3$ y la base $S_2 = \left\{\vec{b}_1, \vec{b}_2\right\}$ de $\mathbb{R}^2$. Usando la definición de L tenemos que:

$$L(1,0,0) = \vec{b}_1, \quad L(0,1,0) = \vec{b}_2, \quad L(0,0,1) = \vec{b}_2.$$

De aquí podemos leer inmediatamente los vectores de coordenadas de $L(1,0,0)$, $L(0,1,0)$, y $L(0,0,1)$ con respecto a S_2, por lo que concluimos que la representación matricial deseada es:

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

□

Ejemplo 4.19. Considere la función $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ dada por:

$$L(x,y,z) = \begin{bmatrix} x + y + z \\ x + 2y + 3z \end{bmatrix}.$$

Buscamos la representación matricial de L respecto a las bases:

$$S_1 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \quad \text{para } \mathbb{R}^3,$$

$$S_2 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}, \quad \text{para } \mathbb{R}^2.$$

Primero que nada, necesitamos encontrar los vectores de coordenadas:

$$(L(1,1,0))_{S_2}, \quad (L(0,1,1))_{S_2}, \quad (L(0,0,1))_{S_2}.$$

Para ésto note que:

$$L(1,1,0) = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} \alpha_1 + \alpha_2 \\ 2\alpha_1 + 3\alpha_2 \end{bmatrix},$$

$$L(0,1,1) = \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \beta_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \beta_2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} \beta_1 + \beta_2 \\ 2\beta_1 + 3\beta_2 \end{bmatrix},$$

$$L(0,0,1) = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \gamma_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \gamma_2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} \gamma_1 + \gamma_2 \\ 2\gamma_1 + 3\gamma_2 \end{bmatrix}.$$

Tenemos entonces que resolver tres sistemas lineales con la misma matriz de coeficientes:

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

y lados derecho $(2,3)^t$, $(2,5)^t$, y $(1,3)^t$ respectivamente. Para hacer ésto note que:

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

De aquí se obtiene que:

$$\alpha_1 = 3, \quad \alpha_2 = -1, \quad \beta_1 = 1, \quad \beta_2 = 1, \quad \gamma_1 = 0, \quad \gamma_2 = 1,$$

por lo que la representación matricial del operador L con respecto a las bases S_1, S_2 está dada por:

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

□

4.3.2 Transformaciones generales

En esta sección vamos a generalizar la discusión de la sección anterior al caso de transformaciones lineales entre espacios vectoriales arbitrarios (no necesariamente $\mathbb{R}^n$ y $\mathbb{R}^m$).

Sean V, W espacios vectoriales con

$$\begin{aligned} S_1 &= \{v_1, v_2, \ldots, v_n\} \text{ una base para } V, \\ S_2 &= \{w_1, w_2, \ldots, w_m\} \text{ una base para } W. \end{aligned}$$

Nota 4.20. El orden de los vectores de las bases se mantiene según se estipula en las definiciones de S_1, S_2. □

Sea $L : V \longrightarrow W$ un operador lineal. Para cada j entre 1 y n, podemos escribir que

$$L(v_j) = \sum_{i=1}^{m} a_{ij} \cdot w_i,$$

por lo que

$$(L(v_j))_{S_2} = \left[a_{1j}, a_{2j}, \ldots, a_{mj}\right]^t.$$

Defina la matriz A de tamaño $m \times n$ por $A = (a_{ij})$. Tenemos entonces el siguiente resultado.

Teorema 4.21. *Sean V, W espacios vectoriales con bases S_1, S_2 respectivamente según se describió anteriormente. Sea*

$$A = (a_{ij}) = \left[(L(v_1))_{S_2}, (L(v_2))_{S_2}, \ldots, (L(v_n))_{S_2}\right]. \tag{4.3}$$

Entonces

$$(L(v))_{S_2} = A(v)_{S_1}, \quad \forall\, v \in V. \tag{4.4}$$

Demostración: Si $v \in V$ y $(v)_{S_1} = (x_1, x_2, \ldots, x_n)^t$, entonces

$$v = \sum_{j=1}^{n} x_j \cdot v_j.$$

Usando ésto podemos ahora escribir que:

$$\begin{aligned} L(v) &= L\left(\sum_{j=1}^{n} x_j \cdot v_j\right) = \sum_{j=1}^{n} x_j \cdot L(v_j), \\ &= \sum_{j=1}^{n} x_j \cdot \sum_{i=1}^{m} a_{ij} \cdot w_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j\right) \cdot w_i, \\ &= \sum_{i=1}^{n} \left[A(v)_{S_1}\right]_i \cdot w_i. \end{aligned}$$

Podemos ahora concluir que (4.4) se cumple. □

La matriz (4.3) se llama la *representación matricial del operador L con respecto a las bases S_1 y S_2.*

Ejemplo 4.22. Defina el operador $L : P_2 \longrightarrow P_3$ por:

$$L(p)(x) = \int_1^x p(t)\,\mathrm{d}t, \quad \forall\, x \in \mathbb{R}.$$

Se puede verificar que L es un operador lineal. Vamos a utilizar las siguientes bases:

$$S_1 = \{1, x\} \text{ para } P_2, \quad S_2 = \left\{1, x, x^2\right\} \text{ para } P_3,$$

para calcular la representación matricial de L con respecto a S_1 y S_2. Si escribimos $p_1(x) = 1, p_2(x) = x$, entonces tenemos que:

$$\begin{aligned}
(L(p_1))_{S_2} &= \left(\int_1^x 1\,\mathrm{d}t\right)_{S_2} = (x-1)_{S_2}, \\
&= ((-1)1 + (1)x + (0)x^2)_{S_2} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \\
(L(p_2))_{S_2} &= \left(\int_1^x t\,\mathrm{d}t\right)_{S_2} = \left(\frac{1}{2}(x^2-1)\right)_{S_2}, \\
&= \left((-\frac{1}{2})1 + (0)x + (\frac{1}{2})x^2\right)_{S_2} = \begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}.
\end{aligned}$$

Tenemos entonces que la representación matricial de L con respecto a estas bases es:

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

Si $p \in P_2$ con $p(x) = a + bx$, entonces

$$(L(p))_{S_2} = A(p)_{S_1} = \begin{bmatrix} -1 & -\frac{1}{2} \\ 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -a - b/2 \\ a \\ b/2 \end{bmatrix},$$

ésto es

$$L(p)(x) = -a - b/2 + ax + (b/2)x^2.$$

$\square$

En el caso especial de (4.3) cuando $V = W$ y $S = S_1 = S_2$, la matriz A se llama la *representación matricial de L con respecto a la base S*.

Ejemplo 4.23. Considere el espacio vectorial $V = \operatorname{span}\{1, \operatorname{sen} x, \cos x\}$. Defina el operador $D : V \longrightarrow V$ por

$$D(f) = f'.$$

Se puede verificar que D es lineal. Buscamos la representación matricial de D con respecto a la base $S = \{1, \operatorname{sen} x, \cos x\}$ de V. Como

$$\begin{aligned}
(D(1))_S &= (0)_S = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \\
(D(\operatorname{sen} x))_S &= (\cos x)_S = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \\
(D(\cos x))_S &= (-\operatorname{sen} x)_S = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix},
\end{aligned}$$

tenemos que

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

Ahora, si $f \in V$ con $f = a + b \operatorname{sen} x + c \cos x$, tenemos que

$$(D(f))_S = A(f)_S = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ -c \\ b \end{bmatrix},$$

es decir que

$$D(f)(x) = (0)1 + (-c) \operatorname{sen} x + (b) \cos x = -c \operatorname{sen} x + b \cos x.$$

$\square$

4.4 Cambio de Bases

Considere el caso especial de la formula (4.3) donde $V = W$, y $L : V \longrightarrow V$ es la transformación identidad, i.e.,

$$L(v) = v, \quad v \in V.$$

Entonces si $S_1 = \{v_1, v_2, \dots, v_n\}$ y $S_2 = \{w_1, w_2, \dots, w_n\}$ son bases de V, la matriz A en la formula (4.3) se llama la *matriz de transición o cambio de base de S_1 a S_2*. En este caso denotamos a A con la letra S y tenemos que

$$S = \Big[(v_1)_{S_2}, (v_2)_{S_2}, \dots, (v_n)_{S_2}\Big], \tag{4.5}$$

y de la formula (4.3) que:

$$(v)_{S_2} = S\,(v)_{S_1}, \quad v \in V. \tag{4.6}$$

Ejemplo 4.24. Considere los conjuntos:

$$S_1 = \{\vec{v}_1, \vec{v}_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}, \quad S_2 = \{\vec{w}_1, \vec{w}_2\} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right\}.$$

Entonces S_1, S_2 son bases de $\mathbb{R}^2$. (¿Por qué?) Ahora

$$(\vec{v}_1)_{S_2} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \text{ si y solo si, } \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 3 \\ -2 \end{bmatrix},$$

$$(\vec{v}_2)_{S_2} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \text{ si y solo si, } \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \beta_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \beta_2 \begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$

Para obtener los α's y β's tenemos que resolver dos sistemas con la misma matriz de coeficientes:

$$\left[\begin{array}{rr|rr} 1 & 3 & 1 & 2 \\ -1 & -2 & 1 & -1 \end{array}\right] \longrightarrow \left[\begin{array}{rr|rr} 1 & 3 & 1 & 2 \\ 0 & 1 & 2 & 1 \end{array}\right].$$

De aquí obtenemos que:

$$(\vec{v}_1)_{S_2} = \begin{bmatrix} -5 \\ 2 \end{bmatrix}, \quad (\vec{v}_2)_{S_2} = \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

por lo que la matriz de transición de S_1 a S_2 está dada por:

$$S = \begin{bmatrix} -5 & -1 \\ 2 & 1 \end{bmatrix}.$$

Si $\vec{x} = 3\vec{v}_1 - 4\vec{v}_2$, es decir $(\vec{x})_{S_1} = (3, -4)^t$, entonces

$$(\vec{x})_{S_2} = S(\vec{x})_{S_1} = \begin{bmatrix} -5 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -4 \end{bmatrix} = \begin{bmatrix} -11 \\ 2 \end{bmatrix}.$$

Esto quiere decir que $\vec{x} = -11\vec{w}_1 + 2\vec{w}_2$. □

Ejemplo 4.25. Para las bases S_1 y S_2 del ejemplo anterior, buscamos $(\vec{x})_{S_2}$ donde $\vec{x} = 6\vec{v}_1 + 2\vec{v}_2$. Como

$$(\vec{x})_{S_1} = \begin{bmatrix} 6 \\ 2 \end{bmatrix},$$

(¿por qué?) tenemos usando la formula (4.6) que

$$(\vec{x})_{S_2} = S\,(\vec{x})_{S_1} = \begin{bmatrix} -5 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \end{bmatrix} = \begin{bmatrix} -32 \\ 14 \end{bmatrix}.$$

Esto significa que $\vec{x}$ se puede escribir en trminos de la base S_2 como:

$$\vec{x} = -32\vec{w}_1 + 14\vec{w}_2.$$

□

Ejemplo 4.26. Considere la base estándar $S_1 = \{\vec{e}_1, \vec{e}_2\}$ de $\mathbb{R}^2$, y el conjunto

$$S_2 = \{\vec{u}_1, \vec{u}_2\} = \{\cos\theta\,\vec{e}_1 + \operatorname{sen}\theta\,\vec{e}_2, -\operatorname{sen}\theta\,\vec{e}_1 + \cos\theta\,\vec{e}_2\}.$$

S_2 es también una base para $\mathbb{R}^2$. Note que S_2 se obtiene rotando los vectores de S_1 por el ángulo θ con respecto al eje positivo de x. Es fácil también ver que:

$$\vec{e}_1 = \cos\theta\,\vec{u}_1 - \operatorname{sen}\theta\,\vec{u}_2, \quad \vec{e}_2 = \operatorname{sen}\theta\,\vec{u}_1 + \cos\theta\,\vec{u}_2.$$

De éstas representaciones podemos leer que:

$$(\vec{e}_1)_{S_2} = \begin{bmatrix} \cos\theta \\ -\operatorname{sen}\theta \end{bmatrix}, \quad (\vec{e}_2)_{S_2} = \begin{bmatrix} \operatorname{sen}\theta \\ \cos\theta \end{bmatrix}.$$

De la formula (4.5) tenemos que la matriz de transición de S_1 a S_2 está dada por:

$$S = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

Si tenemos que $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2$, entonces:

$$(\vec{x})_{S_2} = S(\vec{x})_{S_1} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1\cos\theta + x_2\operatorname{sen}\theta \\ -x_1\operatorname{sen}\theta + x_2\cos\theta \end{bmatrix}.$$

Estas son las conocidas formulas de *rotación de coordenadas* que especifican como cambian las coordenadas de un vector en $\mathbb{R}^2$ cuando los ejes se rotan por el ángulo θ. $\square$

Ejemplo 4.27. Considere las siguientes bases de P_3:

$$S_1 = \left\{1, x+1, x^2-2\right\}, \quad S_2 = \left\{x+1, x+2, x^2+1\right\}.$$

Vamos a calcular la matriz de transición de S_1 a S_2. Note que:

$$\begin{aligned}
(1)_{S_2} &= ((-1)(x+1)+(1)(x+2)+(0)(x^2+1))_{S_2} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \\
(x+1)_{S_2} &= ((1)(x+1)+(0)(x+2)+(0)(x^2+1))_{S_2} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \\
(x^2-2)_{S_2} &= ((3)(x+1)+(-3)(x+2)+(1)(x^2+1))_{S_2} = \begin{bmatrix} 3 \\ -3 \\ 1 \end{bmatrix}.
\end{aligned}$$

Tenemos entonces que la matriz de transición es:

$$S = \begin{bmatrix} -1 & 1 & 3 \\ 1 & 0 & -3 \\ 0 & 0 & 1 \end{bmatrix}.$$

Si $p(x) = (a)(1) + (b)(x+1) + (c)(x^2-2)$, entonces

$$(p)_{S_2} = S(p)_{S_1} = \begin{bmatrix} -1 & 1 & 3 \\ 1 & 0 & -3 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -a+b+3c \\ a-3c \\ c \end{bmatrix}.$$

De aquí que podemos escribir a p en términos de la base S_2 como:

$$p(x) = (-a+b+3c)(x+1) + (a-3c)(x+2) + c(x^2+1).$$

$\square$

Se puede verificar que las matrices de transición en todos estos ejemplos son nosingulares. Estos son solo casos especiales del siguiente resultado.

Proposición 4.28. *Sea S la matriz de transición de la base S_1 a la base S_2 del espacio vectorial V. Entonces S es nosingular y la inversa S^{-1} es la matriz de transición de S_2 a S_1.*

Demostración: Suponga que $S\vec{x} = \vec{0}$, donde $\vec{x} = (x_1, x_2, \ldots, x_n)^t$. Defina

$$v = x_1 \cdot v_1 + x_2 \cdot v_2 + \cdots x_n \cdot v_n,$$

de modo que $(v)_{S_1} = \vec{x}$. Como

$$(v)_{S_2} = S(v)_{S_1} = S\vec{x} = \vec{0},$$

tenemos por la unicidad del vector de coordenadas, que $v = \theta$. Pero ésto implica a su vez que $\vec{x} = (v)_{S_1} = \vec{0}$. Así que $S\vec{x} = \vec{0}$ implica que $\vec{x} = \vec{0}$, por lo que S es nosingular (Teorema 1.37). $\square$

Concluimos este capitulo con un resultado sobre la relación entre las representaciones matriciales de un operador lineal con respecto a distintas bases. En esta discusión nos restringimos a operadores $L : V \longrightarrow V$.

Teorema 4.29. *Sea V un espacio vectorial con bases S_1, S_2, y $L : V \longrightarrow V$ un operador lineal. Defina las matrices A_1, A_2 por:*

A_1 es la representación matricial de L con respecto a S_1,
A_2 es la representación matricial de L con respecto a S_2.

Si S es la matriz de transición de la base S_1 a la base S_2, entonces

$$A_1 = S^{-1} A_2 S. \tag{4.7}$$

Demostración: Usamos la notación (V,S_i) para representar la descripción de V usando la base S_i, donde $i=1,2$. Entonces, por definición:

$$\begin{aligned}
L:(V,S_1)&\longrightarrow(V,S_1), &\quad (L(v))_{S_1}&=A_1(v)_{S_1},\\
L:(V,S_2)&\longrightarrow(V,S_2), &\quad (L(v))_{S_2}&=A_2(v)_{S_2},\\
I:(V,S_1)&\longrightarrow(V,S_2), &\quad (v)_{S_2}&=S(v)_{S_1},
\end{aligned}$$

Tenemos ahora que:

$$S(L(v))_{S_1}=(L(v))_{S_2}=A_2(v)_{S_2}=A_2S(v)_{S_1},$$

es decir,

$$(L(v))_{S_1}=S^{-1}A_2S(v)_{S_1}.$$

Pero $(L(v))_{S_1}=A_1(v)_{S_1}$, de modo que

$$A_1(v)_{S_1}=S^{-1}A_2S(v)_{S_1}.$$

Como $(v)_{S_1}$ es arbitrario, esto implica que $A_1=S^{-1}A_2S$. □

Ejemplo 4.30. Considere el operador $L:P_3\longrightarrow P_3$ definido por

$$L(p)(x)=xp'(x),\quad p\in P_3.$$

En este ejemplo usamos las siguientes bases de P_3:

$$S_1=\left\{1,x+1,x^2-2\right\},\quad S_2=\left\{x+1,x+2,x^2+1\right\}.$$

Del Ejemplo 4.27 sabemos que la matriz de transición de S_1 a S_2 es:

$$S=\begin{bmatrix}-1&1&3\\1&0&-3\\0&0&1\end{bmatrix}.$$

Calculamos ahora las matrices A_1 y A_2 del teorema. Para calcular A_1, observe que:

$$(L(1))_{S_1}\quad=\quad(0)_{S_1}=\begin{bmatrix}0\\0\\0\end{bmatrix},$$

$$
\begin{aligned}
(L(x+1))_{S_1} &= (x)_{S_1} = ((-1)(1) + (1)(x+1) + (0)(x^2-2))_{S_1} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \\
(L(x^2-2))_{S_1} &= (2x^2)_{S_1} = ((4)(1) + (0)(x+1) + (2)(x^2-2))_{S_1} = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}.
\end{aligned}
$$

De estos cálculos tenemos que

$$
A_1 = \begin{bmatrix} 0 & -1 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.
$$

Para la matriz A_2 usamos que:

$$
\begin{aligned}
(L(x+1))_{S_2} &= (x)_{S_2} = ((2)(x+1) + (-1)(x+2) + (0)(x^2+1))_{S_2} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \\
(L(x+2))_{S_2} &= (x)_{S_2} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \\
(L(x^2+1))_{S_2} &= (2x^2)_{S_2} = ((2)(x+1) + (-2)(x+2) + (2)(x^2+1))_{S_2} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}.
\end{aligned}
$$

Tenemos entonces que

$$
A_2 = \begin{bmatrix} 2 & 2 & 2 \\ -1 & -1 & -2 \\ 0 & 0 & 2 \end{bmatrix}.
$$

Multiplicando ahora es fácil verificar que $SA_1 = A_2S$. $\square$

4.5 Ejercicios

Ejercicio 4.1. Verifique que la función $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ dada por:

$$
L(x_1, x_2) = (2x_1 - x_2, x_1 + x_2, x_1 - 2x_2)^t,
$$

es una tranformación lineal.

Ejercicio 4.2. Sea $\vec{u} \in \mathbb{R}^n$ un vector distinto del vector cero. Defina $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ por

$$L(\vec{x}) = \left[\frac{\vec{x}^t\vec{u}}{\vec{u}^t\vec{u}}\right]\vec{u}.$$

Verifique que L es una transformación lineal. L es la transformación que proyecta cualquier vector a la recta que pasa por el origen con dirección dada por $\vec{u}$. Note que L se puede escribir también de la forma $L(\vec{x}) = A\vec{x}$ donde la matriz A esta dada por

$$A = \left[\frac{1}{\vec{u}^t\vec{u}}\right]\vec{u}\vec{u}^t = \left[\frac{1}{\vec{u}^t\vec{u}}\right]\vec{u}\otimes\vec{u}.$$

Ejercicio 4.3. Verifique que la función $L : P_3 \longrightarrow P_2$ dada por:

$$L(p(x)) = p(0)x + p(1),$$

es una tranformación lineal.

Ejercicio 4.4. Verifique que la función $L : P_2 \longrightarrow P_2$ dada por:

$$L(ax+b) = ax + a + b,$$

es una tranformación lineal.

Ejercicio 4.5. Sea $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ una transformación lineal. Suponga que $L(\vec{\mathbf{e}}_1) = (-2,3,2)^t$, $L(\vec{\mathbf{e}}_2) = (3,-4,1)^t$, $L(\vec{\mathbf{e}}_3) = (1,2,-1)^t$, donde $\{\vec{\mathbf{e}}_1,\vec{\mathbf{e}}_2,\vec{\mathbf{e}}_3\}$ es la base estándar de $\mathbb{R}^3$. Halle $L(4,3,6)$.

Ejercicio 4.6. Sea $S_1 = \{\vec{\mathbf{e}}_1,\vec{\mathbf{e}}_2\}$ la base estándar de $\mathbb{R}^2$ y $S_2 = \{\vec{\mathbf{v}}_1,\vec{\mathbf{v}}_2,\vec{\mathbf{v}}_3\}$ una base para $\mathbb{R}^3$. Para $\vec{\mathbf{x}} = x_1\vec{\mathbf{e}}_1 + x_2\vec{\mathbf{e}}_2$ considere la transformación lineal $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ dada por:

$$L(\vec{\mathbf{x}}) = (x_1 - x_2)\vec{\mathbf{v}}_1 + x_1\vec{\mathbf{v}}_2 + (2x_1 - 3x_2)\vec{\mathbf{v}}_3.$$

a) Halle la representación matricial de L con respecto a las bases S_1, S_2.

b) Dado que $\vec{\mathbf{x}} = \vec{\mathbf{e}}_1 - 2\vec{\mathbf{e}}_2$, halle $[L(\vec{\mathbf{x}})]_{S_2}$.

Ejercicio 4.7. Sea $S_1 = \{\vec{v}_1,\vec{v}_2,\vec{v}_3\}$ una base de $\mathbb{R}^3$. Para $\vec{x} = x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3$, considere la transformación lineal $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ dada por:

$$L(\vec{x}) = \begin{bmatrix} x_1 - x_2 + 2x_3 \\ 2x_1 - 3x_2 - x_3 \end{bmatrix}.$$

Usando la base $S_2 = \{\vec{w}_1,\vec{w}_2\}$ de $\mathbb{R}^2$, donde $\vec{w}_1 = [1,2]^t$ y $\vec{w}_2 = [1,-1]^t$, halle:

a) la representación matricial de L con respecto a las bases S_1, S_2.

b) Dado que $\vec{p} = \vec{v}_1 - 2\vec{v}_2 + 5\vec{v}_3$, halle $[L(\vec{p})]_{S_2}$ y escriba $L(\vec{p})$ en términos de $\{\vec{w}_1, \vec{w}_2\}$.

Ejercicio 4.8. Sea $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la transformación lineal dada por $T(x,y) = (2x+3y, 4x-5y)^t$. Halle la representación matricial de T con respecto a la base $S = \{\vec{u}_1, \vec{u}_2\}$ de $\mathbb{R}^2$ donde:

$$\vec{u}_1 = (1,2)^t, \quad \vec{u}_2 = (2,5)^t.$$

Ejercicio 4.9. El operador $L : P_3 \longrightarrow P_2$ dado por $L(p(x)) = p(1)x + p(2)$ es lineal.

a) Halle la representación matricial de L con respecto a las bases (ordenadas) $S_1 = \left\{x, x^2, 1\right\}$ y $S_2 = \{x+1, 1-x\}$ de P_3 y P_2 respectivamente.

b) Halle $(L(2x^2 - x + 3))_{S_2}$ y con ésto escriba $L(2x^2 - x + 3)$ en términos de la base S_2.

Ejercicio 4.10. Considere las bases $S_1 = \left\{1, x, x^2\right\}$ y $S_2 = \{1, x-1, (x-1)(x-2)\}$ de P_3.

a) Halle la matriz de transición para el cambio de base de S_1 a S_2.

b) Calcule $(-2 + 3x + 4x^2)_{S_2}$.

Ejercicio 4.11. Sean $S_1 = \{\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2\}$ y $S_2 = \{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2\}$ bases de $\mathbb{R}^2$. Demuestre que la matriz de transición para el cambio de base de S_1 a S_2 es nosingular. **Ayuda:** Recuerde que una matriz A es nosingular si y solo si la única solución de $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$ es $\vec{\mathbf{x}} = \vec{\mathbf{0}}$.

Ejercicio 4.12. Los siguientes conjuntos:

$$U = \{\vec{u}_1, \vec{u}_2\}, \quad \vec{u}_1 = (1,-2)^t, \quad \vec{u}_2 = (3,-4)^t,$$

$$V = \{\vec{v}_1, \vec{v}_2\}, \quad \vec{v}_1 = (1,3)^t, \quad \vec{v}_2 = (3,8)^t.$$

son bases de $\mathbb{R}^2$.

a) Halle la matriz de transición de la base U a la base V. (10 pts)

b) Dado que $\vec{x} = 3\vec{u}_1 - 2\vec{u}_2$, halle $(\vec{x})_V$ y escriba $\vec{x}$ como una combinación lineal de $\{\vec{v}_1, \vec{v}_2\}$. (5 pts)

Ejercicio 4.13. Considere las bases $S_1 = \{\vec{v}_1, \vec{v}_2\}$ y $S_2 = \{\vec{w}_1, \vec{w}_2\}$, ambas de $\mathbb{R}^2$, donde:

$$\vec{v}_1 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \vec{w}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \vec{w}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

a) Halle la matriz de transición o de cambio de base de S_1 a S_2.

b) Usando la matriz de transición, escriba el vector $\vec{p} = -3\vec{v}_1 + 4\vec{v}_2$ en términos de la base S_2.

c) Para la transformación lineal $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por $L(x, y) = (5x - y, 2x + y)^t$, encuentre la representación matricial A_1 de L con respecto a la base S_1.

d) Usando el resultado del Teorema 4.29, encuentre la representación matricial A_2 de L con respecto a la base S_2.

Ejercicio 4.14. Defina la transformación lineal $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ por $L(\vec{x}) = A\vec{x}$ donde $\vec{x} \in \mathbb{R}^3$ y

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

a) Calcule la matriz de transición de la base estándar de $\mathbb{R}^3$ a la base U de $\mathbb{R}^3$ dada por

$$U = \left\{[1, 1, 0]^t, [0, 1, 1]^t, [1, 2, 2]^t\right\}.$$

b) Usando el resultado del Teorema 4.29, encuentre la representación matricial de L con respecto a la base U. (**Ayuda:** Note que A es la representación matricial de L con respecto a la base estándar de $\mathbb{R}^3$.)

Ejercicio 4.15. Sean V, W espacios vectoriales y $L : V \longrightarrow W$ un operador lineal. Verifique que para cualquier entero $n \geq 1$, escalares $\alpha_1, \ldots, \alpha_n$, y vectores $v_1, \ldots, v_n$ en V, tenemos que

$$L(\alpha_1 \cdot v_1 + \cdots \alpha_n \cdot v_n) = \alpha_1 \cdot L(v_1) + \cdots + \alpha_n \cdot L(v_n).$$

5 Ortogonalidad

En este capitulo vamos a desarrollar la noción de *perpendicularidad* o *ortogonalidad* de vectores en un espacio vectorial. El concepto fundamental para esta discusión es el de *producto escalar o interior*. Con el producto interior se puede entonces definir el largo de un vector lo cual a su vez nos permite caracterizar la noción de ángulo entre vectores desde el punto de vista algebraico (no geométrico). Por supuesto, esta caracterización de perpendicularidad que no utiliza ángulos, reduce al concepto usual en términos de ángulos cuando la dimensión n del espacio es 2 ó 3. Aunque nuestra discusión se limitará al caso de vectores en $\mathbb{R}^n$, la mayoría de los resultados que veremos, pueden generalizarse a espacios vectoriales más generales.

5.1 Producto interior o escalar

Sean $\vec{x} = (x_1, x_2, \ldots, x_n)^t$, $\vec{y} = (y_1, y_2, \ldots, y_n)^t$ vectores en $\mathbb{R}^n$. El *producto interior, o punto, o escalar* de $\vec{x}$ con $\vec{y}$ se representa con $<\vec{x}, \vec{y}>$ y se define por:

$$<\vec{x},\vec{y}> = \vec{x}^t\vec{y} = \sum_{k=1}^{n} x_k y_k.$$

Es costumbre también utilizar la notación $\vec{x}\cdot\vec{y}$ para representar el producto interior.

Ejemplo 5.1. Para los vectores $\vec{x} = [2,-1,6]^t$ y $\vec{y} = [-1,3,5]^t$, tenemos que

$$<\vec{x},\vec{y}> = (2)(-1) + (-1)(3) + (6)(5) = 25.$$

□

El producto interior de vectores tiene las siguientes propiedades:

i) $<\vec{x},\vec{x}> \geq 0$ y $<\vec{x},\vec{x}> = 0$ si y solo si $\vec{x} = \vec{0}$.

ii) $<\vec{x},\vec{y}> = <\vec{y},\vec{x}>$.

iii) $<\alpha\vec{x} + \beta\vec{y}, \vec{z}> = \alpha <\vec{x},\vec{z}> + \beta <\vec{y},\vec{z}>$, $\alpha, \beta \in \mathbb{R}$.

Nota 5.2. Si fijamos $\vec{z} \in \mathbb{R}^n$ y definimos el operador $L_{\vec{z}} : \mathbb{R}^n \longrightarrow \mathbb{R}$ por

$$L_{\vec{z}}(\vec{x}) = <\vec{x}, \vec{z}>,$$

entonces la propiedad (iii) de arriba nos dice que $L_{\vec{z}}$ es un operador lineal. □

La *norma o largo* del vector $\vec{x}$ se define por:

$$\left\|\vec{x}\right\| = \sqrt{<\vec{x}, \vec{x}>} = \sqrt{\sum_{k=1}^{n} x_k^2}.$$

Note que $\left\|\vec{x}\right\|$ representa la distancia del origen en $\mathbb{R}^n$ al punto $[x_1, \ldots, x_n]^t$. Decimos que el vector $\vec{x}$ es *unitario* si $\left\|\vec{x}\right\| = 1$. Si $\vec{x} \neq \vec{0}$ no tiene norma uno, entonces

$$\vec{u} = \frac{\vec{x}}{\left\|\vec{x}\right\|} \equiv \left(\frac{1}{\left\|\vec{x}\right\|}\right)\vec{x},$$

es unitario y decimos que se obtuvo *normalizando* a $\vec{x}$.

Ejemplo 5.3. Para el vector $\vec{x} = [2, -3, 4]^t$ tenemos que

$$\left\|\vec{x}\right\| = \left[2^2 + (-3)^2 + 4^2\right]^{1/2} = \sqrt{29}.$$

El vector

$$\vec{u} = \frac{1}{\sqrt{29}}\vec{x} = \left[\frac{2}{\sqrt{29}}, \frac{-3}{\sqrt{29}}, \frac{4}{\sqrt{29}}\right]^t,$$

es unitario o de largo uno. □

La función $\|\cdot\|$ tiene las siguientes propiedades:

i) $\left\|\vec{x}\right\| \geq 0$ y $\left\|\vec{x}\right\| = 0$ si y solo si $\vec{x} = \vec{0}$.

ii) $\left\|\alpha\vec{x}\right\| = |\alpha| \left\|\vec{x}\right\|$

iii) $\left\|\vec{x} + \vec{y}\right\| \leq \left\|\vec{x}\right\| + \left\|\vec{y}\right\|$ (la desigualdad del triangulo)

El siguiente resultado nos da una relación entre el producto interior de dos vectores y sus respectivos largos.

Teorema 5.4 (Desigualdad de Cauchy–Schwartz)**.** *Sean $\vec{x},\vec{y} \in \mathbb{R}^n$. Entonces*

$$\left| < \vec{x},\vec{y} > \right| \leq \left\| \vec{x} \right\| \left\| \vec{y} \right\|,$$

con igualdad si y solo si los vectores $\vec{x}$, $\vec{y}$ son proporcionales.

Demostración: Sea λ un número real cualquiera. Entonces

$$\begin{aligned} 0 \leq \left\| \vec{x} + \lambda \vec{y} \right\|^2 &= (\vec{x} + \lambda \vec{y})^t (\vec{x} + \lambda \vec{y}), \\ &= \vec{x}^t \vec{x} + \lambda \vec{y}^t \vec{x} + \lambda \vec{x}^t \vec{y} + \lambda^2 \vec{y}^t \vec{y}, \\ &= \left\| \vec{x} \right\|^2 + 2\lambda < \vec{x},\vec{y} > + \lambda^2 \left\| \vec{y} \right\|^2. \end{aligned}$$

El lado derecho de esta desigualdad es una cuadrática en λ. La desigualdad entonces implica que el discriminante de esta cuadrática en λ debe ser menor o igual que cero, es decir:

$$4 < \vec{x},\vec{y} >^2 - 4 \left\| \vec{x} \right\|^2 \left\| \vec{y} \right\|^2 \leq 0,$$

de donde se obtiene luego de despejar y tomar raíces, que

$$\left| < \vec{x},\vec{y} > \right| \leq \left\| \vec{x} \right\| \left\| \vec{y} \right\|.$$

Si hay igualdad en la desigualdad anterior y $\vec{y} \neq \vec{0}$, entonces el discriminante de la cuadrática en λ es cero, y

$$\hat{\lambda} = -\frac{< \vec{x},\vec{y} >}{\left\| \vec{y} \right\|^2},$$

es una raíz de esta cuadrática. Tenemos entonces que $\left\| \vec{x} + \hat{\lambda} \vec{y} \right\|^2 = 0$, es decir que $\vec{x} + \hat{\lambda} \vec{y} = \vec{0}$, lo que implica que $\vec{x}$, $\vec{y}$ son proporcionales. Si $\vec{y} = \vec{0}$, entonces $\vec{x},\vec{y}$ son dependientes, y por lo tanto, proporcionales. □

Si ninguno de los vectores $\vec{x},\vec{y}$ es cero, la desigualdad de Cauchy–Schwartz implica que:

$$-1 \leq \frac{< \vec{x},\vec{y} >}{\left\| \vec{x} \right\| \left\| \vec{y} \right\|} \leq 1.$$

Esto garantiza que la siguiente definición es válida.

Definición 5.5. Sean $\vec{x},\vec{y} \in \mathbb{R}^n$ vectores distintos del vector cero. El *ángulo θ entre estos vectores* se define por:

$$\theta = \cos^{-1}\left[\frac{<\vec{x},\vec{y}>}{\|\vec{x}\|\,\|\vec{y}\|}\right].$$

Nota 5.6. Para los casos $n = 2,3$ se puede verificar que el ángulo θ de la definición, coincide con el ángulo geométrico entre las flechas asociadas con los vectores $\vec{x},\vec{y}$. □

Ejemplo 5.7. Para los vectores $\vec{x} = [2,-1,6]^t$ y $\vec{y} = [-1,3,5]^t$, tenemos que

$$\begin{aligned}
<\vec{x},\vec{y}> &= (2)(-1)+(-1)(3)+(6)(5) = 25,\\
\|\vec{x}\| &= \sqrt{2^2+(-1)^2+6^2} = \sqrt{41},\\
\|\vec{y}\| &= \sqrt{(-1)^2+3^2+5^2} = \sqrt{35}.
\end{aligned}$$

Tenemos entonces que el ángulo entre los vectores $\vec{x}$ y $\vec{y}$ es:

$$\theta = \cos^{-1}\left[\frac{<\vec{x},\vec{y}>}{\|\vec{x}\|\,\|\vec{y}\|}\right] = \cos^{-1}\left[\frac{25}{\sqrt{41}\,\sqrt{35}}\right] \approx 0.8500 \text{ radianes.}$$

□

Note que si en la definición del ángulo entre los vectores $\vec{x},\vec{y}$ ponemos que $<\vec{x},\vec{y}>= 0$, entonces $\theta = \cos^{-1}(0) = \frac{\pi}{2}$. Esto motiva la siguiente definición.

Definición 5.8. Los vectores $\vec{x},\vec{y} \in \mathbb{R}^n$ son *ortogonales o perpendiculares*, si $<\vec{x},\vec{y}>= 0$. En tal caso escribimos que $\vec{x}\perp\vec{y}$.

Nota 5.9. Por definición, el vector cero es ortogonal a cualquier vector $\vec{x} \in \mathbb{R}^n$. □

Ejemplo 5.10. Si $\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\}$ es la base estándar de $\mathbb{R}^n$ (las columnas de la matriz identidad $n\times n$), entonces los vectores $\{\vec{e}_i\}$ son perpendiculares entre si ya que

$$<\vec{e}_i,\vec{e}_j>= \begin{cases} 1 & , \quad i = j,\\ 0 & , \quad i \neq j. \end{cases}$$

□

Ejemplo 5.11. Si $\vec{x} = (3,-2,6)^t$ y $\vec{y} = (1,2,1/6)^t$, entonces:

$$< \vec{x}, \vec{y} >= (3)(1) + (-2)(2) + (6)(1/6) = 0,$$

ésto es, $\vec{x} \perp \vec{y}$. □

Ejemplo 5.12. Dados los vectores $(2,1,2)^t$, $(-1,1,2)^t$, buscamos todos los vectores en $\mathbb{R}^3$ que son perpendiculares a ambos vectores. Si $(x,y,z)^t$ es uno de los vectores que buscamos, entonces la condición de ser perpendicular a los dos vectores dados es equivalente al sistema:

$$\left\{ \begin{array}{rcl} 2x + y + 2z & = & 0, \\ -x + y + 2z & = & 0. \end{array} \right.$$

Este sistema tiene solución $x = 0$, $y = -2z$ por lo que cualquier vector de la forma:

$$(0, -2z, z), \quad z \in \mathbb{R},$$

es perpendicular a los dos vectores dados. □

5.1.1 Producto interior en $\mathbb{C}$

Sean $\vec{x} = (x_1, x_2, \ldots, x_n)^t$, $\vec{y} = (y_1, y_2, \ldots, y_n)^t$ vectores en $\mathbb{C}^n$. El *producto interior, o punto, o escalar* de $\vec{x}$ con $\vec{y}$ se representa con $< \vec{x}, \vec{y} >$ y se define por:

$$< \vec{x}, \vec{y} >= \vec{x}^t \overline{\vec{y}} = \sum_{k=1}^{n} x_k \overline{y}_k,$$

donde la barra representa el conjugado del número complejo o vector en cuestión. Es costumbre también utilizar la notación $\vec{x} \cdot \vec{y}$ para representar el producto interior.

Ejemplo 5.13. Para los vectores $\vec{x} = [3 + 2\imath, -3, 6 + \imath]^t$ y $\vec{y} = [-1, 1 + 2\imath, 5\imath]^t$, tenemos que

$$\begin{aligned} < \vec{x}, \vec{y} > &= (3 + 2\imath)(-1) + (-3)(1 - 2\imath) + (6 + \imath)(-5\imath), \\ &= -3 - 2\imath - 3 + 6\imath - 30\imath + 5 = -1 - 26\imath. \end{aligned}$$

□

El producto interior (complejo) de vectores tiene las siguientes propiedades:

i) $<\vec{x},\vec{x}>\geq 0$ y $<\vec{x},\vec{x}>=0$ si y solo si $\vec{x}=\vec{0}$.

ii) $<\vec{x},\vec{y}>=\overline{<\vec{y},\vec{x}>}$.

iii) $<\alpha\vec{x}+\beta\vec{y},\vec{z}>=\alpha<\vec{x},\vec{z}>+\beta<\vec{y},\vec{z}>,\ \ \alpha,\beta\in\mathbb{R}$.

La *norma o largo* de un vector $\vec{x}\in\mathbb{C}^n$ se define igual que antes y tiene las mismas propiedades que en el caso de vectores reales. También tenemos que la desigualdad de Cauchy–Schwartz es válida para vectores $\vec{x},\vec{y}\in\mathbb{C}^n$.

5.1.2 Proyecciones escalares y vectoriales

Sean $\vec{x},\vec{y}\in\mathbb{R}^n$, $\vec{x},\vec{y}\neq\vec{0}$. Sea

$$\vec{u}=\frac{\vec{y}}{\|\vec{y}\|}.$$

Buscamos un escalar $\alpha\in\mathbb{R}$ y un vector $\vec{z}$, $\vec{z}\perp\vec{y}$, tal que

$$\vec{x}=\alpha\vec{u}+\vec{z}.$$

El escalar α se llama la *proyección escalar* de $\vec{x}$ sobre $\vec{y}$ y está dado por:

$$\alpha=\frac{<\vec{x},\vec{y}>}{\|\vec{y}\|}.$$

El vector $\alpha\vec{u}$ se llama la *proyección vectorial* de $\vec{x}$ sobre $\vec{y}$ y lo denotamos por $\vec{p}_{\vec{y}}(\vec{x})$. Usando las formulas de α y $\vec{u}$, tenemos que

$$\vec{p}_{\vec{y}}(\vec{x})=\frac{<\vec{x},\vec{y}>}{\|\vec{y}\|^2}\vec{y}=\frac{<\vec{x},\vec{y}>}{<\vec{y},\vec{y}>}\vec{y}.$$

Ejemplo 5.14. Buscamos las proyecciones escalar y vectorial de $(2,1,-1)^t$ en $(4,1,-5)^t$. Poniendo $\vec{x}=(2,1,-1)^t$ y $\vec{y}=(4,1,-5)^t$, tenemos que:

$$<\vec{x},\vec{y}>=8+1+5=14,\qquad \|\vec{y}\|=\sqrt{16+1+25}=\sqrt{42}.$$

De aquí que la proyección escalar de $\vec{x}$ en $\vec{y}$ es:

$$\alpha=\frac{14}{\sqrt{42}},$$

y la proyección vectorial de $\vec{x}$ en $\vec{y}$:

$$\vec{p}_{\vec{y}}(\vec{x}) = \frac{14}{42}(4,1,-5)^t = (4/3,1/3,-5/3)^t.$$

$\square$

Ejemplo 5.15. Buscamos el punto en la recta $y = 5x+1$ que esté más cerca del punto $\vec{x} = (4,4)^t$. Si $\vec{u}_1, \vec{u}_2$ son puntos en la recta dada, entonces es fácil ver mediante un diagrama, que el punto de la recta que está más cerca de $\vec{x}$ es el vector $\vec{u}_1$ mas la proyección vectorial de $\vec{x}-\vec{u}_1$ sobre $\vec{u}_2-\vec{u}_1$. Tomando $\vec{u}_1 = (0,1)^t$ y $\vec{u}_2 = (1,6)^t$, tenemos que el punto de la recta más cerca de $\vec{x}$ es:

$$\begin{aligned}
\vec{u}_1 + \frac{<\vec{x}-\vec{u}_1, \vec{u}_2-\vec{u}_1>}{\left\|\vec{u}_2-\vec{u}_1\right\|^2}(\vec{u}_2-\vec{u}_1) &= (0,1)^t + \frac{<(4,3)^t,(1,5)^t>}{\|(1,5)^t\|^2}(1,5)^t, \\
&= (0,1)^t + (19/26, 95/26)^t, \\
&= (19/26, 121/26)^t.
\end{aligned}$$

$\square$

Más general aún, si $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k \in \mathbb{R}^n$ son vectores perpendiculares[1] entre si, donde $k \leq n$, entonces un argumento similar al anterior se puede usar para ver que si $\vec{a} \in \mathbb{R}^n$, la *proyección vectorial de $\vec{a}$ al espacio S =* span$\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ está dada por:

$$\vec{p}_S(\vec{a}) = \sum_{j=1}^{k} \alpha_j \vec{u}_j, \quad \alpha_j = \frac{<\vec{a}, \vec{u}_j>}{<\vec{u}_j, \vec{u}_j>}, \quad j = 1, \ldots, k. \tag{5.1}$$

Note que los términos de la sumatoria son las respectivas proyecciones vectoriales del vector $\vec{a}$ con respecto a cada uno de los vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$. La proyección vectorial de $\vec{a}$ al espacio S tiene la propiedad de que:

$$\left\|\vec{a} - \vec{p}_S(\vec{a})\right\| = \min_{\vec{y} \in S} \left\|\vec{a} - \vec{y}\right\|.$$

[1] El caso en que los vectores $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ son linealmente independientes pero no necesariamente perpendiculares entre si, se puede trabajar usando la técnica de cuadrados mínimos que estudiaremos en la Sección 5.5.

Ejemplo 5.16. Los vectores $\vec{u}_1 = (1,-2,1)^t$, $\vec{u}_2 = (1,3,5)^t$ son perpendiculares entre si. La proyección vectorial de $\vec{a} = (4,-1,2)^t$ al plano S generado por estos dos vectores está dada por:

$$\begin{aligned}\vec{p}_S(\vec{a}) &= \left(\frac{<\vec{a},\vec{u}_1>}{<\vec{u}_1,\vec{u}_1>}\right)\vec{u}_1 + \left(\frac{<\vec{a},\vec{u}_2>}{<\vec{u}_2,\vec{u}_2>}\right)\vec{u}_2,\\ &= \frac{4}{3}(1,-2,1)^t + \frac{11}{35}(1,3,5)^t = \left(\frac{173}{105}, -\frac{181}{105}, \frac{61}{21}\right)^t.\end{aligned}$$

□

5.1.3 Ecuación de un plano

Sean $\vec{x}_0$, $\vec{n}$ vectores en $\mathbb{R}^3$. Buscamos la ecuación del plano que contiene a $\vec{x}_0$ y tiene a $\vec{n}$ como dirección normal al plano. Si $\vec{x}$ es un vector cualquiera con punto terminal en el plano, entonces $\vec{x} - \vec{x}_0$ es ortogonal a $\vec{n}$, i.e.,

$$<\vec{x} - \vec{x}_0, \vec{n}> = 0,$$

lo cual representa una ecuación para el plano. En particular, si ponemos

$$\vec{x} = (x,y,z)^t, \quad \vec{x}_0 = (x_0,y_0,z_0)^t, \quad \vec{n} = (a,b,c)^t,$$

entonces la ecuación de arriba es equivalente a:

$$a(x-x_0) + b(y-y_0) + c(z-z_0) = 0. \tag{5.2}$$

La ecuación (5.2) también se puede escribir como:

$$ax + by + cz + d = 0, \tag{5.3}$$

donde $d = -ax_0 - by_0 - cz_0$. Así que tenemos que todo plano se puede representar por una ecuación lineal de la forma (5.3). Se puede también verificar que cualquier ecuación lineal de la forma (5.3) representa un plano en $\mathbb{R}^3$.

Ejemplo 5.17. Buscamos una ecuación para el plano que contiene el punto $(4,-2,1)^t$ y que es paralelo al plano $3x - 5y + 3z = 10$. Como los planos son paralelos, tienen direcciones normales paralelas o proporcionales. Usando

$\vec{n} = (3, -5, 3)^t$ como dirección normal para el plano que buscamos, tenemos que una ecuación para éste plano es:

$$3(x-4) - 5(y+2) + 3(z-1) = 0,$$

o simplificando que

$$3x - 5y + 3z - 25 = 0.$$

□

Sea P un plano con dirección normal $\vec{n} = (a, b, c)^t$ y que contiene el punto $\vec{x}_0 = (x_0, y_0, z_0)^t$. Sea $\vec{x}^* = (x^*, y^*, z^*)^t$ otro punto no necesariamente en P. Entonces la distancia de $\vec{x}^*$ a P está dada por el valor absoluto de la proyección escalar de $\vec{x}^* - \vec{x}_0$ en $\vec{n}$:

$$\begin{aligned} \text{Distancia} &= \left| \frac{< \vec{x}^* - \vec{x}_0, \vec{n} >}{\|\vec{n}\|} \right|, \\ &= \left| \frac{a(x^* - x_0) + b(y^* - y_0) + c(z^* - z_0)}{\sqrt{a^2 + b^2 + c^2}} \right|, \\ &= \left| \frac{ax^* + by^* + cz^* + d}{\sqrt{a^2 + b^2 + c^2}} \right|. \end{aligned}$$

Ejemplo 5.18. La ecuación del plano que contiene el punto $(1, -1, 3)$ con dirección normal $(2, 3, 6)$ está dada por:

$$2(x-1) + 3(y+1) + 6(z-3) = 0.$$

La distancia del punto $(3, 1, -2)$ al plano es:

$$d = \left| \frac{2(3-1) + 3(1+1) + 6(-2-3)}{\sqrt{2^2 + 3^2 + 6^2}} \right| = \left| \frac{-20}{\sqrt{49}} \right| = \frac{20}{7}.$$

□

5.2 Espacios Ortogonales

Sean X, Y subespacios de $\mathbb{R}^n$. Decimos que X es ortogonal a Y (se escribe $X \perp Y$) si

$$\vec{x} \perp \vec{y}, \quad \forall\, \vec{x} \in X, \quad \vec{y} \in Y.$$

O sea, que todo vector de X es ortogonal a todos los vectores de Y y viceversa. Note que la condición de ortogonalidad para X, Y se puede escribir también como:

$$< \vec{x}, \vec{y} >= 0, \quad \forall\, \vec{x} \in X, \quad \vec{y} \in Y.$$

Ejemplo 5.19. Sea A una matriz $m \times n$. Recuerde que

$$N(A) = \left\{ \vec{x} \in \mathbb{R}^n \, : \, A\vec{x} = \vec{0} \right\},$$

y $R(A^t)$ es el espacio columna de A^t o lo mismo, el espacio fila de A. Vamos a ver que $R(A^t) \perp N(A)$. Para verificar ésto, tomamos un $\vec{x} \in R(A^t)$, y un $\vec{y} \in N(A)$. De aquí tenemos que $\vec{x} = A^t \vec{z}$, para algún $\vec{z} \in \mathbb{R}^m$, y que $A\vec{y} = \vec{0}$. Así que

$$< \vec{x}, \vec{y} >= \vec{x}^t \vec{y} = (A^t \vec{z})^t \vec{y} = \vec{z}^t A \vec{y} = \vec{z}^t \vec{0} = 0.$$

Como $\vec{x} \in R(A^t)$, y $\vec{y} \in N(A)$ son arbitrarios, tenemos que el computo anterior verifica que $R(A^t) \perp N(A)$. Más adelante veremos que la relación entre $R(A^t)$ y $N(A)$ es todavía más estrecha. □

Ejemplo 5.20. Considere los siguientes subespacios de $\mathbb{R}^3$:

$$\begin{aligned} X &= \left\{ (x, 2x, -x)^t \, : \, x \in \mathbb{R} \right\}, \\ Y &= \left\{ (3y, -y, y)^t \, : \, y \in \mathbb{R} \right\}. \end{aligned}$$

Verificamos que $X \perp Y$. Si $\vec{x} \in X$ y $\vec{y} \in Y$ entonces:

$$\vec{x} = (x, 2x, -x)^t, \quad \vec{y} = (3y, -y, y)^t,$$

para algunos $x, y \in \mathbb{R}$. Tenemos pues que

$$\begin{aligned} < \vec{x}, \vec{y} > &= (x)(3y) + (2x)(-y) + (-x)(y), \\ &= 3xy - 2xy - xy = 0. \end{aligned}$$

Como $\vec{x} \in X$, $\vec{y} \in Y$ son arbitrarios, ésto demuestra que $X \perp Y$. □

Note que en el ejemplo anterior, tenemos que el vector $\vec{z} = (2, -1, 0)^t$ es perpendicular a todos los vectores en X pero $\vec{z} \notin Y$. Nos interesa caracterizar el conjunto de todos los vectores perpendiculares a todos los vectores de un subespacio cualquiera X.

Definición 5.21. Sea X un subespacio de $\mathbb{R}^n$. El conjunto de todos los vectores que son perpendiculares a todos los vectores de X se llama el *complemento ortogonal* de X y se denota por $X^\perp$. Esto es

$$X^\perp = \{\vec{y} \in \mathbb{R}^n : <\vec{x}, \vec{y}>= 0, \ \forall \vec{x} \in X\}.$$

Es inmediato de la definición que $X \perp X^\perp$.

Ejemplo 5.22. Considere el subespacio X de $\mathbb{R}^3$ del ejemplo anterior:

$$X = \left\{(x, 2x, -x)^t \, : \, x \in \mathbb{R}\right\}.$$

Buscamos $X^\perp$. Si $\vec{w} = (u, v, w)^t \in X^\perp$, entonces $<\vec{x}, \vec{w}>= 0$ para todo $\vec{x} \in X$. Usando la definición de X, tenemos que esto implica que

$$xu + 2xv - xw = 0.$$

Cancelando la x, si $x \neq 0$, tenemos que ésto implica que $u + 2v - w = 0$. De aquí que

$$X^\perp = \left\{(-2v + w, v, w)^t \, : \, v, w \in \mathbb{R}\right\}.$$

Note que $\vec{z} = (2, -1, 0)^t \in X^\perp$ y que $Y \subset X^\perp$, $Y \neq X^\perp$, donde Y es como en el Ejemplo 5.20. □

Suponga que $X = \text{span}\{\vec{u}_1, \ldots, \vec{u}_k\}$. Es fácil verificar ahora que

$$X^\perp = \{\vec{y} : <\vec{u}_i, \vec{y}>= 0, \ 1 \leq i \leq k\}.$$

Ejemplo 5.23. Sea $X = \text{span}\{(-1, 2, 0, 4)^t, (3, 0, -4, 5)^t\}$. Usando la observación anterior, tenemos que $(a, b, c, d)^t \in X^\perp$ si y solo si

$$\left\{ \begin{array}{rcl} -a + 2b + 4d & = & 0, \\ 3a - 4c + 5d & = & 0. \end{array} \right.$$

Este sistema tiene solución $a = 2b + 4d$, $c = (6b + 17d)/4$, donde $b, d \in \mathbb{R}$. Tenemos entonces que

$$\begin{array}{rcl} X^\perp & = & \{(2b + 4d, b, (6b + 17d)/4, d) \, : \, b, d \in \mathbb{R}\}, \\ & = & \text{span}\left\{(2, 1, 3/2, 0)^t, (4, 0, 17/4, 1)^t\right\}. \end{array}$$

□

El siguiente teorema resume las propiedades más importantes sobre espacios ortogonales y complementos ortogonales.

Teorema 5.24. *Sean X, Y subespacios de $\mathbb{R}^n$, entonces:*

i) *$X^\perp$ es un subespacio de $\mathbb{R}^n$.*

ii) $\dim X + \dim X^\perp = n$

iii) *Si $X \perp Y$, entonces $X \cap Y = \left\{\vec{0}\right\}$.*

iv) *Si $X \perp Y$, entonces $X \subset Y^\perp$.*

v) $(X^\perp)^\perp = X$.

Demostración:

i) Sean $\vec{x}_1, \vec{x}_2 \in X^\perp$. Entonces

$$< \vec{x}_1, \vec{y} >= 0 =< \vec{x}_2, \vec{y} >, \quad \forall \vec{y} \in X.$$

Si $\alpha, \beta \in \mathbb{R}$, tenemos entonces que para cualquier $\vec{y} \in X$:

$$< \alpha\vec{x}_1 + \beta\vec{x}_2, \vec{y} >= \alpha < \vec{x}_1, \vec{y} > + \beta < \vec{x}_1, \vec{y} >= (\alpha)(0) + (\beta)(0) = 0.$$

Podemos concluir entonces que $\alpha\vec{x}_1 + \beta\vec{x}_2, \vec{y} \in X^\perp$, es decir, que $X^\perp$ es un subespacio de $\mathbb{R}^n$.

ii) Sea $\dim X = k$ con $X = \operatorname{span}\{\vec{u}_1, \ldots, \vec{u}_k\}$ donde $k \le n$. Si $k = n$, entonces $X = \mathbb{R}^n$ y $X^\perp = \left\{\vec{0}\right\}$ por lo que el resultado es cierto. Suponga pues que $k < n$. Es fácil ver que

$$X^\perp = \{\vec{y} : < \vec{u}_i, \vec{y} >= 0, \ \ i = 1, \ldots, k\} = N(A^t),$$

donde $A = [\vec{u}_1, \ldots, \vec{u}_k]$. Como A^t es $k \times n$ y $k < n$, el sistema homogéneo $A^t\vec{y} = \vec{0}$ es subdeterminado, por lo que es consistente con $n - k$ variables libres. Tenemos entonces que $\dim X^\perp = \dim N(A^t) = n - k$.

iii) Suponga que $X \perp Y$ y que $\vec{v} \in X \cap Y$. Entonces como $\vec{v} \in Y$ y $X \perp Y$, tenemos que

$$< \vec{x}, \vec{v} >= 0, \quad \forall \vec{x} \in X.$$

Pero también $\vec{v} \in X$. Sustituyendo $\vec{x} = \vec{v}$ en la ecuación anterior, tenemos que $< \vec{v}, \vec{v} >= 0$, es decir que $\vec{v} = \vec{0}$. Así que $X \cap Y = \left\{\vec{0}\right\}$.

iv) Suponga que $X \perp Y$ y que $\vec{v} \in X$. Como $X \perp Y$, tenemos que

$$<\vec{x},\vec{y}>= 0, \quad \forall \vec{x} \in X, \quad \vec{y} \in Y.$$

En particular, tomando $\vec{x} = \vec{v}$, que

$$<\vec{v},\vec{y}>= 0, \quad \forall \vec{y} \in Y.$$

Esto ultimo implica que $\vec{v} \in Y^{\perp}$. Como $\vec{v} \in X$ es arbitrario, podemos concluir que $X \subset Y^{\perp}$.

v) Como $X \perp X^{\perp}$, podemos usar la parte iv) con $Y = X^{\perp}$, para concluir que $X \subset (X^{\perp})^{\perp}$. Si $\dim X = k$, entonces sabemos de la parte ii) que $\dim X^{\perp} = n - k$. Usando la parte ii) nuevamente pero remplazando X con $X^{\perp}$, tenemos ahora que

$$\dim (X^{\perp})^{\perp} = n - \dim X^{\perp} = n - (n-k) = k.$$

En resumen tenemos que $X \subset (X^{\perp})^{\perp}$, $\dim X = \dim (X^{\perp})^{\perp} = k$, lo que implica que $X = (X^{\perp})^{\perp}$.

$\square$

Vamos ahora a estudiar varias consecuencias del teorema anterior. Note que de la parte (ii) del Teorema 5.24, si X es un subespacio de $\mathbb{R}^n$ con $\dim X = k$, entonces $\dim X^{\perp} = n - k$.

Proposición 5.25. *Sea X un subespacio de $\mathbb{R}^n$ con $\dim X = k$. Sean $\{\vec{v}_1, \ldots, \vec{v}_k\}$ y $\{\vec{v}_{k+1} \ldots, \vec{v}_n\}$ bases de X y $X^{\perp}$ respectivamente. Entonces $\{\vec{v}_1, \ldots, \vec{v}_n\}$ es una base para $\mathbb{R}^n$.*

Demostración: Ya que son n vectores en $\mathbb{R}^n$, es suficiente demostrar que $\{\vec{v}_1, \ldots, \vec{v}_n\}$ son linealmente independientes. Suponga que existen escalares $c_1, \ldots, c_n$ tal que

$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}.$$

Podemos escribir esta ecuación como:

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = -c_{k+1}\vec{v}_{k+1} - \cdots - c_n\vec{v}_n.$$

Pero

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k \quad \in \quad \operatorname{span}\{\vec{v}_1, \ldots, \vec{v}_k\} = X,$$

$$-c_{k+1}\vec{v}_{k+1} - \cdots - c_n\vec{v}_n \quad \in \quad \operatorname{span}\{\vec{v}_{k+1}\ldots,\vec{v}_n\} = X^{\perp},$$

lo que combinado con la ecuación anterior implica que

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k \in X \cap X^{\perp}, \quad -c_{k+1}\vec{v}_{k+1} - \cdots - c_n\vec{v}_n \in X \cap X^{\perp}.$$

Pero $X \cap X^{\perp} = \left\{\vec{0}\right\}$ (Teorema 5.24, parte (iii)) lo que implica que

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}, \quad -c_{k+1}\vec{v}_{k+1} - \cdots - c_n\vec{v}_n = \vec{0}.$$

Como los conjuntos $\{\vec{v}_1,\ldots,\vec{v}_k\}$ y $\{\vec{v}_{k+1}\ldots,\vec{v}_n\}$ son linealmente independientes cada uno, las dos ecuaciones anteriores implican que:

$$c_1 = \cdots = c_k = 0, \quad c_{k+1} = \cdots = c_n = 0,$$

lo que finalmente nos da que $\{\vec{v}_1,\ldots,\vec{v}_n\}$ son linealmente independientes.

□

Ejemplo 5.26. Del Ejemplo 5.23 tenemos que para

$$X = \operatorname{span}\left\{(-1,2,0,4)^t,(3,0,-4,5)^t\right\},$$

el complemento ortogonal está dado por:

$$X^{\perp} = \operatorname{span}\left\{(2,1,3/2,0)^t,(4,0,17/4,1)^t\right\}.$$

Como $\{(-1,2,0,4)^t,(3,0,-4,5)^t\}$ y $\{(2,1,3/2,0)^t,(4,0,17/4,1)^t\}$ son cada uno linealmente independientes, tenemos por la Proposición 5.25 que

$$\left\{(-1,2,0,4)^t,(3,0,-4,5)^t,(2,1,3/2,0)^t,(4,0,17/4,1)^t\right\},$$

es una base para $\mathbb{R}^4$. □

Si X y Y son subespacios de $\mathbb{R}^n$, la *suma directa* de X y Y se define por:

$$X \oplus Y = \{\vec{x}+\vec{y} : \vec{x} \in X, \quad \vec{y} \in Y\}.$$

Es fácil verificar que $X \oplus Y$ es un subespacio de $\mathbb{R}^n$. En el caso especial en que $Y = X^{\perp}$, tenemos el siguiente resultado.

Proposición 5.27. *Sea X un subespacio de $\mathbb{R}^n$. Entonces $X \oplus X^{\perp} = \mathbb{R}^n$. En particular, para cualquier vector $\vec{v} \in \mathbb{R}^n$, existen vectores únicos $\vec{s} \in X$, $\vec{s}^{\perp} \in X^{\perp}$ tal que*

$$\vec{v} = \vec{s} + \vec{s}^{\perp},$$

y

$$\left\|\vec{v}\right\|^2 = \left\|\vec{s}\right\|^2 + \left\|\vec{s}^{\perp}\right\|^2.$$

Demostración: Sean $\{\vec{v}_1, \ldots, \vec{v}_k\}$ y $\{\vec{v}_{k+1} \ldots, \vec{v}_n\}$ bases de X y $X^{\perp}$ respectivamente. Por el resultado de la Proposición 5.25, el conjunto $\{\vec{v}_1, \ldots, \vec{v}_n\}$ es una base de $\mathbb{R}^n$. De modo que si $\vec{v}$ es cualquier vector en $\mathbb{R}^n$, entonces existen escalares $c_1, \ldots, c_n$ tal que

$$\vec{v} = c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n.$$

Esta ecuación la podemos escribir como

$$\vec{v} = c_1 \vec{v}_1 + \cdots + c_k \vec{v}_k + c_{k+1} \vec{v}_{k+1} + \cdots + c_n \vec{v}_n.$$

Como

$$\begin{aligned}
\vec{s} &\equiv c_1 \vec{v}_1 + \cdots + c_k \vec{v}_k \in \operatorname{span}\{\vec{v}_1, \ldots, \vec{v}_k\} = X, \\
\vec{s}^{\perp} &\equiv c_{k+1} \vec{v}_{k+1} + \cdots + c_n \vec{v}_n \in \operatorname{span}\{\vec{v}_{k+1} \ldots, \vec{v}_n\} = X^{\perp},
\end{aligned}$$

tenemos que la representación anterior de $\vec{v}$ implica que $\vec{v} = \vec{s} + \vec{s}^{\perp} \in X \oplus X^{\perp}$. Como $\vec{v}$ es arbitrario, tenemos que $\mathbb{R}^n \subset X \oplus X^{\perp}$. Como siempre $X \oplus X^{\perp} \subset \mathbb{R}^n$, podemos concluir que $X \oplus X^{\perp} = \mathbb{R}^n$. También, para el calculo de $\left\|\vec{v}\right\|^2$, tenemos que:

$$\begin{aligned}
\left\|\vec{v}\right\|^2 &= < \vec{s} + \vec{s}^{\perp}, \vec{s} + \vec{s}^{\perp} >, \\
&= < \vec{s}, \vec{s} > + 2 < \vec{s}, \vec{s}^{\perp} > + < \vec{s}^{\perp}, \vec{s}^{\perp} >, \\
&= \left\|\vec{s}\right\|^2 + \left\|\vec{s}^{\perp}\right\|^2,
\end{aligned}$$

donde usamos que $< \vec{s}, \vec{s}^{\perp} >= 0$.

Finalmente, si $\vec{v} = \vec{s}_1 + \vec{s}_1^{\perp} = \vec{s}_2 + \vec{s}_2^{\perp}$ con $\vec{s}_1, \vec{s}_2 \in X$ y $\vec{s}_1^{\perp}, \vec{s}_2^{\perp} \in X^{\perp}$, entonces

$$\vec{s}_1 - \vec{s}_2 = \vec{s}_2^{\perp} - \vec{s}_1^{\perp}.$$

Pero $\vec{s}_1 - \vec{s}_2 \in X$ y $\vec{s}_2^{\perp} - \vec{s}_1^{\perp} \in X^{\perp}$. Usando ahora que $X \cap X^{\perp} = \left\{\vec{0}\right\}$ obtenemos que:

$$\vec{s}_1 - \vec{s}_2 = \vec{0}, \quad \vec{s}_2^{\perp} - \vec{s}_1^{\perp} = \vec{0},$$

es decir que $\vec{s}_1 = \vec{s}_2$, $\vec{s}_1^{\perp} = \vec{s}_2^{\perp}$. $\square$

5.3 Los espacios fundamentales

Sea A una matriz de tamaño $m \times n$. Los *espacios fundamentales* de A son:

$$N(A), \quad R(A), \quad N(A^t), \quad R(A^t).$$

En el Ejemplo 5.19 vimos que $N(A) \perp R(A^t)$. Por el Teorema 5.24, parte (iv), tenemos que

$$N(A) \subset R(A^t)^{\perp}.$$

Ahora, si $\vec{x} \in R(A^t)^{\perp}$, entonces $\vec{x}$ es ortogonal a las columnas de A^t las cuales son las filas de A. Así que $A\vec{x} = \vec{0}$, i.e., $\vec{x} \in N(A)$. Por lo tanto

$$R(A^t)^{\perp} \subset N(A),$$

que combinado con la inclusión anterior demuestra que

$$N(A) = R(A^t)^{\perp}. \tag{5.4}$$

Poniendo A^t por A en la ecuación (5.4) y usando que $(A^t)^t = A$ obtenemos la identidad:

$$N(A^t) = R(A)^{\perp}.$$

Usando ahora el resultado de la parte (v) del Teorema 5.24 en las ultimas dos identidades, obtenemos que:

$$N(A)^{\perp} = R(A^t), \quad N(A^t)^{\perp} = R(A).$$

Combinando estos resultados con otro que vimos en el Capitulo 3, tenemos el siguiente teorema.

Teorema 5.28 (Fundamental del Álgebra Lineal)**.** *Sea A una matriz $m \times n$. Entonces:*

i) $N(A) = R(A^t)^{\perp}, \quad N(A^t) = R(A)^{\perp}$.

ii) $N(A)^{\perp} = R(A^t), \quad N(A^t)^{\perp} = R(A)$.

iii) $\mathit{rank}\, A + \dim N(A) = n$.

De la identidad $N(A^t)^{\perp} = R(A)$, tenemos que el sistema $A\vec{x} = \vec{b}$ tiene solución si y solo si $\vec{b}$ es perpendicular a todos los vectores en $N(A^t)$. Esta observación es parte del resultado que es conocido como el *primer Teorema de la Alternativa de Fredholm* el cual enunciamos a continuación.

Teorema 5.29. *Sea A una matriz $m \times n$ y $\vec{b} \in \mathbb{R}^m$.*

i) Si $R(A) = \mathbb{R}^m$, el sistema $A\vec{x} = \vec{b}$ tiene solución (no necesariamente única) para todo $\vec{b} \in \mathbb{R}^m$.

ii) Si $R(A) \neq \mathbb{R}^m$, entonces el sistema $A\vec{x} = \vec{b}$ tiene solución si y solo si

$$< \vec{b}, \vec{y} >= 0, \quad \forall\, \vec{y} \text{ tal que } A^t \vec{y} = \vec{0}.$$

Ejemplo 5.30. Para la matriz

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

es fácil ver que

$$N(A^t) = \text{span}\left\{(-3,-3,1,0)^t, (-3,-5,0,1)^t\right\} = \text{span}\{\vec{u}_1, \vec{u}_2\}.$$

Note que $R(A) \neq \mathbb{R}^4$. Para decidir si el sistema $A\vec{x} = \vec{b}$ tiene solución o no, donde $\vec{b} = (4,-1,2,0)^t$, calculamos los productos interiores $< \vec{u}_i, \vec{b} >$ para $i = 1,2$:

$$< \vec{u}_1, \vec{b} >= (-3)(4) + (-3)(-1) + (1)(2) + (0)(0) = -7.$$

No es necesario calcular el otro producto interior, ya que por la parte (ii) del Teorema 5.29, podemos concluir que el sistema no tiene solución. $\square$

En el caso en que A es $n \times n$, podemos refinar el teorema de la alternativa de Fredholm aún mas. Este resultado se conoce como el *segundo Teorema de la Alternativa de Fredholm*.

Teorema 5.31. *Sea A una matriz $n \times n$. Entonces:*

i) Si $N(A) = \left\{\vec{0}\right\}$, el sistema $A\vec{x} = \vec{b}$ tiene solución que es única para todo $\vec{b} \in \mathbb{R}^n$.

ii) Si $\dim N(A) = m \geq 1$, entonces $\dim N(A^t) = m$ y el sistema $A\vec{x} = \vec{b}$ tiene solución si y solo si

$$< \vec{b}, \vec{v}_i >= 0, \quad 1 \leq i \leq m,$$

donde $N(A^t) = \operatorname{span}\{\vec{v}_1, \ldots, \vec{v}_m\}$. *En este caso y si* $\vec{x}_p$ *es una solución particular de* $A\vec{x} = \vec{b}$, *entonces la solución general del sistema es:*

$$\vec{x} = \vec{x}_p + \sum_{i=1}^{m} c_i \vec{u}_i, \quad \forall\, c_i \in \mathbb{R}, \quad 1 \leq i \leq m,$$

donde $N(A) = \operatorname{span}\{\vec{u}_1, \ldots, \vec{u}_m\}$.

Demostración:

i) Si $N(A) = \{\vec{0}\}$, entonces la matriz A es nosingular y el sistema $A\vec{x} = \vec{b}$ tiene solución que es única para todo $\vec{b} \in \mathbb{R}^n$.

ii) Si $\dim N(A) = m \geq 1$, entonces como $R(A^t)^\perp = N(A)$ y $\dim R(A^t) + \dim R(A^t)^\perp = n$, tenemos que $\dim R(A^t) = n - m$. Como $\dim N(A^t) + \dim R(A^t) = n$, podemos concluir ahora que $\dim N(A^t) = m$.

Sea $\{\vec{v}_1, \ldots, \vec{v}_m\}$ una base de $N(A^t)$. Entonces como $R(A) = N(A^t)^\perp$, tenemos que el sistema $A\vec{x} = \vec{b}$ tiene solución si y solo si $\vec{b} \in N(A^t)^\perp$, es decir, si y solo si

$$< \vec{b}, \vec{v}_i >= 0, \quad 1 \leq i \leq m.$$

Suponga que $A\vec{x}_p = \vec{b}$ y que $\vec{x}$ es cualquier otra solución, esto es, $A\vec{x} = \vec{b}$. Entonces $A(\vec{x} - \vec{x}_p) = \vec{0}$, es decir que $\vec{x} - \vec{x}_p \in N(A)$. Si $N(A) = \operatorname{span}\{\vec{u}_1, \ldots, \vec{u}_m\}$, entonces podemos concluir que existen escalares c_1, $\ldots$, c_m tal que

$$\vec{x} - \vec{x}_p = c_1\vec{u}_1 + \cdots + c_m\vec{u}_m,$$

es decir, que

$$\vec{x} = \vec{x}_p + c_1\vec{u}_1 + \cdots + c_m\vec{u}_m.$$

□

5.4 Conjuntos Ortonormales

Al hacer cómputos numéricos en espacios vectoriales, usualmente ayuda que los vectores de la base que se este usando, sean perpendiculares entre si. También muchos argumentos teóricos se simplifican si usamos bases con esta propiedad. Esto motiva la siguiente discusión.

Definición 5.32. Un conjunto de vectores $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ (ninguno de ellos igual al vector cero) en $\mathbb{R}^n$ se llama un *conjunto ortogonal* si

$$< \vec{v}_i, \vec{v}_j >= 0, \quad i \neq j. \tag{5.5}$$

Si en adición

$$< \vec{v}_i, \vec{v}_i >= 1, \quad \forall\, i, \tag{5.6}$$

entonces el conjunto se llama *ortonormal*. Una base de $\mathbb{R}^n$ que cumple con (5.5) (y (5.6)) se llama una *base ortogonal (ortonormal)*.

Ejemplo 5.33. La base estándar de $\mathbb{R}^3$ dada por $\vec{e}_1, \vec{e}_2, \vec{e}_3$ es una base ortonormal de $\mathbb{R}^3$. El conjunto:

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

es una base ortogonal de $\mathbb{R}^3$. Si dividimos cada vector por su norma, obtenemos entonces una base ortonormal:

$$\left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

□

Rápidamente tenemos la siguiente propiedad de los conjuntos ortogonales.

Proposición 5.34. *Si $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ son ortogonales en $\mathbb{R}^n$, entonces son linealmente independientes.*

Demostración: Suponga que

$$c_1 \vec{v}_1 + \cdots + c_k \vec{v}_k = \vec{0}.$$

Si multiplicamos interiormente esta ecuación con el vector $\vec{v}_j$ para cualquier j entre 1 y k, obtenemos que

$$< c_1 \vec{v}_1 + \cdots + c_k \vec{v}_k, \vec{v}_j >=< \vec{0}, \vec{v}_j >= 0.$$

Usando la linealidad de $< \cdot, \cdot >$ en su primer argumento y que $< \vec{v}_i, \vec{v}_j >= 0$ si $i \neq j$, ésta ecuación reduce a $c_j < \vec{v}_j, \vec{v}_j >= 0$. Como $\vec{v}_j \neq \vec{0}$, tenemos que $c_j = 0$. Como el índice j es arbitrario, tenemos que todos los c_i's son cero, es decir, que $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ son linealmente independientes. □

Sabemos que cuando $\{\vec{u}_1, \ldots, \vec{u}_n\}$ es una base de $\mathbb{R}^n$, entonces cualquier vector se puede escribir como una combinación lineal de los $\{\vec{u}_i\}$. En general, el computo de los coeficientes en dicha expansión envuelve resolver un sistema de ecuaciones $n \times n$. En el caso en que la base es ortonormal, la matriz de coeficientes de este sistema para los coeficientes en la expansión, es diagonal.

Proposición 5.35. *Sea $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ una base ortonormal de $\mathbb{R}^n$. Entonces para cualquier vector $\vec{x}$ en $\mathbb{R}^n$, tenemos que:*

$$\vec{x} = \sum_{i=1}^{n} < \vec{x}, \vec{v}_i > \vec{v}_i,$$

y que

$$\left\|\vec{x}\right\|^2 = \sum_{i=1}^{n} < \vec{x}, \vec{v}_i >^2 .$$

Nota 5.36. Poniendo $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$, el resultado de la proposición anterior se puede escribir como:

$$(\vec{x})_S = (< \vec{x}, \vec{v}_i >) = B^t \vec{x},$$

donde $B = [\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n]$. □

Demostración: Como $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ una base de $\mathbb{R}^n$, para cualquier $\vec{x} \in \mathbb{R}^n$ existen $c_1, \ldots, c_n$ únicos tal que

$$\vec{x} = \sum_{i=1}^{n} c_i \vec{v}_i.$$

Multiplicando interiormente esta ecuación con cualquier vector $\vec{v}_j$ de la base, tenemos que

$$< \vec{x}, \vec{v}_j >= \left\langle \sum_{i=1}^{n} c_i \vec{v}_i, \vec{v}_j \right\rangle = \sum_{i=1}^{n} c_i < \vec{v}_i, \vec{v}_j >= c_j < \vec{v}_j, \vec{v}_j >= c_j,$$

donde usamos la linealidad de $<\cdot,\cdot>$ en su primer argumento. Tenemos que la expansión de $\vec{x}$ queda ahora como:

$$\vec{x} = \sum_{i=1}^{n} <\vec{x},\vec{v}_i> \vec{v}_i.$$

Usando esta expresión para $\vec{x}$, la linealidad de $<\cdot,\cdot>$ con respecto a cualquiera de sus dos argumentos, y la ortogonalidad de los $\vec{v}_i$'s, tenemos que

$$\begin{aligned}
\left\|\vec{x}\right\|^2 &= \left\langle \sum_{i=1}^{n} <\vec{x},\vec{v}_i> \vec{v}_i, \sum_{j=1}^{n} <\vec{x},\vec{v}_j> \vec{v}_j \right\rangle \\
&= \sum_{i}\sum_{j} <\vec{x},\vec{v}_i><\vec{x},\vec{v}_j><\vec{v}_i,\vec{v}_j> \\
&= \sum_{i=1}^{n} <\vec{x},\vec{v}_i>^2<\vec{v}_i,\vec{v}_i>= \sum_{i=1}^{n} <\vec{x},\vec{v}_i>^2 .
\end{aligned}$$

□

Ejemplo 5.37. Del Ejemplo 5.33 tenemos que

$$S = \{\vec{v}_1,\vec{v}_2,\vec{v}_3\} = \left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

es una base ortonormal para $\mathbb{R}^3$. Dado cualquier vector $\vec{x} = (x_1,x_2,x_3)^t \in \mathbb{R}^3$, tenemos que

$$\begin{aligned}
\vec{x} &= <\vec{x},\vec{v}_1> \vec{v}_1 + <\vec{x},\vec{v}_2> \vec{v}_2 + <\vec{x},\vec{v}_3> \vec{v}_3, \\
&= \left[\frac{x_1+x_2}{\sqrt{2}}\right]\vec{v}_1 + \left[\frac{x_1-x_2}{\sqrt{2}}\right]\vec{v}_2 + x_3\vec{v}_3.
\end{aligned}$$

Tenemos entonces que:

$$(\vec{x})_S = \left[\frac{x_1+x_2}{\sqrt{2}}, \frac{x_1-x_2}{\sqrt{2}}, x_3\right]^t,$$

y que

$$\left\|\vec{x}\right\|^2 = \left[\frac{x_1 + x_2}{\sqrt{2}}\right]^2 + \left[\frac{x_1 - x_2}{\sqrt{2}}\right]^2 + x_3^2.$$

□

Modificando un tanto la demostración para la expansión en el cálculo de $\left\|\vec{x}\right\|^2$ en la proposición anterior, se puede verificar el siguiente resultado.

Proposición 5.38. *Sea $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ una base ortonormal de $\mathbb{R}^n$, y $\vec{x}, \vec{y} \in \mathbb{R}^n$ con*

$$\vec{x} = \sum_{i=1}^{n} b_i \vec{v}_i, \quad \vec{y} = \sum_{i=1}^{n} c_i \vec{v}_i.$$

Entonces

$$< \vec{x}, \vec{y} >= \sum_{i=1}^{n} b_i c_i = \sum_{i=1}^{n} < \vec{x}, \vec{v}_i >< \vec{y}, \vec{v}_i > .$$

Ejemplo 5.39. Suponga que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es una base ortonormal de $\mathbb{R}^3$ y que:

$$\vec{x} = 3\vec{v}_1 - 2\vec{v}_2 + 4\vec{v}_3, \quad \vec{y} = -2\vec{v}_1 + \vec{v}_2 - 5\vec{v}_3.$$

Usando los resultados de las Proposiciones 5.35 y 5.38, tenemos que:

$$\begin{aligned} < \vec{x}, \vec{y} > &= (3)(-2) + (-2)(1) + (4)(-5) = -28, \\ \left\|\vec{x}\right\| &= \sqrt{3^2 + (-2)^2 + 4^2} = \sqrt{29}, \\ \left\|\vec{y}\right\| &= \sqrt{(-2)^2 + 1^2 + (-5)^2} = \sqrt{30}. \end{aligned}$$

De aquí que el ángulo entre los vectores es:

$$\theta = \cos^{-1}\left[\frac{< \vec{x}, \vec{y} >}{\left\|\vec{x}\right\| \left\|\vec{y}\right\|}\right] = \cos^{-1}\left[\frac{-28}{\sqrt{29}\sqrt{30}}\right] \approx 2.8218 \text{ radianes.}$$

□

5.4.1 Matrices ortogonales

La siguiente familia de matrices es de suma importancia en el estudio de métodos numéricos para sistemas lineales y otras aplicaciones. En particular veremos que son muy útiles para la solución de problemas de cuadrados mínimos.

Definición 5.40. Una matriz Q, $m \times n$ se llama ortogonal si las columnas de Q forman un conjunto ortonormal en $\mathbb{R}^m$, esto es

$$Q^t Q = I.$$

Ejemplo 5.41. Para cualquier número real θ, la matriz

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

es ortogonal.

Cualquier matriz que se obtenga reordenando o permutando las columnas de la matriz identidad es ortogonal. Estas matrices se llaman *matrices de permutación*. □

Las siguientes propiedades de las matrices ortogonales son todas consecuencia directa de la ecuación $Q^t Q = I$. En particular, si interpretamos a la matriz Q como una transformación lineal, el resultado nos dice que Q no cambia los largos de los vectores ni los ángulos entre vectores. A este tipo de transformación lineal se le conoce como una *rotación rígida*.

Proposición 5.42. *Sea Q una matriz ortogonal $n \times n$. Entonces*

i) $Q^{-1} = Q^t$.

ii) $< Q\vec{x}, Q\vec{y} > = < \vec{x}, \vec{y} >$.

iii) $\left\| Q\vec{x} \right\| = \left\| \vec{x} \right\|$.

iv) $\det Q = \pm 1$.

5.4.2 El método de Gram–Schmidt

Si un conjunto de vectores linealmente independientes no es ortonormal, existe un método para construir un conjunto ortonormal a partir de estos vectores, y generando el mismo subespacio que el conjunto original de vectores. Este proceso se conoce como el *método de Gram–Schmidt*. Antes de describir este procedimiento en general, vamos a ilustrarlo en el caso de $\mathbb{R}^3$.

Sea $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ una base (no necesariamente ortogonal) de $\mathbb{R}^3$. Definimos primeramente a:

$$\vec{v}_1 = \frac{\vec{u}_1}{\|\vec{u}_1\|}.$$

Note que $\operatorname{span}\{\vec{u}_1\} = \operatorname{span}\{\vec{v}_1\}$. Defina ahora el vector $\vec{p}_1$ como la proyección vectorial de $\vec{u}_2$ en $\vec{v}_1$, esto es:

$$\vec{p}_1 = <\vec{u}_2, \vec{v}_1> \vec{v}_1.$$

Es fácil ver que $\vec{u}_2 - \vec{p}_1$ es perpendicular a $\vec{v}_1$, y es distinto del vector cero (de lo contrario $\vec{u}_1, \vec{u}_2$ serían linealmente dependientes lo cual es imposible). Definimos ahora:

$$\vec{v}_2 = \frac{\vec{u}_2 - \vec{p}_1}{\|\vec{u}_2 - \vec{p}_1\|}.$$

Tenemos pues que $\{\vec{v}_1, \vec{v}_2\}$ son ortonormales y como $\vec{v}_1, \vec{v}_2 \in \operatorname{span}\{\vec{u}_1, \vec{u}_2\}$, tenemos que $\operatorname{span}\{\vec{v}_1, \vec{v}_2\} = \operatorname{span}\{\vec{u}_1, \vec{u}_2\}$. Finalmente definimos a $\vec{p}_2$ como la proyección vectorial de $\vec{u}_3$ sobre $\operatorname{span}\{\vec{v}_1, \vec{v}_2\}$, esto es

$$\vec{p}_2 = <\vec{u}_3, \vec{v}_1> \vec{v}_1 + <\vec{u}_3, \vec{v}_2> \vec{v}_2.$$

Es fácil ver también que $\vec{u}_3 - \vec{p}_2 \in \operatorname{span}\{\vec{v}_1, \vec{v}_2\}^\perp$ y que no es igual al vector cero. Definiendo

$$\vec{v}_3 = \frac{\vec{u}_3 - \vec{p}_2}{\|\vec{u}_3 - \vec{p}_2\|},$$

tenemos que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es ortonormal y $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \operatorname{span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$.

Ejemplo 5.43. Considere la base $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ de $\mathbb{R}^3$ donde

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

Construimos una base ortonormal $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ de $\mathbb{R}^3$ a partir de estos vectores. De acuerdo a la discusión anterior, tomamos

$$\vec{v}_1 = \frac{\vec{u}_1}{\left\|\vec{u}_1\right\|} = \left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right]^t.$$

La proyección vectorial de $\vec{u}_2$ sobre $\vec{v}_1$ es:

$$\vec{p}_1 =< \vec{u}_2, \vec{v}_1 > \vec{v}_1 = \left[\frac{1}{2}, \frac{1}{2}, 0\right]^t.$$

Ahora

$$\vec{v}_2 = \frac{\vec{u}_2 - \vec{p}_1}{\left\|\vec{u}_2 - \vec{p}_1\right\|} = \left[\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right]^t.$$

Para la última etapa, tomamos $\vec{p}_2$ como la proyección vectorial de $\vec{u}_3$ sobre span$\{\vec{v}_1, \vec{v}_2\}$, esto es

$$\vec{p}_2 =< \vec{u}_3, \vec{v}_1 > \vec{v}_1 + < \vec{u}_3, \vec{v}_2 > \vec{v}_2 = \left[\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right]^t.$$

Ahora

$$\vec{v}_3 = \frac{\vec{u}_3 - \vec{p}_2}{\left\|\vec{u}_3 - \vec{p}_2\right\|} = \left[-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right]^t,$$

Tenemos finalmente que

$$\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \right\},$$

es la base ortonormal deseada. □

El siguiente resultado recoge éstas ideas en el caso general de k vectores en $\mathbb{R}^n$. El método o algoritmo que se desprende de este teorema se conoce como el *método de Gram–Schmidt*.

Teorema 5.44. *Sea $\{\vec{u}_1, \ldots, \vec{u}_k\}$ vectores linealmente independientes en $\mathbb{R}^n$ donde $k \leq n$. Defina*

$$\vec{v}_1 = \frac{\vec{u}_1}{\left\|\vec{u}_1\right\|},$$

y para $j = 1,\dots,k-1$:

$$\begin{aligned}\vec{p}_j &= \sum_{i=1}^{j} < \vec{u}_{j+1}, \vec{v}_i > \vec{v}_i,\\ \vec{v}_{j+1} &= \frac{\vec{u}_{j+1} - \vec{p}_j}{\left\|\vec{u}_{j+1} - \vec{p}_j\right\|}.\end{aligned}$$

Entonces $\{\vec{v}_1,\dots,\vec{v}_k\}$ es un conjunto ortonormal y

$$\operatorname{span}\{\vec{v}_1,\dots,\vec{v}_k\} = \operatorname{span}\{\vec{u}_1,\dots,\vec{u}_k\}.$$

Demostración: La demostración es por inducción matemática en la k. Claramente el resultado es cierto si $k = 1$. Suponemos entonces que para $k \geq 1$ el resultado es cierto y tomamos un conjunto de vectores linealmente independientes $\{\vec{u}_1,\dots,\vec{u}_{k+1}\}$. Por la hipótesis de inducción, existen vectores $\{\vec{v}_1,\dots,\vec{v}_k\}$ ortonormales con

$$\operatorname{span}\{\vec{v}_1,\dots,\vec{v}_k\} = \operatorname{span}\{\vec{u}_1,\dots,\vec{u}_k\}.$$

Defina $\vec{p}_k$ por:

$$\vec{p}_k = \sum_{i=1}^{k} < \vec{u}_{k+1}, \vec{v}_i > \vec{v}_i.$$

Note que para cualquier j con $1 \leq j \leq k$ tenemos:

$$\begin{aligned}< \vec{u}_{k+1} - \vec{p}_k, \vec{v}_j > &= \left\langle \vec{u}_{k+1} - \sum_{i=1}^{k} < \vec{u}_{k+1}, \vec{v}_i > \vec{v}_i, \vec{v}_j \right\rangle,\\ &= < \vec{u}_{k+1}, \vec{v}_j > - \sum_{i=1}^{k} < \vec{u}_{k+1}, \vec{v}_i >< \vec{v}_i, \vec{v}_j >,\\ &= < \vec{u}_{k+1}, \vec{v}_j > - < \vec{u}_{k+1}, \vec{v}_j >= 0,\end{aligned}$$

donde usamos que $\{\vec{v}_1,\dots,\vec{v}_k\}$ es ortonormal. Es decir que $\vec{u}_{k+1} - \vec{p}_k \in \operatorname{span}\{\vec{v}_1,\dots,\vec{v}_k\}^{\perp}$. También $\vec{u}_{k+1} - \vec{p}_k \neq \vec{0}$ ya que si $\vec{u}_{k+1} - \vec{p}_k = \vec{0}$, entonces

$$\vec{u}_{k+1} = \vec{p}_k \in \operatorname{span}\{\vec{v}_1,\dots,\vec{v}_k\} = \operatorname{span}\{\vec{u}_1,\dots,\vec{u}_k\}.$$

Pero esto contradice que $\{\vec{u}_1,\ldots,\vec{u}_{k+1}\}$ son linealmente independientes. Así que $\vec{u}_{k+1}-\vec{p}_k \neq \vec{0}$ y podemos ahora definir

$$\vec{v}_{k+1} = \frac{\vec{u}_{k+1}-\vec{p}_k}{\left\|\vec{u}_{k+1}-\vec{p}_k\right\|}.$$

Como $\vec{u}_{k+1}-\vec{p}_k \in \operatorname{span}\{\vec{v}_1,\ldots,\vec{v}_k\}^{\perp}$ tenemos que $\operatorname{span}\{\vec{v}_1,\ldots,\vec{v}_{k+1}\}$ es ortonormal. Por construcción

$$\vec{v}_{k+1} \in \operatorname{span}\{\vec{v}_1,\ldots,\vec{v}_k,\vec{u}_{k+1}\} = \operatorname{span}\{\vec{u}_1,\ldots,\vec{u}_{k+1}\},$$

donde usamos que $\operatorname{span}\{\vec{v}_1,\ldots,\vec{v}_k\} = \operatorname{span}\{\vec{u}_1,\ldots,\vec{u}_k\}$. Tenemos entonces que

$$\operatorname{span}\{\vec{v}_1,\ldots,\vec{v}_{k+1}\} \subset \operatorname{span}\{\vec{u}_1,\ldots,\vec{u}_{k+1}\}.$$

Como claramente, $\vec{u}_{k+1} \in \operatorname{span}\{\vec{v}_1,\ldots,\vec{v}_{k+1}\}$, tenemos usando nuevamente que $\operatorname{span}\{\vec{v}_1,\ldots,\vec{v}_k\} = \operatorname{span}\{\vec{u}_1,\ldots,\vec{u}_k\}$, que

$$\operatorname{span}\{\vec{u}_1,\ldots,\vec{u}_{k+1}\} \subset \operatorname{span}\{\vec{v}_1,\ldots,\vec{v}_{k+1}\},$$

por lo que

$$\operatorname{span}\{\vec{v}_1,\ldots,\vec{v}_{k+1}\} = \operatorname{span}\{\vec{u}_1,\ldots,\vec{u}_{k+1}\}.$$

Esto completa la prueba por inducción. $\square$

Factorización QR

Sea $\{\vec{u}_1,\ldots,\vec{u}_k\}$ un conjunto de vectores linealmente independientes y $\{\vec{v}_1,\ldots,\vec{v}_k\}$ el conjunto ortonormal que se obtiene a partir de los $\vec{u}_i$'s por el método de Gram–Schmidt del Teorema 5.44. Entonces

$$\vec{u}_1 = \alpha_{11}\vec{v}_1, \tag{5.7a}$$

$$\vec{u}_j = \sum_{i=1}^{j} \alpha_{ij}\vec{v}_i, \quad 2 \leq j \leq k, \tag{5.7b}$$

donde

$$\alpha_{11} = \left\|\vec{u}_1\right\|, \tag{5.8a}$$

$$\alpha_{ij} \quad = \quad \begin{cases} <\vec{u}_j,\vec{v}_i> & , \quad 1 \le i < j, \\ \left\|\vec{u}_j - \vec{p}_{j-1}\right\| & , \quad i = j, \end{cases} \qquad 2 \le j \le k. \tag{5.8b}$$

Defina las matrices

$$A = [\vec{u}_1,\ldots,\vec{u}_k], \quad Q = [\vec{v}_1,\ldots,\vec{v}_k], \quad R = [\alpha_{ij}],$$

donde $\alpha_{ij} = 0$ para $i > j$. Entonces Q es ortogonal $n \times k$, R es triangular superior $k \times k$, y las ecuaciones (5.7), (5.8) se pueden escribir como $A = QR$. Además, de la demostración del Teorema 5.44 sabemos que los $\alpha_{ii} \neq 0$, por lo que R es nosingular. Esto completa la demostración del siguiente teorema.

Teorema 5.45 (Factorización QR reducida de una matriz)**.** *Sea A una matriz $n \times k$ de rango k. Entonces existe una matriz ortogonal Q de tamaño $n \times k$ y una matriz R de tamaño $k \times k$, nosingular, y triangular superior, tal que $A = QR$.*

Ejemplo 5.46. Considere la matriz

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

Note que las columnas de A son linealmente independientes. Sean $\{\vec{u}_1,\vec{u}_2\}$ las columnas de A. Usando el método de Gram–Schmidt, tenemos que:

$$\begin{aligned} \left\|\vec{u}_1\right\| &= \sqrt{6}, \quad \vec{v}_1 = \frac{\vec{u}_1}{\left\|\vec{u}_1\right\|} = \left[\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right]^t, \\ \vec{p}_1 &= <\vec{u}_2,\vec{v}_1> \vec{v}_1 = \left[\frac{2}{3}, \frac{4}{3}, -\frac{2}{3}\right]^t, \\ \vec{u}_2 - \vec{p}_1 &= \left[\frac{7}{3}, -\frac{1}{3}, \frac{5}{3}\right]^t, \quad \left\|\vec{u}_2 - \vec{p}_1\right\| = \frac{5\sqrt{3}}{3}, \\ \vec{v}_2 &= \frac{\vec{u}_2 - \vec{p}_1}{\left\|\vec{u}_2 - \vec{p}_1\right\|} = \left[\frac{7\sqrt{3}}{15}, -\frac{\sqrt{3}}{15}, \frac{5\sqrt{3}}{15}\right]^t. \end{aligned}$$

Note que

$$\vec{u}_1 = \sqrt{6}\,\vec{v}_1, \quad \vec{u}_2 = \frac{4}{\sqrt{6}}\,\vec{v}_1 + \frac{5\sqrt{3}}{3}\,\vec{v}_2.$$

Ahora con

$$Q = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{7\sqrt{3}}{15} \\ \frac{2}{\sqrt{6}} & -\frac{\sqrt{3}}{15} \\ -\frac{1}{\sqrt{6}} & \frac{5\sqrt{3}}{15} \end{bmatrix}, \quad R = \begin{bmatrix} \sqrt{6} & \frac{4}{\sqrt{6}} \\ 0 & \frac{5\sqrt{3}}{3} \end{bmatrix},$$

tenemos que $A = QR$ y que $Q^t Q = I$. □

5.5 Problema de Cuadrados Mínimos

Considere el sistema $A\vec{x} = \vec{b}$ donde A es $m \times n$ con $m > n$. Note que en general estos sistemas son inconsistentes. Dado un vector $\vec{x}$, definimos el *vector residual* por:

$$\vec{r}(\vec{x}) = \vec{b} - A\vec{x}.$$

La solución de *cuadrados mínimos* del sistema $A\vec{x} = \vec{b}$ se define como el vector $\vec{x}_{LS}$ tal que:

$$\left\|\vec{r}(\vec{x}_{LS})\right\| = \min_{\vec{x}\in\mathbb{R}^n} \left\|\vec{r}(\vec{x})\right\| = \min_{\vec{x}\in\mathbb{R}^n} \left\|\vec{b} - A\vec{x}\right\|.$$

Este problema es equivalente a buscar $\vec{p}_{LS} \in R(A)$ tal que:

$$\left\|\vec{b} - \vec{p}_{LS}\right\| = \min_{\vec{y}\in R(A)} \left\|\vec{b} - \vec{y}\right\|. \tag{5.9}$$

Por la Proposición 5.27, sabemos que $\mathbb{R}^m = R(A) \oplus R(A)^\perp$, por lo que podemos escribir a $\vec{b}$ como:

$$\vec{b} = \vec{p} + \vec{q}, \quad \vec{p} \in R(A), \quad \vec{q} \in R(A)^\perp.$$

Veamos que $\vec{p}_{LS} = \vec{p}$. Si $\vec{y} \in R(A)$, entonces:

$$\begin{aligned} \left\|\vec{b} - \vec{y}\right\|^2 &= \left\|\vec{b} - \vec{p} + \vec{p} - \vec{y}\right\|^2, \\ &= \left\|\vec{b} - \vec{p}\right\|^2 + \left\|\vec{p} - \vec{y}\right\|^2, \quad \text{ya que } \vec{b} - \vec{p} \in R(A)^\perp \text{ y } \vec{p} - \vec{y} \in R(A), \\ &\geq \left\|\vec{b} - \vec{p}\right\|^2, \end{aligned}$$

con igualdad si y solo si $\vec{y} = \vec{p}$. De modo que el mínimo en (5.9) ocurre cuando $\vec{p}_{LS} = \vec{p}$ y como $\vec{p}_{LS} \in R(A)$, entonces $\vec{p}_{LS} = A\vec{x}_{LS}$. Note también que

$$\vec{r}(\vec{x}_{LS}) = \vec{b} - A\vec{x}_{LS} = \vec{b} - \vec{p}_{LS} \in R(A)^{\perp}.$$

Esto completa la demostración del siguiente teorema.

Teorema 5.47. *Sea A una matriz $m \times n$, $m > n$, y $\vec{b} \in \mathbb{R}^m$. Entonces existe un vector $\vec{x}_{LS} \in \mathbb{R}^n$ tal que*

$$\left\|\vec{r}(\vec{x}_{LS})\right\| = \left\|\vec{b} - A\vec{x}_{LS}\right\| = \min_{\vec{x} \in \mathbb{R}^n} \left\|\vec{b} - A\vec{x}\right\|,$$

donde $\vec{r}(\vec{x}_{LS}) \in R(A)^{\perp}$. Si rank $A = n$, entonces $\vec{x}_{LS}$ es único.

Como

$$\vec{r}(\vec{x}_{LS}) \in R(A)^{\perp} = N(A^t),$$

tenemos que $A^t\vec{r}(\vec{x}_{LS}) = \vec{0}$, es decir que:

$$A^t A\vec{x}_{LS} = A^t\vec{b}.$$

Este sistema lineal de ecuaciones se llaman las *ecuaciones normales* para determinar $\vec{x}_{LS}$.

Ejemplo 5.48. En un experimento de laboratorio con un resorte, se encontró que éste se alarga 4, 7, y 11 pulgadas al colocarle unas pesas de 3, 5, y 8 libras respectivamente. De acuerdo a la Ley de Hooke, la fuerza que ejerce el resorte es proporcional al desplazamiento o alargamiento del resorte. Si k es la constante de proporcionalidad, ésto lo podemos escribir como:

$$F = kx, \quad x \text{ desplazamiento.}$$

Usando los datos dados, tenemos el sistema de ecuaciones:

$$\begin{cases} 4k &= 3, \\ 7k &= 5, \\ 11k &= 8. \end{cases}$$

Note que el sistema no tiene solución. Buscamos la solución de cuadrados mínimos. En este caso

$$A = \begin{bmatrix} 4 \\ 7 \\ 11 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix}, \quad \vec{x}_{LS} = k_{LS}.$$

Tenemos entonces que:

$$A^tA = (4,7,11)\begin{bmatrix} 4 \\ 7 \\ 11 \end{bmatrix} = 186, \quad A^t\vec{b} = (4,7,11)\begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix} = 135,$$

por lo que $A^tA\vec{x}_{LS} = A^t\vec{b}$ reduce a $186k_{LS} = 135$, i.e.

$$k_{LS} = \frac{135}{186} \approx 0.726.$$

□

Ejemplo 5.49. Vamos a calcular la proyección vectorial del vector $\vec{b} = (2,3,2)^t$ sobre $R(A)$ donde

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

La proyección vectorial de $\vec{b}$ sobre $R(A)$ esta dada por el vector $\vec{p} = A\vec{x}_{LS}$ donde $\vec{x}_{LS}$ es la solución en el sentido de los cuadrados minimos del sistema $A\vec{x} = \vec{b}$. Como

$$A^tA = \begin{bmatrix} 6 & 9 \\ 9 & 21 \end{bmatrix}, \quad A^t\vec{b} = \begin{bmatrix} 5 \\ 16 \end{bmatrix},$$

tenemos que $\vec{x}_{LS} = (-13/15, 17/15)^t$. Tenemos entonces que la proyección buscada es:

$$\vec{p} = \begin{bmatrix} -1 & 1 \\ 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} -13/15 \\ 17/15 \end{bmatrix} = \begin{bmatrix} 2 \\ 7/5 \\ 14/5 \end{bmatrix}.$$

□

Ejemplo 5.50. Considere los datos:

x	0	3	6
y	1	4	5

los cuales provienen de algún experimento. Se sospecha que y como función de x se puede representar bien con una relación lineal (una recta). Buscamos la mejor recta en el sentido de los cuadrados mínimos para ajustar los datos. Si $y = mx + b$ es la recta que buscamos, entonces de los datos dados tenemos que:

$$\begin{aligned} m(0)+b &= 1, \\ m(3)+b &= 4, \\ m(6)+b &= 5, \end{aligned}$$

o lo mismo

$$\begin{bmatrix} 0 & 1 \\ 3 & 1 \\ 6 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix}.$$

Tenemos pues que:

$$A^tA = \begin{bmatrix} 0 & 3 & 6 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 3 & 1 \\ 6 & 1 \end{bmatrix} = \begin{bmatrix} 45 & 9 \\ 9 & 3 \end{bmatrix}, \quad A^t\vec{b} = \begin{bmatrix} 0 & 3 & 6 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 42 \\ 10 \end{bmatrix}.$$

Las ecuaciones normales son:

$$\begin{bmatrix} 45 & 9 \\ 9 & 3 \end{bmatrix} \begin{bmatrix} m_{LS} \\ b_{LS} \end{bmatrix} = \begin{bmatrix} 42 \\ 10 \end{bmatrix},$$

que tienen solución $m_{LS} = 2/3$, $b_{LS} = 4/3$. La mejor recta en el sentido de los cuadrados mínimos es

$$y = \frac{2}{3}x + \frac{4}{3}.$$

En la Figura 5.1 se muestra una gráfica de ésta recta junto con los datos originales. □

Ejemplo 5.51. Considere los datos:

$$\begin{aligned} x &= [-1.0000, -0.8000, -0.6000, -0.4000, -0.2000, 0, \\ &\quad\ 0.2000, 0.4000, 0.6000, 0.8000, 1.0000], \\ y &= [75.9950, 91.9720, 105.7110, 123.2030, 131.6690, 150.6970, \\ &\quad\ 179.3230, 203.2120, 226.5050, 249.6330, 281.4220]. \end{aligned}$$

Buscamos el polinomio de grados dos que mejor se ajuste a los datos en el sentido de cuadrados mínimos. Si $y = ax^2 + bx + c$ es la ecuación del dicho

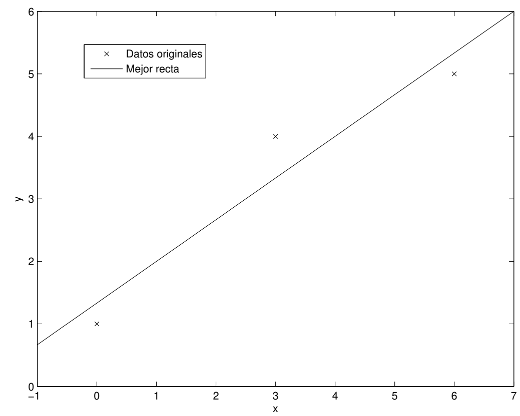


Figura 5.1: Gráfica de la mejor recta para los datos del Ejemplo 5.50.

polinomio, entonces sustituyendo los datos en la ecuación obtenemos el sistema lineal $A\vec{x} = \vec{b}$ donde:

$$A = \begin{bmatrix} 1.0000 & -1.0000 & 1.0000 \\ 0.6400 & -0.8000 & 1.0000 \\ 0.3600 & -0.6000 & 1.0000 \\ 0.1600 & -0.4000 & 1.0000 \\ 0.0400 & -0.2000 & 1.0000 \\ 0 & 0 & 1.0000 \\ 0.0400 & 0.2000 & 1.0000 \\ 0.1600 & 0.4000 & 1.0000 \\ 0.3600 & 0.6000 & 1.0000 \\ 0.6400 & 0.8000 & 1.0000 \\ 1.0000 & 1.0000 & 1.0000 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 75.9950 \\ 91.9720 \\ 105.7110 \\ 123.2030 \\ 131.6690 \\ 150.6970 \\ 179.3230 \\ 203.2120 \\ 226.5050 \\ 249.6330 \\ 281.4220 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Usamos MATLAB ™ para resolver las ecuaciones normales:

```
x=[-1.0000,-0.8000,-0.6000,-0.4000,-0.2000,0,...
    0.2000,0.4000,0.6000,0.8000,1.0000];
y=[75.9950,91.9720,105.7110,123.2030,131.6690,150.6970,...
    179.3230,203.2120,226.5050,249.6330,281.4220];
A=[x'.^2,x',ones(11,1)];
b=y';
x_LS=(A'*A)\(A'*b);
```

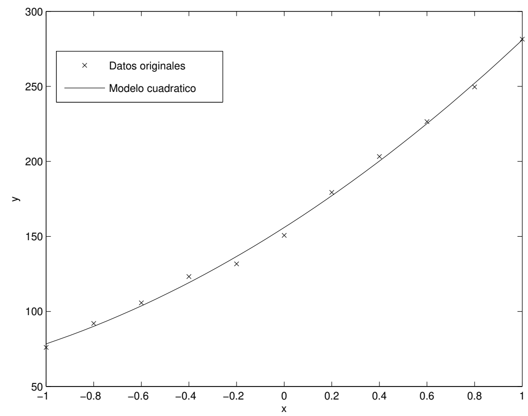


Figura 5.2: Gráfica para el mejor polinomio de grados dos de cuadrados mínimos para los datos del Ejemplo 5.51.

Estas instrucciones producen los resultados:

$$a = 23.7261, \quad b = 101.2651, \quad c = 155.9043.$$

El mejor polinomio cuadrático es (aproximadamente):

$$y = 23.7261x^2 + 101.2651x + 155.9043.$$

En la Figura 5.2 mostramos una gráfica del mejor polinomio cuadrático para los datos junto con los datos originales.

El mismo resultado lo podemos obtener en MATLAB de forma más directa si usamos la función `polyfit`:

```
x=[-1.0000,-0.8000,-0.6000,-0.4000,-0.2000,0,...
    0.2000,0.4000,0.6000,0.8000,1.0000];
y=[75.9950,91.9720,105.7110,123.2030,131.6690,150.6970,...
    179.3230,203.2120,226.5050,249.6330,281.4220];
x_LS=polyfit(x,y,2);
```

□

Ejemplo 5.52. Consideramos el problema de buscar una función de la forma:

$$y = a_1 + a_2 \sin 2x + a_3 \cos 2x,$$

que mejor se ajuste a los datos:

x	0	1	2	-1	-2	3
y	-4	-3	-3	-5	-3	-5

Sustituyendo los datos en la ecuación del modelo obtenemos el sistema lineal $A\vec{x} = \vec{b}$ donde:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & \sin 2 & \cos 2 \\ 1 & \sin 4 & \cos 4 \\ 1 & -\sin 2 & \cos 2 \\ 1 & -\sin 4 & \cos 4 \\ 1 & \sin 6 & \cos 6 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} -4 \\ -3 \\ -3 \\ -5 \\ -3 \\ -5 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

Usamos MATLAB para resolver las ecuaciones normales:

```
x=[0;1;2;-1;-2;3];
y=[-4;-3;-3;-5;-3;-5];
A=[ones(size(x)),sin(2*x),cos(2*x)];
x_LS=(A'*A)\(A'*y);
```

Esto produce los resultados

$$a_1 = -3.8211, \quad a_2 = 0.6853, \quad a_3 = -0.6569.$$

La mejor función para los datos, en el sentido de los cuadrados mínimos, de la forma $y = a_1 + a_2 \sin 2x + a_3 \cos 2x$ está dada entonces por:

$$y = -3.8211 + 0.6853 \sin 2x - 0.6569 \cos 2x.$$

Note que en este problema no se puede utilizar la función `polyfit` de MATLAB ya que el ajuste no es con un polinomio (es con una combinación lineal de funciones trigonométricas), por lo que es necesario montar y resolver las ecuaciones normales directamente. En la Figura 5.3 ilustramos el mejor modelo trigonométrico en el sentido de los cuadrados mínimos junto a los datos originales.

□

Las ecuaciones normales tienen la propiedad que los analistas numéricos llaman "mal acondicionamiento". Esto lo que quiere decir es que al resolver las ecuaciones normales en una computadora digital, con el método de eliminación gaussiana por ejemplo, los errores causados por

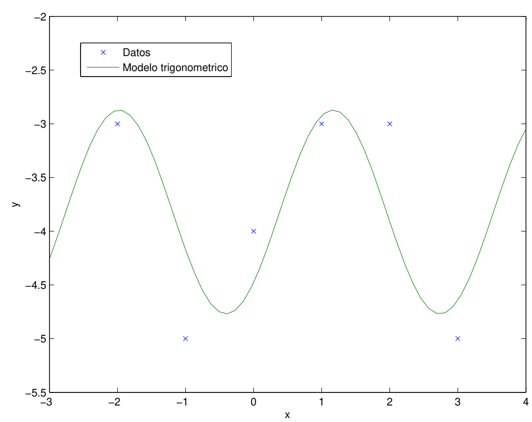


Figura 5.3: Gráfica del mejor modelo trigonométrico y los datos originales del Ejemplo 5.52.

la aritmética finita del computador se pueden propagar sustancialmente llevando a resultados erróneos o con mucho error. Una manera de lidear con esta dificultad es usando la factorización QR del Teorema 5.45. En particular, para el sistema $A\vec{x} = \vec{b}$ donde A es $m \times n$ con $m > n$, si A tiene rango n, entonces por el Teorema 5.45 tenemos que $A = QR$ donde Q es ortogonal y R es triangular superior y nosingular. Usando ésto tenemos que:

$$A^tA = R^tQ^tQR = R^tR, \quad A^t\vec{b} = R^tQ^t\vec{b},$$

por lo que las ecuaciones normales $A^tA\vec{x}_{LS} = A^t\vec{b}$ son equivalentes a:

$$R^tR\vec{x}_{LS} = R^tQ^t\vec{b}.$$

Como R es nosingular, podemos cancelar R^t en ambos lados para obtener que:

$$R\vec{x}_{LS} = Q^t\vec{b}.$$

Note que la matriz de coeficientes de este último sistema es triangular superior.

Ejemplo 5.53. Buscamos la solución de cuadrados mínimos del sistema:

$$\begin{bmatrix} 1 & -2 & -1 \\ 2 & 0 & 1 \\ 2 & -4 & 2 \\ 4 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -2 \end{bmatrix}.$$

Vamos a utilizar la factorización QR de la matriz de coeficientes. Hacemos el calculo con la función qr de MATLAB que calcula factorizaciones QR:

```
A=[1,-2,-1;2,0,1;2,-4,2;4,1,0];
b=[-1,;1;1;-2];
[Q,R]=qr(A,0);
x_LS=R\(Q'*b);
```

La factorización QR calculada por MATLAB es:

$$Q = \begin{bmatrix} -0.2000 & 0.3979 & 0.8347 \\ -0.4000 & -0.1085 & -0.3672 \\ -0.4000 & 0.7959 & -0.3766 \\ -0.8000 & -0.4432 & 0.1632 \end{bmatrix},$$

$$R = \begin{bmatrix} -5.0000 & 1.2000 & -1.0000 \\ 0 & -4.4227 & 1.0853 \\ 0 & 0 & -1.9550 \end{bmatrix},$$

y la solución de cuadrados mínimos:

$$\vec{x}_{LS} = \begin{bmatrix} -0.4013 & -0.0268 & 0.9743 \end{bmatrix}^t.$$

Debemos mencionar que estos cálculos también se puede hacer en MATLAB con las instrucciones:

```
A=[1,-2,-1;2,0,1;2,-4,2;4,1,0];
b=[-1,;1;1;-2];
x_LS=A\b;
```

No obstante en este procedimiento alterno no se utiliza la factorización QR de A lo que puede en general, para problemas más grandes, producir resultados no tan precisos como el método anterior. □

5.6 Ejercicios

Ejercicio 5.1. Considere los vectores $\vec{a} = (2,1,1)^t$, $\vec{b} = (1,-1,2)^t$, $\vec{c} = (8,-3,0)^t$.

a) Halle el ángulo entre los vectores $\vec{a}$ y $\vec{b}$.

b) Halle un vector $\vec{n}$ que sea perpendicular a los vectores $\vec{a} - \vec{c}$ y $\vec{b} - \vec{c}$.

c) Halle la ecuación del plano con normal $\vec{n}$ y que contiene a $\vec{c}$.

d) Verifique que el plano de la parte (c) contiene también a los puntos $\vec{a}$ y $\vec{b}$.

Ejercicio 5.2. Sean $\vec{p}_1 = (1,1,1)^t$, $\vec{p}_2 = (2,4,-1)^t$, $\vec{p}_3 = (0,-1,5)^t$.

a) Halle un vector no cero $\vec{n}$ que sea ortogonal a ambos $\vec{p}_2 - \vec{p}_1$ y $\vec{p}_3 - \vec{p}_1$.

b) Usando el resultado de la primera parte, halle una ecuación para el plano que contiene los puntos $\vec{p}_1, \vec{p}_2, \vec{p}_3$.

Ejercicio 5.3. Sean $\vec{u}_1$ y $\vec{u}_2$ vectores tales que

$$< \vec{u}_1, \vec{u}_1 >= 3, \quad < \vec{u}_1, \vec{u}_2 >= -2, \quad < \vec{u}_2, \vec{u}_2 >= 5.$$

Calcúle $< 5\vec{u}_1 + 8\vec{u}_2, 6\vec{u}_1 - 2\vec{u}_2 >$.

Ejercicio 5.4. Halle el valor de c de modo que $< 2\vec{u}_1 - 3\vec{u}_2, c\vec{u}_1 + 4\vec{u}_2 >= 5$ dado que

$$< \vec{u}_1, \vec{u}_1 >= 3, \quad < \vec{u}_1, \vec{u}_2 >= -4, \quad < \vec{u}_2, \vec{u}_2 >= 7.$$

Ejercicio 5.5. Halle la distancia del punto $\vec{a}$ al plano especificado.

a) $\vec{a} = (-1,-2,1)^t$ al plano $5x - 6y + 3z = 0$.

b) $\vec{a} = (-4,2,1)^t$ al plano $5x - 3y + 10z = 17$.

c) $\vec{a} = (4,3,-2)$ al plano $-6(x-1) + 2(y-3) + 3(z+4) = 0$.

Ejercicio 5.6. Halle la proyección escalar y la vectorial de $\vec{a}$ sobre $\vec{v}$ donde:

a) $\vec{a} = (-1,2,3,-4)^t$ y $\vec{v} = (2,3,2,3)^t$.

b) $\vec{a} = (-5,2,-4)^t$ y $\vec{v} = (1,-1,1)^t$.

Ejercicio 5.7. Calcule la proyección vectorial de $\vec{a} = (-4,2,6,1)^t$ sobre $S = \text{span}\{\vec{u}_1, \vec{u}_2\}$, donde $\vec{u}_1 = (2,-3,5,1)^t$ y $\vec{u}_2 = (-1,2,0,8)^t$.

Ejercicio 5.8. Halle el ángulo entre los vectores:

a) $\vec{a} = (2,1,1)^t$ y $\vec{b} = (1,-1,2)^t$

b) $\vec{a} = (-4,2,-1,3)^t$ y $\vec{b} = (1,-2,2,1)^t$

Ejercicio 5.9. Para los vectores $\vec{v} = (2,1,3)^t$ y $\vec{w} = (6,3,9)^t$, halle el ángulo entre los dos vectores y las proyecciones escalar y vectorial de $\vec{v}$ sobre $\vec{w}$.

Ejercicio 5.10. Halle la distancia del punto $(1,2)$ a la recta $4x - 3y = 0$.

Ejercicio 5.11. Halle una ecuación para el plano que contiene el punto $\vec{c} = (8,-3,0)^t$ y es paralelo al plano $5(x-4) + 10(y+3) - 7(z-2) = 0$.

Ejercicio 5.12. Halle una ecuación para el plano que contiene el punto $(4,2,-5)$ y con dirección normal $(-3,6,2)$.

Ejercicio 5.13. Para el subespacio S especificado, encuentre $S^\perp$ y una base para éste.

a) $S = \text{span}\{(1,-2)^t\} \subset \mathbb{R}^2$.

b) $S = \text{span}\{(1,3,-4)^t\} \subset \mathbb{R}^3$.

c) $S = \text{span}\{(1,-1,1)^t\} \subset \mathbb{R}^3$.

d) $S = \text{span}\{\vec{x}_1, \vec{x}_2\} \subset \mathbb{R}^4$, $\vec{x}_1 = (1,0,-2,1)^t$, $\vec{x}_2 = (0,1,3,2)^t$.

Ejercicio 5.14. Sean $S_1 = \text{span}\{\vec{u}_1\}$ y $S_2 = \text{span}\{\vec{u}_2\}$ donde $\vec{u}_1 = (1,2,1)^t$ y $\vec{u}_2 = (1,-1,1)^t$. Verifique que $S_1 \perp S_2$.

Ejercicio 5.15. La siguiente matriz:

$$B = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

tiene columnas ortonormales que forman una base de $\mathbb{R}^4$.

a) Escriba el vector $\vec{u} = (1,2,3,4)^t$ como una combinación lineal de las columnas de B.

b) ¿Cuál será B^{-1}?

Ejercicio 5.16. Sea $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ una base ortonormal de $\mathbb{R}^3$ y $\vec{x}$ un vector tal que:

$$< \vec{x}, \vec{u}_1 >= -3, \quad < \vec{x}, \vec{u}_2 >= 4, \quad < \vec{x}, \vec{u}_3 >= 2.$$

Escriba $\vec{x}$ como una combinación lineal de $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ y calcule $\left\|\vec{x}\right\|$.

Ejercicio 5.17. Sea $S = \text{span}\{\vec{u}_1, \vec{u}_2\}$ donde $\{\vec{u}_1, \vec{u}_2\}$ es un conjunto ortonormal en $\mathbb{R}^n$. Sea $\vec{y}$ un vector tal que $\langle \vec{y}, \vec{u}_1 \rangle = -4$ y $\langle \vec{y}, \vec{u}_2 \rangle = 6$. Calcule la proyección vectorial de $\vec{y}$ sobre S.

Ejercicio 5.18. Halle todos los valores de a, b que hacen que el conjunto de vectores

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} a \\ b \\ 3 \end{bmatrix} \right\},$$

sea ortogonal.

Ejercicio 5.19. ¿Será posible que una matriz tenga al vector $(3,1,2)^t$ en su espacio fila, y al vector $(2,1,1)^t$ en su espacio nulo? Explique.

Ejercicio 5.20. Sea A una matriz 4×4 de rango 3 y con $N(A^t) = \text{span}\{(-2,4,1,1)^t\}$. Considere el sistema $A\vec{x} = \vec{b}$ donde $\vec{b} = (1,-1,2,4)^t$. ¿Tendrá solución el sistema $A\vec{x} = \vec{b}$? Si tiene solución, ¿será única la solución? Justifique sus contestaciones usando los teoremas discutidos en la clase.

Ejercicio 5.21. Sea A una matriz 4×2. ¿Podría el vector $[-1,4,5,2]^t$ pertenecer a $R(A)$ y $[4,-2,3,1]^t$ a $N(A^t)$? Explique su contestación.

Ejercicio 5.22. Sean

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ -1 & 0 \\ 3 & 1 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}.$$

¿Cuáles son las ecuaciones que deben satisfacer los b_i's para que el sistema $A\vec{x} = \vec{b}$ tenga solución?

Ejercicio 5.23. Sea $S = \text{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$. Verifique que $\vec{y} \in S^{\perp}$ si y solo si $\vec{y} \perp \vec{v}_k$ para $k = 1, 2, \ldots, m$.

Ejercicio 5.24. Sean:

$$\vec{u}_1 = \begin{bmatrix} \frac{1}{3\sqrt{2}} \\ \frac{1}{3\sqrt{2}} \\ -\frac{4}{3\sqrt{2}} \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ -\frac{1}{3} \end{bmatrix}, \quad \vec{u}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}.$$

a) Verifique que $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es una base ortonormal para $\mathbb{R}^3$.

b) Escriba el vector $\vec{x} = (1,1,1)^t$ como una combinación lineal de $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ y calcule $\left\|\vec{x}\right\|$. **Nota:** Use los teoremas para bases ortonormales.

Ejercicio 5.25. Sea $S = \text{span}\{\vec{u}_2, \vec{u}_3\}$ donde $\vec{u}_2, \vec{u}_3$ están dados en el problema anterior. Halle la proyección $\vec{p}$ del vector $\vec{x} = (1,2,2)^t$ sobre S.

Ejercicio 5.26. Sea $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ una base ortonormal de $\mathbb{R}^3$. Suponga que:

$$\vec{u} = \vec{v}_1 + 2\vec{v}_2 + 2\vec{v}_3, \quad \vec{v} = \vec{v}_1 + 7\vec{v}_3.$$

Calcule lo siguiente:

a) $< \vec{u}, \vec{v} >$

b) $\left\|\vec{u}\right\|$ y $\left\|\vec{v}\right\|$

c) el ángulo entre los vectores $\vec{u}$ y $\vec{v}$

Ejercicio 5.27. Calcule la proyección vectorial del vector $\vec{b} = (2,3,6)^t$ sobre $R(A)$ donde

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

Ejercicio 5.28. Considere el sistema:

$$\left\{ \begin{array}{rcl} x_1 + x_2 + x_3 & = & 4, \\ -x_1 + x_2 + x_3 & = & 0, \\ -x_2 + x_3 & = & 1, \\ x_1 \qquad + x_3 & = & 2. \end{array} \right.$$

a) Halle la mejor solución del sistema en el sentido de cuadrados mínimos.

b) Halle la proyección $\vec{p}$ del lado derecho del sistema a $R(A)$, donde A es la matriz de coeficientes del sistema dado. **Nota:** $\vec{p} = A\vec{x}_{LS}$.

c) Calcule el residual $\vec{r}(\vec{x}_{LS})$.

d) Verifique que $\vec{r}(\vec{x}_{LS}) \in N(A^t)$.

Ejercicio 5.29. Halle la mejor recta (en el sentido de cuadrados mínimos) para los datos:

a) $(-1,4)$, $(0,-2)$, $(2,3)$.

b) $(0,2)$, $(1,1)$, $(2,3)$.

Ejercicio 5.30. Halle la mejor recta en el sentido de los cuadrados mnimos para los datos:

x	0	1	-1	2
y	-4	-2	-5	0

Ejercicio 5.31. Halle el mejor polinomio de grado dos, en el sentido de los cuadrados mínimos, para los datos:

x	-1	0	1	2
y	0	1	3	9

Ejercicio 5.32. Halle la ecuación del circulo que da la mejor aproximación en el sentido de cuadrados mínimos para los datos $(-1,-2)$, $(0,2.4)$, $(1.1,-4)$, $(2.4,-1.6)$.

Ejercicio 5.33. Usando el método de Gram-Scmidt, construya un conjunto ortonormal $\{\vec{v}_1, \vec{v}_2\}$ a partir de los vectores

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

de modo que $\operatorname{span}\{\vec{v}_1, \vec{v}_2\} = \operatorname{span}\{\vec{u}_1, \vec{u}_2\}$. Usando ésto encuentre la factorización QR de la matriz $A = [\vec{u}_1, \vec{u}_2]$.

Ejercicio 5.34. Usando el método de Gram-Schmidt, construya un conjunto ortonormal $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ a partir de los vectores:

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

de modo que $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \operatorname{span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$. Usando ésto encuentre la factorización QR de la matriz

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

Ejercicio 5.35. Encuentre una base ortonormal de $\mathbb{R}^4$ que contenga los vectores $[1, 0, 2, 2]^t$ y $[0, 1, 1, -1]^t$.

6 Valores y vectores propios

Los valores propios y vectores propios de matrices o más general aún, de transformaciones lineales, se utilizan para representar las soluciones de un sistema como una suma o expansión en términos de los vectores propios. Estas expansiones son útiles para estudiar diferentes aspectos de las soluciones del sistema. Por ejemplo, en aplicaciones a la biología de poblaciones, es de interés el poder describir el comportamiento de las soluciones del sistema $A^n\vec{x} = \vec{b}$ según la $n \longrightarrow \infty$. Otro ejemplo importante es el de un sistema mecánico donde se desea caracterizar el movimiento o comportamiento del sistema sujeto a un agente externo. Este problema y el anterior se pueden analizar utilizando los valores y vectores propios de los operadores correspondientes.

6.1 Definiciones y conceptos básicos

Un autovector de una matriz es una dirección en el espacio que la matriz, como transformación lineal, deja esencialmente intacta en el sentido de que no cambia la dirección del vector, solo posiblemente su largo o orientación.

Definición 6.1. Sea A una matriz $n \times n$ con entradas en $\mathbb{C}$. Un número $\lambda \in \mathbb{C}$ es un *valor propio* (*valor característico o autovalor*) de A si existe un vector $\vec{x} \in \mathbb{C}^n$, $\vec{x} \neq \vec{0}$ tal que:

$$A\vec{x} = \lambda\vec{x}.$$

El vector $\vec{x}$ se llama *vector propio* (*vector característico* o *autovector*) de A asociado al valor propio λ.

Nota 6.2. El efecto de A, como transformación lineal, sobre el autovector $\vec{x}$ es multiplicar a $\vec{x}$ por el factor λ. □

La ecuación $A\vec{x} = \lambda\vec{x}$ es equivalente a:

$$(A - \lambda I)\vec{x} = \vec{0}.$$

Así que λ es un valor propio de A si y solo si esta ecuación tiene soluciones notriviales, lo cual ocurre si y solo si

$$\det(A - \lambda I) = 0.$$

Note que también esta condición es equivalente a

$$N(A-\lambda I) \neq \left\{\vec{0}\right\}.$$

El conjunto $N(A-\lambda I)$ se llama el *conjunto característico* de A asociado al valor característico λ. Cualquier vector distinto de cero en $N(A-\lambda I)$, es un vector característico correspondiente al valor valor característico λ.

El polinomio de grado n:

$$p_A(\lambda) = \det(A-\lambda I),$$

se llama el *polinomio característico* de A y sus raíces son precisamente los valores propios de A. Por el Teorema Fundamental del Álgebra, este polinomio siempre tiene n raíces complejas contando las multiplicidades de éstas.

Ejemplo 6.3. Para la matriz

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

tenemos que

$$p_A(\lambda) = \det \begin{bmatrix} 1-\lambda & 1 \\ 0 & 2-\lambda \end{bmatrix} = (1-\lambda)(2-\lambda).$$

Así que los autovalores de A son $\lambda = 1, 2$. Ahora buscamos los autovectores de A correspondientes a cada autovalor.

Para $\lambda = 1$:

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

De modo que

$$N(A-I) = \left\{ \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} : \alpha \in \mathbb{R} \right\}.$$

Así que todos los autovectores de A asociados a $\lambda = 1$ son proporcionales al autovector $(1,0)^t$.

Para el autovalor $\lambda = 2$:

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

por lo que:

$$N(A-2I)=\left\{\alpha\begin{bmatrix}1\\1\end{bmatrix} : \alpha\in\mathbb{R}\right\},$$

y todos los autovectores de A asociados a $\lambda=2$ son proporcionales al autovector $(1,1)^t$.

□

Ejemplo 6.4. Para la matriz

$$A=\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix},$$

tenemos que

$$p_A(\lambda)=\lambda^2(3-\lambda).$$

Los autovalores de A son $\lambda=0$ (multiplicidad dos) y $\lambda=3$. Para $\lambda=0$, $A-\lambda I=A$ por lo que si $\vec{x}=(x_1,x_2,x_3)^t$ es un autovector de A asociado a $\lambda=0$, entonces $A\vec{x}=\vec{0}$, i.e.,

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

i.e., $x_1=-x_2-x_3$. Tenemos pues que

$$N(A)=\left\{x_2\begin{bmatrix}-1\\1\\0\end{bmatrix}+x_3\begin{bmatrix}-1\\0\\1\end{bmatrix} : x_2,x_3\in\mathbb{R}\right\}.$$

Esto es, los autovectores $\{(-1,1,0)^t,(-1,0,1)^t\}$ generan (de hecho son una base) el conjunto de autovectores de A asociados a $\lambda=0$. Como $\dim N(A)=2$ decimos que $\lambda=0$ tiene dos autovectores asociados linealmente independientes.

Para $\lambda=3$,

$$A-\lambda I=A-3I=\begin{bmatrix}-2&1&1\\1&-2&1\\1&1&-2\end{bmatrix}.$$

Así que si $\vec{x}=(x_1,x_2,x_3)^t\in N(A-3I)$, entonces

$$\left\{\begin{array}{rcl}-2x_1+x_2+x_3&=&0,\\x_1-2x_2+x_3&=&0,\\x_1+x_2-2x_3&=&0.\end{array}\right.$$

Este sistema tiene solución $x_1 = x_3$, $x_2 = x_3$ por lo que

$$N(A-3I) = \operatorname{span}\left\{(1,1,1)^t\right\}.$$

Así que los autovectores de A asociados a $\lambda = 3$ son todos proporcionales al autovector $(1,1,1)^t$. $\square$

Aún cuando una matriz A tenga todas sus entradas reales, ésta puede tener valores propios complejos. No obstante, como en éste caso $p_A(\lambda)$ tiene coeficientes reales, entonces los valores propios complejos ocurren en pares conjugados. En adición, si $\vec{x}$ es vector propio de A correspondiente al valor propio complejo λ, entonces $\overline{\vec{x}}$ es vector propio de A correspondiente al valor propio $\overline{\lambda}$. Esto es así ya que si A tiene todas sus entradas reales, entonces[1]

$$A\overline{\vec{x}} = \overline{A}\,\overline{\vec{x}} = \overline{A\vec{x}} = \overline{\lambda\vec{x}} = \overline{\lambda}\,\overline{\vec{x}}.$$

Ejemplo 6.5. Considere la matriz

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

Entonces

$$p_A(\lambda) = \det\begin{bmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{bmatrix} = (1-\lambda)^2 + 4.$$

De aquí que los valores propios de A son $\lambda = 1 \pm 2i$.

Para $\lambda = 1+2i$ tenemos que:

$$A - \lambda I = A - (1+2i)I = \begin{bmatrix} -2i & 2 \\ -2 & -2i \end{bmatrix}.$$

Entonces $\vec{x} = (x_1, x_2)^t \in N(A-(1+2i)I)$ si y solo si

$$-2ix_1 + 2x_2 = 0, \quad -2x_1 - 2ix_2 = 0,$$

esto es $x_1 = -ix_2$. De aquí que

$$N(A-(1+2i)I) = \operatorname{span}\left\{(-i,1)^t\right\},$$

[1]Aquí usamos que si B,C son matrices de entradas complejas y de tamaños tales que BC está bien definida, entonces $\overline{BC} = \overline{B}\,\overline{C}$.

por lo que los autovectores de A correspondientes a $\lambda = 1 + 2i$ son proporcionales al autovector $(-i, 1)^t$.

De igual forma se obtiene que los autovectores de A correspondientes a $\lambda = 1 - 2i$ son proporcionales al autovector $(i, 1)^t$. $\square$

El siguiente teorema resume las observaciones generales que hemos hecho hasta ahora sobre los valores propios de una matriz.

Teorema 6.6. *Sea A matriz $n \times n$ con entradas en $\mathbb{C}$. Entonces A tiene exactamente n valores propios contando sus multiplicidades como raíces de $p_A(\lambda)$. Además las siguientes aseveraciones son todas equivalentes:*

i) *λ es un valor propio de A.*

ii) *$(A - \lambda I)\vec{x} = \vec{0}$ tiene solución notrivial.*

iii) *$N(A - \lambda I) \neq \left\{\vec{0}\right\}$.*

iv) *$A - \lambda I$ es singular.*

v) *$\det(A - \lambda I) = 0$.*

Sea $A = (a_{ij})$ una matriz $n \times n$. Entonces la *traza* de A se define por:

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}.$$

Ejemplo 6.7. Para

$$A = \begin{bmatrix} 1 & 4 & -10 \\ 6 & -11 & 2 \\ 3 & -9 & 4 \end{bmatrix},$$

tenemos que $\operatorname{tr} A = 1 - 11 + 4 = -6$. $\square$

Usando inducción en n se puede verificar que:

$$p_A(\lambda) = (-1)^n \left[\lambda^n - (\operatorname{tr} A)\lambda^{n-1} + \cdots + (-1)^n \det A\right].$$

Usando ésto y el Teorema Fundamental del Álgebra, podemos verificar lo siguiente:

Teorema 6.8. *Sea A una matriz $n \times n$ y $\lambda_1, \lambda_2, \ldots, \lambda_n$ los valores propios de A. Entonces:*

$$tr A = \lambda_1 + \lambda_2 + \ldots + \lambda_n, \quad \det A = \lambda_1 \lambda_2 \cdots \lambda_n.$$

Dos matrices A y B de tamaño $n \times n$ son *similares* si existe una matriz nosingular T tal que

$$A = T^{-1} B T.$$

¿Qué podemos decir de los polinomios característicos de dos matrices similares? Si $A = T^{-1} B T$, entonces

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) = \det(T^{-1} B T - \lambda T^{-1} I T), \\ &= \det(T^{-1}(B - \lambda I) T) = (\det T^{-1})(\det(B - \lambda I))(\det T), \\ &= (\det(T^{-1} T))(\det(B - \lambda I)) = \det(B - \lambda I) = p_B(\lambda). \end{aligned}$$

Como los polinomios característicos son iguales, entonces las matrices tienen los mismos valores propios. No obstante, los vectores propios o característicos no son iguales.

Teorema 6.9. *Si A y B son similares, entonces tienen el mismo polinomio característico y por consiguiente los mismos valores propios.*

¡Toda matriz es similar a una matriz triangular superior! Como los valores propios de una matriz triangular son sus elementos de la diagonal, podríamos en principio utilizar éste resultado para calcular los valores propios de cualquier matriz. No obstante, de la demostración de éste teorema no se desprende un procedimiento práctico para calcular la transformación de similitud. Aun así este resultado se utiliza como la base o motivación para el desarrollo de otros métodos computacionales para el cómputo de valores propios.

Teorema 6.10. *Sea A una matriz de tamaño $n \times n$. Entonces existe una matriz nosingular P y una matriz B triangular superior, tal que*

$$B = P^{-1} A P.$$

No vamos a discutir la demostración de este resultado, pero la misma es por inducción en n y utiliza el Teorema 3.67, parte (iv), y el Teorema 4.29.

En todos los ejemplos que hemos visto hasta ahora, los autovectores de las matrices envueltas forman una base de $\mathbb{R}^n$ donde n es el tamaño de A. El siguiente ejemplo muestra que ésto no tiene que ser siempre así.

Ejemplo 6.11. Es fácil verificar que la matriz

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

tiene autovalor $\lambda = 1$ de multiplicidad dos como raíz de $p_A(\lambda)$, y con

$$N(A - \lambda I) = N(A - I) = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\},$$

por lo que no se pueden seleccionar autovectores que generen a $\mathbb{R}^2$. En este caso $\dim N(A - I) = 1$ es estrictamente menor que la multiplicidad como raíz de $p_A(\lambda)$ del autovalor $\lambda = 1$. □

Este ejemplo motiva la siguiente definición y resultado.

Definición 6.12. A la multiplicidad de un valor propio λ^* como raíz de $p_A(\lambda)$, se le llama la *multiplicidad algebraica* del valor propio, mientras que $\dim N(A - \lambda^* I)$ se llama la *multiplicidad geométrica* de λ^*.

La multiplicidad geométrica siempre es menor o igual que la multiplicidad algebraica.

Proposición 6.13. *Sea A una matriz $n \times n$ con valor propio λ^* de multiplicidad algebraica k. Entonces*

$$\dim N(A - \lambda^* I) \leq k.$$

Demostración: Sea $m = \dim N(A - \lambda^* I)$ y $\{\vec{v}_1, \ldots, \vec{v}_m\}$ una base para $N(A - \lambda^* I)$. Por el Teorema 3.67, parte (iv), este conjunto lo podemos completar con vectores $\{\vec{v}_{m+1}, \ldots, \vec{v}_n\}$ de modo que $U = \{\vec{v}_1, \ldots, \vec{v}_n\}$ sea una base de $\mathbb{C}^n$. Como $A\vec{v}_i = \lambda^* \vec{v}_i$, $i = 1, \ldots, m$, la representación matricial con respecto a U de la tranformación lineal generada por A, tiene la forma

$$T = \begin{bmatrix} \lambda^* I & B \\ O & C \end{bmatrix},$$

donde I es la matriz identidad $m \times m$, B tiene dimensiones $m \times (n - m)$, O es la matriz cero de dimensiones $(n - m) \times m$, y C es de tamaño $(n - m) \times (n - m)$. Como A y T son similares (Teorema 4.29), entonces tienen el mismo polinomio característico (Teorema 6.9). Claramente, en el polinomio característico de T, el factor $\lambda - \lambda^*$ aparece al menos m veces. Siendo k la multiplicidad algebraica de λ^*, obtenemos que $m \leq k$. □

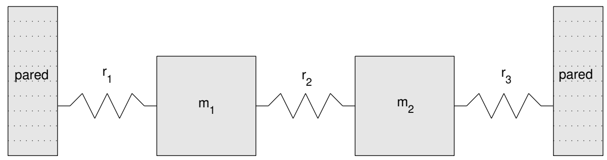


Figura 6.1: Sistema de dos masas y tres resortes.

6.2 Sistemas lineales de ecuaciones diferenciales

Consideramos el problema de un sistema de dos masas m_1, m_2 sujetas a tres resortes según se muestra en la Figura 6.1. Suponemos que las masas se deslizan sin fricción sobre una superficie o mesa, y que las fuerzas de los resortes están dadas por la Ley de Hooke:

$$f_i(x) = -k_i x, \quad i = 1, 2, 3,$$

donde $k_i > 0$, $i = 1, 2, 3$ y x representa el alargamiento del resorte en cuestión.

Si $x_i(t)$ denota el desplazamiento de la masa m_i, i=1,2, entonces usando la segunda ley de Newton obtenemos que las funciones $x_1(t)$ y $x_2(t)$ son soluciones del siguiente *sistema de ecuaciones diferenciales*:

$$\begin{aligned} m_1 \frac{\mathrm{d}^2 x_1}{\mathrm{d}t^2}(t) &= k_2(x_2(t) - x_1(t)) - k_1 x_1(t), \\ m_2 \frac{\mathrm{d}^2 x_2}{\mathrm{d}t^2}(t) &= -k_2(x_2(t) - x_1(t)) - k_3 x_2(t). \end{aligned} \tag{6.1}$$

Si especificamos las posiciones y velocidades iniciales de los resortes, entonces tendríamos que

$$x_1(0) = u_1, \quad \frac{\mathrm{d}x_1}{\mathrm{d}t}(0) = v_1, \quad x_2(0) = u_2, \quad \frac{\mathrm{d}x_2}{\mathrm{d}t}(0) = v_2.$$

Esto se llaman las *condiciones iniciales*.

Si hacemos la sustitución:

$$u_1(t) = x_1(t), \quad u_2(t) = \frac{\mathrm{d}x_1}{\mathrm{d}t}(t), \quad u_3(t) = x_2(t), \quad u_4(t) = \frac{\mathrm{d}x_2}{\mathrm{d}t}(t),$$

entonces el sistema anterior es equivalente a

$$\begin{aligned}
\frac{\mathrm{d}u_1}{\mathrm{d}t}(t) &= u_2(t), \\
\frac{\mathrm{d}u_2}{\mathrm{d}t}(t) &= -\left(\frac{k_1+k_2}{m_1}\right)u_1(t)+\left(\frac{k_2}{m_1}\right)u_3(t), \\
\frac{\mathrm{d}u_3}{\mathrm{d}t}(t) &= u_4(t), \\
\frac{\mathrm{d}u_4}{\mathrm{d}t}(t) &= \left(\frac{k_2}{m_2}\right)u_1(t)-\left(\frac{k_2+k_3}{m_2}\right)u_3(t),
\end{aligned}$$

el cuál se puede escribir como:

$$\frac{\mathrm{d}\vec{u}}{\mathrm{d}t}(t) = A\vec{u}(t), \quad \vec{u}(0) = \vec{u}_0,$$

donde $\vec{u}_0 = (u_1, v_1, u_2, v_2)^t$, y

$$\vec{u}(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \\ u_4(t) \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\dfrac{k_1+k_2}{m_1} & 0 & \dfrac{k_2}{m_1} & 0 \\ 0 & 0 & 0 & 1 \\ \dfrac{k_2}{m_2} & 0 & -\dfrac{k_2+k_3}{m_2} & 0 \end{bmatrix}.$$

El problema de las dos masas y tres resortes es un caso particular de un sistema de la forma:

$$\frac{\mathrm{d}\vec{u}}{\mathrm{d}t}(t) = A\vec{u}(t), \tag{6.2}$$

donde A es una matriz $n \times n$. Este sistema se llama un *sistema lineal de ecuaciones diferenciales de primer orden con coeficientes constantes*. Buscamos soluciones de este sistema de la forma:

$$\vec{u}(t) = \mathrm{e}^{\lambda t}\vec{x},$$

donde $\lambda \in \mathbb{C}$, $\vec{x} \in \mathbb{C}^n$ están por determinarse. Como

$$\frac{\mathrm{d}\vec{u}}{\mathrm{d}t}(t) = \lambda \mathrm{e}^{\lambda t}\vec{x},$$

sustituyendo en la ecuación diferencial (6.2) tenemos que

$$\lambda e^{\lambda t}\vec{x} = A\left(e^{\lambda t}\vec{x}\right),$$

o lo mismo, cancelando el exponencial, que

$$A\vec{x} = \lambda\vec{x}.$$

Así que $\vec{u}(t) = e^{\lambda t}\vec{x}$ es solución del sistema (6.2) si y solo si λ es un autovalor de A y $\vec{x}$ es un autovector de A correspondiente a λ.

Nota 6.14. Por la *linealidad* del sistema (6.2), si $\vec{u}_1(t), \vec{u}_2(t)$ son soluciones del sistema, entonces

$$c_1\vec{u}_1(t) + c_2\vec{u}_2(t),$$

es también solución de (6.2). Esto se conoce como el *principio de superposición* y es valido en general solo para sistemas lineales. □

Ejemplo 6.15. Considere el sistema

$$\begin{aligned}\frac{dx_1}{dt}(t) &= 4x_1(t) - 5x_2(t),\\ \frac{dx_2}{dt}(t) &= 2x_1(t) - 3x_2(t).\end{aligned} \tag{6.3}$$

La matriz de coeficientes es:

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

Es fácil ver que los valores propios y vectores propios de A están dados por:

$$\lambda_1 = -1, \quad \vec{v}_1 = (1,1)^t,$$

$$\lambda_2 = 2, \quad \vec{v}_2 = (5,2)^t.$$

Tenemos pues que

$$\vec{u}_1(t) = e^{-t}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{u}_2(t) = e^{2t}\begin{bmatrix} 5 \\ 2 \end{bmatrix},$$

son ambas soluciones del sistema, y estas soluciones son linealmente independientes. La *solución general* del sistema está dada por:

$$\vec{u}(t) = c_1\vec{u}_1(t) + c_2\vec{u}_2(t) = \begin{bmatrix} c_1 e^{-t} + 5c_2 e^{2t} \\ c_1 e^{-t} + 2c_2 e^{2t} \end{bmatrix}. \tag{6.4}$$

Si en adición tenemos la condición inicial de que

$$x_1(0) = 1, \quad x_2(0) = -1, \tag{6.5}$$

entonces sustituyendo en (6.4) obtenemos el siguiente sistema de ecuaciones lineales para las constantes c_1, c_2:

$$\begin{cases} c_1 + 5c_2 &= 1, \\ c_1 + 2c_2 &= -1, \end{cases}$$

el cual tiene solución $c_1 = -7/3$, $c_2 = 2/3$. Así que la solución del sistema de ecuaciones diferenciales (6.3) y que satisface ta condición inicial (6.5) es:

$$\vec{u}(t) = \begin{bmatrix} -\dfrac{7}{3} e^{-t} + \dfrac{10}{3} e^{2t} \\ \\ -\dfrac{7}{3} e^{-t} + \dfrac{4}{3} e^{2t} \end{bmatrix}.$$

□

6.2.1 Valores propios complejos

Sea A una matriz $n \times n$ con entradas reales y $\lambda = a + bi$ un valor propio (simple) de A con vector propio correspondiente $\vec{x}$. Es fácil ver que $\bar{\lambda} = a - bi$ es también valor propio de A con $\bar{\vec{x}}$ vector propio correspondiente. Partiendo de las soluciones:

$$e^{\lambda t}\vec{x}, \quad e^{\bar{\lambda}}\bar{\vec{x}},$$

del sistema lineal $\vec{u}'(t) = A\vec{u}(t)$, y tomando combinaciones lineales apropiadas, se obtiene que:

$$\vec{u}_1(t) = \mathrm{Re}\left(e^{\lambda t}\vec{x}\right), \quad \vec{u}_2(t) = \mathrm{Im}\left(e^{\lambda t}\vec{x}\right), \tag{6.6}$$

son soluciones linealmente independientes del sistema.

Ejemplo 6.16. Considere el sistema

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}.$$

Los valores propios de la matriz de coeficientes son $\lambda = 1 + 2i$, $\bar{\lambda} = 1 - 2i$, con los vectores propios correspondientes $\vec{x} = (-i, 1)^t$, $\bar{\vec{x}} = (i, 1)^t$. Usando la identidad de Euler: $\mathrm{e}^{i\theta} = \cos\theta + i\sin\theta$, tenemos que

$$\begin{aligned} \mathrm{e}^{\lambda t}\vec{x} &= \mathrm{e}^{(1+2i)t}\begin{bmatrix} -i \\ 1 \end{bmatrix} = \mathrm{e}^{t}\mathrm{e}^{(2t)i}\begin{bmatrix} -i \\ 1 \end{bmatrix}, \\ &= \mathrm{e}^{t}\begin{bmatrix} -i(\cos 2t + i\sin 2t) \\ \cos 2t + i\sin 2t \end{bmatrix}, \\ &= \begin{bmatrix} \mathrm{e}^{t}\sin 2t \\ \mathrm{e}^{t}\cos 2t \end{bmatrix} + i\begin{bmatrix} -\mathrm{e}^{t}\cos 2t \\ \mathrm{e}^{t}\sin 2t \end{bmatrix}. \end{aligned}$$

Usando las formulas (6.6) tenemos que:

$$\vec{u}_1(t) = \mathrm{Re}\left(\mathrm{e}^{\lambda t}\vec{x}\right) = \begin{bmatrix} \mathrm{e}^{t}\sin 2t \\ \mathrm{e}^{t}\cos 2t \end{bmatrix}, \quad \vec{u}_2(t) = \mathrm{Im}\left(\mathrm{e}^{\lambda t}\vec{x}\right) = \begin{bmatrix} -\mathrm{e}^{t}\cos 2t \\ \mathrm{e}^{t}\sin 2t \end{bmatrix},$$

son soluciones linealmente independientes del sistema. La solución general está dada por:

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = c_1\vec{u}_1(t) + \vec{u}_2(t) = \begin{bmatrix} \mathrm{e}^{t}(c_1\sin 2t - c_2\cos 2t) \\ \mathrm{e}^{t}(c_1\cos 2t + c_2\sin 2t) \end{bmatrix}$$

□

Ejemplo 6.17. En el ejemplo de las dos masas y tres resortes al principio de este capitulo, tomamos $m_1 = m_2 = 1$, $k_i = 1$, $i = 1, 2, 3$. El sistema queda de la forma $\vec{u}'(t) = A\vec{u}(t)$ donde

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

Los valores y vectores propios correspondientes de esta matriz son:

$$\lambda_1 = i, \quad \vec{v}_1 = \begin{bmatrix} -i, & 1, & -i & ,1 \end{bmatrix}^t, \quad \lambda_2 = -i, \quad \vec{v}_2 = \bar{\vec{v}}_1,$$

$$\lambda_3 = i\sqrt{3}, \quad \vec{v}_3 = \left[\frac{\sqrt{3}}{3}i, \quad -1, \quad -\frac{\sqrt{3}}{3}i, \quad 1\right]^t, \quad \lambda_4 = -i\sqrt{3}, \quad \vec{v}_4 = \bar{\vec{v}}_3.$$

Tenemos entonces las soluciones:

$$\vec{u}_1(t) = \mathrm{Re}\left(\mathrm{e}^{\lambda_1 t}\vec{v}_1\right) = \begin{bmatrix} \sin t \\ \cos t \\ \sin t \\ \cos t \end{bmatrix}, \quad \vec{u}_2(t) = \mathrm{Im}\left(\mathrm{e}^{\lambda_1 t}\vec{v}_1\right) = \begin{bmatrix} -\cos t \\ \sin t \\ -\cos t \\ \sin t \end{bmatrix},$$

$$\vec{u}_3 = \mathrm{Re}\left(\mathrm{e}^{\lambda_3 t}\vec{v}_3\right) = \begin{bmatrix} -\frac{\sqrt{3}}{3}\sin\sqrt{3}\,t \\ -\cos\sqrt{3}\,t \\ \frac{\sqrt{3}}{3}\sin\sqrt{3}\,t \\ \cos\sqrt{3}\,t \end{bmatrix}, \quad \vec{u}_4 = \mathrm{Im}\left(\mathrm{e}^{\lambda_3 t}\vec{v}_3\right) = \begin{bmatrix} \frac{\sqrt{3}}{3}\cos\sqrt{3}\,t \\ -\sin\sqrt{3}\,t \\ -\frac{\sqrt{3}}{3}\cos\sqrt{3}\,t \\ \sin\sqrt{3}\,t \end{bmatrix}.$$

Se puede verificar que estas cuatro soluciones son linealmente independientes por lo que la solución general del sistema de ecuaciones diferenciales es:

$$\vec{u}(t) = c_1\vec{u}_1(t) + c_2\vec{u}_2(t) + c_3\vec{u}_3(t) + c_4\vec{u}_4(t).$$

Recordando que:

$$u_1(t) = x_1(t), \quad u_2(t) = \frac{\mathrm{d}x_1}{\mathrm{d}t}(t), \quad u_3(t) = x_2(t), \quad u_4(t) = \frac{\mathrm{d}x_2}{\mathrm{d}t}(t).$$

tenemos que la solución general del sistema (sistema (6.1) con $m_1 = m_2 = 1$, $k_i = 1$, $i = 1, 2, 3$):

$$\frac{\mathrm{d}^2x_1}{\mathrm{d}t^2}(t) = -2x_1(t) + x_2(t), \quad \frac{\mathrm{d}^2x_2}{\mathrm{d}t^2}(t) = x_1(t) - 2x_2(t),$$

está dada por:

$$\begin{aligned} x_1(t) &= c_1\sin t - c_2\cos t - c_3\frac{\sqrt{3}}{3}\sin\sqrt{3}\,t + c_4\frac{\sqrt{3}}{3}\cos\sqrt{3}\,t, \\ x_2(t) &= c_1\sin t - c_2\cos t + c_3\frac{\sqrt{3}}{3}\sin\sqrt{3}\,t - c_4\frac{\sqrt{3}}{3}\cos\sqrt{3}\,t. \end{aligned}$$

□

6.3 Matrices diagonalizables

Una matriz es *diagonalizable* si ésta es similar a una matriz diagonal. En esta sección vamos a estudiar condiciones bajo las cuales una matriz es diagonalizable. Para facilitar un tanto la discusión, nos vamos a limitar a matrices con entradas reales.

El siguiente resultado nos da condiciones necesarias y suficientes para que una matriz A sea diagonalizable.

Teorema 6.18. *Una matriz A de tamaño $n \times n$ es diagonalizable si y solo si A tiene n vectores propios linealmente independientes.*

Demostración: Suponga que A es diagonalizable. Entonces existe una matriz nosingular S tal que $AS = DS$, donde D es diagonal. Si escribimos $S = [\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n]$ y $D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, entonces la ecuación $AS = DS$ es equivalente a:

$$A\vec{v}_i = \lambda_i \vec{v}_i, \quad i = 1, 2, \ldots, n,$$

es decir que cada λ_i es valor propio de A con vector propio asociado $\vec{v}_i$. Como S es nosingular, los vectores propios $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ son linealmente independientes.

Suponga $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ son vectores propios de A linealmente independientes asociados a los valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectivamente. Esto es

$$A\vec{v}_i = \lambda_i \vec{v}_i, \quad i = 1, 2, \ldots, n, \tag{6.7}$$

Defina $S = [\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n]$ y $D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Como $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ son linealmente independientes, entonces S es nosingular y las ecuaciones (6.7) son equivalentes a $AS = SD$, lo que implica que A es diagonalizable. $\square$

Tenemos entonces que la pregunta de si una matriz es diagonalizable o no es equivalente a si la matriz tiene o no n vectores propios linealmente independientes. Hay dos casos bien importantes de matrices cuyos vectores propios generan a $\mathbb{R}^n$: cuando la matriz tiene n valores propios distintos, o cuando la matriz es simétrica. Vamos a examinar estos dos casos.

Teorema 6.19. *Sea A una matriz $n\times n$ y sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ sus valores propios. Si los valores propios son todos de multiplicidad algebraica uno, entonces A tiene n vectores propios linealmente independientes.*

Demostración: Procedemos inductivamente, añadiendo un vector propio a la vez. Sean $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ vectores propios asociados a los valores propios $\lambda_1, \lambda_2, \dots, \lambda_n$ respectivamente. Como los vectores propios no son iguales al vector cero, tenemos que $\{\vec{v}_1\}$ es linealmente independiente. Suponga que para algún k con $1 \le k < n$, los vectores $\{\vec{v}_1, \dots, \vec{v}_k\}$ son linealmente independientes. Supongamos que existen escalares no todos cero $c_1, c_2, \dots, c_{k+1}$ tal que

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + c_{k+1}\vec{v}_{k+1} = \vec{0}.$$

Como $\{\vec{v}_1, \dots, \vec{v}_k\}$ son linealmente independientes, tenemos que $c_{k+1} \neq 0$. También, como $\vec{v}_{k+1} \neq \vec{0}$, los escalares $c_1, \dots, c_k$ no pueden ser todos cero. Como $\sum_{j=1}^{k+1} c_j\vec{v}_j = \vec{0}$, tenemos que

$$\vec{0} = A\left[\sum_{j=1}^{k+1} c_j\vec{v}_j\right] - \lambda_{k+1}\left[\sum_{j=1}^{k+1} c_j\vec{v}_j\right] = \sum_{j=1}^{k} c_j(\lambda_j - \lambda_{k+1})\vec{v}_j.$$

Como $\{\vec{v}_1, \dots, \vec{v}_k\}$ son linealmente independientes

$$c_j(\lambda_j - \lambda_{k+1}) = 0, \quad j = 1, \dots, k.$$

Pero como $\lambda_j - \lambda_{k+1} \neq 0$ para $j = 1, \dots, k$, tendríamos entonces que $c_j = 0$ para $j = 1, \dots, k$, lo que es una contradicción ya que estos escalares no pueden ser todos cero. Tenemos pues que $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{k+1}\}$ son linealmente independientes. Continuamos ahora inductivamente hasta que $k = n$. $\square$

Ejemplo 6.20. Considere la matriz

$$A = \begin{bmatrix} 6 & 4 \\ -3 & -1 \end{bmatrix}.$$

Los autovalores de A son 2 y 3, que son distintos, por lo que A es diagonalizable (Teoremas 6.18 y 6.19). Los autovectores correspondientes son

$$\vec{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -4/3 \\ 1 \end{bmatrix}.$$

Tenemos entonces que $AS = SD$ donde

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \quad S = [\vec{v}_1, \vec{v}_2] = \begin{bmatrix} -1 & -4/3 \\ 1 & 1 \end{bmatrix}.$$

$\square$

Veamos ahora el caso de matrices simétricas. Vamos a necesitar el siguiente resultado, que por cierto, no requiere que la matriz sea simétrica.

Lema 6.21. *Sea A una matriz $n \times n$ y U un subespacio de $\mathbb{R}^n$ tal que $A\vec{x} \in U$ para todo $\vec{x} \in U$. Entonces existe un vector $\vec{w} \in U$, $\vec{w} \neq \vec{0}$, y $\lambda \in \mathbb{C}$ tal que $A\vec{w} = \lambda\vec{w}$.*

Demostración: Sea $S = \{\vec{u}_1, \ldots, \vec{u}_k\}$ una base ortonormal para U. Si visualizamos a la matriz A como una transformación lineal, tenemos que $A : U \longrightarrow U$. Sea B la representación matricial de A con respecto a S, es decir

$$B = [(A\vec{u}_1)_S, \ldots, (A\vec{u}_k)_S] \in M_{k\times k}.$$

Como S es ortonormal,

$$(A\vec{u}_j)_S = \left[\sum_{i=1}^{k} < A\vec{u}_j, \vec{u}_i > \vec{u}_i\right]_S = \begin{bmatrix} < A\vec{u}_j, \vec{u}_1 > \\ \vdots \\ < A\vec{u}_j, \vec{u}_k > \end{bmatrix}, \quad j = 1, \ldots, k.$$

Sea $\lambda \in \mathbb{C}$ un valor propio de B con $\vec{v} \in \mathbb{R}^k$ un vector propio asociado. Defina $\vec{w} = \sum_{i=1}^{k} v_i \vec{u}_i \in U$, donde $\vec{v} = (v_i)$. Como $\vec{v} \neq \vec{0}$, entonces $\vec{w} \neq \vec{0}$, y $(\vec{w})_S = \vec{v}$. Tenemos entonces que

$$(A\vec{w})_S = B(\vec{w})_S = B\vec{v} = \lambda\vec{v} = \lambda(\vec{w})_S = (\lambda\vec{w})_S.$$

De aquí que $A\vec{w} = \lambda\vec{w}$ donde $\vec{w} \in U$, $\vec{w} \neq \vec{0}$. $\square$

Vamos ahora a suponer que la matriz A es simétrica.

Lema 6.22. *Sea A una matriz simétrica. Entonces:*

i) $< A\vec{x}, \vec{y} > = < \vec{x}, A\vec{y} >$, *para todo* $\vec{x}, \vec{y} \in \mathbb{C}^n$.

ii) $< A\vec{x}, \vec{x} > \in \mathbb{R}$ *para todo* $\vec{x} \in \mathbb{C}^n$.

iii) Los valores propios de A son reales.

iv) Sean $\{\vec{u}_1, \ldots, \vec{u}_k\}$ vectores propios de A asociados a los valores propios $\lambda_1, \ldots, \lambda_k$ respectivamente. Entonces

$$A\vec{x} \in U^{\perp}, \quad \forall\, \vec{x} \in U^{\perp},$$

donde $U = \operatorname{span}\{\vec{u}_1, \ldots, \vec{u}_k\}$.

Demostración:

i) Usando la definición de producto interior complejo dada en la Sección 5.1.1, tenemos que

$$< A\vec{x}, \vec{y} >= (A\vec{x})^t \overline{\vec{y}} = \vec{x}^t A^t \overline{\vec{y}} = \vec{x}^t A \overline{\vec{y}} = \vec{x}^t \overline{A\vec{y}} =< \vec{x}, A\vec{y} > .$$

ii) Usando el resultado de la parte (i) tenemos que

$$< A\vec{x}, \vec{x} >=< \vec{x}, A\vec{x} >= \overline{< A\vec{x}, \vec{x} >},$$

ésto es, $< A\vec{x}, \vec{x} >$ es real.

iii) De la ecuación $A\vec{x} = \lambda\vec{x}$, $\vec{x} \neq \vec{0}$, tenemos que

$$\lambda = \frac{< A\vec{x}, \vec{x} >}{\|\vec{x}\|^2} \in \mathbb{R},$$

donde usamos el resultado de la parte (ii).

iv) Sea $\vec{y} \in U^\perp$. Entonces $< \vec{y}, \vec{u}_i >= 0$ para todo $i = 1, \ldots, k$. Sea $\vec{x} \in U$ con $\vec{x} = \sum_{i=1}^k \alpha_i \vec{u}_i$. Es fácil ver ahora que

$$< A\vec{y}, \vec{x} >=< \vec{y}, A\vec{x} >= \sum_{i=1}^{k} \alpha_i \lambda_i < \vec{y}, \vec{u}_i >= 0.$$

Como $\vec{x} \in U$ es arbitrario, tenemos que $A\vec{y} \in U^\perp$ para todo $\vec{y} \in U^\perp$.

□

Usando estos resultados intermedios podemos ahora demostrar el siguiente teorema, el cual se conoce como el *teorema de los ejes principales*.

Teorema 6.23 (de los ejes principales)**.** *Sea A una matriz simétrica. Entonces los valores propios de A son reales y los vectores propios de A se pueden seleccionar de modo que formen una base ortonormal de $\mathbb{R}^n$. Si V es la matriz con los vectores propios de A como columnas, entonces V es ortogonal y $A = VDV^t$ donde D es la matriz diagonal que tiene los valores propios de A en la diagonal.*

Demostración: Sabemos que A tiene al menos un valor propio λ_1, que tiene que ser real por el Lema 6.22, con vector propio asociado $\vec{u}_1$ unitario. Para argumentar por inducción, suponga que tenemos $\lambda_1, \ldots, \lambda_k$ valores propios reales de A, $k \geq 1$, con vectores propios asociados $\vec{u}_1, \ldots, \vec{u}_k$ respectivamente, ortonormales. Sea $U = \operatorname{span}\{\vec{u}_1, \ldots, \vec{u}_k\}$. Entonces por el Lema 6.21 y el Lema 6.22, parte (iv), A tiene otro valor propio λ_{k+1}, que tiene que ser real, con vector propio asociado $\vec{u}_{k+1} \in U^\perp$ de norma uno. Como $\vec{u}_{k+1} \in U^\perp$, entonces $\vec{u}_1, \ldots, \vec{u}_k, \vec{u}_{k+1}$ son ortonormales. Podemos ahora repetir el proceso hasta que lleguemos a $k = n$. Tomando $V = [\vec{u}_1, \ldots, \vec{u}_n]$ y $D = \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}$, tenemos que V es ortogonal y es fácil verificar que $AV = VD$. □

Ejemplo 6.24. La matriz

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

es simétrica por lo que es diagonalizable (Teorema 6.23). En el Ejemplo 6.4 vimos que A tiene el valor propio $\lambda_1 = 0$ doble, con autovectores

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

y autovalor $\lambda_2 = 3$ simple con autovector $\vec{u}_3 = [1, 1, 1]^t$. Para construir la matriz ortogonal V del Teorema 6.23, es necesario construir una base ortonormal para $\mathbb{R}^3$ a partir de $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$. Note que $\vec{u}_3$ ya es perpendicular a $\{\vec{u}_1, \vec{u}_2\}$. Así que solo es necesario orto–normalizar $\{\vec{u}_1, \vec{u}_2\}$ usando el método de Gram–Schmidt, y normalizar a $\vec{u}_3$. El método de Gram–Schmidt aplicado a $\{\vec{u}_1, \vec{u}_2\}$ da el siguiente conjunto ortonormal:

$$\vec{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix},$$

Como $\left\|\vec{u}_3\right\| = \sqrt{3}$, tenemos entonces que $AV = VD$, $V^t V = I$, donde

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad V = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}.$$

□

Ejemplo 6.25. La matriz

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \\ 1 & 3 & 1 \end{bmatrix},$$

es simétrica. La instrucción de MATLAB:

```
[V,D]=eig(A);
```

genera dos matrices `V` y `D`, donde `D` es una matriz diagonal con los autovalores de A en la diagonal, y `V` es una matriz con los correspondientes autovectores como columnas. En este caso las instrucciones:

```
A=[1,2,1;2,1,3;1,3,1];
[V,D]=eig(A);
```

generan los siguientes resultados:

$$V = \begin{bmatrix} -0.2734 & 0.8425 & 0.4641 \\ 0.7431 & -0.1214 & 0.6581 \\ -0.6108 & -0.5248 & 0.5929 \end{bmatrix}, \quad D = \begin{bmatrix} -2.2019 & 0 & 0 \\ 0 & 0.0888 & 0 \\ 0 & 0 & 5.1131 \end{bmatrix}.$$

Los autovalores de A redondeados son -2.2019, 0.0888, 5.1131 y las columnas de `V` son los autovectores correspondientes. En MATLAB las instrucción `V'*A*V` y `V'*V` generan `D` y la matriz identidad respectivamente, a la precisión de la computadora. □

6.4 Forma canónica de Jordan

Una matriz A $n \times n$ se llama *defectuosa* si ésta no tiene un conjunto de n vectores propios linealmente independientes, o sea, si no es diagonalizable. Esto ocurre si para algún valor propio

$$\text{multiplicidad geométrica} < \text{multiplicidad algebraica}.$$

Ejemplo 6.26. La matriz

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

tiene autovalor $\lambda = 1$ de multiplicidad algebraica dos, y con

$$N(A - \lambda I) = N(A - I) = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\},$$

que tiene dimensión uno. Así que A es defectuosa por lo que no es diagonalizable. $\square$

En esta sección vamos a estudiar matrices defectuosas y veremos que es posible completar o extender el conjunto de vectores propios de la matriz hasta completar una base de $\mathbb{C}^n$ con propiedades más generales pero parecidas al caso de matrices no–defectuosas. Dos conceptos bien importantes para esta discusión lo son los *vectores propios generalizados* y las *cadenas de vectores propios generalizados*.

Sea λ un valor propio de la matriz A. En lo que resta de esta sección, vamos a denotar con m y p a las multiplicidades algebraica y geométrica respectivamente del valor propio λ. Sabemos que $p \leq m$ y que siempre $p \geq 1$. Recuerde que el valor propio λ es no–defectuoso si $p = m$, y es defectuoso cuando $p < m$.

Definición 6.27. Decimos que $\vec{v}$ es un *vector propio generalizado de rango r asociado al valor λ* para la matriz A si

$$(A - \lambda I)^r \vec{v} = \vec{0}, \quad (A - \lambda I)^{r-1} \vec{v} \neq \vec{0}. \tag{6.8}$$

Note que un vector propio generalizado de rango $r = 1$ es un vector propio. Veamos algunas propiedades de los vectores propios generalizados.

Proposición 6.28. *Sea $\vec{v}$ un vector propio generalizado de rango r asociado a λ para la matriz A. Entonces:*

i) *λ es valor propio de A;*

ii) *los vectores $\left\{\vec{v}, (A - \lambda I)\vec{v}, (A - \lambda I)^2 \vec{v}, \ldots, (A - \lambda I)^{r-1} \vec{v}\right\}$ son linealmente independientes.*

iii) *$(A - \lambda I)^k \vec{v}$ es vector propio generalizado de rango $r - k$ para $0 \leq k \leq r - 1$.*

Demostración: Para la primera parte basta observar que por la definición de vector propio generalizado tenemos que $\vec{u} = (A - \lambda I)^{r-1}\vec{v} \neq \vec{0}$ y que

$$(A - \lambda I)\vec{u} = (A - \lambda I)^r \vec{v} = \vec{0},$$

i.e., $\vec{u}$ es vector propio de A con valor propio λ.

Para la segunda parte, si los vectores

$$\left\{\vec{v}, (A - \lambda I)\vec{v}, (A - \lambda I)^2\vec{v}, \ldots, (A - \lambda I)^{r-1}\vec{v}\right\},$$

son dependientes, entonces para algún j con $0 \leq j \leq r - 1$, tenemos que

$$(A - \lambda I)^j \vec{v} = \sum_{k=0}^{j-1} c_k (A - \lambda I)^k \vec{v} + \sum_{k=j+1}^{r-1} c_k (A - \lambda I)^k \vec{v}$$

Si multiplicamos esta ecuación por $(A - \lambda I)^{r-j}$, obtenemos que

$$\vec{0} = \sum_{k=0}^{j-1} c_k (A - \lambda I)^{r-j+k} \vec{v}.$$

Si $j > 0$, multiplicamos esta expresión, en turnos, por $(A - \lambda I)^{j-1}, \ldots, (A - \lambda I)^1$, uno a uno en forma sucesiva, obtenemos que $c_0 = c_1 = \cdots c_{j-1} = 0$. Quedamos entonces con

$$(A - \lambda I)^j \vec{v} = \sum_{k=j+1}^{r-1} c_k (A - \lambda I)^k \vec{v}.$$

(Esta expresión es válida igualmente en el caso $j = 0$.) Ahora multiplicamos esta ecuación por $(A - \lambda I)^{r-j-1}$ obtenemos que

$$(A - \lambda I)^{r-1} \vec{v} = \sum_{k=j+1}^{r-1} c_k (A - \lambda I)^{r-j-1+k} \vec{v}.$$

Pero $r - j - 1 + k \geq r$ para $j + 1 \leq k \leq r - 1$ y como $(A - \lambda I)^p \vec{v} = \vec{0}$ para $p \geq r$, tenemos que la ecuación de arriba implica que $(A - \lambda I)^{r-1}\vec{v} = \vec{0}$ lo cuál es una contradicción. Por lo tanto los vectores son linealmente independientes.

Por último note que

$$\begin{aligned}(A-\lambda I)^{r-k}(A-\lambda I)^k\vec{v} &= (A-\lambda I)^r\vec{v}=\vec{0},\\ (A-\lambda I)^{r-k-1}(A-\lambda I)^k\vec{v} &= (A-\lambda I)^{r-1}\vec{v}\neq\vec{0},\end{aligned}$$

es decir, $(A-\lambda I)^k\vec{v}$ es vector propio generalizado de rango $r-k$. □

En general decimos que $\{\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k\}$ forman una *cadena de vectores propios generalizados de largo k asociada al valor propio λ* si

$$(A-\lambda I)\vec{v}_1 = \vec{0},\quad \vec{v}_1\neq\vec{0}, \tag{6.9a}$$

$$(A-\lambda I)\vec{v}_j = \vec{v}_{j-1},\quad 2\leq j\leq k. \tag{6.9b}$$

Es fácil ver que (6.9) implica que:

$$(A-\lambda I)^j\vec{v}_j=\vec{0},\quad (A-\lambda I)^{j-1}\vec{v}_j\neq\vec{0},\quad 1\leq j\leq k,$$

es decir, el j–esimo vector en la cadena, es un vector propio generalizado de rango j. Este resultado se puede usar para demostrar que los vectores en la cadena $\{\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k\}$ son linealmente independientes.

Ejemplo 6.29. Los vectores

$$\left\{(A-\lambda I)^{r-1}\vec{v},(A-\lambda I)^{r-2}\vec{v},\ldots,(A-\lambda I)\vec{v},\vec{v}\right\},$$

en la parte (ii) de la Proposición 6.28, forman una cadena de largo r de vectores propios generalizados. □

Un valor propio λ de una matriz A puede tener varias cadenas de vectores propios generalizados asociadas a éste, y estas cadenas pueden ser de largos diferentes. ¿Cómo podemos calcular estas cadenas de vectores generalizados? El siguiente concepto es esencial para poder contestar esta pregunta.

Definición 6.30. Sea A una matriz $n\times n$ con valor propio λ. Los *subespacios propios generalizados* asociados a λ están dados por:

$$E^i(\lambda)=N[(A-\lambda I)^i],\quad i=1,2,\ldots$$

Note que para todo i, tenemos que $E^i(\lambda) \subseteq \mathbb{C}^n$ y que

$$E^1(\lambda) \subseteq E^2(\lambda) \subseteq E^3(\lambda) \subseteq \cdots$$

De modo que existe un entero $k \geq 1$ tal que

$$E^1(\lambda) \subset E^2(\lambda) \subset \cdots \subset E^k, \quad E^r = E^k \text{ para todo } r > k.$$

El conjunto $E^k(\lambda)$ se llama el *subespacio propio generalizado máximo* (o subespacio máximo) y se denota por $M(\lambda)$.

Ejemplo 6.31. La matriz

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

tiene valores propios $\lambda = 1$ de multiplicidad algebraica dos, y $\lambda = 2$ de multiplicidad algebraica uno. Para $\lambda = 2$:

$$N(A - 2I) = \operatorname{span}\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \quad N(A - 2I) = N[(A - 2I)^2],$$

por lo que $M(2) = N(A - 2I)$ y $k = 1$. Para $\lambda = 1$:

$$N(A - I) = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}, \qquad N(A - I) \subset N[(A - I)^2],$$

$$\dim N[(A - I)^2] = 2, \qquad N[(A - I)^2] = N[(A - I)^3],$$

por lo que $M(1) = N[(A - I)^2]$ y $k = 2$. $\square$

Note que en el ejemplo anterior, la dimensión de $M(\lambda)$ coincide con la multiplicidad algebraica de λ. Esto es un caso particular del siguiente teorema.

Teorema 6.32. *Sea A una matriz $n \times n$ con autovalor λ y subespacio máximo asociado $M(\lambda)$. Entonces* $\dim M(\lambda)$ *es igual a la multiplicidad algebraica de λ.*

Vamos ahora a describir el proceso para calcular todas las cadenas de vectores propios generalizados asociadas al autovalor λ de la matriz A.

Este método genera una base de $M(\lambda)$ que consiste de la unión de todas las cadenas asociadas al autovalor λ. Partimos de la siguiente información:

$$E^1(\lambda) \subset E^2(\lambda) \subset \cdots \subset E^k = M(\lambda).$$

Sean

$$\begin{aligned} m_i &= \dim E^i(\lambda), \quad 1 \le i \le k, & (6.10a) \\ r_i &= m_i - m_{i-1}, \quad 2 \le i \le k. & (6.10b) \end{aligned}$$

Tomando $r_1 = m_1$, tenemos que $r_1 + r_2 + \cdots + r_k = m_k$. Vamos a utilizar repetidas veces el dato de que si U es un subespacio del espacio vectorial V, entonces existe otro subespacio W de V tal que $V = U \oplus W$ y $\dim W = \dim V - \dim U$. Usando ésto tenemos que existen vectores $\left\{\vec{v}_1, \ldots, \vec{v}_{r_k}\right\} \subset E^k(\lambda) \setminus E^{k-1}(\lambda)$ tal que

$$E^k(\lambda) = E^{k-1}(\lambda) \oplus \operatorname{span}\left\{\vec{v}_1, \ldots, \vec{v}_{r_k}\right\}.$$

Se puede verificar que los vectores $\left\{(A-\lambda I)\vec{v}_1, \ldots, (A-\lambda I)\vec{v}_{r_k}\right\}$ pertenecen a $E^{k-1}(\lambda) \setminus E^{k-2}(\lambda)$ y que son linealmente independientes. Esto implica que $r_k \le r_{k-1}$. Tenemos ahora que existen vectores $\left\{\vec{v}_{r_k+1}, \ldots, \vec{v}_{r_{k-1}}\right\} \subset E^{k-1}(\lambda) \setminus E^{k-2}(\lambda)$

$$\begin{aligned} E^{k-1}(\lambda) &= E^{k-2}(\lambda) \oplus \operatorname{span}\left\{(A-\lambda I)\vec{v}_1, \ldots, (A-\lambda I)\vec{v}_{r_k}\right\} \\ &\quad \oplus \operatorname{span}\left\{\vec{v}_{r_k+1}, \ldots, \vec{v}_{r_{k-1}}\right\}. \end{aligned}$$

Los vectores

$$\left\{(A-\lambda I)^2\vec{v}_1, \ldots, (A-\lambda I)^2\vec{v}_{r_k}, (A-\lambda I)\vec{v}_{r_k+1}, \ldots, (A-\lambda I)\vec{v}_{r_{k-1}}\right\},$$

pertenecen a $E^{k-2}(\lambda) \setminus E^{k-3}(\lambda)$ y son linealmente independientes. Esto implica a su vez que $r_{k-1} \le r_{k-2}$. Así que existen vectores $\left\{\vec{v}_{r_{k-1}+1}, \ldots, \vec{v}_{r_{k-2}}\right\} \subset E^{k-2}(\lambda) \setminus E^{k-3}(\lambda)$

$$\begin{aligned} E^{k-2}(\lambda) &= E^{k-3}(\lambda) \oplus \operatorname{span}\left\{(A-\lambda I)^2\vec{v}_1, \ldots, (A-\lambda I)^2\vec{v}_{r_k}\right\} \\ &\quad \oplus \operatorname{span}\left\{(A-\lambda I)\vec{v}_{r_k+1}, \ldots, (A-\lambda I)\vec{v}_{r_{k-1}}\right\} \end{aligned}$$

$$\oplus \operatorname{span}\left\{\vec{v}_{r_{k-1}+1},\dots,\vec{v}_{r_{k-2}}\right\}.$$

Esto proceso se continua hasta llegar a una representación similar para $E^1(\lambda)$. Si tomamos $E^0(\lambda) = \left\{\vec{0}\right\}$ y $r_{k+1} = 0$, podemos escribir este proceso por medio de la iteración:

$$E^j(\lambda) = E^{j-1}(\lambda) \bigoplus_{p=j}^{k} \operatorname{span}\left\{(A-\lambda I)^{p-j}\vec{v}_{r_{p+1}+1},\dots,(A-\lambda I)^{p-j}\vec{v}_{r_p}\right\}, \tag{6.11}$$

para $j = k, k-1, \dots, 2, 1$. Las cadenas de vectores propios generalizados asociados al valor propio λ se obtienen de este proceso y están dadas por:

$$\left\{(A-\lambda I)^{p-1}\vec{v}_j, (A-\lambda I)^{p-2}\vec{v}_j, \dots, \vec{v}_j\right\}, \tag{6.12}$$

para $r_{p+1} + 1 \leq j \leq r_p$, $1 \leq p \leq k$. Tenemos también el siguiente resultado cuya demostración se desprende del proceso por el cual obtuvimos (6.11).

Teorema 6.33. *Sea λ un valor propio de la matriz A con subespacio máximo $M(\lambda) = E^k(\lambda)$ y defina los $\{m_i\}$, $\{r_i\}$ según* (6.10) *con $r_1 = m_1$ y $r_{k+1} = 0$. Entonces $r_k \leq r_{k-1} \leq \cdots \leq r_1$ y la unión de todas las cadenas de vectores propios generalizados dadas por* (6.12) *forma una base para $M(\lambda)$.*

En el caso especial en que $m_1 = 1$, esto es, si el autovalor λ tiene multiplicidad geométrica uno, entonces $1 = r_1 = r_2 = \cdots = r_k$ por lo que todas las cadenas en (6.12) son vacías, excepto por la que corresponde a $p = k$ (recuerde que $r_{k+1} = 0$) que reduce a:

$$\left\{(A-\lambda I)^{k-1}\vec{v}_1, (A-\lambda I)^{k-2}\vec{v}_1, \dots, \vec{v}_1\right\}.$$

Si definimos $\vec{w}_1 = (A-\lambda I)^{k-1}\vec{v}_1$, entonces esta cadena se puede escribir como $\{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_k\}$ (cf. (6.9)) donde

$$\begin{aligned}(A-\lambda I)\vec{w}_1 &= \vec{0}, \quad \vec{w}_1 \neq \vec{0},\\ (A-\lambda I)\vec{w}_j &= \vec{w}_{j-1}, \quad 2 \leq j \leq k.\end{aligned}$$

Ejemplo 6.34. La matriz

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

tiene autovalores:

$$\begin{array}{rcl} \lambda & = & 1, \quad \dim N(A-I)=1, \quad M(1)=E^3(1), \quad \dim M(1)=3, \\ \lambda & = & 2, \quad \dim N(A-2I)=2, \quad M(2)=E^2(2), \quad \dim M(2)=3. \end{array}$$

De ésto podemos concluir que 1 tiene solo una cadena de vectores propios generalizados de largo tres. En este caso podemos generar la única cadena asociada al autovalor $\lambda = 1$ resolviendo los sistemas:

$$(A-I)\vec{w}_1 = \vec{0}, \quad (A-I)\vec{w}_2 = \vec{w}_1, \quad (A-I)\vec{w}_3 = \vec{w}_2.$$

Tenemos entonces[2] que:

$$\vec{w}_1 = [0,1,0,0,0,0]^t, \quad \vec{w}_2 = [1,0,0,0,0,1]^t, \quad \vec{w}_3 = [1,0,0,0,0,2]^t.$$

De la información sobre el valor propio 2, podemos concluir que éste tiene una cadena de largo dos y otra de largo uno. Tenemos que $k = 2$, $r_1 = 2$ y que $r_2 = 1$, por lo que (6.11) reduce en este caso a:

$$\begin{array}{rcl} E^2(2) & = & E^1(2) \oplus \operatorname{span}\{\vec{v}_1\}, \\ E^1(2) & = & \operatorname{span}\{\vec{v}_2\} \oplus \operatorname{span}\{(A-2I)\vec{v}_1\}. \end{array}$$

Se puede verificar que con

$$\vec{v}_1 = [0,0,0,1,0,0]^t, \quad \vec{v}_2 = [0,0,0,0,1,0]^t,$$

se cumplen estas dos condiciones[3]. Las dos cadenas de vectores propios generalizados asociadas a $\lambda = 2$ son:

$$\begin{array}{rcl} \{(A-2I)\vec{v}_1, \vec{v}_1\} & = & \Big\{[0,-1,1,0,0,0]^t, [0,0,0,1,0,0]^t\Big\}, \\ \{\vec{v}_2\} & = & \Big\{[0,0,0,0,1,0]^t\Big\}. \end{array}$$

□

Vamos a resumir el proceso para determinar las cadenas de vectores propios generalizados para una matriz A con autovalor defectuoso λ:

[2] Esto lo puede verificar haciendo el calculo a mano o usando algún paquete computacional, preferiblemente que haga álgebra simbólica.

[3] Nuevamente, éstos vectores se calculan a mano o utilizando algún paquete de computación. Note que los vectores $\{\vec{v}_1, \vec{v}_2\}$ se calculan en el orden que indican sus índices, esto es, $\vec{v}_1$ se calcula primero y luego $\vec{v}_2$.

i) El numero de cadenas de autovectores generalizados es igual al numero de autovectores asociados a λ, es decir la multiplicidad geométrica de λ.

ii) La suma de los largos de todas las cadenas asociadas con λ es igual a la multiplicidad algebraica de λ, o lo mismo $\dim M(\lambda)$.

iii) Para un autovector cualquiera $\vec{v}$ asociado a λ, la cadena asociada a este autovector es $\left\{\vec{v}, \vec{w}_2, \ldots, \vec{w}_q\right\}$ donde con $\vec{w}_1 = \vec{v}$, los $\vec{w}$'s se obtienen a partir de la iteración:

$$(A - \lambda I)\vec{w}_j = \vec{w}_{j-1}, \quad 2 \leq j \leq q.$$

iv) Si las dimensiones de todos los subespacios $E^1(\lambda) \subset E^2(\lambda) \subset \cdots E^k(\lambda) = M(\lambda)$ son conocidas, es posible determinar los largos específicos de las cadenas de autovectores generalizados. De lo contrario, hay que considerar todas las posibles combinaciones de largos de cadenas que sumen a la multiplicidad algebraica de λ.

Ejemplo 6.35. Sea A una matriz 5×5 con autovalores λ_1, λ_2 con las siguientes propiedades:

autovalor	mult. geom.	mult. alg.	autovectores
λ_1	1	1	$\vec{v}_1$
λ_2	2	4	$\vec{v}_2, \vec{v}_3$

Como λ_1 es no defectuoso, éste tiene como única cadena asociada $\{\vec{v}_1\}$. Para λ_2, como $\dim N(A - \lambda_2 I) = 2$, este valor propio tiene asociado dos cadenas de vectores propios generalizados. Como los largos de estas cadenas tienen que sumar a $\dim M(\lambda_2) = 4$, tenemos dos posibilidades (los $\vec{w}$'s son vectores propios generalizados):

i) dos cadenas de largo dos, digamos $\{\vec{v}_2, \vec{w}_2\}$ y $\{\vec{v}_3, \vec{w}_3\}$;

ii) una cadena de largo uno y otra de largo tres, digamos $\{\vec{v}_2\}$ y $\{\vec{v}_3, \vec{w}_2, \vec{w}_3\}$.

De la discusión que precede al Teorema 6.33 es fácil ver que estas dos posibilidades corresponden respectivamente a los casos $E^1(\lambda_2) \subset E^2(\lambda_2) = M(\lambda_2)$ ó $E^1(\lambda_2) \subset E^2(\lambda_2) \subset E^3(\lambda_2) = M(\lambda_2)$. □

Es posible que el sistema $(A-\lambda I)\vec{w}=\vec{v}$, donde $\vec{v}$ es un autovector de A asociado a λ, sea inconsistente. En tal caso, si $N(A-\lambda I)=\operatorname{span}\left\{\vec{v}_1,\ldots,\vec{v}_p\right\}$, entonces trabajamos con el sistema

$$(A-\lambda I)\vec{w}=\sum_{i=1}^{p}\alpha_i\vec{v}_i,$$

donde las α's son variables adicionales a los componentes de $\vec{w}$. Este sistema es subdeterminado y consistente por lo que tiene soluciones con $\vec{w}\neq\vec{0}$. Si al menos uno de los α's es distinto de cero, entonces con este $\vec{w}$ se continua la iteración en la parte (iii) del proceso descrito anteriormente.

Ejemplo 6.36. Considere la matriz

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

Es fácil ver que $\lambda=1$ es valor propio de A de multiplicidad algebraica tres, y que

$$\vec{v}_1=\begin{bmatrix}1\\1\\0\end{bmatrix},\quad \vec{v}_2=\begin{bmatrix}1\\0\\1\end{bmatrix},$$

son vectores propios asociados que generan el espacio propio asociado. Se puede ver ahora que ambos sistemas

$$(A-I)\vec{w}=\vec{v}_1,\quad (A-I)\vec{w}=\vec{v}_2,$$

son inconsistentes. Trabajando como se describe en el párrafo anterior, consideramos el sistema $(A-I)\vec{w}=\alpha_1\vec{v}_1+\alpha_2\vec{v}_2$. La matriz aumentada de este sistema es:

$$\left[\begin{array}{rrrrr|r} 1 & -1 & -1 & -1 & -1 & 0 \\ 2 & -2 & -2 & -1 & 0 & 0 \\ -1 & 1 & 1 & 0 & -1 & 0 \end{array}\right],$$

que tiene forma echelon reducida

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

Con $\vec{w} = [a,b,c]^t$ tenemos entonces que

$$\alpha_1 + 2\alpha_2 = 0, \quad a - b - c = \alpha_1 + \alpha_2.$$

Tomando $\alpha_1 = 2$, $\alpha_2 = -1$, $b = c = 0$, y $a = 1$, tenemos $\vec{v} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 = [1,2,-1]^t$ (vector propio) y $\vec{w} = [1,0,0]^t$. Tenemos entonces $\{\vec{v},\vec{w}\}$ como una cadena de largo dos y $\{\vec{v}_1\}$ como una cadena de largo uno. (Pudimos haber usado aquí igualmente a $\vec{v}_2$ ya que tanto $\vec{v}_1$ como $\vec{v}_2$ son linealmente independientes a $\vec{v}$.) □

Estamos ya esencialmente listos para enunciar el teorema sobre la forma canónica de Jordan. Solo necesitamos establecer una relación adicional. Note que (6.9) se puede escribir como

$$A\vec{v}_1 = \lambda\vec{v}_1, \tag{6.13a}$$
$$A\vec{v}_j = \lambda\vec{v}_j + \vec{v}_{j-1}, \quad 2 \le j \le k. \tag{6.13b}$$

Sea $S = \{\vec{v}_1,\ldots,\vec{v}_k\}$ y $V = \operatorname{span} S$. Tenemos entonces que (6.13) implica que $AV \subseteq V$. Defina la transformación lineal $L : V \longrightarrow V$ por $L(v) = Av$. Si $\{\vec{e}_1,\ldots,\vec{e}_k\}$ es la base estándar de $\mathbb{R}^k$, entonces de (6.13) tenemos que

$$\begin{aligned}(L(\vec{v}_1))_S &= \lambda\vec{e}_1,\\ (L(\vec{v}_j))_S &= \lambda\vec{e}_j + \vec{e}_{j-1}, \quad 2 \le j \le k.\end{aligned}$$

Así que la representación matricial de L con respecto a la base S de V está dada por la matriz $k \times k$:

$$J = \begin{bmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 & 0\\ 0 & \lambda & 1 & 0 & \cdots & 0 & 0\\ 0 & 0 & \lambda & 1 & \cdot & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & \lambda & 1\\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}. \tag{6.14}$$

Note que J tiene todas las entradas cero excepto por la diagonal principal que es toda igual a λ, y con la super–diagonal toda igual a uno. La matriz J se llama un *bloque de Jordan de largo k*. Se puede verificar que un bloque de Jordan de largo k a su vez genera o tiene asociada una cadena de vectores propios generalizados de largo k. Tenemos entonces por el Teorema

6.33, que cada autovalor de la matriz A tiene tantas cadenas de vectores propios generalizados (o bloques de Jordan) asociados como su multiplicidad geométrica, y que los largos de estas cadenas (o bloques de Jordan) suman a la multiplicidad algebraica del valor propio. Estas observaciones esencialmente demuestran el siguiente resultado.

Teorema 6.37 (Forma Canónica de Jordan)**.** *Sea A una matriz $n \times n$. Entonces existe una matriz nosingular T tal que $T^{-1}AT = J$ donde J está particionada como:*

$$J = \begin{bmatrix} J_1 & O & O & \cdots & O & O \\ O & J_2 & O & \cdots & O & O \\ O & O & J_3 & \cdots & O & O \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & O & \cdots & J_{r-1} & O \\ O & O & O & \cdots & O & J_r \end{bmatrix}, \tag{6.15}$$

y $J_1, J_2, \ldots, J_r$ son bloques de Jordan. Las columnas de la matriz T se obtienen tomando la unión de todas las cadenas de vectores propios generalizados asociadas a todos los autovalores de A.

Ejemplo 6.38. En el Ejemplo 6.34 vimos que la matriz

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

tiene autovalores $\lambda = 1, 2$ con las siguientes cadenas de vectores propios generalizados:

$$\begin{aligned} \text{para } \lambda &= 1, \quad \left\{[0,1,0,0,0,0]^t, [1,0,0,0,0,1]^t, [1,0,0,0,0,2]^t\right\}, \\ \text{para } \lambda &= 2, \quad \left\{[0,-1,1,0,0,0]^t, [0,0,0,1,0,0]^t\right\}, \quad \left\{[0,0,0,0,1,0]^t\right\}. \end{aligned}$$

Tenemos entonces que $\lambda = 1$ tiene como bloque de Jordan asociado:

$$J_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

y $\lambda = 2$ tiene asociado los bloques de Jordan:

$$J_2 = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \quad J_3 = \begin{bmatrix} 2 \end{bmatrix}.$$

Tenemos ahora que $T^{-1}AT = J$ donde

$$\begin{aligned} J &= \begin{bmatrix} J_1 & O & O \\ O & J_2 & O \\ O & O & J_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}, \\ T &= \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

□

6.5 El exponencial de una matriz

Para cualquier matriz A de tamaño $n \times n$, el *exponencial de A* se define por la serie infinita:

$$\mathrm{e}^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n. \tag{6.16}$$

Usando una prueba de comparación se puede verificar que esta serie es convergente para cualquier matriz A.

Si $A = TJT^{-1}$ es la forma canónica de Jordan de A, entonces

$$\mathrm{e}^A = \sum_{n=0}^{\infty} \frac{1}{n!} (TJT^{-1})^n = \sum_{n=0}^{\infty} \frac{1}{n!} TJ^nT^{-1} = T\mathrm{e}^J T^{-1}.$$

Como J es una matriz particionada diagonalmente en bloques, utilizando la notación (6.15), tenemos que

$$J^n = \begin{bmatrix} J_1^n & O & O & \cdots & O & O \\ O & J_2^n & O & \cdots & O & O \\ O & O & J_3^n & \cdots & O & O \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & O & \cdots & J_{r-1}^n & O \\ O & O & O & \cdots & O & J_r^n \end{bmatrix}.$$

Usando esto en el calculo anterior de e^A, tenemos que

$$\mathrm{e}^A = T \begin{bmatrix} \mathrm{e}^{J_1} & O & O & \cdots & O & O \\ O & \mathrm{e}^{J_2} & O & \cdots & O & O \\ O & O & \mathrm{e}^{J_3} & \cdots & O & O \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & O & \cdots & \mathrm{e}^{J_{r-1}} & O \\ O & O & O & \cdots & O & \mathrm{e}^{J_r} \end{bmatrix} T^{-1}.$$

Así que dada la forma canónica de Jordan de A, el problema de calcular e^A se reduce a calcular e^J para cualquier bloque de Jordan J.

Sea J un bloque de Jordan de tamaño $k \times k$ asociado a un valor propio λ de A. Si $I = [\vec{e}_1, \ldots, \vec{e}_k]$ es la matriz identidad del mismo tamaño que J, entonces $J = \lambda I + N$ donde $N = [\vec{0}, \vec{e}_1, \ldots, \vec{e}_{k-1}]$. Note que para cualquier matriz B, el resultado de BN es una matriz con la primera columna toda cero y con las primeras $k-1$ columnas de B en las columnas dos a la k del resultado. Usando esto es fácil ver ahora que $N^k = O$. Como $NI = IN$, tenemos que para $n \geq 0$:

$$J^n = (\lambda I + N)^n = \sum_{j=0}^{n} \binom{n}{j} \lambda^{n-j} N^j.$$

Usando esto y que $N^k = O$, tenemos que para cualquier $\alpha \in \mathbb{R}$:

$$\begin{aligned} \mathrm{e}^{\alpha J} &= \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} J^n = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} \sum_{j=0}^{n} \binom{n}{j} \lambda^{n-j} N^j, \\ &= \sum_{n=0}^{k-1} \frac{\alpha^n}{n!} \sum_{j=0}^{n} \binom{n}{j} \lambda^{n-j} N^j + \sum_{n=k}^{\infty} \frac{\alpha^n}{n!} \sum_{j=0}^{k-1} \binom{n}{j} \lambda^{n-j} N^j. \end{aligned}$$

Pero
$$\sum_{n=0}^{k-1}\frac{\alpha^n}{n!}\sum_{j=0}^{n}\binom{n}{j}\lambda^{n-j}N^j=\sum_{j=0}^{k-1}\left[\sum_{n=j}^{k-1}\frac{\alpha^n\lambda^{n-j}}{(n-j)!}\right]\frac{1}{j!}N^j.$$
Además
$$\sum_{n=k}^{\infty}\frac{\alpha^n}{n!}\sum_{j=0}^{k-1}\binom{n}{j}\lambda^{n-j}N^j=\sum_{j=0}^{k-1}\left[\sum_{n=k}^{\infty}\frac{\alpha^n\lambda^{n-j}}{(n-j)!}\right]\frac{1}{j!}N^j.$$
Combinando estos dos resultados con el calculo de arriba para $e^{\alpha J}$ llegamos a que

$$e^{\alpha J}=\sum_{j=0}^{k-1}\left[\sum_{n=j}^{\infty}\frac{\alpha^n\lambda^{n-j}}{(n-j)!}\right]\frac{1}{j!}N^j=\sum_{j=0}^{k-1}\left[\sum_{n=0}^{\infty}\frac{\alpha^n\lambda^n}{n!}\right]\frac{\alpha^j}{j!}N^j=e^{\alpha\lambda}\sum_{j=0}^{k-1}\frac{\alpha^j}{j!}N^j.$$

Note que
$$\sum_{j=0}^{k-1}\frac{\alpha^j}{j!}N^j=\begin{bmatrix}1&\alpha&\frac{\alpha^2}{2!}&\cdots&\frac{\alpha^{k-2}}{(k-2)!}&\frac{\alpha^{k-1}}{(k-1)!}\\0&1&\alpha&\cdots&\frac{\alpha^{k-3}}{(k-3)!}&\frac{\alpha^{k-2}}{(k-2)!}\\0&0&1&\cdots&\frac{\alpha^{k-4}}{(k-4)!}&\frac{\alpha^{k-3}}{(k-3)!}\\\vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\0&0&0&\cdots&1&\alpha\\0&0&0&\cdots&0&1\end{bmatrix}.$$

Ejemplo 6.39. Calculamos e^A donde

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

Note que ya A está en la forma canónica de Jordan:

$$A=\begin{bmatrix}J_1&\vec{0}\\\vec{0}^t&J_2\end{bmatrix},$$

donde $\vec{0}=[0,0,0]^t$ y
$$J_1=\begin{bmatrix}2&1&0\\0&2&1\\0&0&2\end{bmatrix},\quad J_2=[1].$$

Usando las formulas para $e^{\alpha J}$ donde J es un bloque de Jordan y tomando $\alpha = 1$, tenemos que

$$e^{J_1} = e^2 \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad e^{J_2} = [e].$$

Combinando estos resultados obtenemos que:

$$e^A = \begin{bmatrix} e^{J_1} & \vec{0} \\ \vec{0}^t & e^{J_2} \end{bmatrix} = \begin{bmatrix} e^2 & e^2 & \frac{e^2}{2} & 0 \\ 0 & e^2 & e^2 & 0 \\ 0 & 0 & e^2 & 0 \\ 0 & 0 & 0 & e \end{bmatrix}.$$

□

6.6 Ejercicios

Ejercicio 6.1. La matriz B se obtiene a partir de A por medio de una transformación elemental de fila, donde

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

¿Cómo comparan los valores propios de A y B?

Ejercicio 6.2. Halle los autovalores y autovectores para las siguientes matrices. Para cada autovalor λ, encuentre $N(A - \lambda I)$. (Escriba su respuesta a esto último utilizando la notación de span.)

$$\begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 9 \\ 6 & -4 \end{bmatrix}.$$

Ejercicio 6.3. Sea A una matriz 3×3 con autovalores $\lambda_1 = -2, \lambda_2 = 1, \lambda_3 = -3$.

a) Halle el polinomio caracterstico de A.

b) Calcule $\operatorname{tr} A$ y $\det A$.

c) Calcule $\det(A+2I)$.

d) ¿ Cuáles serían los autovalores de A^2?

Ejercicio 6.4. La matriz

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

tiene los autovalores y autovectores correspondientes:

$$\lambda_1 = -3, \quad \vec{v}_1 = (1,-1)^t,$$

$$\lambda_2 = 1, \quad \vec{v}_2 = (1,1)^t.$$

Usando esta información, escriba la solución general del sistema:

$$y_1'(t) = -y_1(t) + 2y_2(t), \quad y_2'(t) = 2y_1(t) - y_2(t).$$

Encuentre la solución del sistema que cumple la condición de que $y_1(0) = 1$ y que $y_2(0) = -2$.

Ejercicio 6.5. Sea B una matriz 2×2 que tiene autovalor $\lambda = 4 - 5i$ con vector propio $\vec{v} = (1, 2i)^t$. Considere el sistema de ecuaciones diferenciales $\vec{y}'(t) = B\vec{y}(t)$.

a) Encuentre la solución general de dicho sistema.

b) Calcule la solución particular del sistema que cumple con la condición de que $\vec{y}(0) = (-3,2)^t$.

Ejercicio 6.6. Sea A una matriz 3×3 que tiene autovalores $\lambda_1 = -2$ y $\lambda_2 = 2 - 3i$, ambos simples (de multiplicidad uno). El autovalor λ_1 tiene autovector asociado $\vec{v}_1 = (-2,1,3)^t$ mientras que λ_2 tiene autovector $\vec{v}_2 = (1, i, 2-i)^t$. Encuentre la solución general del sistema $\vec{y}'(t) = A\vec{y}(t)$ donde $\vec{y}(t) = (y_1(t), y_2(t), y_3(t))^t$.

Ejercicio 6.7. Verifique que las siguientes matrices son diagonalizables. En cada caso encuentre una matriz nosingular V y otra matriz diagonal D tal que $AV = VD$. Cuando V se pueda tomar ortogonal, constrúyala de modo que así lo sea.

a) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

b) $\begin{bmatrix} 5 & 6 \\ -2 & -2 \end{bmatrix}$

c) $\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$

Ejercicio 6.8. Sea A una matriz 3×3 con autovalores $\lambda = -1$ de multiplicidad dos, y $\lambda = 5$ de multiplicidad uno (simple). Suponga que:

$$N(A+I) = \operatorname{span}\left\{(-1,2,1)^t\right\}, \quad N(A-5I) = \operatorname{span}\left\{(3,1,5)^t\right\}.$$

Determine si A es diagonalizable o no. Explique.

Ejercicio 6.9. Verifique que para cualesquiera $a, b \in \mathbb{R}$, la matriz

$$A = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & b \end{bmatrix};$$

no es diagonalizable.

Ejercicio 6.10. Para las siguientes matrices, calcule la forma canónica de Jordan de la matriz. Esto es, en cada caso, busque las matrices J y T del Teorema 6.37 tal que $AT = TJ$.

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

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

Ejercicio 6.11. Suponga que una matriz A de tamaño 6×6 tiene valores propios $\lambda_1 = 2$ y $\lambda_2 = -5$ de multiplicidades algebraicas 2 y 4 respectivamente, y con $\dim N(A - \lambda_1 I) = 1$, $\dim N(A - \lambda_2 I) = 2$. Sean $\{\vec{v}_1\}$, $\{\vec{v}_2, \vec{v}_3\}$ vectores propios linealmente independientes correspondientes a λ_1 y λ_2. Describa cuales son las diferentes posibilidades para las cadenas de vectores propios generalizados correspondientes a λ_1 y λ_2.

Ejercicio 6.12. Calcule el exponencial para cada una de las matrices en el Ejercicio 6.10.

Ejercicio 6.13. Verifique que los valores propios de una matriz triangular son los elementos de la diagonal de la matriz.

Ejercicio 6.14. Sea A una matriz simétrica 2×2. Verifique que los valores propios de A son iguales si y solo si $A = aI$ donde I es la matriz identidad 2×2

Ejercicio 6.15. Sea A una matriz simetrica $n \times n$ con $A\vec{x} = \lambda\vec{x}$ y $A\vec{y} = \mu\vec{y}$, donde $\lambda \neq \mu$ y $\vec{x}$, $\vec{y}$ son autovectores de A. Verifique que $\langle \vec{x}, \vec{y} \rangle = 0$.

Ejercicio 6.16. Suponga que $A = T^{-1}BT$ y que $\vec{x}$ es vector propio de A correspondiente al valor propio λ. Verifique que $T\vec{x}$ es vector propio de B correspondiente al valor propio λ.

Ejercicio 6.17. Verifique que para cualquier matriz A de tamaño $n \times n$, A y A^t tienen los mismos valores propios.

Ejercicio 6.18. Sea $\vec{u}$ un vector unitario en $\mathbb{R}^n$. Defina la matriz $A = \vec{u}\vec{u}^t$. Verifique A tiene valor propio simple (de multiplicidad algebraica uno) $\lambda = 1$ con vector propio asociado $\vec{u}$, y que $\lambda = 0$ es valor propio de multiplicidad algebraica $n-1$ con vectores propios asociados $\vec{u}_1, \ldots, \vec{u}_{n-1}$ donde

$$\operatorname{span}\{\vec{u}_1, \ldots, \vec{u}_{n-1}\} = \operatorname{span}\{\vec{u}\}^{\perp}.$$

Ejercicio 6.19. Sea $\{\vec{u}_1, \ldots, \vec{u}_n\}$ una base ortonormal de $\mathbb{R}^n$. Para $c_1, \ldots, c_n \in \mathbb{R}$, defina la matriz A por:

$$A = c_1\vec{u}_1\vec{u}_1^t + \cdots c_n\vec{u}_n\vec{u}_n^t.$$

Verifique que A tiene los autovalores $c_1, \ldots, c_n$ con los vectores propios correspondientes $\vec{u}_1, \ldots, \vec{u}_n$.

Ejercicio 6.20. Suponga que A es similar a B, esto es, existe T nosingular tal que $A = TBT^{-1}$.

a) Verifique que para cualquier entero $n \geq 1$, tenemos que $A^n = TB^nT^{-1}$, es decir A^n es similar a B^n.

b) Si A es nosingular, verifique que B es nosingular y que A^{-1} es similar a B^{-1}.

Ejercicio 6.21. Sea A una matriz diagonalizable con 1 ó -1 como sus únicos valores propios. Verifique que $A^{-1} = A$.

Referencias

[1] Anton, H., Elementary Linear Algebra, 9th, John Wiley & Sons, New York, 2004.

[2] Anton, H. and Rorres, C., Elementary Linear Algebra, Applications Version, 9th ed., John Wiley and Sons, New York, 2004.

[3] Grossman, S. I., Aplicaciones de Algebra Lineal, Grupo Editorial Iberoamericana, Mexico, 1988.

[4] Leon, S. J., Linear Algebra with Applications, 5th ed., PrenticeHall, New Jersey, 1998.

[5] Marcus, M., Matrices and MATLAB: A Tutorial, Prentice-Hall, New York, 1998.

[6] Noble, B. and Daniel, J. W., Applied Linear Algebra, 3rd ed., Prentice-Hall, New York, 1993.

[7] Poole, D., Linear Algebra: A Modern Introduction, 3rd ed., Brooks/Cole Cengage Learning Inc., 2011.

[8] Strang, G., Linear Algebra and its Applications, 4th ed., Brooks Cole, 2005.

[9] Strang, G., Introduction to Linear Algebra, Wellesley Cambridge Press; 4th ed., 2009.

Índice