Álgebra Lineal

Fernando Barrera Mora

GRUPO EDITORIAL PATRIA

ÁLGEBRA LINEAL

GRUPO EDITORIAL PATRIA

ÁLGEBRA LINEAL

Fernando Barrera Mora

PRIMERA EDICIÓN EBOOK
MÉXICO, 2014

GRUPO EDITORIAL PATRIA

Para establecer comunicación con nosotros puede hacerlo por:

correo:
Renacimiento 180, Col. San Juan Tlihuaca, Azcapotzalco, 02400, México, D.F.

fax pedidos:
(01 55) 5354 9109 • 5354 9102

e-mail:
info@editorialpatria.com.mx

home page:
www.editorialpatria.com.mx

Dirección editorial: Ing. Javier Enrique Callejas
Coordinadora editorial: Ing. Estela Delfín Ramírez

Revisión técnica:
Maestro Rogelio Herrera Aguirre
Departamento de Ciencias Básicas
Universidad Autónoma Metropolitana-Azcapotzalco

Diseño de interiores: EG Corporación de Servicios Gráficos
Diseño de portada: Publishare
Ilustraciones: EG Corporación de Servicios Gráficos

Álgebra Lineal

Renacimiento 180, Colonia San Juan Tlihuaca
Delegación Azcapotzalco, Código Postal 02400, México, D.F.

Miembro de la Cámara Nacional de la Industria Editorial Mexicana
Registro Núm. 43

ISBN: 978-607-438-892-3

Impreso en México
Printed in Mexico

Primera edición ebook: 2014

Este trabajo está dedicado a la memoria de mi señor padre, Antonio Barrera Pelcastre (1903-2005), de quien aprendí que el trabajo hace la diferencia entre los individuos y a la vez los hermana.

Prólogo

Uno de los temas de matemáticas más populares del que se han escrito innumerables textos, es el álgebra lineal. Esto no es ninguna casualidad. El álgebra lineal aparece de manera natural en prácticamente todas las disciplinas, tanto de matemáticas como de otras ciencias, inclusive en las ciencias sociales y humanidades, teniendo presencia significativa en las áreas de ingeniería y no digamos la física.

Desde nuestros primeros estudios, digamos a nivel secundaria, el álgebra lineal se estudia, aunque no se use con ese nombre, para la resolución de sistemas de ecuaciones lineales.

Debido a lo anterior, cuando un nuevo texto relacionado con el álgebra lineal aparece en el mercado, uno se pregunta qué puede aportar que no haya sido presentado y estudiado hasta la saciedad en algunos de los innumerables textos que ya existen.

El libro de Fernando Barrera Mora está dedicado al álgebra lineal para un primer curso de licenciatura en matemáticas, ingeniería y áreas afines. El primer punto que me gustaría hacer notar es que el libro privilegia el empezar con problemas concretos que se nos presentan tanto en nuestra vida cotidiana, como en economía, empresas productivas, etcétera. A partir de estos problemas concretos se empieza a elaborar sobre los ingredientes presentes, que facilitan la visualización del estudiante sobre estos componentes cuando son planteados de manera general.

Asimismo, estos problemas concretos que se estudian sirven para establecer tanto los métodos como la teoría necesaria, ya sea para resolverlos, estudiarlos o ubicarlos en un contexto más general.

Un punto de vista valioso a resaltar en este trabajo es el tratamiento que se hace de lo que podríamos llamar la "teoría propia", esto es, la teoría que trata sobre los valores y los vectores propios.

Lo más común para abordar la solución y estudio de los valores y vectores propios es el estudio de la matriz característica, es decir, encontrar los valores para los cuales esta matriz es singular, lo cual nos lleva inmediatamente al cálculo del determinante y por tanto al polinomio característico.

Aunque el análisis de los determinantes es indispensable en el estudio del álgebra lineal, el presentar un estudio exhaustivo de sus propiedades básicas es un problema ya sea laborioso o poco claro, dependiendo del enfoque que seleccionemos.

En este trabajo se selecciona un camino diferente. Se hace énfasis en propiedades inherentes a la matriz que dan origen al problema en estudio. Más precisamente, se estudia el operador asociado a la matriz, respecto a otra base seleccionada adecuadamente, además de la ventaja natural que se tiene al estudiar, de manera intrínseca, al operador. De este modo, se tiene que se hace una presentación sin ninguna necesidad de hacer referencia a los determinantes.

Hay varias otras novedades que diferencian este texto de otros. Por ejemplo, en este trabajo se hace interactuar el álgebra lineal con la geometría analítica; se introducen y se trabajan subespacios sin haber siquiera definido formalmente lo que es un espacio vectorial; hay varias demostraciones novedosas o poco conocidas como por ejemplo la de la existencia del operador adjunto o que cualquier sistema linealmente independiente tiene cardinalidad menor o igual a la cardinalidad de un conjunto de generadores; se construye la base teórica necesaria a partir del espacio dos dimensional, se pasa al tres dimensional y finalmente a cualquier espacio finito dimensional; se presenta un algoritmo para el cálculo del polinomio mínimo de una matriz.

Otros aspectos dignos de mencionar son la forma en que se motiva el producto de matrices, el cual se deriva a partir de un ejemplo concreto sobre producción y que se encuentran varios ejercicios ya sea originales o poco comunes en otros textos.

Un punto final que es necesario enfatizar es que, como se mencionó al principio, el álgebra lineal es de mucha importancia en todo currículum de ciencias y de ingeniería e inclusive de otras áreas. Esta importancia se encuentra en la mente del autor a lo largo de este libro, lo cual se puede percibir por la concepción del álgebra lineal que se presenta durante todo el tratado.

Gabriel D. Villa Salvador,
Departamento de Control Automático,
CINVESTAV del IPN.,
México, D. F.,
Julio de 2007.

Índice general

Introducción ix
Nomenclatura xi

1. Sistemas de ecuaciones lineales 1

1.1. Ejemplos 1
1.2. Sistemas de ecuaciones lineales y su representación geométrica 9
1.3. Conceptos fundamentales y método de reducción de Gauss-Jordan 12
1.3.1. Análisis de las soluciones de un sistema de ecuaciones lineales 19
1.3.2. Ejemplos con *Maple* 25
1.4. Ejercicios 27

2. Matrices 31

2.1. Operaciones con matrices 31
2.1.1. Suma de matrices 32
2.1.2. Producto de matrices 33
2.1.3. Propiedades de la suma y producto de matrices 37
2.2. Matrices elementales e inversas 39
2.2.1. Cálculo de la inversa de una matriz 42
2.3. Aplicaciones 46
2.4. Matrices enteras 53
2.5. Ejercicios 55

3. Espacios vectoriales 61

3.1. Vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$ 61
3.2. Combinaciones lineales y dependencia lineal 64
3.2.1. Ejercicios 68
3.3. Aspectos geométricos de $\mathbb{R}^2$ y $\mathbb{R}^3$ vía álgebra lineal 68
3.3.1. Norma y producto interno 70
3.3.2. Proyección ortogonal de un vector sobre otro 73
3.3.3. Producto cruz de vectores 77
3.3.4. Ecuación de un plano 78
3.3.5. Ejercicios 79
3.4. El espacio vectorial $\mathbb{R}^n$ 81
3.4.1. Subespacios 82
3.4.2. Operaciones con subespacios 86
3.5. Espacios vectoriales generales 88
3.6. Ejercicios 92

4. Transformaciones lineales y matrices 95

4.1. Definiciones y resultados básicos 95
4.2. Transformaciones lineales geométricas 99

4.3. Rango y núcleo de una transformación lineal 103
4.4. Matrices y transformaciones lineales 104
4.4.1. Matrices de cambio de base 107
4.4.2. El espacio de las transformaciones lineales 112
4.5. Ejercicios 113

5. Determinantes 117

5.1. Determinantes y volúmenes de paralelepípedos 117
5.1.1. Propiedades del determinante 124
5.1.2. Existencia y unicidad del determinante 125
5.2. Regla de Cramer, menores y cofactores 126
5.3. Determinantes y ecuaciones diferenciales 128
5.4. Ejercicios 131

6. Eigenteoría: estructura de operadores 135

6.1. Definiciones y resultados básicos 135
6.1.1. El polinomio mínimo 136
6.2. Valores y vectores característicos 144
6.2.1. Calculando el polinomio mínimo 148
6.3. Forma canónica de Jordan 153
6.4. Matrices reales con valores característicos no reales 156
6.4.1. Matrices 2×2 156
6.4.2. Matrices reales con valores característicos diferentes 158
6.5. Aplicaciones 160
6.5.1. Especies que interactúan 160
6.5.2. Sistemas dinámicos lineales discretos 163
6.5.3. Sistema de masas acopladas con resortes 164
6.6. Ejercicios 167

7. Espacios con producto interno 171

7.1. Aspectos geométricos de un espacio vectorial 171
7.1.1. Método de mínimos cuadrados 175
7.2. Espacios vectoriales complejos 176
7.3. Formas cuadráticas y bilineales 178
7.3.1. Formas cuadráticas 181
7.3.2. Teorema de los ejes principales 182
7.3.3. Matrices positivas definidas 183
7.4. Operadores adjuntos y normales 186
7.5. Ejercicios 188

Bibliografía 189
Índice 191

Introducción

Algunos argumentan que Dios es geómetra, enunciado difícil de sostener. Algo más mundano y acorde con la naturaleza lleva a concluir que "La Matemática es la creación suprema de la mente humana".

El álgebra lineal, junto con el cálculo diferencial e integral, constituyen los pilares de la formación matemática de los estudiantes de ciencias e ingeniería. Posiblemente esto explique por qué se han escrito tantos libros de cada una de estas áreas.

Nuestro objetivo al escribir este libro se puede resumir de la manera siguiente. Por un lado, exponer nuestra concepción del álgebra lineal básica; por otro, que esta concepción auxilie a los estudiantes de matemáticas, ingeniería y áreas afines en el proceso de aprendizaje de tan importante área.

Cuando se inicia la discusión de un tema es adecuado aclarar, en la medida de lo posible, cuáles serán los objetos de estudio. Al respecto, queremos señalar que una posible definición del álgebra lineal puede formularse diciendo que es el área de las matemáticas que estudia las ecuaciones:

$$AX = B \text{ y } AX = \lambda X \tag{1}$$

Tomando esto como referente, podemos decir que el presente trabajo se desarrolla en torno al estudio de dichas ecuaciones, en un escenario con tres elementos que consideramos importantes en la actividad matemática: los fundamentos, los métodos y las aplicaciones.

El desarrollo del texto tiene como antecedentes las notas para los cursos de álgebra lineal que he impartido en el programa educativo de Matemáticas Aplicadas que oferta la Universidad Autónoma del Estado de Hidalgo (UAEH), por lo que el enfoque, contenidos y profundidad están relacionados estrechamente con el currículum de dicha licenciatura. Sin embargo, el texto puede ser utilizado como referencia en licenciaturas de matemáticas y áreas afines.

El contenido del libro está estructurado de la siguiente forma. En los primeros cinco capítulos discutimos la solución de sistemas de ecuaciones lineales, los elementos básicos de la teoría de espacios vectoriales y la teoría de determinantes. En el capítulo seis presentamos la teoría de valores y vectores característicos, tomando una ruta diferente a la que usualmente toman los textos de álgebra lineal. Es decir, en un curso usual, la teoría de valores y vectores característicos se inicia con el polinomio característico de una matriz u operador. Una de las desventajas de esta ruta, es que la mayoría de los resultados relacionados con las propiedades fundamentales de un operador en cuanto a diagonalización, triangulación, etcétera, no se obtienen a partir del *polinomio característico*. Para avanzar en esta línea, hace falta introducir el *polinomio mínimo*.

En contraste con lo que hace la mayoría de los textos, la ruta que seguimos en éste inicia con la definición del polinomio mínimo de un operador y a partir de dicho concepto fundamental, se hace un análisis completo de la estructura de un operador, culminando con la forma canónica de Jordan, pasando por la caracterización de los operadores diagonalizables y triangulables. Es importante notar que en la discusión de valores y vectores característicos que estamos presentando, no se utilizan determinantes en ningún momento y partiendo del polinomio mínimo se puede establecer la definición del polinomio característico. Con este enfoque, presentamos una demostración corta del importante teorema de Cayley-Hamilton.

Otro aspecto que nos parece importante en esta ruta es el algoritmo que presentamos para calcular el polinomio mínimo. Éste sólo hace uso de operaciones elementales en la matriz $A - \lambda I$. Otra ventaja que tiene el iniciar la discusión con el polinomio mínimo es su generalidad, pues los resultados principales se pueden formular sobre cualquier campo. Queremos hacer notar que una discusión similar aparece en [2], sin embargo la ruta difiere de la nuestra, dado que allí se toma como punto de partida que el espacio vectorial está definido sobre los números complejos, lo que por sí mismo lleva un precio. Nuestra opinión es que ésta no es una ruta natural para un primer curso de álgebra lineal, pues en muchos de los ejemplos que se discuten allí, los escalares son números reales.

En el capítulo siete, se presenta lo que podría llamarse *aspectos geométricos* de los espacios vectoriales, es decir, allí se discute lo relacionado con propiedades derivadas del producto interno.

Para el desarrollo de este trabajo consultamos varias fuentes, entre éstas se encuentran: [1], [4], [5], [7], [6], [9] y [11].

Finalmente, quiero expresar mi agradecimiento a todas las personas que hicieron posible la realización de este trabajo. Por supuesto, en este proceso están todos los alumnos que han tomado cursos de álgebra lineal conmigo. De manera muy especial, quiero agradecer al Dr. Rubén Martínez Avendaño, profesor investigador del Centro de Investigación en Matemáticas de la UAEH, con quien discutí algunos de los resultados del capítulo seis, además que revisó e hizo observaciones excelentes para mejorar la presentación del texto. También mis agradecimientos especiales van para Fidel Barrera Cruz, egresado de la licenciatura en Matemáticas aplicadas de la UAEH, quien programó el algoritmo que calcula el polinomio mínimo de una matriz.

Pachuca, Hidalgo, julio de 2007.

Nomenclatura

$<$	menor que, página 33
$\leq$	menor o igual que, página 33
$\max\{a_1, a_2, \ldots, a_n\}$	máximo de los elementos $a_1, a_2, \ldots, a_n$, página 75
$U \oplus W$	suma directa de U y W, página 118
$:=$	igualdad por definición, página 34
$\lvert A \rvert$	determinante de la matriz A, página 157
$\lvert x \rvert$	valor absoluto de x, página 95
$\lVert \alpha \rVert$	norma de α, página 92
$\sim$	equivalencia de matrices por filas, página 21
$\geq$	mayor o igual que, página 12
$\Leftrightarrow$	doble implicación, página 21
$\in$	es un elemento de, página 36
$\langle \alpha, \beta \rangle$	producto interno de α y β, página 93
$\mathbb{C}$	campo de los números complejos, página 199
$\mathbb{C}^n$	espacio vectorial complejo de dimensión n, página 199
$\mathbb{N}$	conjunto de los números naturales, página 95
$\mathbb{Q}$	campo de los números racionales, página 107
$\mathbb{R}$	campo de los números reales, página 36
$\mathbb{R}^2$	plano cartesiano, página 79
$\mathbb{R}^n$	espacio vectorial real de dimensión n, página 81
$\mathbb{Z}$	conjunto de los números enteros, página 107
$\neq$	no es igual, página 53
$\notin$	no es elemento de, página 107
$\sqrt{a}$	raíz cuadrada de a, página 80
$\subseteq$	contenido en, página 111
$\mathrm{Proy}_{\vec{v}}\vec{u}$	proyección ortogonal de $\vec{u}$ sobre, $\vec{v}$ página 96
$\mathrm{Adj}(A)$	adjunta clásica de la matriz A, página 162
$\mathrm{ann}\,(\alpha, x)$	T anulador de α, página 181
$\deg(g(x))$	grado del polinomio $g(x)$, página 178
$\mathrm{diag}\{d_1, d_2, \ldots, d_r\}$	matriz diagonal, página 72

$\dim(\mathbb{R}^n)$	dimensión de $\mathbb{R}^n$, página 109
$\mathrm{Im}(z)$	parte imaginaria del complejo z, página 220
$\mathrm{Re}(z)$	parte real del complejo z, página 220
$\mathrm{tr}(A)$	traza de la matriz A, página 61
$\sum_{i=1}^{n} a_i$	suma de los elementos $a_1, a_2, \ldots, a_n$, página 61
$\vec{u} \times \vec{v}$	producto vectorial de $\vec{u}$ y $\vec{v}$, página 100
$A \cap B$	intersección de los conjuntos A y B, página 107
$A \cup B$	unión de los conjuntos A y B, página 107
A^*	adjunta de la matriz A, página 232
A^t	transpuesta de la matriz A, página 61
A^{-1}	inverso multiplicativo de A, página 55
$C(\beta, T)$	subespacio T-cíclico generado por β, página 182
$f: X \to Y$	función con dominio X y contradominio Y, página 121
I_n	matriz identidad $n \times n$, página 53
$m_T(\mathrm{x})$	polinomio mínimo de T, página 175
$n!$	factorial de n, página 161
N_T	núcleo de T, página 133
R_T	rango de T, página 133
T^*	adjunto del operador T, página 232
T^n	composición de la función T consigo misma n veces, página 146
$T_1 \circ T$	composición de las funciones T_1 y T, página 137
$U \setminus W$	diferencisa de conjuntos, página 112
$V \cong W$	isomorfismo entre V y W, página 134
V^*	espacio dual de V, página 148
$W(g_1, g_2, \ldots, g_n)(x)$	wronskiano de las funciones $g_1, g_2, \ldots, g_n$, página 165
$W^{\perp}$	complemento ortogonal de W, página 217
$\mathcal{L}(S)$	subespacio generado por S, página 108
$\mathcal{L}(V; W)$	espacio de las transformaciones lineales de V en W, página 144
$\mathcal{M}_{m \times n}(\mathcal{R})$	conjunto de matrices de m filas y n columnas con entradas en los reales, página 52

Capítulo 1

Sistemas de ecuaciones lineales

En una gama amplia de problemas, ya sean teóricos o aplicados, su formulación lleva al estudio de un sistema de ecuaciones lineales, tema central en álgebra lineal. Por ejemplo, Wassily W. Leontief,[1] usando sistemas de ecuaciones lineales, desarrolló un modelo económico para describir la actividad económica de los Estados Unidos de América. A grandes rasgos el modelo de Leontief consistió en dividir la economía en 500 sectores, tales como la industria eléctrica, la automotriz, la de las comunicaciones, etcétera, y a partir de esto formuló una ecuación lineal que describe la forma en la que cada sector distribuía su producción entre los restantes, ver ejemplo 1.1.2.

En el ámbito puramente matemático, algunos problemas se formulan mediante sistemas de ecuaciones lineales y otros pueden representarse mediante una o varias funciones, las cuales bajo hipótesis adecuadas pueden ser aproximadas por funciones lineales, llevando el problema al ámbito del álgebra lineal.

Estos elementos muestran la importancia que tiene el hacer un estudio sistemático y profundo de los sistemas de ecuaciones lineales. Para lograr este objetivo se desarrollarán conceptos fundamentales como: espacio vectorial, dependencia e independencia lineal, base, dimensión, transformación lineal, valores y vectores característicos, determinantes, entre otros.

Iniciamos la discusión en este capítulo presentando algunos ejemplos que ilustran el uso de sistemas de ecuaciones lineales para abordar una situación.

1.1. Ejemplos

Ejemplo 1.1.1. *Supongamos que se tienen dos empresas* E_1 y E_2, *y en cada una se producen los bienes* B_1 *y* B_2. *Supongamos que por cada unidad monetaria que se invierte en las empresas, la producción es como se describe en la tabla 1.1*.

	E_1	E_2
B_1	.8	.6
B_2	.4	.7

Tabla 1.1. Relación de producción.

La segunda fila de la tabla significa que por cada unidad monetaria, la empresa E_1 produce .8 del bien B_1; la empresa E_2 produce .6 del mismo bien. De esto se tiene

[1] Wassily W. Leontief, *Input-Output Economics, Scientific American*, octubre de 1951, pp. 15-21. Leontief ganó el premio Nobel de Economía en 1973 por sus contribuciones a la teoría económica, usando sistemas de ecuaciones lineales.

que si cada empresa invierte un peso, las empresas producen .8 + .6 pesos del bien B_1. Asimismo, la tercera fila indica que por cada unidad monetaria la empresa E_1 produce .4 del bien B_2 y la empresa E_2 produce .7 de ese bien; con esta información se tiene que ambas producen en total .4 + .7 del bien B_2, en el caso de invertir cada una un peso.

Por ejemplo, si en las empresas E_1 y E_2 se invierten 20 y 18.5 millones respectivamente, entonces el valor de los productos en millones es:

$$
\begin{aligned}
.8(20) + .6(18.5) &= 27.1: \quad \text{valor de } B_1\\
.4(20) + .7(18.5) &= 20.95: \quad \text{valor de } B_2
\end{aligned}
$$

¿Cuánto hay de ganancia total? Notemos que la ganancia es igual al valor de los bienes menos lo que se invirtió en producirlos.

Generalizando, si en las empresas E_1 y E_2 se invierten x y y pesos respectivamente, y representamos el valor total de los bienes B_1 y B_2 por b_1 y b_2 en aquel mismo orden, entonces se tiene:

$$
\begin{aligned}
.8x + .6y &= b_1\\
.4x + .7y &= b_2.
\end{aligned}
\tag{1.1}
$$

Si en el sistema anterior los decimales se transforman a cocientes de enteros y se resuelve para y en cada una de las ecuaciones, éstas se pueden representar en forma equivalente mediante el sistema:

$$
\begin{aligned}
y &= -\frac{4}{3}x + \frac{5b_1}{3}\\
y &= -\frac{4}{7}x + \frac{10b_2}{7}
\end{aligned}
\tag{1.2}
$$

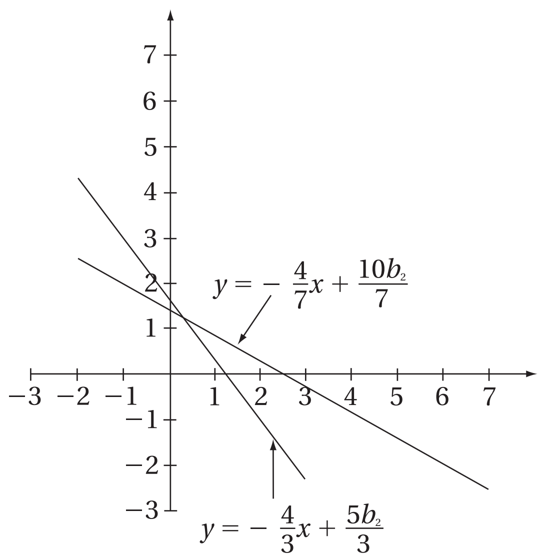


Figura 1.1. Representación geométrica de las ecuaciones 1.2 para $b_1 = b_2 = 1$.

Una pregunta que puede ser de importancia es: ¿cómo debe ser la inversión en cada empresa para que se obtenga una producción de los bienes B_1 y B_2 con valores

b_1 y b_2, respectivamente? Dado que las inversiones se representan por cantidades no negativas, una manera de formular la pregunta anterior es: ¿para qué valores de b_1 y b_2 el sistema 1.1 tiene soluciones no negativas?

En la figura 1.1 se ha representado al sistema 1.1 para el caso $b_1 = b_2 = 1$ y se observa que tiene solución positiva. La interpretación geométrica de la pregunta anterior es: ¿para cuáles valores no negativos de b_1 y b_2 las rectas representadas por el sistema 1.1 se intersecan en el primer cuadrante?

Como los sistemas (1.1) y (1.2) son equivalentes, además de que en el segundo y está despejada, una forma de resolverlo es igualando las expresiones de y; haciendo esto y simplificándola se obtiene el valor de x y después el valor de y, de manera explícita:

$$\begin{aligned} x &= \frac{35b_1 - 30b_2}{16} \\ y &= \frac{10b_2 - 5b_1}{4}. \end{aligned} \tag{1.3}$$

Notemos que las soluciones se han obtenido en términos de b_1 y b_2, los cuales se suponen conocidos. Las condiciones de la pregunta original implican que $x, y \geq 0$, es decir, se deben satisfacer las condiciones:

$$\begin{aligned} x &= \frac{35b_1 - 30b_2}{16} \geq 0 \\ y &= \frac{10b_2 - 5b_1}{4} \geq 0. \end{aligned} \tag{1.4}$$

Éste es un sistema de desigualdades en b_1 y b_2, el cual equivale a

$$\begin{aligned} 7b_1 - 6b_2 &\geq 0 \\ 2b_2 - \mathrm{b}_1 &\geq 0, \end{aligned} \tag{1.5}$$

y se puede representar como:

$$\begin{aligned} \frac{7b_1}{6} &\geq b_2 \\ b_2 &\geq \frac{b_1}{6}. \end{aligned} \tag{1.6}$$

En la figura 1.2 se representa a la región del plano b_1 b_2 en la cual se satisfacen las desigualdades 1.6. Esa región también puede ser interpretada como la imagen del primer cuadrante bajo la función que se describe abajo.

Desde una perspectiva puramente matemática, el proceso de producción lo podemos formular mediante una función y su interpretación es: la función

$$(x, y) \to T(x, y) = (0.8x + 0.4y, 0.4x + 0.7y) = (b_1, b_2)$$

transforma cada punto (x, y) del primer cuadrante (plano de la inversión), en un punto de la región sombreada de la figura 1.2. Funciones con las características de T serán consideradas ampliamente cuando se discutan transformaciones lineales.

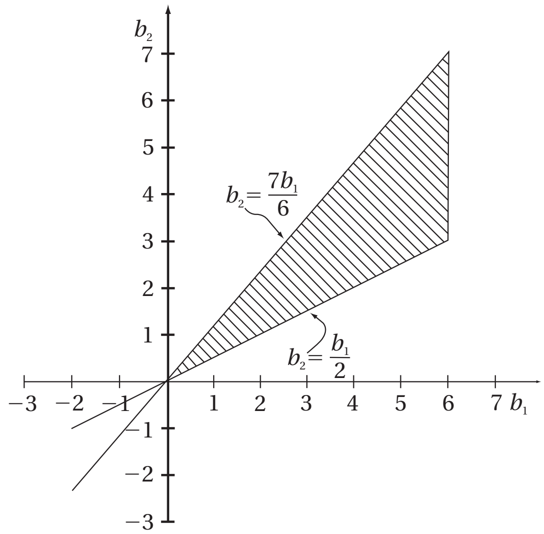


Figura 1.2. Imagen del primer cuadrante bajo *T*.

Conteste las siguientes preguntas y use diferentes representaciones (geométricas, algebraicas, verbales, etc.) en su discusión.

1. ¿Qué cantidades se deben invertir si se desea obtener valores de los bienes B_1 y B_2 iguales a 4 y 3 millones de pesos, respectivamente?
2. ¿Qué significado tiene, en términos de inversiones, que $b_1 = 4$ y $b_2 = 2$? ¿Tiene solución el sistema para este caso?
3. ¿Se puede invertir de manera que b_2 sea el doble de b_1? Explique numérica y geométricamente.
4. ¿Cuál es el resultado de invertir a partes iguales en las dos empresas?
5. ¿Cómo es el valor numérico de la pendiente de las rectas que pasan por el origen y tienen al menos dos puntos en la región descrita por las desigualdades anteriores?
6. ¿Puede proponer un modelo de producción como el del ejemplo anterior, de manera que $b_2 = 2b_1$ sea posible? Haga una discusión geométrica y algebraica, e interprete sus resultados desde el punto de vista económico.

Ejemplo 1.1.2. (Modelo de Leontief). *Supongamos que una economía consiste de* n *sectores, donde cada uno consume parte de lo que produce y parte de cada producto es consumido por el público demandante. El modelo que propuso Leontief tiene un par de hipótesis.*

1. *Cada sector tiene ganancias, es decir, la producción de cada sector es mayor que lo que requiere para producir.*
2. *La economía está en equilibrio, esto significa que el total de cada producto es igual a lo que consumen los sectores, más lo que consume el público demandante.*

Ilustraremos este modelo con un caso especial. Supongamos que se tienen solamente tres sectores: transporte, energético y agrícola. Denotemos por x, y y z a la cantidad en pesos que produce respectivamente cada sector. Los requerimientos de éstos se establecen en la tabla 1.2.

	Transporte	Energía	Agricultura
Transporte	$.2x$	$.1y$	$.2z$
Energía	$0.4x$	$.3y$	$.3z$
Agricultura	$.3x$	$.5y$	$.35z$

Tabla 1.2. Requerimientos por sector.

La segunda columna de la tabla 1.2 significa que el sector transporte para producir x pesos requiere: $.2x$ de su propio producto; $.4x$ del sector energía y $.3x$ del sector agricultura. Note que el sector transporte para producir x pesos requiere $.9x$ pesos, es decir, espera ganancias de 10%. Las columnas 3 y 4 se interpretan de la misma manera.

Si el público consume 2, 3 y 5 unidades monetarias de los sectores transporte, energía y agricultura respectivamente, ¿será satisfecha la demanda?

La segunda condición del modelo, es decir, la condición de equilibrio establece que para cada uno de los sectores se cumple:

$$\text{Producción} = \text{consumo}$$

Esta condición y los requerimientos de la tabla 1.2 llevan al sistema:

$$\begin{aligned} x &= .2x + .1y + .2z + 2, \\ y &= .4x + .3y + .3z + 3, \\ z &= .3x + .5y + .35z + 5. \end{aligned} \tag{1.7}$$

Una forma de intentar resolver el sistema anterior es mediante *prueba y error*; otra es mediante argumentos de tipo económico.

Ejercicio 1.1.1. *Dado el sistema*

$$\begin{aligned} x &= ax + by + c \\ y &= a_1x + b_1y + c_1, \end{aligned} \tag{1.8}$$

con todos los coeficientes positivos de y, *además* $a + a_1 < 1$ y $b + b_1 < 1$. *Demuestre que el sistema tiene solución positiva.*

Ejemplo 1.1.3. *En una región la población se mantiene constante*[2] *y se divide en rural y urbana. Se ha observado que cada año,* 25% *de habitantes de la zona rural pasa a la urbana y* 5% *de la urbana se cambia a la rural. Si al inicio de un experimento para determinar el movimiento de la población se tienen* 8 *millones en la zona rural y* 2 *en la urbana, ¿cuántos habitantes habrá en cada zona al cabo de* 1, 2, 3, 4, 5 *años?*

Discusión: Al analizar una situación, es de importancia identificar y representar la información usando diferentes medios. En este caso usaremos una tabla para representar la forma en que la población migra con el paso del tiempo.

Tabla 1.3. Cambio de población con el tiempo.

Año	Población rural	Población urbana
0	8	2
1	(.75)8 + (.05)2 = 6.10	(.25)8 + (.95)2 = 3.9
2	(.75)(6.10) + (.05)(3.9) = 4.77	(.25)(6.1) + (.95)(3.9) = 5.23
3	(.75)(4.77) + (.05)(5.23) = 3.839	(.25)(4.77) + (.95)(5.23) = 6.161
4	(.75)(3.839) + (.05)(6.161) = 3.1873	(.25)(3.839) + (.95)(6.161) = 6.8127
5	(.75)(3.1873) + (.05)(6.8127) = 2.73111	(.25)(3.1873) + (.95)(6.8127) = 7.26889

[2] El crecimiento de población en Alemania durante 2005 fue cero. (*http://www.indexmundi.com/germany/population_growth_rate.html*)

De acuerdo con la información de la tabla 1.3, ¿podría predecir cuál será la distribución de la población en 20 años?

Si en el año n denotamos por u_n y r_n al número de habitantes en la zona urbana y rural respectivamente, entonces u_{n+1} y r_{n+1} están dados por:

$$\begin{aligned} u_{n+1} &= .25r_n + .95u_n \\ r_{n+1} &= .75r_n + 0.0.5u_n. \end{aligned} \tag{1.9}$$

De acuerdo con la hipótesis, el total de la población es 10 millones, ¿podrá distribuirse la población de manera que cada año la cantidad de habitantes en la zona rural sea la misma, y de igual forma ocurra en la zona urbana? Si al inicio el número de habitantes en la zona rural y urbana los denotamos por r_0 y u_0, entonces deseamos que $r_1 = r_0$ y $u_1 = u_0$, es decir, en general $r_n = r_0$ y $u_n = u_0$. Las primeras ecuaciones equivalen a:

$$\begin{aligned} u_0 &= .25r_0 + .95u_0 \\ r_0 &= .75r_0 + .05u_0. \end{aligned} \tag{1.10}$$

Simplificando términos en el sistema, notamos que éste se reduce a una sola ecuación: $.25r_0 - 0.05u_0 = 0$; y esta última equivale a $u_0 = 5r_0$, cuyo significado es que la población rural debe ser una quinta parte de la urbana para mantenerse en equilibrio.

Ejemplo 1.1.4. *Supongamos que una empresa administra tres refinerías de petróleo y cada una produce tres derivados: gasolina, diesel y aceite lubricante. Supongamos también que por cada barril de petróleo (aproximadamente 159 litros) la producción, en galones,*[3] *es como se indica en la tabla 1.4.*

	Refinería 1	Refinería 2	Refinería 3
Gasolina	20	21	19
Diesel	11	12	13
Aceite lubricante	9	8	8

Tabla 1.4. Producción de cada refinería.

La información de la tabla 1.4 se interpreta de la manera siguiente: por cada barril de petróleo la refinería 1 produce 20 galones de gasolina, 11 de diesel y 9 de aceite lubricante; la refinería 2 produce 21 galones de gasolina, 12 de diesel y 8 de aceite lubricante; la refinería 3 produce 19 galones de gasolina, 13 de diesel y 8 de aceite lubricante.

Supongamos que se desea una producción de 1 250 galones de gasolina, 750 de diesel y 520 de aceite lubricante. ¿Cuántos barriles de petróleo debe procesar cada refinería para satisfacer esa demanda?

Discusión

Identificando información importante

1. *Cantidades conocidas.*
   a) Se deben producir:
      1) 1 250 galones de gasolina.

[3] Un galón es aproximadamente 3.87 litros.

2) 750 galones de diesel.

3) 520 galones de aceite lubricante.

b) Producción por refinería. Note que esta información está descrita y organizada en la tabla anterior.

2. *Cantidades desconocidas*. Las cantidades desconocidas son el número de barriles que debe procesar cada una de las refinerías. Denotemos por x_1, x_2 y x_3 a las cantidades de barriles que debe procesar la refinería 1, 2 y 3, respectivamente.

Relacionando datos y variables

Notemos que la producción total de gasolina, diesel y aceite lubricante, es la suma de la producción de cada refinería, así que con la notación introducida se tiene:

Total de gasolina:	$20x_1 + 21x_2 + 19x_3$
Total de diesel:	$11x_1 + 12x_2 + 13x_3$
Total de aceite lubricante:	$9x_1 + 8x_2 + 8x_3$

Dadas las condiciones de la situación, deseamos saber cuáles son los valores de x_1, x_2 y x_3, de tal forma que se cumplan las siguientes ecuaciones:

$$\begin{aligned} 20x_1 + 21x_2 + 19x_3 &= 1\,250 \\ 11x_1 + 12x_2 + 13x_3 &= 750 \\ 9x_1 + 8x_2 + 8x_3 &= 520. \end{aligned} \tag{1.11}$$

Usando algunas estrategias para resolver un sistema de ecuaciones

Una forma de aproximarse a la solución del sistema 1.11 es mediante *prueba y error*, es decir, proponer valores para las variables x_1, x_2 y x_3 además de calcular el valor de las expresiones en las ecuaciones anteriores. Por ejemplo, si $x_1 = x_2 = x_3 = 20$, entonces se tienen los siguientes resultados

$$\begin{aligned} 20(20) + 21(20) + 19(20) &= 1\,200 \\ 11(20) + 12(20) + 13(20) &= 720 \\ 9(20) + 8(20) + 8(20) &= 500. \end{aligned}$$

Nótese que estas cantidades necesitan ser incrementadas para satisfacer lo demandado. ¿Qué criterio se usaría para incrementar los valores de las variables? Probemos un caso más:

$$\begin{aligned} 20(20) + 21(20) + 19(21) &= 1\,219 \\ 11(20) + 12(20) + 3(21) &= 733 \\ 9(20) + 8(20) + 8(21) &= 508. \end{aligned}$$

Como se puede observar en estas pruebas, es muy difícil encontrar la solución del sistema a base de prueba y error. Con esto surge la pregunta: ¿existe un método para resolver el sistemas de ecuaciones anterior? En general, ¿se puede resolver un sistema de m ecuaciones en n incógnitas? Para dar respuesta a estas preguntas iniciemos analizando el sistema a la luz de algunas propiedades que se tienen al "operar" con igualdades.

1. **Propiedad 1.** Si los miembros de una ecuación se multiplican por un mismo número, el resultado es otra ecuación.

   Ejemplo. Dada la ecuación $4x + 2y = 2$, al multiplicar por $1/2$ se tiene $2x + y = 1$.
2. **Propiedad 2.** Si se tienen dos ecuaciones y una de ellas se multiplica por un número y se suma a la otra, el resultado es otra ecuación.

Ejemplo. Dadas las ecuaciones $x + y = 3$ y $3x + y = -1$, al multiplicar a la primera por -3 y sumarla a la segunda se tiene $-3(x + y) + (3x + y) = -3(3) + (-1)$. Esta ecuación equivale a $-2y = -10$.

La Propiedad 1 permite resolver ecuaciones del tipo $ax = b$, con $a \neq 0$, es decir, multiplicando por $\frac{1}{a}$ se tiene: $x = \frac{b}{a}$.

Apliquemos estas propiedades para analizar el sistema (1.11).

$$20x_1 + 21x_2 + 19x_3 = 1\,250 \tag{1.12}$$

$$11x_1 + 12x_2 + 13x_3 = 750 \tag{1.13}$$

$$9x_1 + 8x_2 + 8x_3 = 520. \tag{1.14}$$

Multiplicando la ecuación (1.13) por -20, la ecuación (1.12) por 11, sumando los resultados y simplificando se tiene:

$$-9x_2 - 51x_3 = -1\,250 \tag{1.15}$$

Multiplicando la ecuación (1.14) por –20, la ecuación (1.12) por 9, sumando los resultados y simplificando se tiene:

$$29x_2 + 11x_3 = 850 \tag{1.16}$$

Con las ecuaciones (1.15) y (1.16) formamos el sistema:

$$9x_2 + 51x_3 = 1\,250 \tag{1.17}$$

$$29x_2 + 11x_3 = 850 \tag{1.18}$$

Multiplicando la ecuación (1.18) por -9, la ecuación (1.17) por 29, sumando los resultados y simplificando se tiene:

$$x_3 = \frac{1\,430}{69}$$

Sustituyendo este valor de x_3 en la ecuación (1.17), obtenemos

$9x_2 + 51\left(\frac{1\,430}{69}\right) = 1{,}250$; de esta última ecuación llegamos a:

$9x_2 = 1\,250 - 51\left(\frac{1\,430}{69}\right) = \frac{86\,250 - 72\,930}{69} = \frac{13\,320}{69} = \frac{4\,440}{23}$ concluyendo que

$$x_2 = \frac{1\,480}{69}$$

Sustituyendo los valores de x_3 y x_2 en la ecuación (1.12) obtenemos

$20x_1 + 21\left(\frac{1\,480}{69}\right) + 19\left(\frac{1\,430}{69}\right) = 1\,250$, y de esto se concluye:

$$20x_1 = 1\,250 - 21\left(\frac{1\,480}{69}\right) - 19\left(\frac{1\,430}{69}\right) = \frac{86\,250 - 31\,080 - 27\,170}{69} = \frac{28\,000}{69}$$

Finalmente,

$$x_1 = \frac{1\,400}{69} \cong 20.290, \quad x_2 = \frac{1\,480}{69} \cong 21.449 \quad \text{y} \quad x_3 = \frac{1\,430}{69} \cong 20.725$$

son los valores buscados.

Comprobando resultados

Una componente muy importante en el proceso de solución de un problema es verificar que los resultados obtenidos satisfagan las condiciones del problema. En el caso que hemos discutido debemos verificar que los valores de las incógnitas satisfacen el sistema 1.11, es decir, debemos verificar que las siguientes igualdades se cumplan.

$$\begin{aligned} 20\left(\frac{1\,400}{69}\right) + 21\left(\frac{1\,480}{69}\right) + 19\left(\frac{1\,430}{69}\right) &= 1\,250 \\ 11\left(\frac{1\,400}{69}\right) + 12\left(\frac{1\,480}{69}\right) + 13\left(\frac{1\,430}{69}\right) &= 750 \\ 9\left(\frac{1\,400}{69}\right) + 8\left(\frac{1\,480}{69}\right) + 8\left(\frac{1\,430}{69}\right) &= 520. \end{aligned}$$

Desarrollando los cálculos indicados en cada ecuación se verifican las igualdades.

Ejercicio 1.1.2. *Considere las mismas condiciones del problema anterior pero cambie las demandas al doble. ¿Cuál es la solución? Si las cantidades de gasolina, diesel y aceite lubricante son 5 000, 3 000 y 1 500 galones respectivamente, tiene solución el problema?*

Ahora que hemos encontrado una solución de un sistema de tres ecuaciones en tres variables, surge la pregunta: ¿se puede resolver cualquier sistema de tres ecuaciones lineales en tres variables? Antes de intentar responder, parece natural preguntar lo que ocurre con sistemas de ecuaciones en dos variables, es decir, ¿se puede resolver cualquier sistema de ecuaciones en dos incógnitas?

1.2. Sistemas de ecuaciones lineales y su representación geométrica

Recordemos que en geometría analítica, una ecuación lineal en dos variables representa una recta en el plano cartesiano y un sistema representa una colección de rectas. Por ejemplo, las ecuaciones $x + y = 1$ y $2x - 3y + 4 = 0$ representan a dos rectas, como se muestra en la figura 1.3.

En el caso de ecuaciones que representan líneas rectas, encontrar valores de las variables que satisfagan a cada una de dichas ecuaciones significa encontrar las coordenadas de los puntos de intersección.

Recordemos que una recta l, representada por una ecuación de la forma $ax + by = c$, interseca a los dos ejes $\Leftrightarrow ab \neq 0$. Desde el punto de vista algebraico, la condición $ab = 0$ se traduce a que una de las variables que aparece en la ecuación está prácticamente despejada y el resolver un sistema se reduce a sustituir y despejar la otra variable. Tomando esta observación como punto de partida, podemos suponer que estudiaremos rectas que intersecan a los dos ejes, es decir, consideremos el sistema:

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned} \tag{1.19}$$

con la hipótesis $a_1a_2b_1b_2 \neq 0$.

Nótese que el sistema anterior queda completamente determinado por los coeficientes de las variables y los términos independientes. Esto significa que el sistema se puede representar mediante un arreglo rectangular en el que las filas representan a las ecuaciones. Con esta aclaración, el sistema se representa en la forma:

$$\left[\begin{array}{cc|c} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{array}\right]$$

La línea vertical es para separar los coeficientes de las variables de los términos constantes.

Ejemplo 1.2.1. *Consideremos el sistema*

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

para este caso su representación en forma de arreglo es:

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & -3 & -4 \end{array}\right]$$

Representado de esta forma el sistema, procederemos a la aplicación sucesiva de las propiedades 1 y 2, enunciadas en la página 7, que se aplican cuando se opera con ecuaciones.

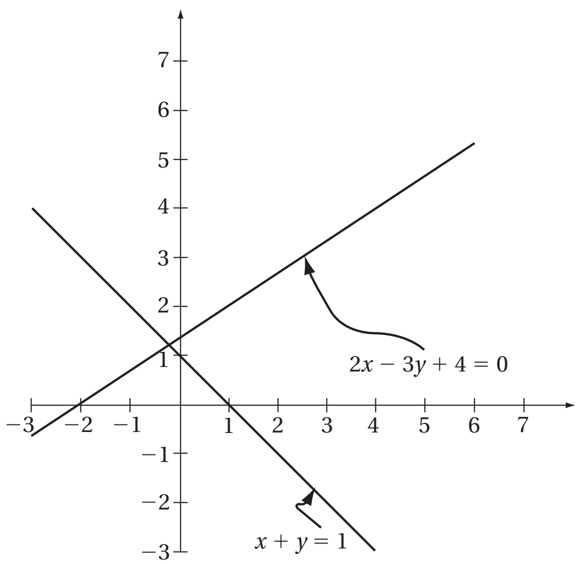


Figura 1.3. Representación gráfica de las ecuaciones $x + y = 1$ y $2x - 3y + 4 = 0$.

Multiplicando la primera fila por -2 y sumando el resultado a la fila dos se obtiene:

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & -3 & -4 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -5 & -6 \end{array}\right]$$

Aquí el símbolo $\sim$ denota la equivalencia de los arreglos. Note que la primera fila solamente funcionó como "pivote" y permanece sin cambio. El arreglo de la derecha representa al sistema:

$$\begin{aligned} x + y &= 1 \\ 0x - 5y &= -6 \end{aligned}$$

De la segunda ecuación de este sistema se tiene que el valor de y se obtiene multiplicando por $-\frac{1}{5}$. Este paso lo podemos representar en el último arreglo, multiplicando la segunda fila por $-\frac{1}{5}$.

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & -3 & -4 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -5 & -6 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & 6/5 \end{array}\right]$$

El último arreglo representa al sistema:

$$\begin{aligned} x + y &= 1 \\ 0x + y &= \frac{6}{5} \end{aligned}$$

Si en este último sistema restamos la segunda ecuación de la primera y todos los pasos anteriores los representamos, se tiene:

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & -3 & -4 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -5 & -6 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & 6/5 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & -1/5 \\ 0 & 1 & 6/5 \end{array}\right]$$

El último arreglo representa al sistema:

$$\begin{aligned} x + 0y &= -\frac{1}{5} \\ 0x + y &= \frac{6}{5}. \end{aligned}$$

en el cual se tienen los valores de x y y.

Ejercicio 1.2.1. *En cada uno de los casos que siguen represente los sistemas como arreglos rectangulares, resuélvalos y represéntelos gráficamente.*

1. $\begin{aligned} x - y &= 5 \\ x + y &= 6 \end{aligned}$ $\qquad$ $\begin{aligned} x + 3y &= 3 \\ 3x + y &= 4 \end{aligned}$ $\qquad$ $\begin{aligned} x + y &= 3 \\ x - 3y &= 2 \\ x - 2y &= -1 \end{aligned}$ $\qquad$ $\begin{aligned} x - 5y &= 7 \\ 3x + 2y &= 2 \end{aligned}$

2. *Encuentre los valores de* a *de tal forma que el siguiente sistema tenga solución:*

$$\begin{aligned} ax + y &= 3 \\ 3x - ay &= 4. \end{aligned}$$

3. *Represente el sistema cuyas ecuaciones están dadas por (1.19) mediante un arreglo rectangular y use el procedimiento del ejemplo discutido para resolverlo.*
4. *¿Qué notación usaría para representar un sistema de* 10 *ecuaciones en* 30 *variables?*
5. *¿Puede proponer una situación de la cual se obtenga un sistema de ecuaciones lineales como el de la pregunta anterior?*

1.3. Conceptos fundamentales y método de reducción de Gauss-Jordan

Antes de presentar un método general para resolver sistemas de ecuaciones lineales, es importante definir algunos términos que hagan la discusión más precisa.

Definición 1.3.1. *Por una ecuación lineal en las variables* $x_1, x_2, \ldots, x_n$ *entenderemos una ecuación del tipo:*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{1.20}$$

en donde $a_1, a_2, \ldots, a_n$ *y* b *no dependen de las variables* $x_1, x_2, \ldots, x_n$.

Ejemplo 1.3.1. *Los siguientes son ejemplos de ecuaciones lineales en* x, y *y* z*:*

- $3x + 6y - z = 0$
- $x + y = 4z$
- $x + t^2y + z = t^3$, *t independiente de x, y, z.*

Ejercicio 1.3.1. *Decida en cuáles variables son lineales las siguientes ecuaciones:*

- $x^2y + t^3z = w$
- $ax + by = w, b = x^2$
- $t^3 + x + y + z = 0$

Definición 1.3.2. *Un sistema de ecuaciones lineales es una colección de ecuaciones del tipo (1.20). Más precisamente, un sistema de m ecuaciones lineales en las variables* $x_1, x_2, \ldots, x_n$ *es una colección de ecuaciones del tipo:*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1.21}$$

en donde a_{ij} *y* b_i *son independientes de* $x_1, x_2, \ldots, x_n$, *para* $i = 1, 2, \ldots, m$ *y* $j = 1, 2, \ldots, n$.

Dado el sistema 1.21, éste se puede representar mediante un arreglo rectangular:

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

en donde las filas del arreglo representan a las ecuaciones. Por ejemplo, la fila 1 representa a la primera ecuación del sistema. Note que en esta representación hemos omitido la línea vertical que separa a los coeficientes de las variables de los términos independientes.

La importancia de la representación anterior es tal que recibe un nombre, se llama la *matriz aumentada* del sistema 1.21. Cuando se suprime la columna de términos independientes, al arreglo se le llama *matriz de coeficientes* del sistema, la cual es:

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

Antes de continuar la discusión es pertinente precisar lo que entenderemos por solución de un sistema de ecuaciones lineales.

Definición 1.3.3. *Por solución del sistema (1.21) entenderemos una n-ada* $(c_1, c_2, \ldots, c_n)$ *tal que* $a_{i1}c_1 + a_{i2}c_2 + \cdots + a_{in}c_n = b_i$ *para cada* $i = 1, 2, \ldots, n$. *Cuando el sistema tiene solución se dice consistente, de otra forma es inconsistente.*

Si todos los b_i en el sistema 1.21 son cero, entonces (0, 0, …, 0) es solución, es decir, el sistema es consistente, compruébelo.

Ejemplo 1.3.2. C*onsidere el sistema:*

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

Geométricamente, las ecuaciones anteriores representan rectas paralelas; algebraicamente esto significa que el sistema no tiene solución, pues si (c_1, c_2) fuese una solución, entonces $c_1 + 2c_2 = 3$ y $2c_1 + 4c_2 = 1$, pero estas ecuaciones son incompatibles, pues el primer miembro de la segunda es el doble del primer miembro de la primera, mientras que el segundo miembro de la segunda no es el doble del segundo miembro de la primera.

En el proceso de solución que seguimos en el ejemplo de las refinerías usamos esencialmente las siguientes "operaciones".

1. Intercambiar dos ecuaciones.
2. Multiplicar una ecuación por una constante no cero.
3. Multiplicar una ecuación por una constante y sumarla a otra.

Las operaciones anteriores cuando se aplican a la representación de un sistema mediante la matriz aumentada reciben un nombre: *operaciones elementales* en las filas de la matriz y se especifican a continuación.

1. Intercambiar dos filas.
2. Multiplicar una fila por una constante no cero.
3. Multiplicar una fila por una constante y sumarla a otra.

Note que las operaciones elementales tienen inversa. Por ejemplo, si se intercambió la fila *i* con la *j*, la operación inversa consiste en intercambiar la fila *j* con la *i*.

Definición 1.3.4. *Dos sistemas de ecuaciones*

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

y

$$\begin{aligned} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n &= d_1 \\ c_{21}x_1 + c_{22}x_2 + \cdots + c_{2n}x_n &= d_2 \\ &\vdots \\ c_{m1}x_1 + c_{m2}x_2 + \cdots + c_{mn}x_n &= d_m \end{aligned}$$

se dicen equivalentes si tienen las mismas soluciones.

Teorema 1.3.1. *Si un sistema de ecuaciones lineales se obtiene de otro por medio de operaciones elementales, entonces los sistemas son equivalentes.*

Demostración. La demostración que haremos supone que se aplica una sola operación elemental, pues el caso general se obtiene de éste aplicando sucesivamente el mismo argumento. Sea:

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

un sistema de ecuaciones. Es claro que si en este sistema se intercambian dos ecuaciones, o una ecuación se multiplica por un real no cero, el nuevo sistema tiene las mismas soluciones que el original. Resta demostrar que las soluciones no cambian cuando el nuevo sistema se obtiene del original, multiplicando una ecuación por un real y sumándola a otra. Supongamos que la ecuación i se multiplica por a y se suma a la ecuación j, entonces la ecuación j en el nuevo sistema es:

$$(a_{j1} + aa_{i1})\, x_1 + (a_{j2} + aa_{i2})x_2 + \cdots + (a_{jn} + aa_{in})x_n = b_j + ab_i \tag{1.22}$$

y todas las restantes son las mismas que las del original.

Si $(c_1, c_2, \ldots, c_n)$ es solución del sistema original, entonces:

$$a_{k1}c_1 + a_{k2}c_2 + \cdots + a_{kn}c_n = b_k, \text{ para todo } k = 1, 2, \ldots, m \tag{1.23}$$

en particular se cumplen las ecuaciones:

$$a_{i1}c_1 + a_{i2}c_2 + \cdots + a_{in}c_n = b_i \tag{1.24}$$

y

$$a_{j1}c_1 + a_{j2}c_2 + \cdots + a_{jn}c_n = b_{\mathrm{j}} \tag{1.25}$$

Multiplicando la ecuación 1.24 por a, sumándola a (1.25) y agrupando se obtiene la ecuación 1.22 evaluada en $(c_1, c_2, \ldots, c_n)$, concluyendo de esto que $(c_1, c_2, \ldots, c_n)$ es solución del nuevo sistema.

Recíprocamente, si $(c_1, c_2, \ldots, c_n)$ es solución del nuevo sistema, entonces:

$$a_{k1}c_1 + a_{k2}c_2 + \cdots + a_{kn}c_n = b_k, \quad \text{para todo} \quad k \neq j \tag{1.26}$$

y

$$(a_{j1} + aa_{i1})c_1 + (a_{j2} + aa_{i2})c_2 + \cdots + (a_{jn} + aa_{in})c_n = b_j + ab_i \tag{1.27}$$

Tomando $k = i$ en (1.26), multiplicándola por a y restándola a (1.27) se obtiene:

$$a_{j1}c_1 + a_{j2}c_2 + \cdots + a_{jn}c_n = b_j \tag{1.28}$$

concluyendo con esto que $(c_1, c_2, \ldots, c_n)$ es solución del sistema original.

En la discusión del ejemplo de las refinerías usamos esencialmente lo que se conoce como el *Método de Reducción de Gauss-Jordan*. Lo que haremos abajo es representar el sistema de ecuaciones mediante una matriz y aplicar el método usando operaciones elementales en las filas.

Recordemos que el sistema discutido es:

$$\begin{aligned} 20x_1 + 21x_2 + 19x_3 &= 1\,250 \\ 11x_1 + 12x_2 + 13x_3 &= 750 \\ 9x_1 + 8x_2 + 8x_3 &= 520, \end{aligned}$$

por lo que su matriz aumentada resulta ser: $\begin{bmatrix} 20 & 21 & 19 & 1\,250 \\ 11 & 12 & 13 & 750 \\ 9 & 8 & 8 & 520 \end{bmatrix}$

En lo que sigue aplicaremos operaciones elementales a esta matriz para resolver el sistema de ecuaciones. Pedimos al lector que identifique el tipo de operaciones que se aplican al pasar de una matriz a la siguiente.

$$\begin{bmatrix} 20 & 21 & 19 & 1\,250 \\ 11 & 12 & 13 & 750 \\ 9 & 8 & 8 & 520 \end{bmatrix} \sim \begin{bmatrix} 20 & 21 & 19 & 1\,250 \\ 2 & 4 & 5 & 230 \\ 9 & 8 & 8 & 520 \end{bmatrix} \sim \begin{bmatrix} 0 & -19 & -31 & -1\,050 \\ 2 & 4 & 5 & 230 \\ 1 & -8 & -12 & -400 \end{bmatrix} \sim$$

$$\begin{bmatrix} 0 & 19 & 31 & 1\,050 \\ 0 & 20 & 29 & 1\,030 \\ 1 & -8 & -12 & -400 \end{bmatrix} \sim \begin{bmatrix} 0 & 19 & 31 & 1\,050 \\ 0 & 1 & -2 & -20 \\ 1 & -8 & -12 & -400 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 69 & 1\,430 \\ 0 & 1 & -2 & -20 \\ 1 & 0 & -28 & -560 \end{bmatrix} \sim$$

$$\begin{bmatrix} 0 & 0 & 1 & 430/69 \\ 0 & 1 & -2 & -20 \\ 1 & 0 & -28 & -560 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 1 & 1\,430/69 \\ 0 & 1 & 0 & 1\,480/69 \\ 1 & 0 & 0 & 1\,400/69 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1\,430/69 \\ 0 & 1 & 0 & 1\,480/69 \\ 0 & 0 & 1 & 1\,400/69 \end{bmatrix}$$

Al estudiar un sistema de ecuaciones lineales, de manera natural surgen dos preguntas fundamentales.

1. ¿Cuándo tiene solución un sistema de ecuaciones lineales?
2. ¿Cómo encontrar las soluciones de un sistema de ecuaciones lineales?

Iniciamos la discusión de las preguntas planteadas presentando algunos ejemplos.

Ejemplo 1.3.3. *La matriz* $\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{bmatrix}$ *representa un sistema de ecuaciones resuelto.*

La interpretación es la siguiente. Si las variables son x, y *y* z, *la primera, segunda y tercera filas representan a las ecuaciones* $x + 0y + 0z = 4$, $0x + y + 0z = -1$ *y* $0x + 0y + z = 3$ *respectivamente, en otras palabras la solución del sistema es la triada* $(x, y, z) = (4, -1, 3)$.

Ejemplo 1.3.4. *Si la matriz anterior se cambia a* $\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, *entonces el sistema no tiene solución, pues la última ecuación no se satisface para cualesquiera que sean los valores de las variables.*

Ejemplo 1.3.5. *La siguiente matriz representa un sistema con muchas soluciones* (*¿por qué?*).

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

Los ejemplos anteriores ilustran cómo la matriz aumentada de un sistema representa la información relacionada con las soluciones, entonces, decidir si un sistema tiene solución se reduce a establecer algunas propiedades que debe tener la matriz después de haberle aplicado sucesivamente operaciones elementales. Note que las matrices de los ejemplos tienen ciertas características que precisamos en la siguiente definición.

Definición 1.3.5. *Una matriz se dice que está en forma escalonada reducida, si cumple las siguientes condiciones:*

1. *Las filas cero, de existir, se encuentran al final.*
2. *La primera entrada no cero, de izquierda a derecha, de una fila no cero es uno* (*entrada principal*).
3. *La columna que contiene a la entrada principal de una fila, tiene solamente ceros, excepto la entrada principal.*
4. *Las entradas principales aparecen en forma escalonada, es decir, si hay* r *filas no cero y la entrada principal de la fila número* i *se encuentra en la columna* k_i, *entonces se tiene* $k_1 < k_2 < \cdots < k_r$.

Si una matriz satisface las condiciones estipuladas en la definición 1.3.5, diremos que se encuentra en *forma escalonada reducida*.

Ejercicio 1.3.2. *La matriz* $A = \begin{pmatrix} 1 & 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 4 & 7 \\ 0 & 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$ *no se encuentra en forma escalonada reducida, ¿cuáles condiciones de la definición no cumple? Aplique operaciones elementales en sus filas para llevarla a forma escalonada reducida.*

Ejercicio 1.3.3. *Suponga que se tiene un sistema de* n *ecuaciones lineales en* n *variables y que la matriz aumentada del sistema se ha llevado, por medio de operaciones elementales, a una escalonada reducida en la que no hay filas cero, ¿tiene solución el sistema?*

En el proceso de resolver un sistema de ecuaciones, las respuestas a las siguientes preguntas son importantes y se relacionan con las planteadas anteriormente. ¿Qué condición deben cumplir las filas de una matriz escalonada reducida para que el sistema que representa tenga solución? ¿Cómo saber si un sistema tiene muchas soluciones? ¿Cómo saber si un sistema tiene una única solución? ¿Se puede llevar cualquier matriz no cero, mediante operaciones elementales, a una en forma escalonada reducida?

Ejercicio 1.3.4. *Construya ejemplos que ilustren las posibles respuestas a las preguntas planteadas. Compare sus conclusiones con las del siguiente teorema.*

Teorema 1.3.2. (método de reducción de Gauss-Jordan) *Toda matriz* $A \neq 0$ *se puede llevar, mediante operaciones elementales en sus filas, a una matriz escalonada reducida.*

Demostración. Como A es no cero, algún $a_{ij} \neq 0$; podemos suponer que j es el menor índice que cumple la condición, para algún i, esto significa que las columnas con índice menor que j son cero, es decir, la matriz A luce como:

$$A = \begin{bmatrix} 0 & \cdots & 0 & a_{1j} & \cdots & a_{1n} \\ 0 & \cdots & 0 & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

Si $i = 1$, multiplicando la primera fila por a_{1j}^{-1}, el elemento que resulta en la posición $(1, j)$ es 1. Ahora, para $i = 2, \ldots, m$ multiplicamos la primer fila por $-a_{ij}$ y la sumamos a la fila i obteniendo la matriz:

$$A_1 = \begin{bmatrix} 0 & \cdots & 1 & b_{1j+1} & \cdots & b_{1n} \\ 0 & \cdots & 0 & b_{2j+1} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & b_{mj+1} & \cdots & b_{mn} \end{bmatrix}$$

Si $i \neq 1$, intercambiamos la fila i con la primera y aplicamos el mismo procedimiento, obteniendo una matriz de la misma forma que A_1.

El siguiente paso es examinar la submatriz de A_1, que tiene entradas b_{ik} con $i \geq 2$ y $k > j$.

Si esta submatriz es cero hemos terminado, en caso contrario hay un $k > j$ y un $i \geq 2$, tales que $b_{ik} \neq 0$. Podemos suponer que k es el menor de tales índices con esta propiedad.

Si $i = 2$ multiplicamos la fila 2 de A_1 por b_{2k}^{-1} y después de esto, para $i \neq 2$, multiplicamos la fila 2 por $-b_{ik}$ y la sumamos a la fila i obteniendo la matriz:

$$A_2 = \begin{bmatrix} 0 & \cdots & 1 \cdots 0 & c_{1k+1} & \cdots & c_{1n} \\ 0 & \cdots & 0 \cdots 1 & c_{2k+1} & \cdots & c_{2n} \\ \vdots & \vdots & \vdots \; \vdots \; \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \cdots 1 & c_{mk+1} & \cdots & c_{mn} \end{bmatrix}$$

Si $i \neq 2$, intercambiamos la fila dos con la i-ésima y aplicamos el procedimiento descrito.

Note que las primeras columnas de A_2 están en forma escalonada. Continuando el proceso con A_2, se concluye que termina en un número finito de etapas, produciendo una matriz en la forma deseada.

Ejemplo 1.3.6. *Haga una discusión del tipo de soluciones que tiene el sistema de ecuaciones:*

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

Discusión. Primeramente construimos la matriz aumentada del sistema; después la llevamos a forma escalonada reducida, y finalmente en esta matriz "leemos" el tipo de soluciones que tiene el sistema.

La matriz aumentada es:

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

Aplicando operaciones elementales en las filas de B se tiene la sucesión de matrices que aparecen en las siguientes líneas. Identifique el tipo de operaciones elementales efectuadas en cada caso.

$$\begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 0 & 3 & 4 \\ -1 & 1 & 3 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & -4 & -3 & 6 \\ 0 & 3 & 6 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & -4 & -3 & 6 \\ 0 & -1 & 3 & 5 \end{bmatrix} \sim$$

$$\begin{bmatrix} 1 & 0 & 9 & 9 \\ 0 & 0 & -15 & -14 \\ 0 & -1 & 3 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 9 & 9 \\ 0 & 0 & 1 & 14/15 \\ 0 & -1 & 3 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 3/5 \\ 0 & 0 & 1 & 14/15 \\ 0 & -1 & 0 & 11/5 \end{bmatrix} \sim$$

$$\begin{bmatrix} 1 & 0 & 0 & 3/5 \\ 0 & 1 & 0 & -11/15 \\ 0 & 0 & 1 & 14/5 \end{bmatrix}$$

De acuerdo con los cálculos efectuados, de esta última matriz se tiene que el sistema de ecuaciones tiene solución única: $C = (3/5, -11/5, 14/15)$. Para tener certeza de que se trata de una solución única, efectuaremos los cálculos necesarios para verificar que C es solución y de esto se concluirá que C es la única solución (explique). Sustituyendo las coordenadas de C en las ecuaciones se tiene:

$$\begin{aligned} 3/5 + 2(-11/5) + 3(14/15) &= -1 \\ 2(3/5) + 3(14/15) &= 4 \\ -3/5 - 11/5 + 3(14/15) &= 0 \end{aligned}$$

concluyendo lo esperado.

1.3.1. Análisis de las soluciones de un sistema de ecuaciones lineales

Conociendo la forma escalonada reducida de la matriz aumentada de un sistema de ecuaciones lineales, es relativamente sencillo determinar si el sistema tiene solución y de ser así, determinar si tiene muchas o solamente una.

Sea:

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

la matriz aumentada de un sistema y R su forma escalonada reducida. En lo que sigue se hará un análisis del tipo de soluciones que tiene el sistema a partir de R.

El sistema no tiene solución

Si la entrada principal de una fila no cero de R se encuentra en la columna $n + 1$, el sistema no tiene solución, pues en este caso dicha fila representaría una ecuación en la que los coeficientes de las variables son todos cero y el término independiente es no cero, lo cual significa que esa ecuación no se satisface para cualquier valor de las variables, concluyéndose que el sistema no tiene solución.

El sistema tiene solución

Supongamos que la entrada principal de la última fila no cero de R se encuentra en la columna $j < n + 1$, entonces el sistema representado por A tiene solución. Más precisamente, si R tiene l filas no cero, con entradas principales en las columnas $k_1 < k_2 < \cdots < k_l$, entonces las ecuaciones que representa R son de la forma:

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

en donde las variables $u_1, u_2, \ldots, u_{n-l}$, llamadas *variables libres*, son un reordenamiento de las originales que son diferentes de $x_{k1}, x_{k2}, \ldots, x_{kl}$. Todas las soluciones del sistema representado por A se obtienen dando valores a $u_1, u_2, \ldots, u_{n-l}$ y despejando $x_{k1}, \ldots, x_{kl}$ de las ecuaciones correspondientes.

Ejemplo 1.3.7. *Sea* $A = \begin{bmatrix} 1 & 0 & 0 & 0 & 3 & 4 \\ 0 & 1 & 2 & 0 & 2 & 5 \\ 0 & 0 & 0 & 1 & -1 & 6 \end{bmatrix}$ *la matriz aumentada de un sistema de ecuaciones en las variables* x_1, x_2, x_3, x_4 *y* x_5. *Escriba el sistema representado por A y diga si tiene soluciones, en caso afirmativo encuéntrelas todas.*

Discusión. El sistema representado por A es:

$$\begin{aligned} x_1 + 0x_2 + 0x_3 + 0x_4 + 3x_5 &= 4 \\ 0x_1 + x_2 + 2x_3 + 0x_4 + 2x_5 &= 5 \\ 0x_1 + 0x_2 + 0x_3 + x_4 - x_5 &= 6 \end{aligned} \tag{1.29}$$

Note que la matriz A está en forma escalonada reducida; la entrada principal de la última fila no cero de A está antes de la última columna, por lo que el sistema tiene solución. Las variables libres son x_3 y x_5 (¿por qué?). Entonces, todas las soluciones se obtienen despejando x_1, x_2 y x_4 del sistema 1.29 y dando valores libremente a x_3 y a x_5.

Solución única

De la discusión anterior se tiene que si el sistema tiene solución, ésta será única $\Leftrightarrow$ no hay variables libres, y esto último ocurre $\Leftrightarrow$ el número de filas no cero es igual al número de variables, en particular, si $n = m$, el sistema tiene solución única $\Leftrightarrow$ las entradas principales de la forma escalonada reducida de A aparecen en la posición (i, i), para $i = 1, 2, \ldots, n$.

Sistema homogéneo

Un caso de importancia en el estudio de sistemas de ecuaciones lineales ocurre cuando todos los términos independientes son cero, es decir, el sistema es de la forma:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned} \tag{1.30}$$

Note que el sistema 1.30 siempre tiene solución, entonces es importante saber cuándo tiene más de una solución. El siguiente resultado establece condiciones al respecto.

Teorema 1.3.3. *Sea* A *la matriz aumentada de un sistema homogéneo de* m *ecuaciones en* n *variables.*

1. *Si* m < n*, entonces el sistema tiene más de una solución.*
2. *Si el sistema tiene solución única, entonces* n ≤ m.

Demostración. Cuando decimos que un sistema es en n variables, esto significa que cada variable aparece en al menos una ecuación, lo cual se traduce diciendo que en la matriz A, ninguna de las n primeras columnas es cero. Note que si una columna es no cero, entonces por medio de operaciones elementales en las filas tampoco puede transformarse en cero. Esto lleva a que la forma escalonada reducida de A tiene todas las primeras n columnas diferentes de cero.

Si $m < n$, por el principio de las casillas,[4] al menos una fila de la forma escalonada reducida de A contiene dos o más entradas no cero, es decir, el sistema tiene al menos una variable libre, concluyéndose que tiene más de una solución.

Solución general de un sistema no homogéneo

Una ecuación típica del sistema representado por A es:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \tag{1.31}$$

Supongamos que el sistema tiene solución y denotemos a una de éstas por $C = (c_1, c_2, \ldots, c_n)$. De la definición de solución se tiene:

$$a_{i1}c_1 + a_{i2}c_2 + \cdots + a_{in}c_n = b_i \tag{1.32}$$

para todo $i = 1, 2, \ldots, m$.

Sea A_1 la matriz que tiene sus primeras n columnas iguales a las de A, y la última es cero. Podemos considerar que A_1 representa un sistema homogéneo. La ecuación número i de este sistema es:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0 \tag{1.33}$$

Sea $D = (d_1, d_2, \ldots, d_n)$ cualquier solución del sistema representado por A_1, es decir,

$$a_{i1}d_1 + a_{i2}d_2 + \cdots + a_{in}d_n = 0 \tag{1.34}$$

para todo $i = 1, 2, \ldots, m$.

Sumando las ecuaciones (1.32) y (1.34) se tiene que $C + D := (c_1 + d_1, c_2 + d_2, \ldots, c_n + d_n)$ es solución del sistema representado por A.

Si $K = (k_1, k_2, \ldots, k_n)$ es otra solución del sistema representado por A, entonces:

$$a_{i1}k_1 + a_{i2}k_2 + \cdots + a_{in}k_n = b_i \tag{1.35}$$

para todo $i = 1, 2, \ldots, m$. Restando la ecuación (1.32) de la ecuación (1.35) y agrupando se tiene:

$$a_{i1}(k_1 - c_1) + a_{i2}(k_2 - c_2) + \cdots + a_{in}(k_n - c_n) = 0 \tag{1.36}$$

es decir, $K - C := (k_1 - c_1, k_2 - c_2, \ldots, k_n - c_n)$ es solución del sistema homogéneo representado por A_1.

[4] El principio de las casillas establece que si hay n objetos que se han de colocar en m casillas y $m < n$, entonces al menos una casilla contiene más de un objeto.

La discusión anterior es la demostración del siguiente teorema.

Teorema 1.3.4. *Sea* A *la matriz aumentada de un sistema de* m *ecuaciones en* n *variables y* A_1 *la matriz cuyas primeras* n *columnas coinciden con las de* A *y la última es cero. Si el sistema representado por* A *tiene una solución particular* $C = (c_1, c_2, \ldots, c_n)$ *y el sistema homogéneo representado por* A_1 *tiene por conjunto solución a* W, *entonces el conjunto solución del sistema representado por* A *es*

$$W + C := \{w + C \mid w \in W\}$$

Ejemplo 1.3.8. *Determine si el sistema:*

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 &= 10 \\ 2x_1 + 4x_2 + 6x_3 + 8x_4 + 10x_5 &= 20 \\ -x_1 - 2x_2 - 2x_3 - x_4 + 3x_5 &= 4 \end{aligned}$$

tiene solución. En caso afirmativo, obténgalas todas.

Discusión. La matriz aumentada es $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 10 \\ 2 & 4 & 6 & 8 & 10 & 20 \\ -1 & -2 & -2 & -1 & 3 & 4 \end{pmatrix}$

Aplicando operaciones elementales se tiene:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 10 \\ 2 & 4 & 6 & 8 & 10 & 20 \\ -1 & -2 & -2 & -1 & 3 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 10 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 8 & 14 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 10 \\ 0 & 0 & 1 & 3 & 8 & 14 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 & -5 & -19 & -32 \\ 0 & 0 & 1 & 3 & 8 & 14 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

La última matriz representa al sistema:

$$\begin{aligned} x_1 + 2x_2 + 0x_3 - 5x_4 - 19x_5 &= -32 \\ 0x_1 + 0x_2 + x_3 + 3x_4 + 8x_5 &= 14 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 &= 0 \end{aligned}$$

Es inmediato verificar que $C = (-32, 0, 14, 0, 0)$ es solución de este sistema.

De acuerdo con el teorema 1.3.4, todas las soluciones del sistema anterior se obtienen sumando $C = (-32, 0, 14, 0, 0)$ a las soluciones del sistema homogéneo:

$$\begin{aligned} x_1 + 2x_2 + 0x_3 - 5x_4 - 19x_5 &= 0 \\ 0x_1 + 0x_2 + x_3 + 3x_4 + 8x_5 &= 0 \end{aligned}$$

En este último sistema hay tres variables libres, x_2, x_4 y x_5 por lo que sus soluciones están dadas por:

$$\begin{aligned} x_2 &= x_2 \\ x_4 &= x_4 \\ x_5 &= x_5 \\ x_1 &= -2x_2 + 5x_4 + 19x_5 \\ x_3 &= -3x_4 - 8x_5 \end{aligned}$$

En forma de conjunto, estas condiciones se expresan por:

$$W = \{(-2x_2 + 5x_4 + 19x_5, x_2, -3x_4 - 8x_5, x_4, x_5) \mid x_2, x_3, x_5 \in \mathbb{R}\}$$

Resumiendo, la soluciones del sistema:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 &= 10 \\ 2x_1 + 4x_2 + 6x_3 + 8x_4 + 10x_5 &= 20 \\ -x_1 - 2x_2 - 2x_3 - x_4 + 3x_5 &= 4 \end{aligned}$$

están dadas por:

$$\begin{aligned} W &= \{(-2x_2 + 5x_4 + 19x_5, x_2, -3x_4 - 8x_5, x_4, x_5) \mid x_2, x_3, x_5 \in \mathbb{R}\} + (-32, 0, 14, 0, 0) \\ &= \{(-2x_2 + 5x_4 + 19x_5 - 32, x_2 + x_4 - 8x_5 + 14, x_4, x_5) \mid x_2, x_3, x_5 \in \mathbb{R}\} \end{aligned}$$

Ejemplo 1.3.9. *Tráfico vehicular en calles de un solo sentido con cuatro puntos de intersección.*

Supongamos que se tiene un circuito formado por cuatro calles como se ilustra en la figura 1.4. Adicionalmente, supongamos que las cantidades de vehículos que entran a dicho circuito, son b_1 *y* b_3 *y las que salen son* b_2 *y* b_4*. Es natural suponer que la cantidad de vehículos que entran al circuito y la cantidad que salen es la misma, es decir,* $b_1 + b_3 = b_2 + b_4$*. Nos interesa determinar la forma en que se distribuye el tráfico en el circuito.*

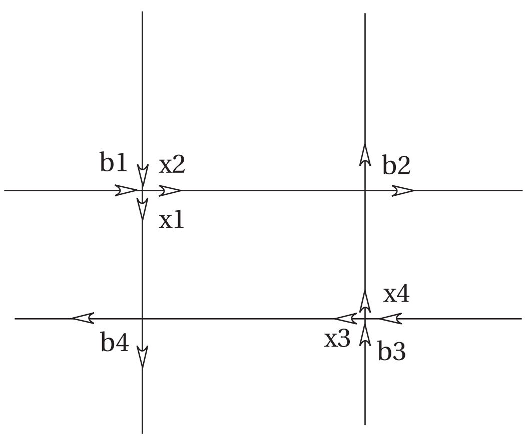


Figura 1.4. Distribución de tráfico en los nodos del circuito.

Discusión. Si denotamos a las cantidades de vehículos que circulan en el circuito por x_1, x_2, x_3 y x_4, el análisis en cada nodo da lugar a una ecuación lineal, es decir, se tiene el siguiente sistema de ecuaciones.

$$\begin{aligned} x_1 + x_2 &= b_1 \\ x_2 + x_4 &= b_2 \\ x_3 + x_4 &= b_3 \\ x_1 + x_3 &= b_4 \end{aligned} \tag{1.37}$$

La matriz aumentada del sistema 1.37 es:

$$\begin{pmatrix} 1 & 1 & 0 & 0 & b_1 \\ 0 & 1 & 0 & 1 & b_2 \\ 0 & 0 & 1 & 1 & b_3 \\ 1 & 0 & 1 & 0 & b_4 \end{pmatrix}$$

Aplicando el método de reducción de Gauss-Jordan se tienen las siguientes matrices equivalentes por filas. Identifique el tipo de operación que se efectuó al pasar de una matriz a la siguiente.

$$\begin{aligned} \begin{pmatrix} 1 & 1 & 0 & 0 & b_1 \\ 0 & 1 & 0 & 1 & b_2 \\ 0 & 0 & 1 & 1 & b_3 \\ 1 & 0 & 1 & 0 & b_4 \end{pmatrix} &\sim \begin{pmatrix} 1 & 1 & 0 & 0 & b_1 \\ 0 & 1 & 0 & 1 & b_2 \\ 0 & 0 & 1 & 1 & b_3 \\ 0 & -1 & 1 & 0 & b_4 - b_1 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 1 & 0 & 0 & b_1 \\ 0 & 0 & 1 & 1 & b_2 + b_4 - b_1 \\ 0 & 0 & 1 & 1 & b_3 \\ 0 & -1 & 1 & 0 & b_4 - b_1 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 1 & 0 & 0 & b_1 \\ 0 & 0 & 1 & 1 & b_2 + b_4 - b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_1 - b_2 - b_4 \\ 0 & -1 & 1 & 0 & b_4 - b_1 \end{pmatrix} \end{aligned} \tag{1.38}$$

La hipótesis $b_1 + b_3 = b_2 + b_4$ implica que la tercera fila de la última matriz es cero, por lo que intercambiando dicha fila con la última se obtiene la matriz:

$$\begin{pmatrix} 1 & 1 & 0 & 0 & b_1 \\ 0 & 0 & 1 & 1 & b_2 + b_4 - b_1 \\ 0 & -1 & 1 & 0 & b_4 - b_1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Después de aplicar algunas operaciones elementales (¿cuáles?) a esta última matriz se llega a:

$$\begin{pmatrix} 1 & 0 & 0 & -1 & b_4 - b_3 \\ 0 & 1 & 0 & 1 & b_2 \\ 0 & 0 & 1 & 1 & b_3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

que está en forma escalonada reducida, por lo que el sistema inicial equivale a:

$$\begin{aligned} x_1 - x_4 &= b_4 - b_3 \\ x_2 + x_4 &= b_2 \\ x_3 + x_4 &= b_3 \end{aligned} \tag{1.39}$$

y tiene muchas soluciones. En términos de tráfico, ¿cuál es la interpretación de este resultado?

Ejercicio 1.3.5. *En cada uno de los siguientes casos determine si los sistemas representados por las correspondientes matrices aumentadas tienen solución única, muchas soluciones o son inconsistentes.*

1. $\begin{pmatrix} 2 & -1 & 2 \\ -4 & 1 & 5 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 2 & -1 & 0 \\ -2 & 3 & 2 & -1 \\ 5 & 2 & 0 & 4 \end{pmatrix}$

3. $\begin{pmatrix} -1 & 1 & 1 & -1 & 2 \\ 1 & 0 & 3 & -3 & 1 \\ 2 & 1 & 0 & 1 & 0 \\ -1 & 2 & 0 & 1 & 0 \end{pmatrix}$

4. *Determine los valores de* a *de forma que el sistema representado por la matriz* $\begin{pmatrix} a & -1 & 1 & 0 \\ 1 & 1-a & -1 & 1 \\ -1 & 1 & 1 & 2 \end{pmatrix}$, *sea consistente, inconsistente, tenga solución única o infinidad de soluciones.*

1.3.2. Ejemplos con *Maple*

El siguiente ejemplo ilustra el uso de *Maple* para aplicar operaciones elementales en las filas de una matriz y llevarla a una en forma escalonada reducida.

Recuerde que para efectuar ciertos comandos en *Maple*, primero hay que *cargar* el paquete correspondiente; en este caso, el que usaremos es *linalg* y se activa con el comando:

```
>   with(linalg):
```

Para declarar una matriz en *Maple*, puede usar la instrucción:

```
>   A := array( [[3,2,0,4],[1,4,sqrt(2),3],[2,3,4,5]] );
```

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

Los siguientes comandos producen operaciones elementales en las filas de una matriz. Por ejemplo, el siguiente intercambia las filas uno y dos en la matriz *A*.

```
>   A1:=swaprow(A, 1, 2);
```

$$A1 := \begin{bmatrix} 1 & 4 & \sqrt{2} & 3 \\ 3 & 2 & 0 & 4 \\ 2 & 3 & 4 & 5 \end{bmatrix}$$

El siguiente comando produce una matriz que se obtiene de *A*l multiplicando la fila uno por -3 y sumándola a la dos:

```
>   A2:=addrow(A1,1,2,-3);
```

$$A2 := \begin{bmatrix} 1 & 4 & \sqrt{2} & 3 \\ 0 & -10 & -3\sqrt{2} & -5 \\ 2 & 3 & 4 & 5 \end{bmatrix}$$

La matriz *A*3 se obtiene de *A*2 multiplicando la fila uno por -2 y sumándola a la fila tres.

```
>   A3:=addrow(A2,1,3,-2);
```

$$A3 := \begin{bmatrix} 1 & 4 & \sqrt{2} & 3 \\ 0 & -10 & -3\sqrt{2} & -5 \\ 0 & -5 & -2\sqrt{2}+4 & -1 \end{bmatrix}$$

Multiplicando la fila tres de *A*3 por -2 y sumándola a la fila dos se obtiene *A*4.

```
>   A4:=addrow(A3,3,2,-2);
```

$$A4 := \begin{bmatrix} 1 & 4 & \sqrt{2} & 3 \\ 0 & 0 & \sqrt{2}-8 & -3 \\ 0 & -5 & -2\sqrt{2}+4 & -1 \end{bmatrix}$$

La matriz $A5$ se obtiene de $A4$ multiplicando la fila tres por $-1/5$.

```
>  A5:=mulrow(A4, 3, -1/5);
```

$$A5 := \begin{bmatrix} 1 & 4 & \sqrt{2} & 3 \\ 0 & 0 & \sqrt{2}-8 & -3 \\ 0 & 1 & \frac{2}{5}\sqrt{2}-\frac{4}{5} & \frac{1}{5} \end{bmatrix}$$

Multiplicando la fila tres de $A5$ por -4 y sumándola a la fila uno se obtiene:

```
>  A6:=addrow(A5,3, 1, -4);
```

$$A6 := \begin{bmatrix} 1 & 0 & -\frac{3}{5}\sqrt{2}+\frac{16}{5} & \frac{11}{5} \\ 0 & 0 & \sqrt{2}-8 & -3 \\ 0 & 1 & \frac{2}{5}\sqrt{2}-\frac{4}{5} & \frac{1}{5} \end{bmatrix}$$

```
>  A7:=swaprow(A6, 3, 2);
```

$$A7 := \begin{bmatrix} 1 & 0 & -\frac{3}{5}\sqrt{2}+\frac{16}{5} & \frac{11}{5} \\ 0 & 1 & \frac{2}{5}\sqrt{2}-\frac{4}{5} & \frac{1}{5} \\ 0 & 0 & \sqrt{2}-8 & -3 \end{bmatrix}$$

¿Cómo se obtuvo $A7$? Continúe el proceso hasta llevar la matriz original a una en forma escalonada reducida y conteste las siguientes preguntas.

1. ¿Cuál es el sistema que representa la matriz A?
2. ¿Cuál es el sistema que representa la forma escalonada reducida?
3. ¿Tiene solución el sistema que representa A?

En la discusión de los siguientes ejercicios se sugiere usar *Maple*.

1.4. Ejercicios

1. En una oficina se dispone de \$100 000 pesos para adquirir equipo de cómputo. Si el precio de una computadora es de \$5 000 pesos, y el de una impresora es \$1 500 pesos, ¿cuántas computadoras y cuántas impresoras se pueden comprar? Una posibilidad es que se compren 20 computadoras y cero impresoras, ¿es ésta una respuesta razonable? Proponga una solución que usted considere apropiada.

2. Resuelva los siguientes sistemas mediante operaciones elementales en la matriz aumentada.

$$\begin{aligned} x+y&=3\\ 3x+y&=0 \end{aligned}\qquad \begin{aligned} x-y&=3\\ 3x+2y&=2 \end{aligned}\qquad \begin{aligned} x-y+z&=3\\ 3x+2y-z&=2\\ x-y+2z&=0 \end{aligned}\qquad \begin{aligned} 2x-3y+8z&=4\\ 3x+2y-2z&=5\\ 4x-y+2z&=-1 \end{aligned}$$

3. Suponga que $\begin{bmatrix} 1 & 0 & \lambda & 1\\ 0 & \lambda & \lambda & 0\\ \lambda & 2 & 1 & -1 \end{bmatrix}$ es la matriz aumentada de un sistema. Determine los valores de λ de tal forma que el sistema:

a) Tenga solución única.

b) Tenga infinidad de soluciones.

c) Sea inconsistente.

4. Suponga que las siguientes son matrices aumentadas de sistemas de ecuaciones. Determine si los sistemas son consistentes o inconsistentes.

$$\begin{pmatrix} 1 & 3 & 0 & 5 & -5\\ 0 & 1 & -6 & 9 & 0\\ 0 & 0 & 2 & 7 & 1\\ 0 & 0 & 1 & 4 & -2 \end{pmatrix},\ \begin{pmatrix} 1 & 3 & 0 & 5 & -5\\ 0 & 3 & -6 & 9 & 0\\ 0 & 0 & 2 & 0 & 8\\ 0 & 0 & 1 & 4 & -16 \end{pmatrix},\ \begin{pmatrix} 3 & 1 & -1 & 2 & 4\\ 2 & -1 & 2 & 5 & 0\\ 5 & 4 & 4 & -2 & 6\\ 7 & 0 & 1 & -1 & 0 \end{pmatrix}$$

5. Una compañía dispone de un presupuesto de \$280 000 dólares para equipo de cómputo. Se han de adquirir tres tipos de productos: computadoras personales a \$500 dólares cada una; impresoras a \$2 000 dólares cada una y computadoras portátiles a \$5 000 dólares cada una. Se deben comprar cinco computadoras personales por cada impresora y por cada computadora portátil dos computadoras personales. Determine si es posible gastar el presupuesto de forma exacta.

6. Decida si los siguientes enunciados son verdaderos, construyendo ejemplos o contraejemplos que le convenzan; después convenza a alguien de lo válido de sus argumentos.

a) En un sistema de ecuaciones lineales en las variables x, y y z las soluciones, de existir, son triadas.

b) Un sistema de ecuaciones lineales es inconsistente si tiene más de una solución.

c) Hay sistemas de ecuaciones lineales con más de una solución.

d) Para resolver un sistema de ecuaciones se requiere que el número de ecuaciones sea igual al número de variables.

e) Si la matriz aumentada de un sistema tiene en su última columna solamente ceros el sistema es inconsistente.

f) Si la forma escalonada reducida de una matriz aumentada tiene una fila con entrada principal en la última columna, el sistema es consistente.

g) Un sistema de ecuaciones lineales puede tener exactamente dos soluciones.

h) Si el número de ecuaciones es menor que el número de incógnitas el sistema tiene infinidad de soluciones.

i) Si los coeficientes de un sistema lineal consistente son racionales, entonces las soluciones son racionales.

j) Si un sistema de ecuaciones lineales tiene solución única, entonces el número de ecuaciones es igual al número de variables.

7. Las siguientes son matrices aumentadas de un sistema de ecuaciones lineales. Determine los valores de h de tal forma que sea consistente.

a) $\begin{pmatrix} 1 & 4 & 2 \\ 5 & h & 0 \end{pmatrix}$ *b*) $\begin{pmatrix} 1 & 4 & 2 \\ h & 1 & -1 \end{pmatrix}$ *c*) $\begin{pmatrix} 1 & -2+h & 2 \\ 5 & 0 & 0 \end{pmatrix}$

8. Determine los valores de h y k de tal forma que las matrices representen sistemas con: solución única, con muchas soluciones, e inconsistentes.

a) $\begin{pmatrix} 1 & h & 1 \\ 2 & 3 & k \end{pmatrix}$ *b*) $\begin{pmatrix} 1 & -3 & 1 \\ 2 & -h & k \end{pmatrix}$ *c*) $\begin{pmatrix} k & 4 & 2 \\ 5 & -h & 1 \end{pmatrix}$

9. Escriba un programa, en el lenguaje de su preferencia, para resolver sistemas de dos ecuaciones en dos variables.

10. Determine los coeficientes de un polinomio de grado tres cuya gráfica pase por los puntos (1, 2), (2, 0), (−1, 0) y (−2, 4). Explique por qué se necesitan cuatro puntos para determinar un polinomio de grado tres y solamente dos para determinar un polinomio de grado uno.

11. Se tiene una red de distribución de agua como se ilustra en la figura 1.5. Suponga que la cantidad que entra es la misma que sale, es decir, se tiene $f_1 + f_2 = f_3 + f_4$. Determine la cantidad de agua que fluye en cada tubo.

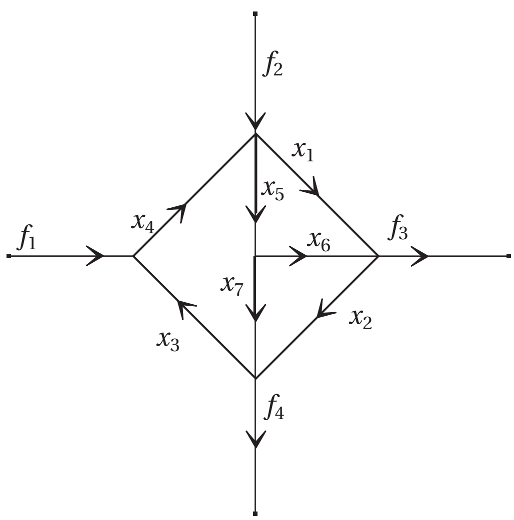


Figura 1.5. Red de agua.

12. Suponga que una economía tiene tres sectores, agricultura, minería y manufactura. Agricultura vende 5% de su producción a minería, 30% a manufactura y retiene el resto. Minería vende 20% de su producción a agricultura, 70% a manufactura y retiene el resto. Manufactura vende 20% de su producción a agricultura, 30% a minería y retiene el resto. Encuentre un conjunto de precios de equilibrio para la economía, es decir, precios para los cuales los ingresos de cada sector sean iguales a sus gastos.

13. Suponga que en el modelo propuesto por Leontief hay n sectores, el valor de lo que produce el sector i es denotado por x_i y el público demanda, de ese sector, b_i pesos. Proponga una notación adecuada de las cantidades que requiere cada sector, de los restantes, para llevar a cabo su producción; formule el modelo mediante un sistema de ecuaciones. ¿Tendrá solución el sistema?

14. Suponga que la matriz aumentada de un sistema de ecuaciones es de tamaño $n \times (n+1)$ y su forma escalonada reducida no tiene filas de ceros. ¿Es consistente el sistema representado por la matriz? Proporcione ejemplos para $n = 2, 3, 4, 5$.

15. Proponga ejemplos en los que se muestre que el orden en que se apliquen operaciones elementales produce resultados diferentes.

16. Suponga que se tiene un sistema inconsistente de m ecuaciones en n variables. Haciendo los términos constantes cero, resulta un sistema homogéneo. ¿Cuántas soluciones tiene el nuevo sistema?

Capítulo 2

Matrices

En matemáticas, los diferentes sistemas numéricos se han creado por la necesidad de resolver problemas concretos, ya sea en la propia disciplina o en otras relacionadas con ésta. Por ejemplo, los números complejos son introducidos por Cardano (1545) al resolver ecuaciones cúbicas. De igual forma, las matrices aparecen al resolver problemas relacionados con sistemas de ecuaciones lineales. Entre los iniciadores del estudio sistemático de las matrices se cuenta a: Cramer, Bezout, Vandermonde, Lagrange, Gauss, Cauchy, Cayley, Jordan, Sylvester y Frobenius.

En el capítulo 1 usamos las matrices para resolver sistemas de ecuaciones lineales. Esto nos permitió contestar, de manera eficiente, preguntas fundamentales de álgebra lineal relacionadas con la solución de sistemas de ecuaciones lineales. En lo que sigue, las matrices serán usadas, entre otras cosas, para representar y "operar" con información. Esto llevará a considerar a las matrices no solamente relacionadas con sistemas de ecuaciones, sino como un sistema algebraico, lo cual permitirá lograr un entendimiento más profundo de sus propiedades y de lo que pueden representar.

2.1. Operaciones con matrices

Un dicho popular establece que "*el orden de los factores no altera el producto*", sin embargo, cuando se estudian operaciones con entes que no son números reales es necesario ser cuidadoso; pues, como se verá en esta sección, el orden de los factores sí puede alterar el producto cuando no se opera con números reales. En esta sección se discuten las propiedades básicas de las operaciones de suma y producto de matrices y habrá algunas sorpresas. Por ejemplo, el producto no es conmutativo; se pueden tener dos matrices no cero cuyo producto sea cero. De hecho, esta última propiedad es de las más interesantes, pues da lugar a que un sistema de ecuaciones lineales homogéneo tenga muchas soluciones.

Artículo no.	Mes 1	Mes 2	Mes 3	Mes 4	Mes 5	Mes 6
1	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}
2	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}
3	a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	a_{36}
4	a_{41}	a_{42}	a_{43}	a_{44}	a_{45}	a_{46}
5	a_{51}	a_{52}	a_{53}	a_{54}	a_{55}	a_{56}

Tabla 2.1. Producción de la empresa *A*.

2.1.1. Suma de matrices

Supongamos que se tienen dos empresas que producen cinco diferentes artículos en común. El valor de la producción se calcula mensualmente y se harán observaciones de la producción durante seis meses. A las empresas les denotaremos empresa A y empresa B. A los artículos que producen les llamaremos artículo 1, artículo 2, etcétera. Para representar el valor de los productos usaremos variables con doble índice. Por ejemplo, los valores de los artículos que produce la empresa A en el primer mes los representaremos por a_{11}, a_{21}, a_{31}, a_{41} y a_{51}; así mismo, los valores de los artículos en el segundo mes se representan por a_{12}, a_{22}, a_{32}, a_{42} y a_{52}. Note que el primer subíndice se refiere al número de artículo y el segundo especifica el número de mes. Una manera adecuada de representar toda esta información es mediante una tabla, es decir, la producción de la empresa A, en los seis meses, la podemos representar mediante la tabla 2.1.

Procediendo de manera análoga con la empresa B, su producción la podemos representar mediante la tabla 2.2.

Artículo no.	Mes 1	Mes 2	Mes 3	Mes 4	Mes 5	Mes 6
1	b_{11}	b_{12}	b_{13}	b_{14}	b_{15}	b_{16}
2	b_{21}	b_{22}	b_{23}	b_{24}	b_{25}	b_{26}
3	b_{31}	b_{32}	b_{33}	b_{34}	b_{35}	b_{36}
4	b_{41}	b_{42}	b_{43}	b_{44}	b_{45}	b_{46}
5	b_{51}	b_{52}	b_{53}	b_{54}	b_{55}	b_{56}

Tabla 2.2. Producción de la empresa *B*.

Con esta representación de la información es claro como se obtiene la cantidad total de cada producto, simplemente se suman las cantidades del producto correspondiente al mes que se requiere. Por ejemplo, la cantidad total del producto 2 en el mes 4 es $a_{24} + b_{24}$.

Transformando las representaciones de las tablas en arreglos se tiene que:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \end{pmatrix} \text{ y } \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} & b_{16} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} & b_{26} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & b_{36} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} & b_{46} \\ b_{51} & b_{52} & b_{53} & b_{54} & b_{55} & b_{56} \end{pmatrix} \text{ representan la}$$

producción de las empresas A y B. Para obtener la producción total de las empresas, por producto y por mes, simplemente se suman las correspondientes entradas de los arreglos. La situación anterior se puede generalizar a m productos y n meses. En este caso, la producción de las empresas se representa por arreglos rectangulares de m filas y n columnas. A un arreglo de este tipo le llamaremos matriz de orden $m \times n$. El ejemplo discutido se puede abstraer y formular la definición de suma de matrices.

Definición 2.1.1. *Dadas dos matrices* A *y* B *de orden* m × n*, se define la suma de* A *y* B*, denotada* A + B = C*, cuyas entradas son obtenidas de las entradas de* A *y* B *sumando entrada a entrada. Es decir, si las entradas de* A *y* B *se denotan por* a_{ij} *y* b_{ij} *respectivamente, entonces las entradas de* A + B *son* $a_{ij} + b_{ij} = c_{ij}$*, para* i = 1, 2, ..., m *y* j = 1, 2, ..., n.

Ejemplo 2.1.1. *Dadas las matrices* $A = \begin{pmatrix} 1 & -2 & 4 & 5 \\ 4 & 5 & 0 & \pi \\ 4 & 0 & 3 & 4 \end{pmatrix}$ *y* $B = \begin{pmatrix} 0 & -5 & 4 & 5 \\ -4 & 15 & 10 & -\pi \\ -3 & 1 & -4 & 3 \end{pmatrix}$

la suma $A + B$ es:

$$\begin{pmatrix} 1 & -2 & 4 & 5 \\ 4 & 5 & 0 & \pi \\ 4 & 0 & 3 & 4 \end{pmatrix} + \begin{pmatrix} 0 & -5 & 4 & 5 \\ -4 & 15 & 10 & -\pi \\ -3 & 1 & -4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & -7 & 8 & 10 \\ 0 & 20 & 10 & 0 \\ 1 & 1 & -1 & 7 \end{pmatrix}$$

Ejemplo 2.1.2. *Dadas las matrices*:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} \quad y\, B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{pmatrix}$$

la suma de A y B es:

$$C = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} & a_{14}+b_{14} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} & a_{24}+b_{24} \\ a_{31}+b_{31} & a_{32}+b_{32} & a_{33}+b_{33} & a_{34}+b_{34} \end{pmatrix}$$

2.1.2. Producto de matrices

Una de las operaciones con matrices que es difícil de explicar, es el *producto*, pues su definición, desligada de otros contextos, siempre causa “dificultades”.

El producto de matrices se puede “motivar” introduciendo conceptos relacionados con espacios vectoriales y transformaciones lineales. Esta forma de “justificar” la definición de producto de matrices tiene por lo menos un inconveniente: hay que presentar primero la discusión de transformaciones lineales y considerar que una matriz representa a una transformación lineal respecto a una base.

En la discusión que estamos presentando, la multiplicación de matrices se motiva con un ejemplo de un “sistema económico”.

Ejemplo 2.1.3. *Considere un consorcio con* n *empresas que producen* m *diferentes productos o bienes. Supongamos que todas las empresas producen los mismos bienes. Con estas consideraciones, la producción de cada bien se describe abajo.*

Por cada millón de pesos que se invierte en la empresa 1, la producción es:

a_{11} pesos del producto 1

a_{21} pesos del producto 2

$$\vdots$$

$$a_{m1} \text{ pesos del producto } m$$

En general, por cada millón de pesos que se invierte en la empresa j-ésima, ésta produce:

$$a_{1j} \text{ pesos del producto 1}$$

$$a_{2j} \text{ pesos del producto 2}$$

$$\vdots$$

$$a_{mj} \text{ pesos del producto } m$$

La información anterior la podemos representar mediante la tabla 2.3, en donde E_j y P_i representan a la empresa j-ésima y el producto i-ésimo, respectivamente.

Si en cada una de las empresas se invierte un millón de pesos, entonces el valor total del producto 1 es:

$$a_{11} + a_{12} + \cdots + a_{1n}$$

	E_1	E_2	E_3	E_4	$\cdots$	E_n
P_1	a_{11}	a_{12}	a_{13}	a_{14}	$\cdots$	a_{1n}
P_2	a_{21}	a_{22}	a_{23}	a_{24}	$\cdots$	a_{2n}
P_3	a_{31}	a_{32}	a_{33}	a_{34}	$\cdots$	a_{3n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
P_m	a_{m1}	a_{m2}	a_{m3}	a_{m4}	$\cdots$	a_{mn}

Tabla 2.3. Representación de la producción.

En general, el valor total del producto i es:

$$a_{i1} + a_{i2} + \cdots + a_{in}$$

Si denotamos por x_j la cantidad de dinero que se invierte en la empresa j, entonces el valor total del producto i es:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = c_i$$

Esta información la podemos representar de la siguiente forma:

$$\begin{array}{ccc} \text{Datos de producción} & \text{Inversión} & & \text{Producción en pesos} \\ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} & \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} & \Rightarrow & \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_m \end{bmatrix} \end{array}$$

o en la forma

$$
\begin{array}{ccc}
\text{Datos de producción} & \text{Inversión} & & \text{Producción en pesos} \\
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} & \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} & = & \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_m \end{bmatrix}
\end{array}
$$

en donde la igualdad se interpreta como:

$$c_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \qquad \text{para } i = 1, 2, \ldots, m$$

Con esta notación se tiene:

$$
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + & a_{12}x_2 + & \cdots + & a_{1n}+x_n \\ a_{21}x_1 + & a_{22}x_2 + & \cdots + & a_{2n}+b_n \\ a_{31}x_1 + & a_{32}x_2 + & \cdots + & a_{3n}+b_n \\ & \vdots & & \\ a_{m1}x_1 + & a_{m2}x_2 + & \cdots + & a_{mn}+x_n \end{bmatrix}
$$

Esta ecuación es la que define el producto de la matriz $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$

por la matriz $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$ dando como resultado la matriz del lado derecho.

Supóngase que el consorcio que administra las n empresas quiere analizar el comportamiento de la producción por un periodo de p años, manteniéndose fijas las condiciones de producción durante ese periodo, ¿se puede usar notación de matrices para analizar el proceso durante esos p años?

Iniciemos introduciendo algo de notación para representar lo que se invierte cada año en cada empresa. Por ejemplo, denotemos por: b_{11} la cantidad que se invierte en la empresa uno en el primer año; b_{12} la cantidad que se invierte en la empresa uno en el segundo año. En general, denotemos por b_{1j} la cantidad que se invierte en la empresa uno durante el año j-ésimo. ¿Cómo se representan, mediante una matriz, las inversiones de todas las empresas durante los p años? ¿Cómo se representa, mediante una matriz, la producción durante ese periodo?

La siguiente matriz representa las inversiones de las empresas por año. La interpretación es como sigue: la fila uno representa las inversiones en la empresa uno, la fila dos las inversiones en la empresa dos, etcétera.

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ b_{31} & b_{32} & \cdots & b_{3p} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}$$

Dado que en cada año se producirán los m bienes, denotemos por: c_{11} el valor total del producto uno en el año uno, c_{12} el valor total del producto uno en el año dos. En general, denotemos por c_{1j} el valor total del producto uno en el año j-ésimo. De esto se tiene que el valor total del producto i-ésimo en el año j-ésimo lo podemos denotar por c_{ij}.

Con esta notación, la matriz de valores de la producción durante el periodo es:

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ c_{31} & c_{32} & \cdots & c_{3p} \\ \vdots & \vdots & \cdots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix}$$

Si

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

representa las condiciones de producción, entonces el proceso de producción se puede representar mediante el producto de matrices $AB = C$ y se ilustra en seguida

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ b_{31} & b_{32} & \cdots & b_{3p} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{p2} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ c_{31} & c_{32} & \cdots & c_{3p} \\ \vdots & \vdots & \cdots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix}$$

Los símbolos anteriores significan, de acuerdo con las condiciones de producción:

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1}$$
$$c_{12} = a_{11}b_{12} + a_{12}b_{22} + \cdots + a_{1n}b_{n2}$$

En general,

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \tag{2.1}$$

para todos $i = 1, \ldots, m$ y $j = 1, \ldots, p$.

Lo anterior nos lleva a definir el producto de dos matrices de órdenes "adecuados".

Definición 2.1.2. *Si* A *y* B *son matrices de orden* m × n *y* n × p *respectivamente, se define el producto* AB : = C*, con* C *la matriz de orden* m × p *cuyas entradas están definidas por la ecuación 2.1.*

Con el lenguaje de producto de matrices, podemos formular los problemas relacionados con la solución de sistemas de ecuaciones lineales.

Si tenemos el sistema

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

entonces definiendo A, X y B como:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad X := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \quad B := \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{pmatrix}$$

el sistema se representa por $AX = B$. Esta manera de formularlo tiene varias ventajas, una de ellas es que podemos pensar que se trata de una sola ecuación cuya variable es X. De particular importancia es el caso $n = m$, en esta situación la matriz A es cuadrada y, como veremos, las preguntas relacionadas con las soluciones del sistema se reducen a decidir si A tiene *inversa*.

2.1.3. Propiedades de la suma y producto de matrices

En los sistemas de números conocidos en los niveles básicos, se ha enfatizado que las operaciones en éstos satisfacen ciertas propiedades como son: existencia de neutros o identidades, asociatividad, conmutatividad, distributividad del producto respecto a la suma. Teniendo esto en mente, es natural preguntarse si el producto y suma de matrices satisfacen algunas de esas propiedades.

En la definición de suma de matrices se ha enfatizado que para poder sumar dos matrices, ésas deben tener el mismo orden, es decir, las dos deben ser de orden $m \times n$. Al conjunto de matrices de orden $m \times n$, con entradas en los números reales, lo denotaremos por $\mathcal{M}_{m\times n}(\mathbb{R})$. Para multiplicar dos matrices A y B se necesita que el número de columnas del factor izquierdo sea igual al número de filas del factor derecho. Con la notación anterior, el producto AB estará definido solamente si $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ y $B \in \mathcal{M}_{n\times p}(\mathbb{R})$. Cuando esta condición es satisfecha estaremos en condiciones de averiguar el tipo de propiedades que cumple el producto.

Propiedades de la suma

Para analizar las propiedades de la suma supondremos que las matrices A, B y C pertenecen al conjunto $\mathcal{M}_{m\times n}(\mathbb{R})$.

1. Si $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$, digamos que las entradas de A son a_{ij} y las de B son b_{ij}, entonces la matriz $A + B$ tiene por entradas a $a_{ij} + b_{ij}$. Como a_{ij} y b_{ij} son números reales y en los reales la suma es conmutativa, concluimos que $a_{ij} + b_{ij} = b_{ij} + a_{ij}$. Por otro lado, los elementos $b_{ij} + a_{ij}$ son las entradas de $B + A$. De esto se ha probado que $A + B = B + A$; *conmutatividad de la suma*.
2. Existe una matriz en $\mathcal{M}_{m\times n}(\mathbb{R})$ que denotaremos 0, cuyas entradas son todas cero y satisface, $A + 0 = 0 + A = A$, para toda $A \in \mathcal{M}_{m\times n}(\mathbb{R})$; *existencia de neutro aditivo*.
3. Dada la matriz $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, existe $-A \in \mathcal{M}_{m\times n}(\mathbb{R})$ tal que $A + (-A) = 0$. Si las entradas de A son a_{ij}, las entradas de $-A$ las tomamos como $-a_{ij}$. Claramente se cumple $a_{ij} + (-a_{ij}) = 0$. *Existencia de inversos aditivos*.
4. Si A, B y C son matrices en $\mathcal{M}_{m\times n}(\mathbb{R})$, entonces $(A + B) + C = A + (B + C)$. La igualdad se obtiene escribiendo los elementos de $(A + B) + C$ y haciendo uso de la asociatividad de la suma en los números reales. *Asociatividad de la suma*.

Propiedades del producto de matrices

En lo que sigue enunciamos las propiedades que satisface el producto de matrices y demostramos solamente la primera, las restantes quedan como ejercicio para el lector.

1. Si $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $B \in \mathcal{M}_{n\times p}(\mathbb{R})$ y $C \in \mathcal{M}_{p\times q}(\mathbb{R})$, entonces $(AB)C = A(BC)$; *asociatividad del producto*.
2. Si A y B pertenecen a $\mathcal{M}_{\mathrm{m}\times n}(\mathbb{R})$ y C pertenece a $\mathcal{M}_{\mathrm{n}\times p}(\mathbb{R})$, entonces $(A + B)C = AC + BC$; *distributividad por la izquierda del producto respecto de la suma*.
3. Si A y B pertenecen a $\mathcal{M}_{m\times n}(\mathbb{R})$ y C pertenece a $\mathcal{M}_{p\times m}(\mathbb{R})$, entonces $C(A + B) = CA + CB$; *distributividad por la derecha del producto respecto de la suma*.
4. Existe una matriz de orden $m \times m$ denotada I_m, cuyas entradas c_{ij} satisfacen
   $c_{ij} = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases}$, tal que $I_m A = A$ para toda matriz $A \in \mathcal{M}_{m\times n}(\mathbb{R})$; *existencia de identidad por la izquierda*.
5. Existe una matriz de orden $n \times n$ denotada I_n, cuyas entradas c_{ij} satisfacen
   $c_{ij} = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases}$ tal que $AI_n, = A$ para toda matriz $A \in \mathcal{M}_{m\times n}(\mathbb{R})$; *existencia de identidad por la derecha*.

Discusión. Para demostrar la primera propiedad, notemos que los dos productos $(AB)C$ y $A(BC)$ están definidos y son matrices de orden $m \times q$. Pongamos $(AB)C = D$ y $A(BC) = E$, llamando d_{ij} a los elementos de D; asimismo, llamemos e_{ij} a los elementos de E. Para demostrar que $D = E$, debemos probar que $d_{ij} = e_{ij}$ para todo $i = 1, 2, \ldots, m$ y para todo $j = 1, 2, \ldots, q$.

Recordemos cómo se define el producto de matrices. Por ejemplo, si deseamos obtener el elemento que se encuentra en la fila i y en la columna j de AB, se toma la fila i-ésima de A y se multiplica por la columna j-ésima de B. Dicho elemento es

$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$. Si deseamos todos los elementos de la fila i-ésima de AB, entonces hacemos que j tome los valores 1, 2, ..., p en la expresión anterior, es decir, la fila i-ésima en AB es:

$$[a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{in}b_{n1}, a_{i1}b_{12} + a_{i2}b_{22} + \cdots + a_{in}b_{n2}, ..., a_{i1}b_{1p} + a_{i2}b_{2p} + \cdots + a_{in}b_{np}]$$

De esto y la definición del producto de matrices se tiene:

$$\begin{aligned} d_{ij} &= (a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{in}b_{n1})c_{1j} + (a_{i1}b_{12} + a_{i2}b_{22} + \cdots + a_{in}b_{n2})c_{2j} \\ &+ \cdots + (a_{i1}b_{1p} + a_{i2}b_{2p} + \cdots + a_{in}b_{np})c_{pj} \end{aligned}$$

De manera análoga, para obtener a e_{ij} se toma la fila i-ésima de A y la columna j-ésima de (BC).

La columna j-ésima de (BC) es:

$$\begin{bmatrix} b_{11}c_{1j} + b_{12}c_{2j} + \cdots + b_{1p}c_{pj} \\ b_{21}c_{1j} + b_{22}c_{2j} + \cdots + b_{2p}c_{pj} \\ \vdots \\ b_{n1}c_{1j} + b_{n2}c_{2j} + \cdots + b_{np}c_{pj} \end{bmatrix}$$

Aplicando nuevamente la definición del producto de matrices se tiene:

$$\begin{aligned} e_{ij} &= a_{i1}(b_{11}c_{1j} + b_{12}c_{2j} + \cdots + b_{1p}c_{pj}) + a_{i2}(b_{21}c_{1j} + b_{22}c_{2j} + \cdots + b_{2p}c_{pj}) \\ &+ \cdots + a_{in}(b_{n1}c_{1j} + b_{n2}c_{2j} + \cdots + b_{np}c_{pj}) \end{aligned}$$

Invitamos al lector a comparar las expresiones anteriores que representan a d_{ij} y a e_{ij}. Con esto tendrá elementos para corroborar que se tiene igualdad.

Observación 2.1.1. *Cuando* A *es una matriz cuadrada, digamos de orden* $n \times n$*, la identidad izquierda y derecha coinciden, le llamamos simplemente matriz identidad. En este caso denotaremos a la matriz identidad por* I_n *y se tiene* $I_nA = AI_n = A$.

Observación 2.1.2. *Note que si* A *y* B *son matrices y* AB *está definido,* BA *no necesariamente lo está. Puede ocurrir que los productos* AB *y* BA *estén definidos y sin embargo* $AB \neq BA$. *Otra propiedad que se tiene en el producto de matrices es que* AB *puede ser la matriz cero sin que* A *y* B *sean cero.*

2.2. Matrices elementales e inversas

En esta sección presentaremos una discusión de las operaciones elementales en las filas de una matriz, dándoles una interpretación en términos de producto de matrices. También caracterizaremos a las matrices cuadradas que tienen inversa, esto lo haremos en términos de matrices elementales. Una vez que sepamos cuándo tiene inversa una matriz, y cómo calcularla, regresaremos a estudiar un sistema de ecuaciones lineales, el cual sabemos se puede escribir de forma $AX = B$. Si A tiene inversa, entonces multiplicando la ecuación anterior por A^{-1} tendremos $A^{-1}AX = A^{-1}B$, de esto y las propiedades: $I_nX = X$ y $A^{-1}A = I_n$ se tendrá que $X = A^{-1}B$, es decir, el sistema de ecuaciones se habrá resuelto.

Cuando se discutió el método de reducción de Gauss-Jordan definimos lo que llamamos operaciones elementales en las filas de una matriz. Nos dimos cuenta que en esencia se tienen tres tipos de operaciones para eliminar elementos en las filas de una matriz.

En lo que sigue mostraremos que esas operaciones se pueden efectuar por medio de un producto de matrices. Recordemos cuáles son las operaciones elementales que podemos aplicar en las filas de una matriz.

1. Intercambiar cualesquiera dos filas de la matriz.
2. Multiplicar una fila por un escalar no cero.
3. Multiplicar una fila por un escalar y sumarlo a otra fila.

Supongamos que se tiene una matriz A de orden $m \times n$.

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

1. Sea E_{ij} la matriz que se obtiene de la identidad I_m intercambiando la fila i con la j. ¿Cuál es el resultado al multiplicar E_{ij} por la izquierda con A?
2. Sea $r \neq 0$ y $E_i(r)$ la matriz que se obtiene de la identidad I_m al multiplicar la fila i-ésima por r. ¿Cuál es el resultado al multiplicar $E_i(r)$ por la izquierda con A?
3. Sea $E_{ij}(r)$ la matriz que se obtiene de la identidad I_m al multiplicar la fila j-ésima por r y sumarla a la fila i-ésima. ¿Cuál es el resultado al multiplicar $E_{ij}(r)$ por la izquierda con A?

Actividades para fundamentar las respuestas a las preguntas planteadas

1. Tome $m = 2$ y $n = 3$ en las preguntas anteriores y haga los cálculos. Sean

   $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ y $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$, entonces $E_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

   y $E_{12}A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{pmatrix}$. Haga los cálculos para los casos dos y tres de las preguntas planteadas.

2. Tome $m = 3$ y $n = 4$ en las preguntas anteriores y haga los cálculos. En este caso hay varias posibilidades para definir las matrices E_{ij}, más precisamente, tenemos

   las siguientes matrices. $E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $E_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ y $E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

   Si A $= \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}$ entonces se tiene

$$E_{12}A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} & a_{24} \\ a_{11} & a_{12} & a_{13} & a_{14} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}$$

$$E_{13}A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} = \begin{pmatrix} a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{11} & a_{12} & a_{13} & a_{14} \end{pmatrix}$$

$$E_{23}A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{pmatrix}$$

3. ¿Tiene alguna conclusión general para poder contestar a las preguntas planteadas arriba?

Definición 2.2.1. *Una matriz* $n \times n$ *se dice inversible, si existe una matriz* B, $n \times n$ *tal que* $AB = BA = I_n$.

Observación 2.2.1. *La inversa de una matriz es única, es decir, si existen dos matrices* B *y* C *tales que* $AB = BA = I_n = AC = CA$ *entonces,* $C = CI_n = CAB = (CA)B = I_nB = B$, *probando que* $B = C$. *A la inversa de* A *la denotaremos por* A^{-1}.

Ejercicio 2.2.1. *Demuestre que las matrices elementales tienen inversa.*

En el lenguaje de matrices elementales, el teorema 1.3.2 puede ser enunciado de la siguiente forma. *Dada una matriz* $m \times n$ *no cero, existen matrices elementales* $E_1, E_2, \ldots, E_t$ tales que $E_t \ldots E_2E_1A = R$, *en donde* R *es la forma escalonada reducida de la matriz* A.

Un caso de mucha importancia ocurre cuando la matriz A es cuadrada, es decir, $m = n$. Bajo este supuesto analizaremos las condiciones que debe cumplir A para que exista su inversa; también se propondrá un método para encontrarla. Con esto y la observación 2.2.1 tendremos un método para resolver un sistema de n ecuaciones en n variables.

Supongamos que A es una matriz $n \times n$, y sean $E_1, \ldots E_t$ matrices elementales tales que $E_t \ldots E_1A = R$ es la matriz escalonada reducida de A. Recordemos que R es una matriz cuyas filas no cero tienen entrada principal igual a 1, y la columna en la que se encuentra tiene solamente esa entrada diferente de cero. La siguiente observación será importante en el análisis que haremos para decidir si una matriz tiene inversa.

Observación 2.2.2. *Sea* R *la forma escalonada reducida de* A, *matriz* $n \times n$. *Entonces* R *no tiene filas cero* $\Leftrightarrow$ *es igual a* I_n.

En efecto, si $R = I_n$, claramente R no tiene filas cero. Recíprocamente, si R no tiene filas ceros, entonces R tiene n filas no cero y las condiciones de R (siendo escalonada reducida), fuerzan $R = I_n$. Hay otras observaciones que son importantes:

Observación 2.2.3:

1. *Si dos matrices* A *y* B *tienen inversa, el producto* AB *tiene inversa y* $(AB)^{-1} = B^{-1}A^{-1}$.
2. *Si una matriz* A *tiene una fila de ceros y* B *es otra matriz tal que* AB *está definida, entonces* AB *tiene una fila de ceros, en particular si* A *es cuadrada y tiene una fila de ceros, entonces* A *no tiene inversa. ¿Cuál es el resultado si* B *tiene una columna de ceros?*

De estas observaciones se obtiene lo siguiente:

Teorema 2.2.1. *Sea* A *una matriz* n × n, $E_1, \ldots, E_t$ *matrices elementales tales que* $E_t \ldots E_1 A = R$ *es la forma escalonada reducida de* A. *Entonces* A *tiene inversa* ⇔ $R = I_n$ *y en este caso* $A^{-1} = E_t \cdots E_1$.

Demostración. Si $R = I_n$, claramente la inversa de A es $E_t \cdots E_1$. Supongamos que A tiene inversa, entonces por la observación 2.2.3(1) y el hecho que las matrices elementales tienen inversa, ejercicio 2.2.1, se llega a que $R = E_t \cdots E_1 A$ tiene inversa. Ahora aplicando la observación 2.2.3(2) se concluye que las filas de R son diferentes de cero. La conclusión final se obtiene de la observación 2.2.2.

El teorema 2.2.1 proporciona uno de los métodos más eficientes para calcular la inversa de una matriz. En la siguiente sección presentaremos la justificación del método y algunos ejemplos.

2.2.1. Cálculo de la inversa de una matriz

Supongamos que A es una matriz $n \times n$. Formemos la matriz de orden $n \times 2n$ cuyas primeras n columnas son las de A y las restantes son las de la identidad de orden n. Abusando un poco de la notación, esta nueva matriz la representaremos por $[A : I_n]$. Si $E_1, \ldots, E_t$ son matrices elementales tales que $E_t \ldots E_1 A = R$, entonces al aplicar las operaciones elementales, representadas por las matrices $E_1, \ldots, E_t$, en las filas de $[A : I_n]$ terminaremos con la matriz $[R : E_1 \ldots E_t]$. Por el teorema 2.2.1, A tiene inversa ⇔ $R = I_n$ y A^{-1} es la matriz que aparece a la derecha en $[R : E_1 \ldots E_t]$.

Ejemplo 2.2.1. *Consideremos la matriz*

$$A := \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

Determinaremos si tiene inversa y en caso afirmativo, la encontraremos.

Discusión. Aplicación del método.

A partir de A y la identidad I_3, formemos la matriz:

$$A_1 := \begin{bmatrix} 1 & 4 & 7 & 1 & 0 & 0 \\ 2 & 5 & 8 & 0 & 1 & 0 \\ 3 & 6 & 9 & 0 & 0 & 1 \end{bmatrix}$$

Aplicando operaciones elementales en las filas de A_1 se obtiene sucesivamente.

$$A_1 := \begin{bmatrix} 1 & 4 & 7 & 1 & 0 & 0 \\ 0 & -3 & -6 & -2 & 1 & 0 \\ 3 & 6 & 9 & 0 & 0 & 1 \end{bmatrix} \to A_2 := \begin{bmatrix} 1 & 4 & 7 & 1 & 0 & 0 \\ 0 & -3 & -6 & -2 & 1 & 0 \\ 0 & -6 & -12 & -3 & 0 & 1 \end{bmatrix} \to$$

$$A_3 := \begin{bmatrix} 1 & 4 & 7 & 1 & 0 & 0 \\ 0 & 1 & 2 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & -6 & -12 & -3 & 0 & 1 \end{bmatrix} \to A_4 := \begin{bmatrix} 1 & 4 & 7 & 1 & 0 & 0 \\ 0 & 1 & 2 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{bmatrix}$$

En este punto podemos detener los cálculos y concluir que A no tiene inversa, pues la última fila de lo que será R es cero y de acuerdo con la observación 2.2.2, y con el teorema 2.2.1, A no tiene inversa.

Ejemplo 2.2.2. *Determine si la matriz:*

$$A := \begin{bmatrix} 1 & 2 & 0 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

tiene inversa y en caso afirmativo, encuéntrela.

Discusión. Aplicación del método.

A partir de A y la identidad I_3, formamos la matriz:

$$A1 := \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 5 & 8 & 0 & 1 & 0 \\ 3 & 6 & 9 & 0 & 0 & 1 \end{bmatrix}$$

Aplicando operaciones elementales en las filas de A_1 se obtiene sucesivamente:

$$A_1 := \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 5 & 8 & 0 & 1 & 0 \\ 3 & 6 & 9 & 0 & 0 & 1 \end{bmatrix} \to A_2 := \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 8 & -2 & 1 & 0 \\ 3 & 6 & 9 & 0 & 0 & 1 \end{bmatrix} \to$$

$$A_2 := \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 8 & -2 & 1 & 0 \\ 3 & 6 & 9 & 0 & 0 & 1 \end{bmatrix} \to A_3 := \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 8 & -2 & 1 & 0 \\ 0 & 0 & 1 & -\frac{1}{3} & 0 & \frac{1}{9} \end{bmatrix}$$

En este momento podemos decir que A tiene inversa (¿por qué?), y solamente hace falta determinarla. Esto se logra aplicando operaciones elementales hasta que el bloque 3×3 de la izquierda esté en forma escalonada reducida. Efectúe los cálculos requeridos para llegar a la matriz:

$$A_6 := \begin{bmatrix} 1 & 0 & 0 & -\frac{1}{3} & -2 & \frac{16}{9} \\ 0 & 1 & 0 & \frac{2}{3} & 1 & -\frac{8}{9} \\ 0 & 0 & 1 & -\frac{1}{3} & 0 & \frac{1}{9} \end{bmatrix}$$

de la cual, usando el método, la inversa de A es:

$$A^{-1} = \begin{bmatrix} -\frac{1}{3} & -2 & \frac{16}{9} \\ \frac{2}{3} & 1 & -\frac{8}{9} \\ -\frac{1}{3} & 0 & \frac{1}{9} \end{bmatrix}$$

Ejemplo 2.2.3. *Resuelva el siguiente sistema de ecuaciones, calculando la inversa de la matriz de coeficientes.*

$$x + 2y = 3$$
$$2x + 5y + 8z = 1$$
$$3x + 6y + 9z = 2$$

Discusión. El sistema se puede representar en la forma $AX = B$, en donde

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ y } \quad B = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

Como A es la matriz del ejemplo anterior y tiene inversa, la ecuación $AX = B$ es equivalente a $X = A^{-1}B$. Usando el resultado del ejemplo citado se tiene:

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & -2 & \frac{16}{9} \\ \frac{2}{3} & 1 & -\frac{8}{9} \\ -\frac{1}{3} & 0 & \frac{1}{9} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{5}{9} \\ \frac{11}{9} \\ -\frac{7}{9} \end{bmatrix}$$

Esta ecuación matricial es equivalente a: $x = \dfrac{5}{9}$, $y = \dfrac{11}{9}$ y $z = -\dfrac{7}{9}$, es decir, el sistema dado tiene solución única: $\left(\dfrac{5}{9}, \dfrac{11}{9} - \dfrac{7}{9}\right)$

Ejercicio 2.2.2. *En los siguientes casos determine si* A *tiene inversa y, de ser así, encuéntrela.*

- *Sean* a *y* b *números reales diferentes. y* A $= \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix}$
- *Sean* a, b *y* c *tres números reales diferentes a pares y* A $= \begin{pmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{pmatrix}$

Ejercicio 2.2.3. *Sea* A $= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ *tal que* $\Delta = a_{11}a_{22} - a_{12}a_{21} \neq 0$. *Determine si existe* A^{-1} *y en caso afirmativo, obténgala.*

Definición 2.2.2. *Sean* A *y* B *dos matrices* $m \times n$. *Decimos que* A *es equivalente por filas a* B, *si existen matrices elementales* $E_1, E_2, \ldots, E_t$ *tales que* $A = E_1E_2 \ldots E_tB$.

El siguiente es uno de los teoremas más importantes de álgebra lineal. Pedimos al lector que lo demuestre, con eso seguramente lo estará incorporando a sus conocimientos.

Teorema 2.2.2. *Sea* A *una matriz* n × n. *Demuestre que las siguientes condiciones son equivalentes.*

1. *La matriz* A *tiene inversa.*
2. *La matriz* A *es equivalente por filas a* I_n.
3. *La matriz* A *es producto de matrices elementales.*
4. *El sistema* AX = 0 *tiene solución única.*
5. *Para todo* $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, *el sistema* AX = B *tiene solución única.*

Definición 2.2.3. *Sea* $A = (a_{ij})$ *una matriz* m × n. *Se define la transpuesta de* A, *denotada* $A^t = (b_{ij})$, *como la matriz que satisface* $b_{ij} = a_{ji}$ *para todos* i = 1, 2, ..., n *y* j = 1, 2, ..., m. *Si* A *es una matriz tal que* $A^t = A$ $(A^t = -A)$, *diremos que* A *es simétrica* (*antisimétrica*). *Note que en estos casos necesariamente* A *es cuadrada.*

Observación 2.2.4. *Para que una matriz* A *sea simétrica o antisimétrica es necesario que sea cuadrada. Si* A *es antisimétrica, entonces los elementos de la diagonal* a_{ii} *son cero. Si* A *es cualquier matriz cuadrada, entonces las matrices* $B = A^t + A$ *y* $C = A^t - A$ *son simétrica y antisimétrica respectivamente. ¿Se puede expresar* A *en términos de* B *y* C?

Ejemplo 2.2.4.

$$Si\ \boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix},\ entonces\ \boldsymbol{A}^t = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \\ a_{14} & a_{24} & a_{34} \end{pmatrix}$$

Algunas propiedades de la transpuesta y de la traza

Los siguientes enunciados y preguntas conciernen a la transpuesta y traza de una matriz. Sugerimos al lector que discuta lo que se plantea.

1. Sean A y B matrices tales que el producto AB está definido. Demuestre que $(AB)^t = B^tA^t$.
2. Si A es una matriz cuadrada que tiene inversa, ¿tiene inversa A^t?
3. Para una matriz cuadrada A se define su *traza*, denotada $\mathrm{tr}(A) := \sum_{i=1}^{n} a_{ii}$. ¿Cuándo es cero la traza de AA^t?
4. Si A y B son matrices cuadradas, calcule $\mathrm{tr}(AB)$ y $\mathrm{tr}(BA)$. ¿Hay alguna relación entre estos números?
5. ¿Puede tener traza cero una matriz que tiene inversa?
6. ¿Es correcta la ecuación $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$?
7. ¿Existen matrices $n \times n$, A y B tales que $AB - BA = I_n$?

2.3. Aplicaciones

En esta sección presentamos algunas aplicaciones de matrices y sistemas de ecuaciones que tienen lugar en ingeniería y otras áreas.

Deformación de vigas

Supongamos que se tiene una viga elástica en la que se aplican pesos de magnitudes $W_1, W_2, \ldots, W_n$ en los puntos $P_1, P_2, \ldots, P_n$, respectivamente (ver la figura 2.1). Denotemos por y_i a la magnitud del desplazamiento del punto P_i por la acción de las cargas. Supongamos que la viga obedece a la ley de Hooke[1] para materiales elásticos. De acuerdo con dicha ley, el peso W_j produce un desplazamiento igual a $d_{ij}W_j$ en el punto P_i, en donde d_{ij} es una constante que depende de las propiedades de elasticidad de la viga y de la distancia del punto P_i al punto P_j.

De estas consideraciones se tiene que y_i es la suma de los desplazamientos originados por cada uno de los pesos, es decir, se tiene:

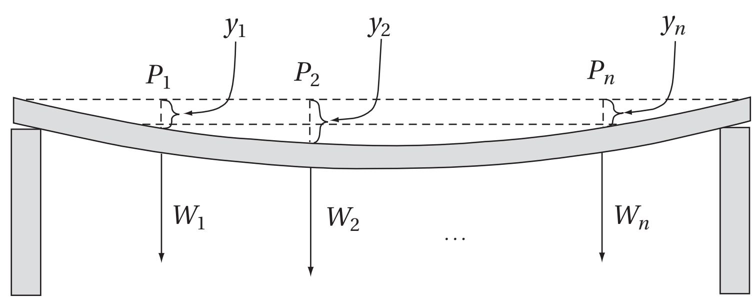


Figura 2.1. Deformación de una viga bajo la acción de n cargas.

$$y_i = d_{i1}W_1 + d_{i2}W_2 + \cdots + d_{in}W_n, \quad i = 1, 2, \ldots, n. \tag{2.2}$$

En forma matricial las ecuaciones 2.2 se pueden escribir como:

$$DF = Y, \text{ en donde } Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \text{ y } F = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{bmatrix} \tag{2.3}$$

A la matriz D se le llama la *matriz de flexibilidad* y sus columnas se interpretan de la siguiente forma.

Supongamos que en el punto P_i se aplica un peso unitario y en los restantes puntos el peso es cero, entonces $F = \begin{bmatrix} 0 \\ \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, por lo que

[1] La Ley de Hooke establece que la magnitud de la deformación de un cuerpo elástico bajo la acción de una fuerza F, es directamente proporcional a la magnitud de F.

$$Y = \begin{bmatrix} d_{11} & \cdots & d_{1n} \\ d_{21} & \cdots & d_{2n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ d_{n1} & \cdots & d_{nn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} d_{1i} \\ d_{2i} \\ \vdots \\ \vdots \\ d_{ni} \end{bmatrix}$$

El significado de la ecuación anterior es que la columna i-ésima de la matriz D representa la deformación cuando se aplica un peso unitario en el punto P_i, es decir, un peso unitario aplicado en el punto i produce desplazamientos d_{1i}, d_{2i}, ..., d_{ni} en los correspondientes puntos P_1, P_2, ..., P_n.

Cuando D tiene inversa, conociendo los desplazamientos y_1, y_2, ..., y_n se pueden conocer los pesos W_1, W_2, ..., W_n que actúan sobre la viga, pues de la ecuación $DF = Y$ se obtiene $F = D^{-1}Y$. Las columnas de D^{-1} se interpretan como sigue.

Cuando se toma un desplazamiento de una unidad en el punto i-ésimo y cero en los restantes, se tiene $Y = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, por lo que al efectuar el producto $D^{-1}Y$, el resultado es la columna i-ésima de D^{-1} cuyas entradas son las magnitudes de los pesos W_1, W_2, ... W_n que producen un desplazamiento de una unidad en el punto P_i y cero en los restantes.

Circuitos eléctricos

Aplicando las leyes de Kirchhoff junto con la ley de Ohm para circuitos eléctricos, se puede resolver una gama amplia de problemas que los involucran. Una herramienta importante en el proceso de solución de tales problemas, son los sistemas de ecuaciones lineales.

Ley de Kirchhoff para la corriente. En cualquier nodo de un circuito donde la densidad de carga no cambia con el tiempo, la suma de las corrientes que entran es igual a la suma de las corrientes que salen.

Ley de Kirchhoff para el voltaje. La suma algebraica de las diferencias de potencial eléctrico en un circuito es cero.

Ley de Ohm. La diferencia de potencial (voltaje) a través de un conductor ideal es directamente proporcional a la intensidad de corriente que circula en éste. La constante de proporcionalidad se llama resistencia del conductor.

Si V, I y R denotan la diferencia de potencial, intensidad de corriente y resistencia respectivamente, entonces la ley de Ohm se expresa por medio de la ecuación: $V = RI$.

Supongamos que se tiene un circuito eléctrico como se ilustra en la figura 2.2. Determine la intensidad de corriente que circula en cada rama del circuito.

Antes de iniciar el análisis del circuito, recordamos las convenciones que se usan para considerar la diferencia de potencial debida a una fuente. Si la corriente pasa de menor a mayor potencial, éste será considerado negativo.

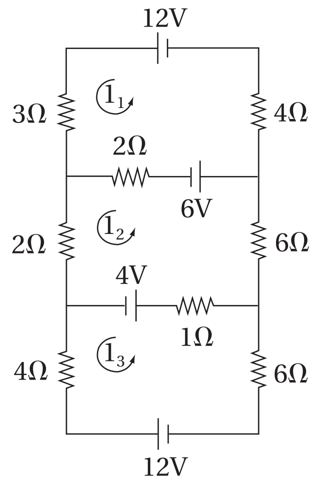


Figura 2.2. Circuito eléctrico con tres ramas.

En la rama superior del circuito se tiene que I_1 pasa por tres resistores por lo que la caída de potencial debido a la corriente I_1 es $3I_1 + 2I_1 + 4I_1$. La corriente I_2 del circuito central también fluye en parte del circuito superior, de hecho lo hace en sentido contrario a I_1, al pasar por el resistor que tiene valor 2Ω, entonces esta corriente produce una caída de potencial igual a $-2I_2$. Aplicando la ley de Kirchhoff para el voltaje obtenemos:

$$9I_1 - 2I_2 - 12 - 6 = 0 \tag{2.4}$$

En el circuito central fluyen las tres corrientes, por lo que sumando las caídas de potencial debido a cada una de esas corrientes se tiene $11I_2 - 2I_1 - I_3$. Nuevamente, aplicando la ley de Kirchhoff para el voltaje se llega a:

$$-2I_1 + 11I_2 - I_3 + 6 - 4 = 0 \tag{2.5}$$

Haciendo un análisis similar para el circuito inferior se tiene la ecuación:

$$-I_2 + 11I_3 + 12 + 4 = 0 \tag{2.6}$$

Con las ecuaciones anteriores obtenemos el sistema:

$$\begin{aligned} 9I_1 - 2I_2 &= 18 \\ -2I_1 + 11I_2 - I_3 &= -2 \\ -I_2 + 11I_3 &= -16 \end{aligned}$$

el cual puede ser resuelto usando el método de reducción de Gauss-Jordan.

La primera de las siguientes matrices es la aumentada del sistema; las restantes corresponden a diferentes etapas del proceso de reducción. Identifique las operaciones elementales efectuadas y verifique los cálculos.

$$\begin{pmatrix} 9 & -2 & 0 & 18 \\ -2 & 11 & -1 & -2 \\ 0 & -1 & 11 & -16 \end{pmatrix} \sim \begin{pmatrix} 1 & 42 & -4 & -10 \\ -2 & 11 & -1 & -2 \\ 0 & -1 & 11 & -16 \end{pmatrix} \sim \begin{pmatrix} 1 & 42 & -4 & 10 \\ 0 & 95 & -9 & 18 \\ 0 & -1 & 11 & -16 \end{pmatrix} \sim$$

$$\begin{pmatrix} 1 & 0 & 458 & -662 \\ 0 & 0 & 1036 & -1502 \\ 0 & -1 & 11 & -16 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 458 & -662 \\ 0 & 0 & 1 & -751/518 \\ 0 & -1 & 11 & -16 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 521/259 \\ 0 & 0 & 1 & -751/518 \\ 0 & 1 & 0 & 27/518 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 521/259 \\ 0 & 1 & 0 & 27/518 \\ 0 & 0 & 1 & -751/518 \end{pmatrix}$$

La última matriz representa al sistema resuelto:

$$I_1 = \frac{521}{259}\,\text{Amps}$$

$$I_2 = \frac{27}{518}\,\text{Amps}$$

$$I_3 = -\frac{751}{518}\,\text{Amps}$$

Dada la estructura del circuito, explique el significado de los signos en las magnitudes de las corrientes.

Distribución de temperaturas en placas delgadas

Para hacer un análisis de la conducción de calor en un medio isotrópico y homogéneo, es necesario usar la ecuación de calor:

$$U_t = k(U_{xx} + U_{yy} + U_{zz}) \tag{2.7}$$

en donde k es una constante que depende de la conductividad térmica, de la densidad y de la capacidad calorífica del material; U_{xx}, U_{yy}, y U_{zz} son la segunda derivada parcial de U respecto de x, y y z respectivamente y U_t es la derivada parcial de U respecto de t, que mide la variación de temperatura en el tiempo t.

El ejemplo que discutimos tiene otras hipótesis. Supongamos que se tiene una placa rectangular delgada como se muestra en la figura 2.3. Adicionalmente, asuma que se conocen las temperaturas en los bordes de la placa y la temperatura en cada uno de los nodos interiores de la malla es el promedio de las temperaturas de los puntos más cercanos. Por ejemplo, si denotamos por t_i a la temperatura en el nodo i, para el caso $i = 1$ se tiene $t_1 = \frac{1}{4}(40 + t_2 + t_4 + 15)$.

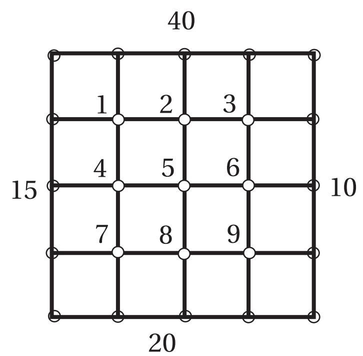


Figura 2.3. Distribución de temperaturas en una placa delgada.

Continuando de esta manera con cada uno de los nodos restantes se tiene un sistema de nueve ecuaciones en nueve variables como se describe en las siguientes líneas.

$$
\begin{aligned}
t_1 &= \frac{1}{4}(40 + t_2 + t_4 + 15)\\
t_2 &= \frac{1}{4}(40 + t_1 + t_3 + t_5)\\
t_3 &= \frac{1}{4}(40 + t_2 + t_6 + 10)\\
t_4 &= \frac{1}{4}(15 + t_1 + t_5 + t_7)\\
t_5 &= \frac{1}{4}(t_2 + t_4 + t_6 + t_8)\\
t_6 &= \frac{1}{4}(10 + t_3 + t_5 + t_9)\\
t_7 &= \frac{1}{4}(15 + 20 + t_4 + t_8)\\
t_8 &= \frac{1}{4}(20 + t_5 + t_7 + t_9)\\
t_9 &= \frac{1}{4}(10 + 20 + t_6 + t_8)
\end{aligned}
\tag{2.8}
$$

Procederemos a resolver este sistema usando *Maple*. Para resolver un sistema de ecuaciones usando *Maple* siga los pasos que se describen.

Declare a cada una de las ecuaciones como se indica en los siguientes comandos.

```
> e1:=t1=(1/4)*(40+t2+t4+15);
```

$$e1 := t1 = \frac{55}{4} + \frac{1}{4}\ t2 + \frac{1}{4}\ t4$$

```
> e2:=t2=(1/4)*(40+t1+t3+t5);
```

$$e2 := t2 = 10 + \frac{1}{4}\ t1 + \frac{1}{4}\ t3 + \frac{1}{4}\ t5$$

```
> e3:=t3=(1/4)*(40+t2+t6+10);
```

$$e3 := t3 = \frac{25}{2} + \frac{1}{4}\ t2 + \frac{1}{4}\ t6$$

```
> e4:=t4=(1/4)*(15+t1+t5+t7);
```

$$e4 := t4 = \frac{15}{4} + \frac{1}{4}\ t1 + \frac{1}{4}\ t5 + \frac{1}{4}\ t7$$

```
> e5:=t5=(1/4)*(t2+t4+t6+t8);
```

$$e5 := t5 = \frac{1}{4}\ t2 + \frac{1}{4}\ t4 + \frac{1}{4}\ t6 + \frac{1}{4}\ t8$$

```
> e6:=t6=(1/4)*(10+t3+t5+t9);
```

$$e6 := t6 = \frac{5}{2} + \frac{1}{4}\ t3 + \frac{1}{4}\ t5 + \frac{1}{4}\ t9$$

```
> e7:=t7=(1/4)*(15+20+t4+t8);
```

$$e7 := t7 = \frac{35}{4} + \frac{1}{4}\ t4 + \frac{1}{4}\ t8$$

```
> e8:=t8=(1/4)*(20+t5+t7+t9);
```

$$e8 := t8 = 5 + \frac{1}{4}\ t5 + \frac{1}{4}\ t7 + \frac{1}{4}\ t9$$

```
> e9:=t9=(1/4)*(20+10+t6+t8);
```

$$e9 := t9 = \frac{15}{2} + \frac{1}{4}\ t6 + \frac{1}{4}\ t8$$

Una vez declaradas las ecuaciones puede proceder con el comando *solve*, como se indica en la siguiente línea

```
> solve({el,e2,e3,e4,e5,e6,e7,e8,e9},{t1,t2,t3,t4,t5,
> t6,t7,t8,t9});
```

Maple produce el resultado:

$$\left\{ t1 = , \frac{180}{7}, t2 = \frac{3\,105}{112}, t3 = \frac{335}{14}, t4 = \frac{2\,255}{112}, t5 = \frac{85}{4}, t6 = \frac{2\,015}{112}, t7 = \frac{130}{7}, t8 = \frac{2\,145}{112}, t9 = \frac{235}{14} \right\}$$

Note que con el auxilio de un sistema computacional como *Maple* se pueden resolver sistemas de ecuaciones relativamente grandes.

Conservación de bosques

La explotación de los recursos forestales se debiera hacer con estricto apego a las reglas ecológicas establecidas, pues de otra forma ponen en grave riesgo la supervivencia de una gran cantidad de especies que viven en los bosques. El ejemplo que presentamos, si bien hipotético, tiene como finalidad, ilustrar en lo posible, la forma de resolver un problema de explotación maderera bajo un plan sustentable ecológicamente.

Una compañía administra tres regiones para explotación de madera en las que se cortan pinos, encinos y robles. La compañía tiene un contrato con un aserradero para proveerle de 2 000 m^3 de madera de pino; 1 500 m^3 de madera de encino y 800 m^3 de madera de roble, todo esto por mes. El volumen de explotación en cada región está descrito en la tabla 2.4.

Región	Explotación/ha	% Pinos	% Encinos	% Robles
región 1	310 m^3/ha	80	10	10
región 2	350 m^3/ha	10	80	10
región 3	280 m^3/ha	10	10	80

Tabla 2.4. Explotación de madera por región.

¿Cuántas hectáreas deben ser explotadas por mes en cada región para satisfacer la demanda del aserradero?

Primeramente, la información de la tabla se interpreta así. En la región 1 se tiene permitido cortar 310 m^3 por hectárea, de los cuales 80% es de pinos, 10% de encinos y el otro 10% de robles. La interpretación para las restantes regiones es análoga. Denotemos por x, y y z a las cantidades de hectáreas que se pueden explotar en las regiones 1, 2 y 3 respectivamente. Con estas consideraciones en mente, las cantidades de pino que se pueden cortar en la regiones 1, 2 y 3 son .8 (310)x, .1(350)y y .1(280)z, respectivamente. Como la demanda es de 2 000 m^3 de pino por mes, se tiene la ecuación:

$$.8(310)x + .l(350)y + .1(280)z = 2\,000 \tag{2.9}$$

Análogamente, para la madera de encino se llega a la ecuación:

$$.1(310)x + .8(350)y + .1(280)z = 1\,500 \tag{2.10}$$

De la misma manera que en los casos anteriores, para la madera de roble se tiene:

$$.1(310)x + .1(350)y + .8(280)z = 800 \tag{2.11}$$

Con estas tres ecuaciones se forma el sistema:

$$.8(310)x + .l(350)y + .1(280)z = 2\,000$$

$$.1(310)x + .8(350)y + .1(280)z = 1\,500$$

$$.1(310)x + .1(350)y + .8(280)z = 800$$

Multiplicando cada ecuación por 10, haciendo los cambios s $= 350y$, $t = 310x$ y $u = 280z$, el sistema anterior equivale a:

$$8s + t + u = 15\,000$$

$$s + 8t + u = 20\,000$$

$$s + t + 8u = 8\,000$$

en el cual se facilitan los cálculos aritméticos en el proceso de solución.

Aplicando el método de reducción de Gauss-Jordan para resolverlo se llega a:

$$s = \frac{10\,700}{7}$$

$$t = \frac{15\,700}{7}$$

$$u = \frac{3\,700}{7}$$

Finalmente, usando estos valores en $s = 350y$, $t = 310x$, $u = 280z$ y simplificando se tienen los valores de x, y y z, es decir:

$$x = \frac{1\,570}{217}$$

$$y = \frac{1\,070}{245}$$

$$z = \frac{370}{196}$$

Observación 2.3.1. *Note que una ventaja adicional al hacer el cambio de las variables* x, y *y* z *a las variables* s, t *y* u *es que el sistema en estas últimas variables tiene solución positiva, por lo que no importa cuál sea la cuota permitida por región, siempre se podrá satisfacer la demanda contratada.*

2.4. Matrices enteras

Esta sección tiene como finalidad ilustrar las dificultades que surgen al estudiar sistemas de ecuaciones lineales con coeficientes enteros y se pide que las soluciones también sean enteras.

Un taller de costura produce camisas, playeras y chamarras usando tres tipos diferentes de máquinas de coser. La producción por semana es como se muestra en la tabla 2.5. ¿Cuántas máquinas deben trabajar para coser 1 200 camisas, 900 playeras y 600 chamarras por semana?

	Máquina tipo 1	Máquina tipo 2	Máquina tipo 3
Camisas	100	120	90
Playeras	70	80	70
Chamarras	50	60	40

Tabla 2.5. Maquila por máquina.

Si x representa el número de máquinas del tipo uno que trabajarán por semana, su producción será: $100x$ camisas, $70x$ playeras y $50x$ chamarras. Así mismo, si y denota al número de máquinas del tipo 2 que se usarán por semana, su producción será: $120y$ camisas, $80y$ playeras y $60y$ chamarras. De manera semejante, si z representa al número de máquinas del tipo tres, entonces su producción será: $90z$ camisas, $70z$ playeras y $40z$ chamarras por semana.

De esta información se tiene que el número total de camisas que produce el taller por semana es:

$$\text{total de camisas/semana: } 100x + 120y + 90z$$

Expresiones similares se tienen para playeras y chamarras; de manera explícita:

$$\text{total de playeras/semana: } 70x + 80y + 70z$$

$$\text{total de chamarras/semana: } 50x + 60y + 40z$$

Para contestar la pregunta es necesario discutir el sistema de ecuaciones que se obtiene al igualar las expresiones anteriores con las cantidades que se requiere producir, es decir, se debe abordar el sistema de ecuaciones:

$$\begin{aligned} 100x + 120y + 90z &= 1\,200 \\ 70x + 80y + 70z &= 900 \\ 50x + 60y + 40z &= 600 \end{aligned}$$

el cual equivale a:

$$\begin{aligned} 10x + 12y + 9z &= 120 \\ 7x + 8y + 7z &= 90 \\ 5x + 6y + 4z &= 60 \end{aligned} \tag{2.12}$$

Iniciamos la discusión respecto a las soluciones, construyendo la matriz aumentada del sistema, la cual resulta ser:

$$A=\begin{bmatrix} 10 & 12 & 9 & 120 \\ 7 & 8 & 7 & 90 \\ 5 & 6 & 4 & 60 \end{bmatrix}$$

Una de las características importantes de esta matriz es que todas sus entradas son números enteros. De las condiciones del ejemplo se requiere que las soluciones, de existir, sean números enteros no negativos.

Recordemos que al aplicar operaciones elementales en las filas de la matriz aumentada de un sistema, la matriz que resulta representa a un sistema equivalente, teorema 1.3.1, página 13. Tomando esto como punto de referencia y dado que estamos buscando soluciones enteras, es de importancia considerar que en cada uno de los pasos del proceso para resolver el sistema, se efectúen operaciones elementales que garanticen soluciones enteras.

Dado que una de las operaciones elementales consiste en multiplicar una fila por una constante no cero, y las operaciones elementales tienen inversa, entonces los únicos valores que podemos tomar para multiplicar una fila son ± 1, pues éstos son los únicos enteros que tienen inverso multiplicativo.

En otras palabras, el tipo de operaciones elementales permitidas cuando se trate de matrices con entradas enteras son:

1. Intercambiar dos filas cualesquiera.
2. Multiplicar una fila por ± 1.
3. Multiplicar una fila por una constante y sumarla a otra.

Aplicaremos este tipo de operaciones a la matriz A para discutir las soluciones de (2.12). Pedimos al lector identificar las operaciones efectuadas al pasar de una matriz a la siguiente.

$$\begin{bmatrix} 10 & 12 & 9 & 120 \\ 7 & 8 & 7 & 90 \\ 5 & 6 & 4 & 60 \end{bmatrix} \sim \begin{bmatrix} 10 & 12 & 9 & 120 \\ 2 & 2 & 3 & 30 \\ 5 & 6 & 4 & 60 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 1 & 0 \\ 2 & 2 & 3 & 30 \\ 5 & 6 & 4 & 60 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 1 & 0 \\ 2 & 2 & 3 & 30 \\ 1 & 2 & -2 & 0 \end{bmatrix} \sim$$

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & -2 & 7 & 30 \\ 1 & 2 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & -2 & 7 & 30 \\ 1 & 0 & 5 & 30 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & -2 & 0 & 30 \\ 1 & 0 & 0 & 30 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 30 \\ 0 & 2 & 0 & -30 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Notemos que la solución no tiene sentido desde el punto de vista de la pregunta planteada, pues el valor de y es negativo y representa el número de máquinas del tipo 2. Sin embargo, algo de importancia que se debe notar respecto al tipo de operaciones elementales que se utilizaron en la discusión, es su "efectividad" para determinar las soluciones.

El ejemplo anterior y su discusión conducen a formular una serie de preguntas relacionadas con soluciones de sistemas de ecuaciones con coeficientes enteros.[2]

---

[2] A las ecuaciones que tienen coeficientes enteros o racionales se les llama ecuaciones diofánticas o diofantinas en honor al matemático de la antigua Grecia, Diofanto, quien vivió más o menos en los años 200-284 de nuestra era y fue uno de los primeros en estudiar de manera sistemática este tipo de ecuaciones. Algunos historiadores consideran a Diofanto el padre del Álgebra. Actualmente las ecuaciones diofantinas son una rama muy importante de la teoría de números. Un ejemplo de ecuaciones diofantinas es la muy famosa ecuación de Fermat: $x^n + y^n = z^n$.

Para precisar un poco, daremos nombre a las matrices en las que todas sus entradas son números enteros.

Definición 2.4.1. *Una matriz* A *se dice entera, si todas sus entradas son números enteros.*

¿Cuándo tiene solución entera un sistema de ecuaciones con coeficientes enteros? ¿Cómo encontrar todas las soluciones de un sistema de ecuaciones con coeficientes enteros? ¿Se puede ajustar el método de reducción de Gauss-Jordan para matrices enteras?

La discusión de estas preguntas lleva a lo que se conoce como *teoría de matrices enteras*.

Uno de los resultados fundamentales en la teoría de matrices enteras, que además aporta elementos para contestar a las preguntas planteadas es el siguiente:

Teorema 2.4.1. *Para toda matriz entera* A *de orden* $m \times n$, *existen matrices enteras* L *y* R *con inversas enteras, de órdenes* $m \times m$ *y* $n \times n$ *respectivamente tales que* $LAR = diag\{d_1, d_2, ..., d_r, 0 ..., 0\}$, *en donde* d_i *es un número entero positivo para* $i = 1, 2, ..., r$ *y* d_i *divide a* d_{i+1} *para todo* $i = 1, 2, ..., (r - 1)$.

Sugerimos al lector que demuestre el resultado anterior para el caso en que A es 2×2. Un posible procedimiento para demostrar el teorema en este caso, consiste en aplicar operaciones elementales en las filas y en las columnas de A. La demostración del caso general queda fuera del alcance del texto; el lector interesado en matrices enteras puede consultar [10], en donde encontrará una demostración del teorema anterior.

Definición 2.4.2. *A la matriz* $D = diag\{d_1, d_2, ..., d_r, 0 ..., 0\}$ *se le llama la forma normal de Smith.*

2.5. Ejercicios

1. Enuncie los teoremas y definiciones que se han discutido en este capítulo.

2. *a*) Sea $A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ -1 & 0 & 3 & 5 \\ 1 & -2 & 1 & 1 \end{bmatrix}$. Determine la forma escalonada reducida R de A y una matriz inversible P tal que $PA = R$.

b) Para cada una de las siguientes matrices determine si tiene inversa, en caso afirmativo, encuéntrela.

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

3. En este ejercicio A y B son dos matrices $n \times n$.

a) Si A es inversible y $AB = O$. Demuestre que $B = O$.

b) AB es inversible $\Leftrightarrow$ A y B son inversibles.

c) Si A no es inversible construya una matriz $B \neq O$ tal que $AB = O$.

d) Encuentre dos matrices A y B tales que $AB = 0$ y $BA \neq 0$.

e) Sean $A \neq 0$ y $B \neq 0$ tales que $AB = 0$. Demuestre que A y B no tienen inversa.

4. Sea $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Demuestre que A tiene inversa $\Leftrightarrow ad - bc \neq 0$. Si las entradas de A son enteros, entonces A^{-1} tiene entradas enteras $\Leftrightarrow ad - bc = \pm 1$.
5. Sean A y B matrices $m \times n$ y $n \times m$ respectivamente tales que $AB = I_m$ y $BA = I_n$. Demuestre que $n = m$.
6. Sea A una matriz $n \times n$.
   a) Suponga que existe k entero positivo tal que $A^k = O$. ¿Es posible que A tenga inversa?
   b) Para $n = 2, 3, 4, 5, 6$ construya ejemplos de matrices $A \neq 0$, tales que $A^k = O$ para algún entero $k > 1$.
   c) Suponga que existe k entero positivo tal que $A^k = I_n$. ¿Tiene inversa A?
   d) Encuentre ejemplos de matrices que satisfagan la parte *c*) y no sean la identidad.
7. Sea A una matriz $n \times n$ que no es inversible. Demuestre que existe B matriz $n \times 1$ tal que el sistema $AX = B$ es inconsistente.
8. Demuestre que las matrices elementales son inversibles.
9. Sea $A = \begin{bmatrix} 1 & 0 & \lambda \\ 0 & \lambda & \lambda \\ \lambda & 2 & 1 \end{bmatrix}$

Determine los valores de λ de manera que el sistema $AX = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$
   a) Tenga solución única.
   b) Tenga infinidad de soluciones.
   c) Sea inconsistente.
10. Sea A una matriz $n \times n$. Demuestre que A no tiene inversa si y sólo si existe una matriz B, $n \times n$ no cero tal que $AB = 0$.
11. Sean P_1, P_2 y P_3 tres puntos del plano cartesiano tales que cualesquiera dos no están en una línea vertical. Demuestre que hay un único polinomio de segundo grado cuya gráfica pasa por ellos. ¿Cuál debe ser la condición para que por cuatro puntos pase la gráfica de un único polinomio de grado tres? ¿Puede generalizar las situaciones anteriores al caso de n puntos?
12. Una matriz A, $n \times n$, se dice triangular inferior (superior) si $a_{ij} = 0$ para todo $i < j \leq n$ ($a_{ij} = 0$ para todo $j < i \leq n$). Sea A una matriz triangular inferior con entradas no cero en la diagonal. Demuestre que A tiene inversa y su inversa también es triangular inferior. ¿Se cumple lo mismo si A es triangular superior?
13. Sea $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ tal que $x_k \neq 0$ para algún k. Se define $M_k = I_n - \Omega E_k$, en donde

$$\Omega = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_{k+1}/x_k \\ \vdots \\ x_n/x_k \end{bmatrix}$$ y E_k es la matriz $1 \times n$ con cero en sus entradas salvo la k-ésima,

cuyo valor es 1. Calcule M_kX. ¿Tiene inversa M_k? La matriz M_k se llama *matriz de Gauss*.

14. Sea A una matriz $n \times n$ tal que $a_{11} \neq 0$. Construya una matriz de Gauss M tal que MA tenga ceros en la primer columna, salvo la primer entrada, la cual debe ser a_{11}. Sea A una matriz 2×2 ¿qué condiciones debe cumplir A para que existan matrices de Gauss $M_1, \ldots, M_k$ tales que $M_k \ldots M_1A$ sea triangular superior?

15. Sea $A = diag\{A, A, \ldots, A\}$, con A_i matriz cuadrada para todo $i = 1, 2, \ldots, k$. Demuestre que $A^n = diag\{A_1^n, A_2^n, \ldots, A_k^n\}$ para todo $n \geq 1$. ¿Cuáles son las condiciones para que A tenga inversa?

16. Sea $A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$, con A_{11} y A_{22} matrices cuadradas. Demuestre que A es inversible

$\Leftrightarrow A_{11}$ y A_{22} lo son. En este caso determine una expresión para la inversa de A.

17. Sean A y B matrices de orden $m \times n$ y $n \times p$ respectivamente; denotemos por A_1, A_2, ..., A_n a las columnas de A y por $B_1, B_2, \ldots, B_n$ a las filas de B. ¿Tiene sentido la ecuación $AB = A_1B_1 + A_2B_2 + \cdots + A_nB_n$? Argumente.

18. En este ejercicio presentamos los elementos matemáticos del modelo de Leontief, ejemplo 1.1.2, página 4. Para una discusión desde el punto de vista económico le sugerimos al lector que consulte [8]. Supongamos que la economía de México se divide en n sectores y denotemos por X el vector de producción, por B el vector de consumo y por A la matriz de entrada-salida en el modelo. La matriz A determina la demanda de los sectores para satisfacer la demanda final. Con esta interpretación y la hipótesis sobre el sistema económico en equilibrio, los datos anteriores se relacionan por medio de la ecuación

$$X = AX + B \tag{2.13}$$

Una pregunta natural es: ¿existe un vector de producción que satisfaga la demanda? Recordemos que una de las hipótesis del modelo de Leontief es que la suma de los elementos de cada columna de A es < 1. En este ejercicio daremos algunas indicaciones que permitan contestar la pregunta.

Dada la matriz $A = (a_{ij})$, $n \times n$. Se define la norma de A denotada $\|A\|$, como

$\|A\| := \max. \left[\sum_{i=1}^{n} |a_{ij}| \right]_{j=1}^{n}$ Demuestre:

a) Para todo par de matrices A y B se cumple $\|AB\| \leq \|A\|\,\|B\|$, en particular $\|A^n\| \leq \|A\|^n$ para todo $n = 1, 2, 3\ldots$

b) Sea A una matriz cuadrada tal que $\|A\| < 1$. Demuestre que $A^k \to 0$ cuando $k \to \infty$. Usando la igualdad $I - A^k = (I - A)(I + A + A^2 + \cdots + A^{k-1})$ y lo probado antes concluya que la matriz $I - A$ tiene inversa y se tiene $(I - A)^{-1} = I + A + A^2 + \cdots + A^{k-1} + \cdots$. De esto se infiere que la ecuación 2.13 tiene solución positiva para todo B con entradas positivas. Una aproximación de la solución es $X = (I + A + A^2 + \cdots + A^{k-1})B$.

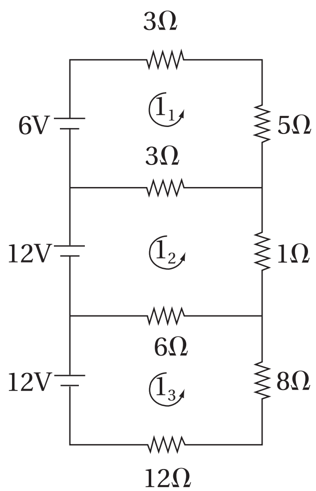
19. Use los resultados del ejercicio anterior y algún sistema computacional para encontrar una aproximación con tres decimales de la solución del sistema $X = AX + B$, en donde:[3]

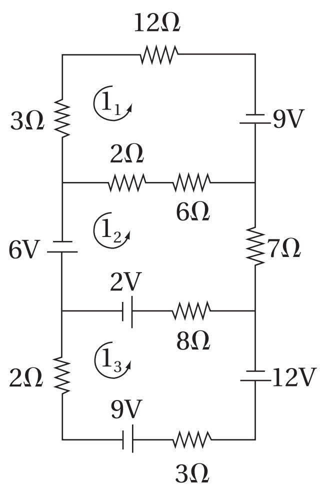
$$A = \begin{bmatrix} .1588 & .0064 & .0025 & .0304 & .0014 & .0083 & .1594 \\ .0057 & .2645 & .0436 & .0099 & .0083 & .0201 & .3413 \\ .0264 & .1506 & .3557 & .0139 & .0142 & .0070 & .0236 \\ .3299 & .0565 & .0495 & .3636 & .0204 & .0483 & .0649 \\ .0089 & .0081 & .0333 & .0295 & .3412 & .0237 & .0020 \\ .1190 & .0901 & .0996 & .1260 & .1722 & .2368 & .3369 \\ .0063 & .0126 & .0196 & .0098 & .0064 & .0132 & .0012 \end{bmatrix}, \quad B = \begin{bmatrix} 74\,000 \\ 56\,000 \\ 10\,500 \\ 25\,000 \\ 17\,500 \\ 196\,999 \\ 5\,000 \end{bmatrix}$$

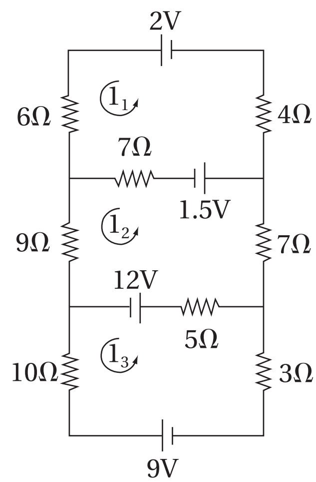
las unidades monetarias están en millones de dólares.

20. Sea A una matriz 2×2 con entradas positivas. Demuestre que existe un $c > 0$ tal que la matriz $A - cI$ no tiene inversa. ¿Se tiene el mismo resultado para matrices de orden tres?

21. Resuelva los circuitos que se representan en las figuras siguientes.







[3] Los datos de las matrices son tomados de [7], ejercicio 13, página 154.

22. Suponga que $D = \begin{bmatrix} .0090 & .0070 & .0040 & .0010 \\ .0070 & .0095 & .0070 & .0040 \\ .0040 & .0070 & .0095 & .0070 \\ .0010 & .0040 & .0070 & .0090 \end{bmatrix}$ es la matriz de flexibilidad para una viga elástica sometida a cuatro pesos uniformemente distribuidos. Las unidades son en centímetros por newton. Conteste las siguientes preguntas.

a) ¿Cuál es el significado físico de la condición $D = D^t$?

b) ¿Tiene inversa la matriz *D*? Use algún sistema de cómputo para contestar esta pregunta y la siguiente.

c) ¿Cuáles son las magnitudes de los pesos que actúan en la viga de modo que producen flexiones de .3, .25, .4, .2 centímetros en los correspondientes puntos?

23. ¿Bajo qué condiciones físicas de una viga, la matriz de flexibilidad es simétrica?

24. ¿Puede ser antisimétrica la matriz de flexibilidad de una viga?

25. Suponga que toda matriz es la suma de una simétrica y una antisimétrica. ¿Cuál es el significado físico de este hecho si la matriz de flexibilidad no es simétrica?

Capítulo 3

Espacios vectoriales

Los conceptos relacionados con el término *vector*, aparecen en diferentes áreas de física, ingeniería y matemáticas. Por mencionar unos cuantos, conceptos importantes en física como fuerza, aceleración y velocidad se pueden representar de manera adecuada por medio de ideas geométricas y algebraicas relacionadas con vectores. Lo concerniente a cuestiones geométricas se logra al identificar a un vector con un segmento dirigido cuyo origen coincide con el de un sistema coordenado. La representación algebraica se logra al asociar a cada punto P del sistema coordenado, distinto del origen, el vector que tiene a P por extremo. Esta última representación de un vector tiene la ventaja de expresarlo en términos de *componentes* en las direcciones de los ejes coordenados. Por ejemplo, en la figura 3.1, el vector representado tiene componente de magnitud 2 en la dirección del eje x y componente de magnitud 5 en la dirección del eje y.

Esta representación de vectores permite interpretar la suma de la siguiente forma. Al sumar dos vectores $\vec{\alpha}$ y $\vec{\beta}$, el resultado debe ser un vector cuyas componentes se obtengan de la "contribución" que cada uno hace en las correspondientes direcciones.

Para formular estas ideas, de manera precisa, se requiere introducir notación y terminología, lo cual hacemos en la siguiente sección.

3.1. Vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$

El plano cartesiano lo entenderemos algebraicamente como el conjunto de parejas ordenadas de números reales y será denotado por $\mathbb{R}^2$. En símbolos, el plano cartesiano se representa por: $\mathbb{R}^2 = \{(x, y) \mid x, y, \in \mathbb{R}\}$.

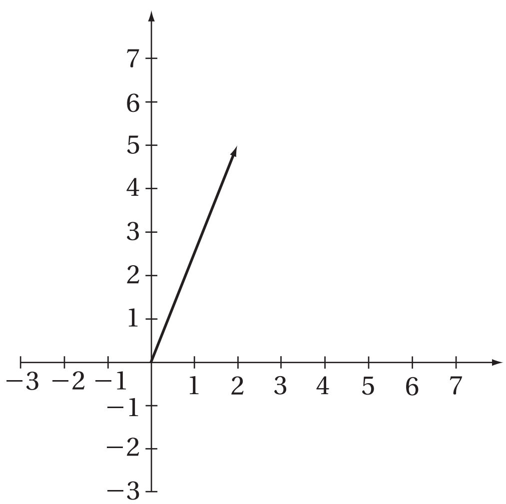


Figura 3.1. Representación gráfica de un vector.

Definición 3.1.1. *Dados* $\alpha = (a, b)$ *y* $\beta = (c, d)$ *en* $\mathbb{R}^2$ *se define su suma como:* $\alpha + \beta := (a + c, b + d)$. *Si* $\lambda \in \mathbb{R}$ *se define el producto de* $\alpha = (a, b)$ *por el escalar* λ *como* $\lambda(a, b) := (\lambda a, \lambda b)$.

La suma y multiplicación por escalar definidas anteriormente tienen un significado geométrico que se ilustra en lo que sigue.

Suma: La representación geométrica de la suma de vectores se obtiene trasladando a uno de ellos al extremo del otro, formando un paralelogramo una de cuyas diagonales representa el resultado. Ver la figura 3.2. A esta representación de la suma de vectores es a lo que se le llama la *ley del paralelogramo*.

Ejemplos

1. Represente geométricamente los vectores (1, 2), $(-2, \sqrt{2})$, $(\pi, -3)$.
2. Represente geométricamente la suma de los pares de vectores siguientes:
   - $\vec{u} = (2, 3)$ y $\vec{v} = (-1, -3)$
   - $\vec{u} = (-2, 4)$ y $\vec{v} = (-\pi, -3)$
   - $\vec{u} = (0, 3)$ y $\vec{v} = (-\sqrt{3}, -3)$

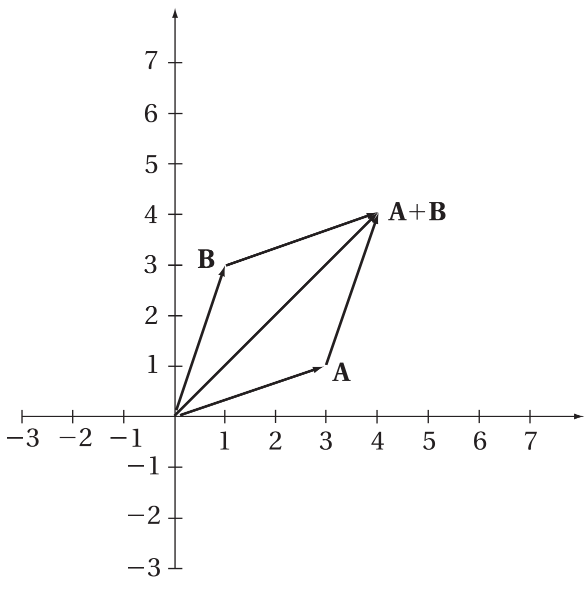


Figura 3.2. Representación gráfica de la suma de vectores: ley del paralelogramo.

3. Dado un vector no cero $\vec{u} = (a, b)$, represente geométricamente al vector $\vec{U} = \frac{1}{\lambda}\vec{u}$ en donde $\lambda = \sqrt{a^2+b^2}$. Demuestre que las componentes de $\vec{U}$ satisfacen la ecuación $x^2 + y^2 = 1$. El escalar λ se llama la *norma* de $\vec{u}$. Note que λ es la distancia del origen al punto que determina al vector.

Producto por escalar

El producto de un escalar por un vector se interpreta así: si el escalar es positivo y diferente de uno, el vector cambia su *magnitud*; si el escalar es negativo y diferente de menos uno, el vector cambia su *magnitud* y *sentido*; si el escalar es cero, el vector se hace cero. En la figura 3.3 se ilustra el caso en que el escalar es mayor que uno.

La formulación de la suma y multiplicación por escalar que hemos hecho en el plano cartesiano se puede extender al espacio cartesiano $\mathbb{R}^3$; de manera más general, al espacio $\mathbb{R}^n$, el cual será definido más adelante.

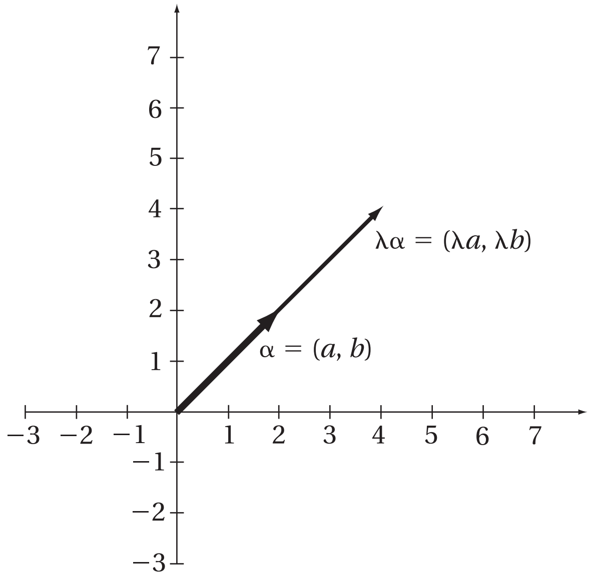


Figura 3.3. Representación gráfica del producto por escalar.

La representación geométrica de $\mathbb{R}^3$ se hace mediante un sistema de tres ejes perpendiculares a pares como se ilustra en la figura 3.4. Tal representación puede ser interpretada como la construcción de un tercer eje perpendicular al plano determinado por los ejes x y y. Esto se puede describir diciendo que los elementos del espacio son triadas de números reales. Más precisamente, $\mathbb{R}^3 := \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$.

Para definir la suma y producto por escalar en $\mathbb{R}^3$ procedemos como en $\mathbb{R}^2$, es decir, la definición 3.1.1 se extiende a $\mathbb{R}^3$ de forma natural, agregando una tercera coordenada. La suma de vectores y producto por escalar en $\mathbb{R}^3$ se interpreta de manera análoga a lo que hemos hecho en el plano.

Ejercicio 3.1.1. *Haga representaciones de suma de vectores y producto por escalar en* $\mathbb{R}^3$, *haciendo énfasis en lo que representa cada coordenada.*

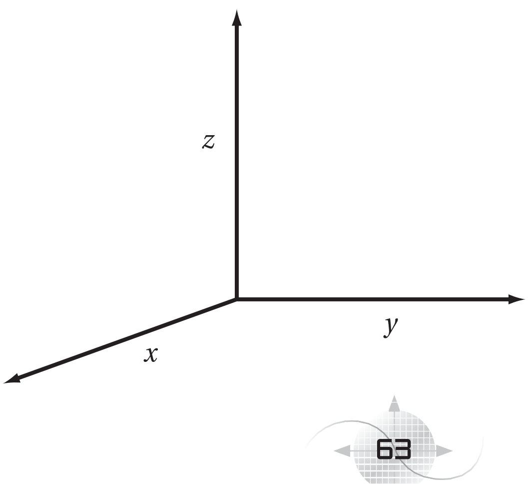


Figura 3.4. Representación de un sistema de ejes coordenados en $\mathbb{R}^3$.

3.2. Combinaciones lineales y dependencia lineal

Con el lenguaje y operaciones entre vectores podemos dar otra interpretación a un sistema de ecuaciones. Dicha interpretación resultará muy útil para determinar soluciones, así como para representar información.

Consideremos el sistema:

$$\begin{aligned} 2x + y &= 3 \\ x + 2y &= 4 \end{aligned} \tag{3.1}$$

Usando notación de vectores, estas ecuaciones se pueden representar como:

$$(2x + y, x + 2y) = (3, 4) \text{ o en forma equivalente } x(2, 1) + y(l, 2) = (3, 4)$$

A los vectores de la forma $x(2, 1) + y(1, 2)$ les llamaremos *combinación lineal* de los vectores (2, 1) y (1, 2). Note que estos vectores se obtienen de las columnas de la matriz de coeficientes del sistema 3.1. Una forma de interpretar que el sistema 3.1 tiene solución, es que el vector (3, 4) sea combinación lineal de los vectores (2, 1) y (1, 2).

Para valores positivos de los escalares x, y, el miembro de la izquierda en cualquiera de las ecuaciones anteriores representa a todos los puntos de la región sombreada que se ilustra en la figura 3.5.

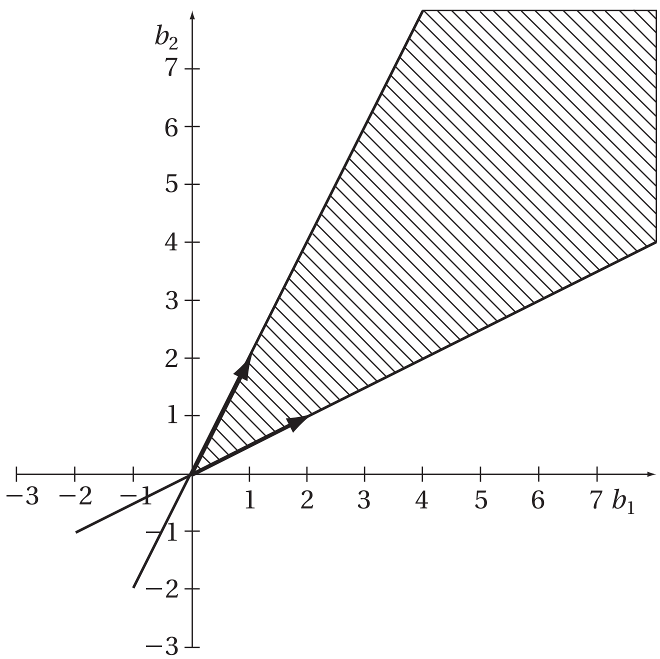


Figura 3.5. Representación de $x(1, 2) + y(2, 1)$, para $x, y \geq 0$.

Si a un sistema de ecuaciones se le representa en forma vectorial, como en el ejemplo que estamos discutiendo, el hecho de tener soluciones significa que el vector de términos constantes se expresa como combinación lineal de las columnas de la matriz de coeficientes.

Como el vector (3, 4) se encuentra en la región sombreada, sin resolver el sistema de ecuaciones podemos decir que tiene solución positiva.

Haciendo un análisis más general podemos llegar a que la ecuación $(2x + y, x + 2y) = (a, b)$ tiene solución para todo (a, b). Una forma geométrica de este análisis consiste en notar que los vectores (1, 2) y (2, 1) determinan a las rectas $L_1 : y = 2x$ y $L : y = \frac{x}{2}$ respectivamente. Para alcanzar el punto (a, b), partiendo del origen, puede uno desplazarse en direcciones paralelas a tales rectas hasta alcanzar dicho punto. Ver la figura 3.6.

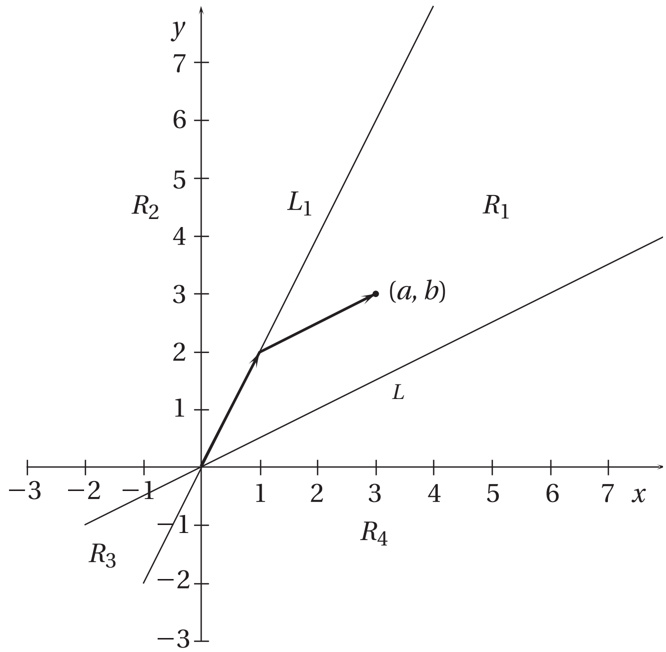


Figura 3.6. Representación de $x(1, 2) + y(2, 1) = (a, b)$.

De acuerdo con la figura 3.6, el plano queda dividido en cuatro regiones que hemos denotado R_1, R_2, R_3 y R_4.

Sin resolver la ecuación $x(1, 2) + y(2, 1) = (a, b)$, podemos determinar los signos de x, y de manera que la satisfagan.

1. Para $(a, b) \in R_1$, los valores de x, y son positivos. $(x, y \geq 0)$
2. Para $(a, b) \in R_2$, se tiene $x \geq 0$ y $y \leq 0$.
3. Para $(a, b) \in R_3$, ambos son negativos. $(x, y \leq 0)$
4. Para $(a, b) \in R_4$, se tiene $x \leq 0$ y $y \geq 0$.

En la discusión anterior, a un sistema de ecuaciones lo hemos representado como una sola ecuación. Esto se puede extender a cualquier sistema de ecuaciones, por ejemplo, el sistema:

$$a_1x + b_1y = c_1 \tag{3.2}$$

$$a_2x + b_2y = c_2 \tag{3.3}$$

lo podemos representar en la forma $x(a_1, a_2) + y(b_1, b_2) = (c_1, c_2)$ y esta ecuación tendrá solución exactamente cuando (c_1, c_2) sea combinación lineal de (a_1, a_2) y (b_1, b_2). Note que si $c_1 = c_2 = 0$ el sistema siempre tiene solución y cuando tiene más de una diremos que los vectores (a_1, a_2) y (b_1, b_2) son *linealmente dependientes*. En caso contrario diremos que son *linealmente independientes* . Dado que el concepto de independencia lineal es uno de los más importantes en álgebra lineal lo formulamos en una definición.

Definición 3.2.1. *Los vectores* $\alpha_1, \alpha_2, \ldots, \alpha_k$ *son linealmente independientes si la única solución de la ecuación* $x_1\alpha_1 + x_2\alpha_2 + \cdots + x_k\alpha_k = 0$ *es* $x_1 = x_2 = \cdots = x_k = 0$. *En caso contrario son linealmente dependientes.*

Observación 3.2.1

- *Si* $k = 2$ *y* $\alpha_1, \alpha_2 \in \mathbb{R}^2$ *son linealmente dependientes, entonces existen escalares* a *y* b, *no ambos cero, tales que* $a\alpha_1 + b\alpha_2 = 0$. *Por ejemplo, si* $a = 0$ *entonces de la ecuación anterior tenemos* $\alpha_1 = -\frac{b}{a}\alpha_2$. *Recíprocamente, si existe una* c *tal que* $\alpha_1 = c\alpha_2$, *entonces la ecuación* $x\alpha_1 + y\alpha_2 = 0$ *tiene más de una solución:* $x = y = 0$ *y* $x = 1$, $y = -c$ *son soluciones. Lo que hemos discutido lo podemos formular diciendo que los vectores* α_1 *y* α_2 *son linealmente dependientes* $\Leftrightarrow$ *uno es múltiplo del otro. La interpretación geométrica de esto es:* α_1 *y* α_2 *son linealmente dependientes* $\Leftrightarrow$ *están en una recta que pasa por el origen. Ver figura 3.7.*
- *Si uno de los vectores en la definición 3.2.1 es cero, entonces son linealmente dependientes. Por ejemplo. si* $\alpha_1 = 0$, *entonces además de la solución que consiste de ceros se tiene la solución* $x_1 = 1, x_2 = x_3 = \cdots = x_k = 0$.

Ejercicio 3.2.1. *Demuestre lo siguiente.*

- *Cualesquiera tres vectores o más en* $\mathbb{R}^2$ *son linealmente dependientes. Este enunciado equivale a: dados cualesquiera tres vectores o más en* $\mathbb{R}^2$, *al menos uno de ellos es combinación lineal de los restantes.*
- *Cualesquiera cuatro vectores o más en* $\mathbb{R}^3$ *son linealmente dependientes.*

Ejercicio 3.2.2. *Sean* A *una matriz* 3×3, X *y* B *matrices* 3×1. *Represente al sistema* $AX = B$ *como una ecuación en donde* B *es combinación lineal de las columnas de* A.

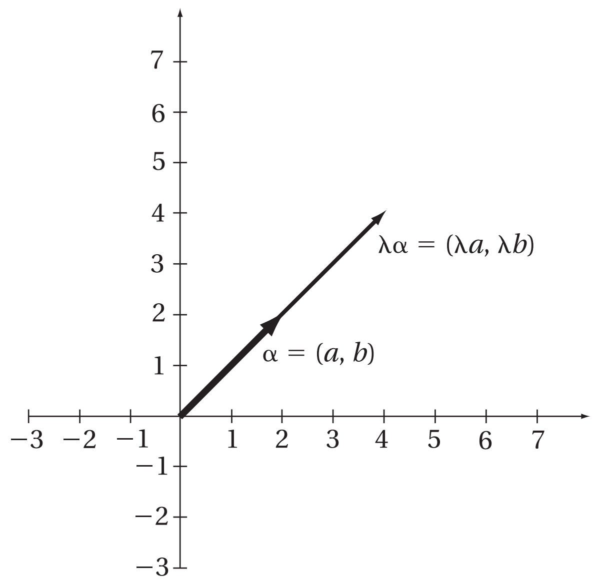


Figura 3.7. Dependencia lineal de vectores.

Ejemplo 3.2.1. *El ejemplo de las refinerías reformulado.*
Consideremos nuevamente el ejemplo de las refinerías, ejemplo 1.1.4, página 16. Recordemos que por cada barril de petróleo la producción de las refinerías está determinada por la tabla 3.1.

La segunda columna de la tabla contiene las cantidades de gasolina, diesel y aceite que produce la refinería 1 por barril; de manera similar, la tercera y cuarta columnas representan la producción de las refinerías 2 y 3 respectivamente. La producción de cada una de las refinerías la podemos organizar en forma de vectores:

Producción de la refinería 1: (20, 11, 9).
Producción de la refinería 2: (21, 12, 8).
Producción de la refinería 3: (19, 13, 8).

Al vector (20, 11, 9) le llamamos el vector de producción de la refinería 1, por barril de crudo; entonces multiplicar este vector por un escalar $x \geq 0$, significa que la refinería 1 procesará x barriles de crudo y se tendrá un vector de producción $x(20, 11, 9) = (20x, 11x, 9x)$. Su interpretación geométrica la visualizamos como una semirrecta que tiene

	Refinería 1	Refinería 2	Refinería 3
Gasolina	20	21	19
Diesel	11	12	13
Aceite lubricante	9	8	8

Tabla 3.1. Representación de la producción de las refinerías.

su origen en (0, 0, 0) y pasa por el punto (20, 11, 9). En términos económicos, el conjunto $\{(20x, 11x, 9x) : x \geq 0\}$ se llama *actividad de producción* de la refinería 1. Para cada una de las refinerías se puede hacer la misma formulación y obtener las correspondientes actividades de producción. Combinando las actividades de producción de las tres refinerías definimos el conjunto de producción de las tres como:

$$P := \{x(20, 11, 9) + y(21, 12, 8) + z(19, 13, 8) : x, y, z \geq 0\}$$

Una de las preguntas planteadas en la discusión fue: ¿cuántos barriles debe procesar cada una de las refinerías para tener una producción de 1 250 galones de gasolina, 750 galones de diesel y 520 de aceite lubricante? En lenguaje puramente matemático la pregunta se puede formular diciendo: ¿cuáles son los valores no negativos de x, y y z que justifican $(1\,250, 750, 520) \in P$? En forma equivalente, ¿cuáles son los valores no negativos de x, y y z para que la ecuación:

$$x(20, 11, 9) + y(2l, 12, 8) + z(19, 13, 8) = (1\,250, 750, 520) \tag{3.4}$$

tenga solución?

Haciendo las operaciones indicadas en la ecuación 3.4 e igualando coordenada a coordenada esta ecuación se transforma en el sistema:

$$\begin{aligned} 20x + 21y + 19z &= 1\,250 \\ 11x + 12y + 13z &= 750 \\ 9x + 8y + 8x &= 520 \end{aligned}$$

De lo discutido en el primer capítulo, la solución de este sistema es: $x = \frac{1\,430}{69}$, $y = \frac{1\,480}{69}$ y $z = \frac{1\,400}{69}$. Entonces se tiene:

$$\frac{1\,430}{69}(20, 11, 9) + \frac{1\,480}{69}(21, 12, 8) + \frac{1\,400}{69}(19, 13, 8) = (1\,250, 750, 520)$$

Ejercicio 3.2.3. *Represente geométricamente al conjunto* P *y use argumentos geométricos para decidir si los siguientes vectores están en* P. *Note que* P *es una pirámide infinita con vértice en* (0, 0, 0).

1. (200, 30, 50)

2. (1 000, 300, 400)

Use argumentos algebraicos para comprobar sus conclusiones.

3.2.1. Ejercicios

1. Dados los vectores $\alpha = (l, 2)$ y $\beta\ (2, -3)$.

   a) Bosqueje el conjunto $\{x\alpha + y\beta\ :\ x, y \geq 0\}$.

   b) Determine si cualquier vector de $\mathbb{R}^2$ es combinación lineal de α y β.

   c) Bosqueje el conjunto $\{x\alpha + y\beta\ :\ x > 0, y \leq 0\}$.

   d) Explique, en términos de combinaciones lineales, el significado que damos a que el sistema:

$$\begin{aligned} x + 2y &= 1 \\ 2x - 3y &= 6 \end{aligned}$$

   tenga solución.

2. Exprese el sistema:

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

   como una combinación lineal de vectores.

3. Represente el cuadrado con vértices en los puntos: (0, 0), (1, 0), (1, 1) y (0, 1), como combinación lineal de los vectores (1, 0) y (0, 1).

4. Determine cuáles vectores son combinación lineal de las columnas de la matriz:

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

5. ¿Cómo representa a una recta y a una semirrecta que pasan por el origen usando vectores?

6. ¿Qué significa que el vector (1, 2, 3) se exprese como combinación lineal de las columnas de la matriz

$$\begin{pmatrix} 2 & 0 & 2 \\ 5 & 2 & -1 \\ 1 & 2 & 3 \end{pmatrix}?$$

3.3. Aspectos geométricos de $\mathbb{R}^2$ y $\mathbb{R}^3$ vía álgebra lineal

Antes de formular los conceptos que permitirán discutir aspectos geométricos de $\mathbb{R}^2$ y $\mathbb{R}^3$, es conveniente presentar una discusión de la ecuación de una recta y un plano.

Del curso de geometría analítica sabemos que la ecuación general de una recta L, referida al plano cartesiano, es de la forma $Ax + By + C = 0$; en donde A, B y C son constantes reales y al menos uno de A y B es no cero. Considerando a esta recta como un subconjunto del plano cartesiano, se puede decir que los puntos de L son precisamente aquellos cuyas coordenadas satisfacen la ecuación $Ax + By + C = 0$. Este conjunto de puntos se puede describir de otra forma: *existe un vector* $\alpha \neq 0$ *y un punto* $\mathrm{P} \in \mathrm{L}$ *tales que los puntos de* L *son precisamente de la forma* $\mathrm{P} + \mathrm{t}\vec{\alpha}$, *en donde* t *varía en* $\mathbb{R}$. A esta representación de L se le llama *la forma vectorial* o *paramétrica*. Ver la figura 3.8.

En efecto, sea $P = (x_0, y_0)$ tal que $Ax_0 + By_0 + C = 0$, entonces $-Ax_0 - By_0 = C$. Sustituyendo C en la ecuación de la recta y agrupando se tiene $A(x - x_0) + B(y - y_0) = 0$, es decir, las coordenadas de un punto (x, y) satisfacen la ecuación $Ax + By + C = 0 \Leftrightarrow A(x - x_0) + B(y - y_0) = 0$. Por otro lado, para cada $t \in \mathbb{R}$ se tiene $A(-tB) + B(tA) = 0$, entonces haciendo $x - x_0 = -tB$ y $y - y_0 = tA$, estas dos ecuaciones se pueden representar como: $(x, y) = (x_0, y_0) + t(-B, A)$.

Proponiendo $\vec{\alpha} = (-B, A)$ y $P = (x_0, y_0) \in L$ se tiene que un punto (x, y) pertenece a L si y sólo si $(x, y) = (x_0, y_0) + t(-B, A)$.

Usando la representación vectorial de una recta y tomado en cuenta que un plano se "genera" por rectas, es "razonable" definir un plano en $\mathbb{R}^3$ como:

Definición 3.3.1. *Dados:* $\mathrm{P} = (\mathrm{x}_0, \mathrm{y}_0, \mathrm{z}_0) \in \mathbb{R}^3$ *y dos vectores linealmente independientes,* $\vec{\alpha}$, $\vec{\beta} \in \mathbb{R}^3$, *al subconjunto* $\{\mathrm{P} + \mathrm{s}\vec{\alpha} + \mathrm{t}\vec{\beta} : \mathrm{s}, \mathrm{t} \in \mathbb{R}\}$ *le llamaremos el plano que contiene a* P *y es generado por* $\vec{\alpha}$ *y* $\vec{\beta}$.

¿Qué ocurre si en la definición anterior no se pide que los vectores sean linealmente independientes? Por ejemplo, tome los vectores $\vec{\alpha} = (1, 1, 2)$, $\vec{\beta} = (3, 3, 6)$ y el punto $P = (0, 0, 0)$. Bosqueje los elementos de la forma $P + s\vec{\alpha} + t\vec{\beta}$, con $s, t \in \mathbb{R}$.

Ejercicio 3.3.1. *Haga un bosquejo geométrico de un plano, de acuerdo con la definición anterior.*

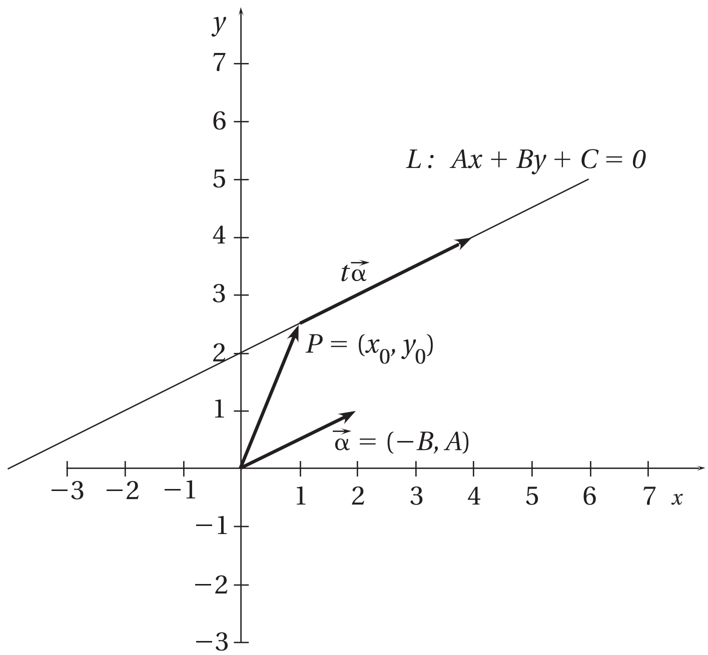


Figura 3.8. Representación vectorial de la recta $L : Ax + By + C = 0$.

Ejercicio 3.3.2. *Describa el plano que contiene al punto* P = (1, 2, 0) *y es generado por los vectores* $\vec{\alpha} = (1, 2, 0)$ *y* $\vec{\beta} = (1, 1, 1)$.

Ejercicio 3.3.3. *Describa el plano que contiene a los puntos* (1, 0, 0), (0, 1, 0) y (1, 1, 1).

Ejercicio 3.3.4. *Determine si los siguientes conjuntos son planos.*

- $\{(1 + 2t + s, 3 + 4t + s, 0) \in \mathbb{R}^3 : s, t \in \mathbb{R}\}$
- $\{(t, s, 3) \in \mathbb{R}^3 : s, t \in \mathbb{R}\}$

3.3.1. Norma y producto interno

Dos conceptos fundamentales en geometría euclidiana son: distancia entre dos puntos y ángulo entre rectas. Estos conceptos se pueden formular en un espacio vectorial que tiene condiciones apropiadas, como es el caso de $\mathbb{R}^2$, $\mathbb{R}^3$ y en general $\mathbb{R}^n$. Para tal efecto se requiere definir el concepto de norma o longitud de un vector y ángulo entre vectores. Iniciamos con la idea de norma o magnitud de un vector.

Una aplicación del Teorema de Pitágoras en el espacio permite definir la distancia del origen de coordenadas al punto que determina un vector $\vec{\alpha} = (a, b, c)$. Esto se precisa en la:

Definición 3.3.2. *Dado un vector* $\vec{\alpha} = (\mathrm{a, b, c}) \in \mathbb{R}^3$ *se define y se denota su norma o longitud como*: $\|\vec{\alpha}\| := \sqrt{a^2 + b^2 + c^2}$.

Ejercicio 3.3.5. *Haga un dibujo que ilustre la definición de norma de un vector.*

Dados los vectores $\vec{\alpha} = (x, y)$ y $\vec{\beta} = (a, b)$ no cero en $\mathbb{R}^2$, éstos forman un ángulo θ como se ilustra en la figura 3.9.

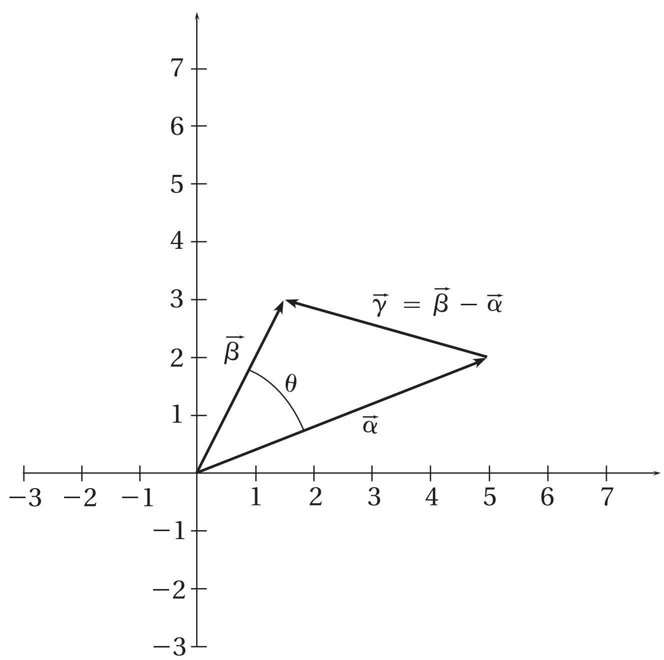


Figura 3.9. Ángulo entre vectores en $\mathbb{R}^2$.

Sea $\vec{\gamma} = \vec{\beta} - \vec{\alpha} = (a - x, b - y)$. De acuerdo con la ley de los cosenos tenemos:

$$\|\vec{\gamma}\|^2 = \|\vec{\alpha}\|^2 + \|\vec{\beta}\|^2 - 2\|\vec{\alpha}\|\,\|\vec{\beta}\|\cos(\theta) \tag{3.5}$$

Usando las coordenadas de los vectores y la definición de norma, esta ecuación se puede escribir como:

$$(a-x)^2+(b-y)^2=x^2+y^2+a^2+b^2-2\sqrt{x^2+y^2}\sqrt{a^2+b^2}\ \cos(\theta)$$

Desarrollando los cuadrados y simplificando, la ecuación anterior conduce a:

$$\cos(\theta)=\frac{ax+by}{\sqrt{x^2+y^2}\sqrt{a^2+b^2}} \tag{3.6}$$

Notemos que la ecuación 3.6, establece que los vectores $\vec{\alpha}$ y $\vec{\beta}$ son perpendiculares $\Leftrightarrow ax+by=0$, pues para $0\le\theta\le\pi$, $\cos(\theta)=0\Leftrightarrow\theta=\frac{\pi}{2}$.

Definición 3.3.3. *Dados los vectores* $\vec{\alpha}=(\mathrm{x},\mathrm{y})$ *y* $\vec{\beta}=(\mathrm{a},\mathrm{b})$, *se define su producto interno como* $\langle\vec{\alpha},\ \vec{\beta}\rangle := \mathrm{xa}+\mathrm{yb}$.

Con esta notación, y usando la ecuación 3.6, el ángulo θ entre dos vectores $\vec{\alpha}$ y $\vec{\beta}$ se puede obtener a través de la ecuación:

$$\cos(\theta)=\frac{\langle\vec{\alpha},\ \vec{\beta}\rangle}{\|\vec{\beta}\|\,\|\vec{\alpha}\|} \tag{3.7}$$

De la discusión anterior, una pregunta natural será si hay una expresión análoga a la ecuación 3.7 para vectores en $\mathbb{R}^3$, es decir, si $\vec{\alpha}=(x,y,z)$ y $\vec{\beta}=(a,b,c)$ son elementos en $\mathbb{R}^3$, ¿cómo se determina el ángulo entre ellos? Puesto que si $\vec{\alpha}=(x,y,z)$ y $\vec{\beta}=(a,b,c)$ son vectores en $\mathbb{R}^3$, éstos y el origen pertenecen a un mismo plano. Ver figura 3.10.

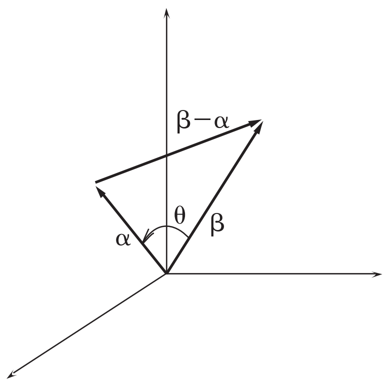


Figura 3.10. Ángulo entre vectores en $\mathbb{R}^3$.

Podemos proceder de manera análoga a lo que hicimos en $\mathbb{R}^2$ y aplicar la ley de los cosenos, es decir, se tiene:

$$\|\vec{\beta}-\vec{\alpha}\|^2=\|\vec{\alpha}\|^2+\|\vec{\beta}\|^2-2\|\vec{\alpha}\|\,\|\vec{\beta}\|\cos(\theta)$$

Usando las coordenadas de α y β, así como la definición de norma de un vector, la ecuación anterior equivale a:

$$(a-x)^2+(b-y)^2+(c-z)^2=x^2+y^2+z^2+a^2+b^2+c^2-2\sqrt{x^2+y^2+z^2}\sqrt{a^2+b^2+c^2}\ \cos(\theta)$$

Desarrollando el cuadrado de los binomios y simplificando, esta ecuación conduce a:

$$-2ax - 2by - 2cz = -\,2\sqrt{x^2+y^2+z^2}\,\sqrt{a^2+b^2+c^2}\,\cos(\theta)$$

y de esta última se obtiene:

$$\cos(\theta) = \frac{ax+by+cz}{\sqrt{x^2+y^2+z^2}\sqrt{a^2+b^2+c^2}} \tag{3.8}$$

De igual forma que en $\mathbb{R}^2$, los vectores $\vec{\alpha} = (x, y, z)$ y $\vec{\beta} = (a, b, c)$ son perpendiculares, es decir, el ángulo entre ellos es $\frac{\pi}{2} \Leftrightarrow \cos(\theta) = 0$. Usando la ecuación 3.8, esto último ocurre $\Leftrightarrow ax + by + cz = 0$. Como ya fue hecho para el caso de $\mathbb{R}^2$, se define el producto interno de $\vec{\alpha}$ y $\vec{\beta}$ agregando una coordenada: $\langle \vec{\alpha}, \vec{\beta} \rangle := xa + yb + zc$, teniéndose que $\vec{\alpha}$ es perpendicular a $\vec{\beta} \Leftrightarrow \langle \vec{\alpha}, \vec{\beta} \rangle = 0$.

Con los términos y notación introducidos, el ángulo entre vectores en $\mathbb{R}^3$ se expresa mediante la ecuación:

$$\cos(\theta) = \frac{\langle \vec{\alpha}, \vec{\beta} \rangle}{\|\vec{\beta}\|\,\|\vec{\alpha}\|} \tag{3.9}$$

en analogía al caso de $\mathbb{R}^2$.

Ejercicio 3.3.6. *Calcule el ángulo entre las parejas de vectores* (1, 2) *y* (−1, 3); (1, 2, 3) *y* (−1, 3, 4).

Ejercicio 3.3.7. *Determine el ángulo entre las rectas cuyas ecuaciones son:* $2x + y = 1$, $x + y = 3$.

Ejercicio 3.3.8. (Identidad de Lagrange). *Parte de este ejercicio consiste en que usted proponga una definición de* $\mathbb{R}^n$.

1. *Defina el producto interno para dos elementos de* $\mathbb{R}^n$.
2. *Dados* $\vec{u} = (a_1, a_2, ..., a_n)$ *y* $\vec{v} = (b_1, b_2, ..., b_n)$ en $\mathbb{R}^n$, *demuestre:*

   a) $\langle \vec{u}, \vec{u} \rangle \langle \vec{v}, \vec{v} \rangle = \sum_{i=1}^{n} a_i^2 b_i^2 + \sum_{j=1}^{n}\left(\sum_{i\neq j} a_i^2 b_j^2\right)$

   b) $\langle \vec{u}, \vec{v} \rangle^2 = \sum_{i=1}^{n} a_i^2 b_i^2 + 2\sum_{1\leq i<j\leq n} a_i b_i a_j b_j$

   c) *Usando los incisos anteriores demuestre que:*

   $$\langle \vec{u}, \vec{u} \rangle \langle \vec{v}, \vec{v} \rangle - \langle \vec{u}, \vec{v} \rangle^2 = \sum_{1\leq i<j\leq n} (a_i b_j - a_j b_i)^2 \qquad (\textit{Identidad de Lagrange})$$

3. *Usando la Identidad de Lagrange, obtenga una demostración de la importante desigualdad:*

   $$|x_1y_1 + x_2y_2 + \cdots + x_ny_n| \leq \sqrt{x_1^2+x_2^2+\cdots+x_n^2}\sqrt{y_1^2+y_2^2+\cdots+y_n^2}$$

   llamada Desigualdad de Cauchy-Schwarz.

4. *Sea* S = $\{a_1^2 + a_2^2 + \cdots + a_n^2 : a_i \in \mathbb{R},\ n \in \mathbb{N}\}$. *Usando la Identidad de Lagrange, demuestre que el conjunto* S *es cerrado bajo multiplicación.*
5. *Demuestre que el producto interno satisface las siguientes propiedades:*

   i. $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$ (*Simetría*).

   ii. $\langle \vec{u}, \vec{u} \rangle \geq 0$ y $\langle \vec{u}, \vec{u} \rangle = 0 \Leftrightarrow \vec{u} = 0$, (*Positividad*).

   iii. $\langle \lambda\vec{u} + \mu\vec{v}, \vec{w} \rangle = \lambda\langle \vec{u}, \vec{w} \rangle + \mu\langle \vec{v}, \vec{w} \rangle$ (*Linealidad por la izquierda*).

   iv. $\langle \vec{u}, \lambda\vec{v} + \mu\vec{w} \rangle = \lambda\langle \vec{u}, \vec{v} \rangle + \mu\langle \vec{u}, \vec{w} \rangle$ (*Linealidad por la derecha*). *Note que esta propiedad se obtiene de la primera y la tercera.*

3.3.2. Proyección ortogonal de un vector sobre otro

En varios problemas geométricos y de física, es importante determinar la proyección ortogonal de un vector en la dirección de otro. Por ejemplo, esa es la idea que se usa para determinar la distancia de un punto a una recta. Esta idea se puede usar para determinar la distancia de un punto a un plano. Dados dos vectores no cero, $\vec{u}$ y $\vec{v}$, podemos descomponer a $\vec{u}$ como suma de dos vectores: uno ortogonal a $\vec{v}$ que le denotaremos $\vec{x}$ y otro paralelo a $\vec{v}$ que es de la forma $\lambda\vec{v}$. Ver figura 3.11.

Las condiciones anteriores se pueden describir como sigue: $\vec{u} = \vec{x} + \lambda\vec{v}$. De esto se tiene $\vec{x} = \vec{u} - \lambda\vec{v}$. Ahora, de la hipótesis sobre $\vec{x}$ y $\vec{v}$ tenemos $0 = \langle \vec{x}, \vec{v} \rangle = \langle \vec{u} - \lambda\vec{v}, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle - \lambda\langle \vec{v}, \vec{v} \rangle$. De esta última ecuación se obtiene el valor de λ, pues como $\vec{v} \neq \vec{0}$, entonces su norma tampoco es cero y se tiene, $\lambda = \dfrac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}$.

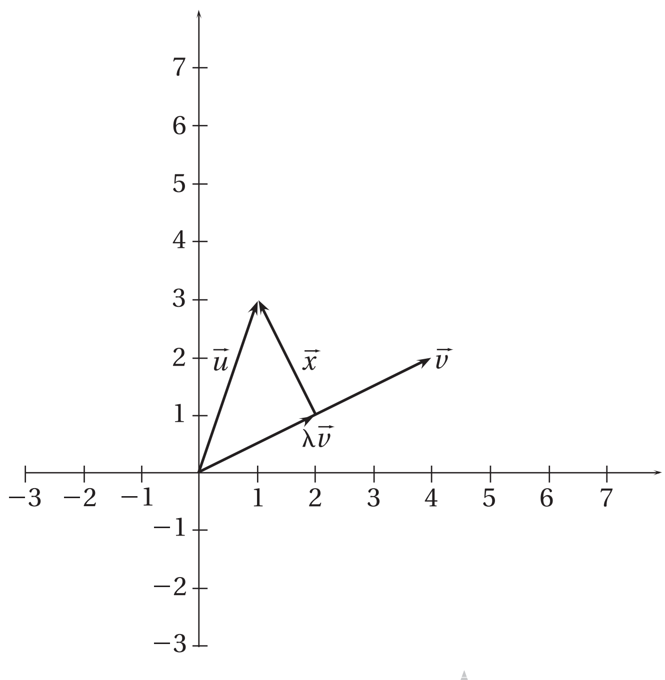


Figura 3.11. Proyección ortogonal de un vector sobre otro.

Definición 3.3.4. *Con las condiciones como en la discusión anterior, al vector* $\frac{\langle \vec{u},\vec{v}\rangle}{\langle \vec{v},\vec{v}\rangle}\vec{v}$ *se le llama la proyección ortogonal del vector* $\vec{u}$ *sobre el vector* $\vec{v}$ *y suele denotársele por* $\operatorname{Proy}_{\vec{v}}\vec{u} = \frac{\langle \vec{u},\vec{v}\rangle}{\langle \vec{v},\vec{v}\rangle}\vec{v}$

Note que el vector $\frac{\vec{v}}{\|\vec{v}\|}$ tiene norma uno, por lo que el vector $\operatorname{Proy}_{\vec{v}}\vec{u}$ tiene longitud $\frac{\langle \vec{u},\vec{v}\rangle}{\|\vec{v}\|}$ en la dirección de $\vec{v}$.

Ejemplo 3.3.1. *Encuentre la proyección ortogonal del vector* $\vec{u} = (1, 2, 3)$ *en la dirección de* $\vec{v} = (2, 0, 1)$. *Haga un bosquejo geométrico.*

Discusión. La proyección de $\vec{u}$ en la dirección de $\vec{v}$ se obtiene aplicando la ecuación de la definición 3.3.4, es decir, $\operatorname{Proy}_{\vec{v}}\vec{u} = \frac{\langle (1,2,3),(2,0,1)\rangle}{\langle (2,0,1),(2,0,1)\rangle}(2,0,1) = \frac{2+3}{4+1}(2,0,1) = (2,0,1)$. Note que, en este caso, la proyección $\vec{u}$ de sobre $\vec{v}$ es el mismo $\vec{v}$.

Ejercicio 3.3.9. *En cada uno de los siguientes casos exprese a* $\vec{u}$ *como la suma de su proyección ortogonal sobre* $\vec{v}$ *y otro ortogonal a* $\vec{v}$. *Haga un bosquejo geométrico de los resultados.*

1. $\vec{u} = (3, 8)$, $\vec{v} = (1, 0)$
2. $\vec{u} = (1, 2, 3)$, $\vec{v} = (1, 1, 0)$
3. $\vec{u} = (2, 1, 1)$, $\vec{v} = (1, 2, 0)$

Una aplicación a geometría

Dada una recta L de ecuación $Ax + By + C = 0$ y un punto $P = (x_0, y_0)$ que no pertenece a ella, determine la distancia de P a L.

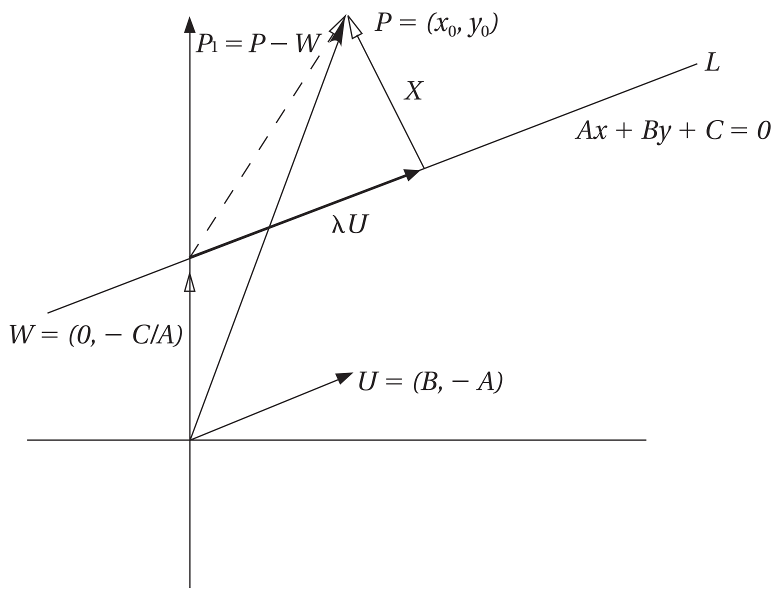


Figura 3.12. Distancia de un punto a una recta.

La figura 3.12 ilustra la situación a discutir.

Notemos que la distancia de P a L es la norma del vector X, entonces debemos calcular $\|X\|$.

Sin perder generalidad, podemos suponer que $AB \neq 0$, pues si alguno de A o B es cero, la recta es vertical u horizontal y el problema se resuelve fácilmente.

Para determinar la norma de X, construimos un vector en la dirección de L; por ejemplo, $U = (B, -A)$ tiene esta condición, pues la pendiente de este vector es la misma que la de L. El extremo del vector $W = (0, -C/B)$ se encuentra en la recta L, pues sus coordenadas satisfacen la ecuación $Ax + By + C = 0$. El vector $P_1 = P - W$ se expresa como $P_1 = X + \lambda U$, pues el segundo sumando es la proyección de P_1 sobre U. Sabemos que el valor de λ es: $\lambda = \dfrac{\langle P_1, U\rangle}{\|U\|^2}$. Como X y λU son perpendiculares, aplicando el Teorema de Pitágoras se obtiene:

$$\|P_1\|^2 = \|X\|^2 + \|\lambda U\|^2 = \|X\|^2 + |\lambda|^2\|U\|^2 \tag{3.10}$$

Por otro lado, tenemos que las coordenadas de P_1 son $(x_0, y_0 + C/B)$. De la ecuación anterior, usando coordenadas, y sustituyendo el valor de λ se tiene:

$$\begin{aligned}
\|X\|^2 &= x_0^2 + (y + C/B)^2 - \frac{\left[Bx_0 - A(y_0 + C/B)\right]^2}{(A^2+B^2)^2}(A^2+B^2)\\
&= x_0^2 + (y + C/B)^2 - \frac{\left[Bx_0 - A(y_0 + C/B)\right]^2}{(A^2+B^2)}\\
&= x_0^2 + \frac{(A^2+B^2)(y_0+C/B)^2 - \left[B^2x_0^2 - 2ABx_0(y+C/B) + A^2(y_0+C/B)^2\right]}{A^2+B^2}\\
&= \frac{x_0^2(A^2+B^2) + A^2(y_0+C/B)^2 + B^2(y_0+C/B)^2}{A^2+B^2}\\
&+ \frac{-B^2x_0^2 + 2ABx_0(y_0+C/B) - A^2(y_0+C/B)^2}{A^2+B^2}\\
&= \frac{A^2x_0^2 + (By_0+C)^2 + 2Ax_0(By_0+C)}{A^2+B^2}\\
&= \frac{(Ax_0+By_0+C)^2}{A^2+B^2}
\end{aligned}$$

De esta ecuación se llega a:

$$\|X\| = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2+B^2}} \tag{3.11}$$

que es la bien conocida fórmula que expresa la distancia de la recta de ecuación $Ax + By + C = 0$ al punto $P = (x_0, y_0)$.

Dado un vector $\vec{u}$ en $\mathbb{R}^2$, nos interesa encontrar a todos aquellos vectores que son ortogonales a $\vec{u}$. Si $\vec{u} = (0, 0)$, entonces todos los vectores del plano son ortogonales a $\vec{u}$. En caso contrario, si $\vec{u} = (a, b)$ estamos buscando a todos aquellos vectores $\vec{X} = (x, y)$ tales que $\langle \vec{u}, \vec{X} \rangle = ax + by = 0$. En esta ecuación se reconoce que los vectores que se buscan son precisamente aquellos que se localizan sobre la recta L que tiene por ecuación: $ax + by = 0$. Si $\vec{v}$ es un vector no cero en L, entonces los vectores que se buscan son precisamente los que son de la forma $\lambda\vec{v}$, con $\lambda \in \mathbb{R}$. Ver figura 3.13.

Si $\vec{u} = (a, b, c) \in \mathbb{R}^3$, entonces los vectores $\vec{X} = (x, y, z)$ que son ortogonales a $\vec{u} = (a, b, c)$ son aquellos que satisfacen: $(\vec{u}, \vec{X}) = ax + by + cz = 0$, y como en el caso anterior, tenemos dos posibilidades: si $\vec{u} = (0, 0, 0)$, entonces todos los elementos de $\mathbb{R}^3$ satisfacen la condición requerida. En otro caso, los vectores que se buscan son precisamente aquellos que satisfacen la ecuación $\langle \vec{u}, \vec{X} \rangle = ax + by + cz = 0$. ¿Qué objeto geométrico representa esa ecuación?

Consideremos la ecuación $ax + by + cz = 0$ y supongamos, por ejemplo, que $a \neq 0$, entonces la podemos escribir en la forma $x = -\frac{b}{a}y - \frac{c}{a}z$. Tomando $y = 0$ y $z = -a$ se tiene el vector $\vec{v}_1 = (c, 0, -a)$. Intercambiando papeles entre y y z en el caso anterior se tiene el vector $\vec{v}_2 = (b, -a, 0)$.

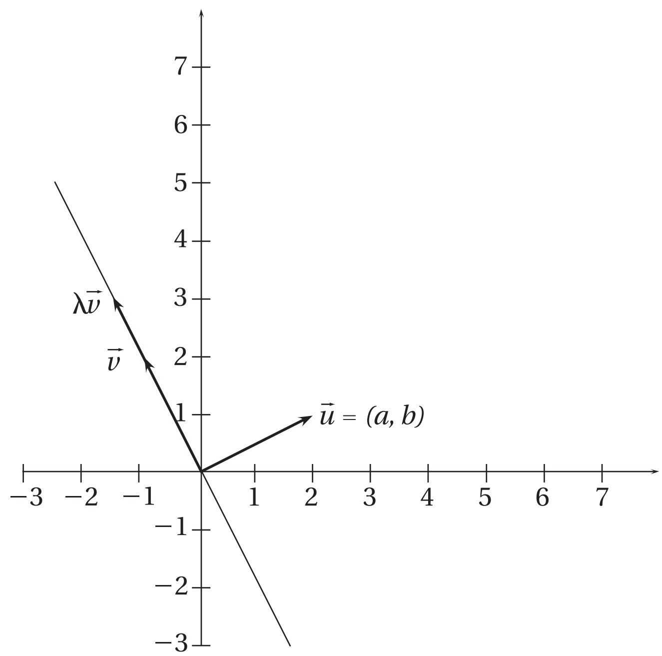


Figura 3.13. Vectores ortogonales a uno dado.

Afirmación 1: $\vec{v}_1$ y $\vec{v}_2$ son linealmente independientes.

Si $x\vec{v}_1 + y\vec{v}_2 = (xc + by, -ax, -ay) = (0, 0, 0)$, la hipótesis sobre a implica que $x = y = 0$, probando lo afirmado.

Afirmación 2: Los conjuntos $\{(x, y, z) : ax + by + cz = 0\}$ y $\{\mu\vec{v}_1 + \lambda\vec{v}_2 : \mu, \lambda, \in \mathbb{R}\}$ son iguales.

Demostración. Ejercicio.

De la discusión anterior tenemos que el conjunto $\{(x, y, z) : ax + by + cz = 0\}$ es un plano que pasa por el origen y es ortogonal al vector $\overrightarrow{N} = (a, b, c)$.

Ejemplo 3.3.2. *Encuentre la ecuación del plano que es ortogonal al vector* $(1, -1, 2)$ *y que pasa por el origen. También encuentre dos vectores que lo generan*.

Discusión. El plano que es ortogonal a $N = (1, -1, 2)$ y que pasa por el origen es el conjunto de puntos $P = \{(x, y, z) \in \mathbb{R}^3\}$ que satisfacen $\langle N, X\rangle = \langle(1, -1, 2), (x, y, z)\rangle = x - y + 2z = 0$. Para encontrar dos vectores que lo generan debemos encontrar dos vectores $\alpha, \beta \in P$ que sean linealmente independientes. De la ecuación que define al plano se tiene que $\alpha = (1, 1, 0)$ y $\beta = (2, 0, -1)$ pertenecen a P y son linealmente independientes.

3.3.3. Producto cruz de vectores

Sean $\vec{u}$ y $\vec{v}$ dos vectores linealmente independientes, ¿cuáles son los vectores $\overrightarrow{X} = (x, y, z)$ que son ortogonales a ellos? Supongamos que $\vec{u} = (a, b, c)$ y $\vec{v} = (a_1, b_1, c_1)$ entonces la condición de ortogonalidad da lugar a las siguientes ecuaciones:

$$ax + by + cz = 0 \tag{3.12}$$

$$a_1x + b_1y + c_1z = 0 \tag{3.13}$$

Multiplicando la ecuación 3.12 por a_1, la ecuación 3.13 por $-a$, sumando y factorizando se obtiene $(a_1b - ab_1)y + (a_1c - ac_1)z = 0$. Por observación se pueden obtener algunas soluciones de esta última ecuación. Tomando $z = -(a_1b - ab_1)$, $y = (a_1c - ac_1)$ y sustituyendo en (3.12) obtenemos $ax - b(a_1c - ac_1) - c(a_1b - ab_1) = 0$ y de esto, $ax = a(bc_1 - b_1c)$. Tomando $x = bc_1 - b_1c$ se tiene una solución del sistema formado por las ecuaciones (3.12) y (3.13), es decir, el vector $\overrightarrow{X} = (bc_1 - b_1c, a_1c - ac_1, ab_1 - a_1b)$ es perpendicular a $\vec{u}$ y a $\vec{v}$.

Observación 3.3.1. *El vector* $\overrightarrow{X}$ *determinado antes es diferente de cero* $\Leftrightarrow$ *los vectores* $\vec{u}$ *y* $\vec{v}$ *son linealmente independientes*.

Demostración. Ejercicio.

Definición 3.3.5. *Dados los vectores* $\vec{u} = (a, b, c)$ y $\vec{v} = (a_1, b_1, c_1)$ *se define el* **producto cruz** *o* **producto vectorial** *de* $\vec{u}$ *y* $\vec{v}$ *como* $\vec{u} \times \vec{v} := (bc_1 - b_1c, a_1c - ac_1, ab_1 - a_1b)$.

Como regla mnemotécnica usaremos la siguiente representación:

$$\vec{u} \times \vec{v} := \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a & b & c \\ a_1 & b_1 & c_1 \end{vmatrix} := (bc_1 - b_1c)\vec{i} + (a_1c - ac_1)\vec{j} + (ab_1 - a_1b)\vec{k},$$

en donde $\vec{i} = (1, 0, 0)$, $\vec{j} = (0, 1, 0)$ y $\vec{k} = (0, 0, 1)$.

Ejemplo 3.3.3. *Dados* $(1, 2, 0)$ *y* $(1, 3, 0)$ *calcule* $(1, 2, 0) \times (1, 3, 0)$.

Discusión. De acuerdo con la definición del producto cruz tenemos:

$$(1, 2, 0) \times (1, 3, 0) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 2 & 0 \\ 1 & 3 & 0 \end{vmatrix} = (2 \cdot 0 - 3 \cdot 0)\vec{i} + (1 \cdot 0 - 1 \cdot 0)\vec{j} + (3 - 2)\vec{k} = (0, 0, 1)$$

Interpretación geométrica de la magnitud de $\vec{u} \times \vec{v}$

Sabemos que el ángulo θ entre los vectores no cero $\vec{u}$ y $\vec{v}$ se obtiene de la ecuación $\cos(\theta) = \dfrac{\langle \vec{u}, \vec{v} \rangle}{\| \vec{u} \| \| \vec{v} \|}$. Por otro lado tenemos la bien conocida identidad $\cos^2(\theta) + \operatorname{sen}^2(\theta) = 1$, de lo cual se obtiene:

$$\operatorname{sen}^2(\theta) = 1 - \cos^2(\theta) = \frac{\langle \vec{u}, \vec{u} \rangle \langle \vec{v}, \vec{v} \rangle - \langle \vec{u}, \vec{v} \rangle^2}{\langle \vec{u}, \vec{u} \rangle \langle \vec{v}, \vec{v} \rangle} \tag{3.14}$$

Asignando coordenadas a los vectores $\vec{u} = (a_1, a_2, a_3)$ y $\vec{v} = (b_1, b_2, b_3)$ y usando la Identidad de Lagrange (2*c*), se tiene:

$$\langle \vec{u}, \vec{u} \rangle \langle \vec{v}, \vec{v} \rangle - \langle \vec{u}, \vec{v} \rangle^2 = (a_1b_2 - a_2b_1)^2 + (a_1b_3 - a_3b_1)^2 + (a_2b_3 - a_3b_2)^2 = \| \vec{u} \times \vec{v} \|^2$$

De esto último y de la ecuación 3.14 obtenemos $\| \vec{u} \times \vec{v} \|^2 = \| \vec{u} \|^2 \| \vec{v} \|^2 = \operatorname{sen}^2(\theta)$. Si $0 \leq \theta \leq \pi$, entonces $\operatorname{sen}(\theta) \geq 0$, por lo que:

$$\| \vec{u} \times \vec{v} \| = \| \vec{u} \| \, \| \vec{v} \| \operatorname{sen}(\theta) \tag{3.15}$$

La expresión de la derecha en 3.15 es el área de un paralelogramo que tiene lados $\vec{u}$ y $\vec{v}$, como usted se puede convencer haciendo un dibujo.

Observación 3.3.2. *La ecuación 3.15 proporciona una forma de calcular el área de un paralelogramo determinado por los vectores* $\vec{u}$ *y* $\vec{v}$ *(explique). También proporciona una alternativa para calcular el ángulo entre dos vectores. La desventaja de esta fórmula para calcular el ángulo entre vectores, es que no se puede extender a* $\mathbb{R}^n$*, sin embargo, en la ecuación 3.14 sí se puede usar para calcular el ángulo entre cualquiera de dos vectores de* $\mathbb{R}^n$ *usando la función seno.*

Ejemplo 3.3.4. *Calcule el área del paralelogramo determinado por los vectores* $\vec{u} = (2, 1, 0)$ *y* $\vec{v} = (1, 3, 0)$.

Discusión. Se tiene $\vec{u} \times \vec{v} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 2 & 1 & 0 \\ 1 & 3 & 0 \end{vmatrix} = (0, 0, 5)$. De donde se obtiene que el área buscada es $\|(0, 0, 5)\| = 5$.

3.3.4. Ecuación de un plano

Sabemos, de la geometría euclidiana, que tres puntos no colineales determinan a un único plano. Si tales puntos P, Q y R pertenecen a $\mathbb{R}^3$, deseamos encontrar la ecuación del plano que los contiene. El procedimiento que utilizaremos se basa en lo siguiente. Dado un vector $\vec{N} \neq 0$, hay un único plano ortogonal a $\vec{N}$ que pasa por el origen, de esto se tiene que dado un vector $\vec{N} \neq 0$ y un punto $S \in \mathbb{R}^3$, hay un único plano ortogonal a $\vec{N}$ y que pasa por S.

Si P, Q y R son puntos no colineales en $\mathbb{R}^3$, los vectores $\alpha = Q - P$ y $\beta = R - P$ son linealmente independientes (demostrarlo), por lo que $\vec{N} = \alpha \times \beta = (a, b, c)$ es un vector no cero ortogonal a α y β. Si $X = (x, y, z)$ es un vector en el plano que pasa por P, entonces $\overrightarrow{N}$ es ortogonal a $X - P$. Supongamos que las coordenadas de P son $P = (a_1, b_1, c_1)$, entonces la condición de ortogonalidad descrita antes se traduce a:

$$\begin{aligned}\langle (a, b, c), (x - a_1, y - b_1, z - c_1)\rangle &= a(x - a_1) + b(y - b_1) + c(z - c_1)\\ &= ax + by + cz + d = 0 \qquad (3.16)\end{aligned}$$

en donde $d = -(aa_1 + bb_1 + cc_1)$.

Resumiendo, la ecuación del plano que pasa por los puntos no colineales $P = (a_1, b_1, c_1)$, $Q = (a_2, b_2, c_2)$ y $R = (a_3, b_3, c_3)$, está dada por:

$$ax + by + cz + d = 0$$

en donde $d = -(aa_1 + bb_1 + cc_1)$, y $(a, b, c) = (Q - P) \times (R - P)$.

Ejemplo 3.3.5. *Encuentre la ecuación del plano que pasa por los puntos* (1, 2, 0), (0, 1, 1) y (1, 2, 3).

Discusión. Tomemos como $P = (1, 2, 0)$, entonces de acuerdo con la notación anterior se tiene, $\alpha = (0, 1, 1) - (1, 2, 0) = (-1, -1, 1)$ y $\beta = (1, 2, 3) - (1, 2, 0) = (0, 0, 3)$. De esto, obtenemos que el producto cruz de α y β está dado por:

$$\overrightarrow{N} = \begin{vmatrix} i & j & k \\ -1 & -1 & +1 \\ 0 & 0 & 3 \end{vmatrix} = -3i + 3j = (-3, 3, 0)$$

Siguiendo el procedimiento descrito antes, la ecuación del plano buscado es:

$$\langle (-3, 3, 0), (x - 1, y - 2, z)\rangle = -3(x - 1) + 3(y - 2) = -3x + 3y - 3 = 0$$

De la discusión anterior se tiene que la ecuación de un plano que es normal a un vector no cero $\overrightarrow{N}$ y pasa por un punto P es de la forma $ax + by + cz + d = 0$. Recíprocamente, una ecuación de la forma $ax + by + cz + d = 0$, con al menos uno de a, b o c diferente de cero, representa un plano que tiene por vector normal a $\overrightarrow{N} = (a, b, c)$. Pues como al menos uno de los coeficientes de las variables es distinto de cero, existe un punto $P = (x_0, y_0, z_0)$ cuyas coordenadas satisfacen a la ecuación dada, es decir, $ax_0 + by_0 + cz_0 + d = 0$. Esta ecuación se puede reescribir como $d = -ax_0 - by_0 - cz_0$. Sustituyendo y agrupando en la ecuación inicial se tiene $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$ y se interpreta así: los puntos del espacio que satisfacen a esta última ecuación son todos aquellos que al restarles $P = (x_0, y_0, z_0)$ son ortogonales a $\overrightarrow{N} = (a, b, c)$.

Lo anterior lo podemos resumir diciendo:

Observación 3.3.3. *Un plano es el conjunto de puntos de* $\mathbb{R}^3$ *que satisfacen una ecuación de la forma* ax + by + cz + d = 0, *con* a, b, c *y* d *constantes reales*.

3.3.5. Ejercicios

1. Haga un dibujo a escala y conteste las siguientes preguntas.

a) ¿Puede uno alcanzar el punto (.5, .7) "caminando" en direcciones paralelas a las líneas determinadas por los vectores (1, 2) y (2, 1)? ¿Caminará más de una unidad en cada dirección?

b) ¿Cuáles son los puntos del plano que puede alcanzar "caminando" una unidad en la dirección del vector (2, 1)?

c) ¿Puede alcanzar cualquier punto del plano desplazándose en direcciones paralelas a las direcciones determinadas por los vectores (1, 2) y (2, 1)?

d) ¿Cómo deben ser dos vectores en $\mathbb{R}^2$ para que se alcance cualquier punto del plano desplazándose en direcciones paralelas a ellos?

2. Dada una circunferencia y tres puntos sobre ella *P*, *Q* y *R*, tales que exactamente dos de ellos se encuentran en un diámetro, digamos *P* y *Q*, demuestre que el ángulo *PRQ* es recto.

3. Dados los vectores $\vec{u}$, $\vec{v}$ y $\vec{w}$ en $\mathbb{R}^3$. Demuestre que $(\vec{u} \times \vec{v}) \times \vec{w} = \langle \vec{u}, \vec{w} \rangle \vec{v} - \langle \vec{v}, \vec{w} \rangle \vec{u}$ **(triple producto vectorial).** Use esta identidad para demostrar que en general no se cumple $(\vec{u} \times \vec{v}) \times \vec{w} = \vec{u} \times (\vec{v} \times \vec{w})$. ¿Bajo qué condiciones se cumpliría la igualdad anterior?

4. Demuestre que el volumen del paralelepípedo determinado por los vectores $\vec{u}$, $\vec{v}$ y $\vec{w}$, está dado por $|\langle \vec{u}, \vec{v} \times \vec{w} \rangle|$. Al número $\langle \vec{u}, \vec{v} \times \vec{w} \rangle$, se le llama **triple producto escalar**.

5. Demuestre que los vectores $\vec{u}$, $\vec{v}$ y $\vec{w}$, son linealmente dependientes $\Leftrightarrow$ $\langle \vec{u}, \vec{v} \times \vec{w} \rangle = 0$. ¿Cuál es la interpretación geométrica de este resultado a la luz del ejercicio 4?

6. Dados los vectores $\vec{u}$, $\vec{v}$ y $\vec{w}$ en $\mathbb{R}^3$, demuestre las siguientes identidades.

a) $\langle (\vec{u} + \vec{v}), (\vec{u} \times \vec{w}) \times (\vec{u} + \vec{v}) \rangle = 0$

b) $\vec{u} \times (\vec{u} \times (\vec{u} \times \vec{v})) = \langle \vec{u}, \vec{u} \rangle (\vec{v} \times \vec{u})$

c) $(\vec{u} \times \vec{v}) \times (\vec{u} \times \vec{w}) = \langle \vec{u}, \vec{v} \times \vec{w} \rangle \vec{u}$

d) $\vec{u} \times (\vec{v} \times \vec{w}) + \vec{v} \times (\vec{w} \times \vec{u}) + \vec{w} \times (\vec{u} \times \vec{v}) = \vec{0}$

7. Sea $\vec{u}$ un vector no cero en $\mathbb{R}^3$, α, β y γ los ángulos que forma este vector con los ejes *x*, *y* y *z* respectivamente. A estos ángulos se les llama *ángulos directores* de $\vec{u}$. Demuestre que se tiene $\cos^2(\alpha) + \cos^2(\beta) + \cos^2(\gamma) = 1$. Sugerencia: inicie con el caso en que el vector está sobre uno de los planos coordenados.

8. Determine si los siguientes puntos son los vértices de un paralelogramo.

$$(3, 7, -2),\ (5, 5, 1),\ (4, 0, -1),\ (6, -2, 2).$$

9. Calcule el volumen del paralelepípedo determinado por los vectores (1, 1, 1), (1, 2, 3) y (0, 1, 0).

10. Encuentre la ecuación del plano que satisface las condiciones estipuladas.

a) Pasa por los puntos (1, 0, –1), (0, 1, 2) y (1, 0, –2).

b) Es normal a (1, 2, 4) y pasa por (1, 5, 9).

c) Pasa por los puntos (*a*, 0, 0), (0, *b*, 0) y (0, 0, *c*). Si $abc \neq 0$. Demuestre que la ecuación del plano se puede escribir como $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$.

11. Encuentre una fórmula para calcular la distancia de un punto a un plano. Sugerencia: revise el procedimiento empleado para obtener la distancia de una recta a un punto.

12. Usando terminología y notación vectorial, defina una esfera y exprésela mediante una ecuación.

13. ¿Cuál es la ecuación de un elipsoide centrado en el origen? ¿Cuál es la ecuación de un paraboloide con vértice en el origen? Exponga todas las posibilidades, tomando como guía el caso de una parábola en el plano cartesiano con vértice en el origen.

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

Después de haber discutido las propiedades básicas de $\mathbb{R}^2$ y $\mathbb{R}^3$ como espacios vectoriales, nos disponemos a efectuar lo análogo en $\mathbb{R}^n$, con lo cual estaremos en posibilidad de formular de manera precisa la conexión entre un sistema de ecuaciones lineales y el concepto de dependencia lineal. Para tal efecto, iniciamos con la discusión del espacio $\mathbb{R}^n$, así como la suma y producto por escalar.

Definición 3.4.1. *Para cada entero positivo* n *definimos:*

$$\mathbb{R}^n := \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R} \textit{ para todo } i = 1, 2, \ldots, n\}$$

La operación de suma en $\mathbb{R}^n$ y el producto por escalar se formulan como:

- dados $(x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, se define

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

- dados $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ y $c \in \mathbb{R}$, se define

$$c\,(x_1, x_2, \ldots, x_n) := (cx_1, cx_2, \ldots, cx_n)$$

Note que la suma y producto por escalar en $\mathbb{R}^n$ son simplemente las extensiones de la suma y producto por escalar en $\mathbb{R}^2$ y $\mathbb{R}^3$. Las representaciones geométricas de la suma y producto por escalar pueden ser "pensadas" de la misma forma que en $\mathbb{R}^2$ y $\mathbb{R}^3$.

Supongamos que A es una matriz $m \times n$, entonces a las columnas de A las podemos considerar como elementos de $\mathbb{R}^m$ y a sus filas como elementos de $\mathbb{R}^n$, con esta convención al sistema de ecuaciones lineales $AX = B$ lo podemos interpretar en términos de combinaciones lineales en $\mathbb{R}^m$. Más precisamente, si $AX = B$ lo escribimos en la forma

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{3.17}$$

entonces este sistema se puede escribir como

$$x_1A_1 + x_2A_2 + \cdots + x_nA_n = B$$

en donde hemos elegido la notación: $A_j = (a_{1j}, a_{2j}, \ldots, a_{mj})$ y $B = (b_1, b_2, \ldots, b_m)$. Si el sistema tiene solución decimos que B es combinación lineal de $A_1, A_2, \ldots, A_n$.

Antes se definieron en $\mathbb{R}^2$ y $\mathbb{R}^3$ los conceptos de combinación lineal y dependencia lineal. Dada la importancia de tales conceptos los recordamos para el caso de $\mathbb{R}^m$.

Definición 3.4.2. *Sean* $A_1, A_2, ..., A_n$ *como antes. Las expresiones* $x_1A_1 + x_2A_2 + \cdots + x_nA_n$ *son llamadas combinaciones lineales de* $A_1, A_2, ..., A_n$ *y al conjunto de todas las combinaciones lineales les llamamos el espacio columna de* A.

Observación 3.4.1. *Note que el sistema* $Ax = B$ *tiene solución* $\Leftrightarrow$ B *está en el espacio columna de* A. *Si B es cero, la ecuación* $ax = 0$ *siempre tiene solución*.

Puede suceder que la ecuación $AX = 0$ tenga más de una solución, en este caso diremos que las columnas de A son linealmente dependientes; la importancia de este caso es tal que lo formulamos de manera general.

Definición 3.4.3. *Sean* $A_1, A_2, ..., A_k$ *vectores en* $\mathbb{R}^m$. *Decimos que son linealmente independientes si la única solución de la ecuación* $x_1A_1 + x_2A_2 + \cdots + x_kA_k = 0$ *es* $x_1 = x_2 = \cdots = x_k = 0$. *En caso que la ecuación anterior tenga más de una solución, decimos que los vectores son linealmente dependientes*.

Teorema 3.4.1. *Los vectores* $\vec{u}_1, \vec{u}_2, ..., \vec{u}_p$, *con* $p \geq 2$, *son linealmente dependientes* $\Leftrightarrow$ *al menos uno de ellos es combinación lineal de los restantes. Más precisamente, si* $\vec{u}_1$ *no es cero, entonces los vectores son linealmente dependientes* $\Leftrightarrow$ *algún* $\vec{u}_j (j > 1)$ *es combinación lineal de los previos*.

Demostración. Es claro que si uno es combinación lineal de los restantes, entonces son linealmente dependientes.

Supongamos que los vectores son linealmente dependientes, entonces la ecuación $x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_p\vec{u}_p = \vec{0}$ tiene una solución con al menos uno de los escalares no cero. Sea k el mayor índice tal que $x_k \neq 0$, entonces la ecuación $x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_p\vec{u}_p = \vec{0}$ se reduce a $x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_k\vec{u}_k = \vec{0}$ y $k > 1$, pues como $\vec{u}_1$ no es cero, k no puede ser 1. De la última ecuación despeje $\vec{u}_k$ y se tiene el resultado.

Observación 3.4.2. *En* $\mathbb{R}^m$ *cualquier colección de* $l \geq m + 1$ *vectores es linealmente dependiente*.

Demostración. La conclusión se obtiene del teorema 1.3.3, página 20.

3.4.1. Subespacios

En muchos problemas que se abordan usando vectores, es importante saber qué subconjuntos del espacio satisfacen las mismas propiedades del espacio total, respecto a las operaciones. Esto se precisa en la siguiente definición.

Definición 3.4.4. *Sea* W *un subconjunto no vacío de* $\mathbb{R}^n$. *Se dice que* W *es su subespacio, si satisface las siguientes propiedades*.

1. *Para todos* $\alpha, \beta \in W$ *se tiene* $\alpha + \beta \in W$, *es decir,* W *es cerrado bajo la suma*.
2. *Para todo* $\alpha \in W$ *y para todo* $\lambda \in \mathbb{R}$, $\lambda\alpha \in W$, *es decir,* W *es cerrado bajo producto por escalar*.

Ejemplo 3.4.1. *Determinaremos cuáles de los siguientes conjuntos son subespacios.*

1. *Sea* $W = \{(x, y) \in \mathbb{R}^2 \mid x, y \in \mathbb{Z}\}$. *Notemos que al sumar dos elementos de* W, *el resultado también es un elemento de* W, *pues las coordenadas de los elementos de* W *son enteros y la suma de enteros es nuevamente un entero. Sin embargo, al tomar* $\lambda = \sqrt{2}$ *y* $\alpha = (1, 2) \in W$, *el resultado* $\lambda\alpha = (\sqrt{2}, 2\sqrt{2}) \notin W$, *es decir,* W *no es subespacio*.

2. *Sea* $\mathrm{W} = \{(\mathrm{x}, \mathrm{y}) \in \mathbb{R}^2 \mid 0 \leq \mathrm{x} \leq 1, 0 \leq \mathrm{y} \leq 1\}$. *Los elementos* (1, 1) (1, .5) *pertenecen a* W, *pero* (1, 1) + (1, .5) = (2, 1.5) *no pertenece a* W, *por lo que* W *no es subespacio*.
3. *Sea* $\mathrm{W} = \{(\mathrm{x}, \mathrm{y}) \in \mathbb{R}^2 \mid \mathrm{y} = \mathrm{ax}$, a *un real fijo*$\}$. *Dados* (x, y) *y* (v, w) *en* W, *se tiene que* y = ax *y* w = av. *Sumando estas dos ecuaciones tenemos* y + w = ax + av = a(x + v), *es decir,* (x, y) + (v, w) *pertenece a* W. *También se tiene que* $\lambda(\mathrm{x}, \mathrm{y}) = (\lambda\mathrm{x}, \lambda\mathrm{y})$ *pertenece a* W, *pues como* y = ax *entonces* $\lambda\mathrm{y} = \mathrm{a}\lambda\mathrm{y}$, *concluyendo que* W *si es subespacio*.

Ejercicio 3.4.1. *Determine cuáles de los siguientes subconjuntos son subespacios.*

1. $W = \{(x, y) \in \mathbb{R}^2 \mid y = 2x\} \cup \{(x, y) \in \mathbb{R}^2 \mid y = 3x\}$
2. $W = \{(x, y) \in \mathbb{R}^2 \mid y = 2x\} \cap \{(x, y) \in \mathbb{R}^2 \mid y = 3x\}$
3. $W = \{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$
4. $W = \{(x_1, x_2, ..., x_n) \in \mathbb{R}^n \mid x_i \in \mathbb{Z}$ *para todo* $i = 1, 2, ..., n\}$
5. $W = \{(x_1, x_2, ..., x_n) \in \mathbb{R}^n \mid x_i \in \mathbb{Q}$ *para todo* $i = 1, 2, ..., n\}$

De los ejemplos anteriores, se tiene que hay subconjuntos que no son subespacios, entonces surge una pregunta, ¿cómo se pueden "generar" subespacios a partir de subconjuntos? Para que un subconjunto sea un subespacio debe cumplir las propiedades estipuladas en la definición 3.4.4, por lo que se requiere que al multiplicar elementos del subconjunto por escalares, el resultado sea un elemento del subconjunto y al sumar elementos del subconjunto, el resultado también pertenezca al subconjunto, entonces, un subespacio generado a partir de un subconjunto debe contener elementos de la forma $a_1\alpha_1 + a_2\alpha_2 + \cdots + a_k\alpha_k$, en donde $k = 1, 2, 3, ...$; α_i pertenece al conjunto generador y los coeficientes de los α son escalares. ¿Es suficiente tomar esos elementos para obtener un subespacio? En otras palabras, ¿la colección de elementos de la forma descrita antes es un subespacio? Si tomamos dos elementos del tipo descrito antes y los sumamos, el resultado es del mismo tipo, gracias a que podemos tomar cualquier cantidad de sumandos. Si multiplicamos un elemento del tipo descrito, por un escalar, el resultado también es de ese tipo, por lo que la respuesta es afirmativa, entonces la discusión anterior la podemos formalizar en la siguiente:

Definición 3.4.5. *Dado un subconjunto no vacío* **S** *de* $\mathbb{R}^n$, *el subespacio generado por* **S**, denotado $\mathcal{L}(\mathbf{S})$, *se define como:*

$$\mathcal{L}(S) := \left\{ \sum_{i=1}^{k} a_i\alpha_i \middle| a_i \in R,\ \alpha_i \in S \ y \ k = 1,2,3,... \right\}$$

Cuando $\mathcal{L}(\mathrm{S}) = \mathbb{R}^{\mathrm{n}}$, *decimos que* S *genera a* $\mathbb{R}^{\mathrm{n}}$.

Observación 3.4.3. *Un subconjunto* S *genera a* $\mathbb{R}^{\mathrm{n}} \Leftrightarrow$ *para todo* $\alpha = (\mathrm{a}_1, \mathrm{a}_2, ..., \mathrm{a}_{\mathrm{n}}) \in \mathbb{R}^{\mathrm{n}}$, *la ecuación:*

$$(a_1, a_2, ..., a_n) = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_k\alpha_k$$

tiene solución para algunos escalares $\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_{\mathrm{k}}$ *y para algunos vectores* $\alpha_1, \alpha_2, ..., \alpha_{\mathrm{k}}$ *en* S.

Después de haber introducido la idea de subespacio generado por un conjunto, y conjunto generador de $\mathbb{R}^n$, surge una pregunta, ¿cuál es la cantidad mínima de vectores que se requieren para generar a $\mathbb{R}^2$, $\mathbb{R}^3$ y en general a $\mathbb{R}^n$? Note que $\mathbb{R}^2$ puede ser generado por los elementos $e_1 = (1, 0)$ y $e_2 = (0, 1)$; el espacio $\mathbb{R}^3$ puede ser generado por $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ y $e_3 = (0, 0, 1)$. ¿Qué elementos generan a $\mathbb{R}^n$? Por analogía a lo que ocurre en $\mathbb{R}^2$ y $\mathbb{R}^3$ se tiene que los vectores $e_1^0 := (0, 0, ..., 0, 1, 0. . . , 0, 0)$, $i = 1, 2, ..., n$ generan a $\mathbb{R}^n$. Una vez encontrados los elementos que generan a $\mathbb{R}^n$, surgen varias preguntas: ¿cuál es la cantidad mínima de elementos que generan a $\mathbb{R}^n$? ¿Será linealmente independiente un conjunto que tiene una cantidad mínima de generadores? Si dos conjuntos S y T generan y son linealmente independientes, ¿tendrán la misma cantidad de elementos? En la discusión que sigue responderemos estas preguntas.

Teorema 3.4.2. *Sean $\alpha_1, \alpha_2, ..., \alpha_l$ generadores de $\mathbb{R}^n$ y $\beta_1, \beta_2, ..., \beta_k \in \mathbb{R}^n$ elementos linealmente independientes, entonces* $k \leq l$.

Demostración. Como $\{\alpha_1, \alpha_2, ..., \alpha_l\}$ genera, el conjunto $\{\beta_1, \alpha_1, \alpha_2, ..., \alpha_l\}$ también genera a $\mathbb{R}^n$ (explique) y es linealmente dependiente, por lo que podemos aplicar el teorema 3.4.1 y concluir que existe un $i \geq 1$ tal que α_i se expresa como combinación lineal de $\{\beta_1, \alpha_1, ..., \alpha_{i-1}\}$. De esto, $\{\beta_1, \alpha_1, \alpha_2, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_l\}$ también genera a $\mathbb{R}^n$ (¿por qué?). Anexando β_2 a este conjunto y procediendo como lo hicimos antes, concluimos que $\{\beta_1, \beta_2, \alpha_1, \alpha_2, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_l\}$ es linealmente dependiente. Aplicando nuevamente el teorema 3.4.1 se tiene que existe $j \geq 1$ tal que α_j se expresa como combinación lineal de $\{\beta_1, \beta_2, \alpha_1, \alpha_2, ..., \alpha_{j-1}\}$ por lo que el conjunto $\{\beta_1, \beta_2, \alpha_1, \alpha_2, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_{j-1}, \alpha_{j+1}, ..., \alpha_l\}$ genera a $\mathbb{R}^n$ (mismo argumento que antes). Continuando el proceso anterior, es decir, anexando un β_i en cada paso y retirando un correspondiente α_i, se tiene que $k \leq l$.

Corolario 3.4.1. *En $\mathbb{R}^n$ cualquiera dos conjuntos de generadores que son linealmente independientes tienen el mismo número de elementos*.

Demostración. Sean $\{\alpha_1, \alpha_2, ..., \alpha_l\}$ y $\{\beta_1, \beta_2, ..., \beta_k\}$ conjuntos de elementos en $\mathbb{R}^n$ que generan y son linealmente independientes. Como el conjunto $\{\alpha_1, \alpha_2, ..., \alpha_l\}$ genera y el conjunto $\{\beta_1, \beta_2, ..., \beta_k\}$ es linealmente independiente, aplicando el teorema anterior se tiene $k \leq l$. La hipótesis permite intercambiar papeles, es decir, ahora suponemos que $\{\beta_1, \beta_2, ..., \beta_k\}$ genera y que $\{\alpha_1, \alpha_2, ..., \alpha_l\}$ es linealmente independiente, entonces aplicando nuevamente el teorema anterior se concluye que $l \leq k$, con lo que obtenemos $k = l$.

Sean $e_1 = (1, 0, 0, ..., 0)$, $e_2 = (0, 1, 0, ..., 0)$, ..., $e_n = (0, 0, 0, ..., 1)$, elementos de $\mathbb{R}^n$. Un argumento sencillo justifica que estos elementos generan y son linealmente independientes. Aplicando el corolario anterior se tiene que cualquier conjunto de generadores que es linealmente independiente, tiene n elementos. A este número le llamaremos la **dimensión** de $\mathbb{R}^n$ y la denotaremos por dim $(\mathbb{R}^n)$.

Definición 3.4.6. *A un conjunto de vectores que genera y es linealmente independiente le llamaremos una base. A la cantidad de elementos que tiene una base le llamaremos la dimensión del espacio.*

Teorema 3.4.3. *Sea $\{\beta_1, \beta_2, ..., \beta_n\}$ un subconjunto de $\mathbb{R}^n$, entonces los siguientes enunciados son equivalentes.*

1. *El conjunto $\{\beta_1, \beta_2, ..., \beta_n\}$ es una base.*
2. *El conjunto $\{\beta_1, \beta_2, ..., \beta_n\}$ es linealmente independiente.*
3. *El conjunto $\{\beta_1, \beta_2, ..., \beta_n\}$ genera a $\mathbb{R}^n$.*

Demostración

1. $\Rightarrow$ 2. Por definición de base, el conjunto $\{\beta_1, \beta_2, ..., \beta_n\}$ es linealmente independiente.
2. $\Rightarrow$ 3. Si $\{\beta_1, \beta_2, ..., \beta_n\}$ no genera, entonces existe un γ que no es combinación lineal de los β por lo que el conjunto $\{\beta_1, \beta_2, ..., \beta_n, \gamma\}$ es linealmente independiente y de esto se concluiría que hay un conjunto con $n + 1$ elementos en $\mathbb{R}^n$ que es linealmente independiente, contradiciendo al teorema 3.4.2.
3. $\Rightarrow$ 1. Solamente falta demostrar que el conjunto $\{\beta_1, \beta_2, ..., \beta_n\}$ es linealmente independiente. Si no fuese linealmente independiente uno de los β es combinación lineal de los restantes y esos $n - 1$ restantes también generan a $\mathbb{R}^n$, contradiciendo que el mínimo número de generadores de $\mathbb{R}^n$ es n.

El teorema anterior es muy útil para construir bases, pues lo que se debe hacer, conociendo la dimensión del espacio, es proponer tantos elementos como la dimensión y verificar que son linealmente independientes.

Ejemplo 3.4.2. *Sabemos que los elementos* (1, 0, 0), (0, 1, 0), (0, 0, 1) *son una base de* $\mathbb{R}^3$; *encuentre otra base que no incluya a los elementos anteriores.*

Discusión. Por el teorema anterior debemos proponer un conjunto con tres elementos y verificar que sea linealmente independiente. Podemos iniciar con cualquier vector no cero. Por ejemplo, proponemos al conjunto $\{(1, 1, 0)\}$, el cual es linealmente independiente. Como (1, 0, 1) no es múltiplo de (1, 1, 0), entonces el conjunto $\{(1, 1, 0), (1, 0, 1)\}$ es linealmente independiente. Verificaremos que el conjunto $\{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$ es linealmente independiente. Para tal efecto propongamos una combinación lineal de sus elementos e igualémosla a cero, es decir, escribamos $x(1, 1, 0) + y(1, 0, l) + z(0, 1, 1) = (x + y, x + z, y + z) = (0, 0, 0)$. Esta ecuación equivale al sistema:

$$\begin{aligned} x + y &= 0 \\ x + z &= 0 \\ y + z &= 0 \end{aligned}$$

que tiene solamente la solución (0, 0, 0), es decir, los vectores son una base de $\mathbb{R}^3$.

Por la importancia que tiene el teorema 2.2.2 lo enunciamos nuevamente, anexando otra condición.

Teorema 3.4.4. *Sea* A *una matriz* n $\times$ n. *Las siguientes condiciones son lógicamente equivalentes.*

1. *La matriz* A *tiene inversa.*
2. *La matriz* A *es equivalente por filas a* I_n.
3. *La matriz* A *es producto de matrices elementales.*
4. *El sistema* AX = 0 *tiene solución única.*
5. *Para todo* $\mathrm{B} = \begin{bmatrix} \mathrm{b}_1 \\ \mathrm{b}_2 \\ \vdots \\ \mathrm{b}_\mathrm{n} \end{bmatrix}$, *el sistema* AX = B *tiene solución única.*
6. *Las columnas de* A, *consideradas como elementos de* $\mathbb{R}^\mathrm{n}$, *son linealmente independientes.*

La forma de aplicar el teorema anterior para verificar si un conjunto de n vectores en $\mathbb{R}^n$ es linealmente independiente, puede ser construyendo una matriz, tomándolos como columnas y decidiendo si la matriz tiene inversa.

Usaremos los vectores del ejemplo 3.4.2 para ilustrar esta idea. Con los vectores (1, 1, 0), (1, 0, 1), (0, 1, 1), construimos la matriz $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. Aplicando operaciones elementales en las filas de A se llega a que su forma escalonada reducida es la identidad, es decir, A tiene inversa y por el teorema anterior los vectores propuestos son linealmente independientes.

Cuando se tiene un conjunto de vectores y se desea saber si es una base, algunas posibilidades pueden ocurrir. Puede ser linealmente independiente, generar o ninguna de éstas. Si el conjunto genera o es linealmente independiente, surge una pregunta, ¿se puede obtener una base a partir de él? La respuesta la da el siguiente:

Teorema 3.4.5.

1. *Sea* $\{\beta_1, \beta_2, \ldots, \beta_k\}$ *un conjunto de vectores linealmente independiente en* $\mathbb{R}^n$*, entonces este conjunto se puede extender a una base.*
2. *Sea* $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ *un conjunto de vectores no cero que generan a* $\mathbb{R}^n$*, entonces de este conjunto se puede extraer una base.*

Demostración. Ejercicio.

3.4.2. Operaciones con subespacios

Dados dos subespacios W y U, se pueden formar los conjuntos $W \cap U$ y $W \cup U$ y preguntarse, ¿cuál de estos conjuntos es subespacio?

Si $\alpha, \beta \in W \cap U$, entonces $\alpha, \beta \in W$ y $\alpha, \beta \in U$, como W y U son subespacios, $\alpha + \beta \in W$ y $\alpha + \beta \in U$, es decir, $\alpha + \beta \in W \cap U$.

Si $\alpha \in W \cap U$ y $x \in \mathbb{R}$, entonces $\alpha \in W$ y $\alpha \in U$, nuevamente, usando el hecho que W y U son subespacios se tiene que $x\alpha \in W$ y $x\alpha \in U$, es decir, $x\alpha \in W \cap U$.

De lo anterior se concluye que $W \cap U$ es un subespacio. De hecho se prueba algo más general, la intersección de cualquier colección de subespacios es un subespacio.

Ejercicio 3.4.2. *Encuentre condiciones necesarias y suficientes de tal forma que* $W \cup U$ *sea un subespacio.*

Supongamos que W y U son subespacios, definimos la suma de W y U como $W + U := \{\alpha + \beta \mid \alpha \in W, \beta \in U\}$. Se tiene que $W + U$ es un subespacio, llamado la suma de W y U.

En efecto, si $\alpha, \beta \in W + U$, se tiene que $\alpha = \alpha_1 + \alpha_2$ y $\beta = \beta_1 + \beta_2$, con $\alpha_1, \beta_1 \in W$ y $\alpha_2, \beta_2 \in U$, por tanto $\alpha + \beta = (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) = (\alpha_1 + \beta_1) + (\alpha_2 + \beta_2) \in W + U$.

Si $c \in \mathbb{R}$, entonces $c\alpha = c(\alpha_1 + \alpha_2) = c\alpha_1 + c\alpha_2 \in W + U$, como se afirmó.

Dado un subespacio $W \neq \{0\}$ de $\mathbb{R}^n$, surge una pregunta: ¿admite una base?

Dado que W no es el subespacio cero, W contiene un elemento α_1 no cero y como W es un subespacio, W contiene al subespacio generado por α_1, es decir, $\{x\alpha_1 \mid x \in \mathbb{R}\} \subseteq W$. Si $W = \{x\alpha_1 \mid x \in \mathbb{R}\}$, el conjunto $\{\alpha_1\}$ es una base de W, pues es linealmente independiente (¿por qué?) y genera a W. Si $W \neq \{x\alpha_1 \mid x \in \mathbb{R}\}$, entonces existe un $\alpha_2 \in W$ que no es múltiplo de α_1 por lo que $\{\alpha_1, \alpha_2\}$ es linealmente independiente. Si $W = \{x\alpha_1 + y\alpha_2 \mid x, y \in \mathbb{R}\}$, entonces el conjunto $\{\alpha_1, \alpha_2\}$ es una base (misma razón que antes). Si W $\neq \{x\alpha_1 + y\alpha_2 \mid x, y \in \mathbb{R}\}$, existe un $\alpha_3 \in W$ y α_3 no es combinación lineal de α_1, α_2, por lo que el conjunto $\{\alpha_1, \alpha_2, \alpha_3\}$ es linealmente independiente. Este proceso termina necesariamente en a lo más n pasos, pues no hay más de n vectores linealmente independientes en $\mathbb{R}^n$, es decir, W admite una base. La discusión anterior la podemos resumir en el siguiente:

Teorema 3.4.6. *Todo subespacio no cero de* $\mathbb{R}^n$ *admite una base que tiene* k *elementos y* $k \leq n$.

Ejercicio 3.4.3. *Decida si los siguientes conjuntos son subespacios de los correspondientes espacios. En caso afirmativo, encuentre una base y la dimensión.*

1. $W = \{(x, y) \in \mathbb{R}^2 \mid 2x + 3y = 0\}$
2. $W = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + 3y - z = 0\}$
3. $W = \{(x, y, z, w) \in \mathbb{R}^4 \mid 2x + 3y - z + w + 1 = 0\}$
4. $W = \{(x, y, z, w) \in \mathbb{R}^4 \mid 2x + 3y - z + w = 0\}$
5. $W = \{(x_1, x_2, \ldots x_{n-1}, x_n) \in \mathbb{R}^n \mid x_1 + 2x_2 + 3x_3 + \cdots + nx_n = 0\}$
6. $W = \{(x_1, x_2, \ldots x_{n-1}, x_n) \in \mathbb{R}^n \mid x_1x_2x_3 \cdots x_n = 0\}$

Como fue visto antes, a partir de los subespacios W y U se pueden formar los subespacios $W + U$ y $W \cap U$. También se tiene $W, U \subseteq W + U$. ¿Cómo se relacionan las dimensiones de los subespacios, W, U, $W + U$ y $W \cap U$? Para adquirir alguna idea de la respuesta, supongamos que W y U son subespacios de $\mathbb{R}^3$ de dimensión dos y consideremos las siguientes posibilidades.

Si $U = W$, entonces $U + W = U$ y $U \cap W = U$. De estas condiciones se tiene $\dim(W + U) = \dim(W) + \dim(U) - \dim(W \cap U)$.

Supongamos ahora que $U \neq W$, entonces ninguno está contenido en el otro, pues si por ejemplo $U \subseteq W$, entonces necesariamente son iguales, pues una base de U también es base de W. De este argumento podemos encontrar $\alpha \in U \setminus W$. Si β y γ forman una base de W, entonces los vectores α, β y γ son linealmente independientes (explique esto). También se tiene $\{\alpha, \beta, \gamma\} \subseteq U + W \subseteq \mathbb{R}^3$, de lo que se concluye que $U + W$ tiene dimensión tres. Como $U \neq W$, el subespacio $U \cap W$ tiene dimensión 0 o 1. La dimensión de este último subespacio no puede ser cero, pues de ser así, el subespacio $U + W$ tendría dimensión cuatro, lo cual es imposible porque $U + W$ es subespacio de $\mathbb{R}^3$. De los argumentos anteriores se tiene: $\dim(W + U) = 3$, $\dim(W \cap U) = 1$. De esto último concluimos que la ecuación $\dim(W + U) = \dim(W) + \dim(U) - \dim(W \cap U)$ tiene lugar.

Este caso particular sugiere que la ecuación anterior puede cumplirse en general, en efecto, ese es el caso y lo enunciamos en el siguiente:

Teorema 3.4.7. *Sean* W *y* U *subespacios de* $\mathbb{R}^n$, *entonces se cumple lo siguiente.*

$$dim(W + U) = dim(W) + dim(U) - dim(W \cap U) \tag{3.18}$$

Demostración. La demostración será efectuada tomando los dos únicos casos posibles.

Primer caso: $W \cap U = \{0\}$. Si alguno de los subespacios es cero, el resultado es claro. Podemos suponer que ninguno de los subespacios es cero. Sean $\{\alpha_1, \ldots, \alpha_l\}$ y $\{\beta_1, \ldots, \beta_r\}$ bases de W y U respectivamente. Afirmamos que $\{\alpha_1, \ldots, \alpha_l, \beta_1, \ldots, \beta_r\}$ es base de $W + U$. En efecto, pues si $a_1\alpha_1 + \cdots + a_l\alpha_l + b_1\beta_1 + \cdots + b_r\beta_r = 0$, entonces $a_1\alpha_1 + \cdots + a_l\alpha_l = -(b_1\beta_1 + \cdots + b_r\beta_r)$ pertenece a la intersección, por lo que se tiene $a_1\alpha_1 + \cdots a_l\alpha_l = -(b_1\beta_1 + \cdots + b_r\beta_r) = 0$. Ahora, usando el hecho que $\{\alpha_1, \ldots, \alpha_l\}$ y

$\{\beta_1, \ldots, \beta_r\}$ son base se tiene que los coeficientes son cero, probando que el conjunto es linealmente independiente. Cualquier $\alpha + \beta \in W + U$, con $\alpha \in W$ y $\beta \in U$ se escribe en la forma $a_1\alpha_1 + \cdots + a_l\alpha_l + b_1\beta_1 + \cdots + b_r\beta_r = \alpha + \beta$, probando que el conjunto $\{\alpha_1, \ldots, \alpha_l, \beta_1, \ldots, \beta_r\}$ genera.

Segundo caso: $W \cap U \neq \{0\}$, sea $\{\gamma_1, \ldots, \gamma_s\}$ una base de $W \cap U$, en particular éste es un subconjunto linealmente independiente de W y U. Por el teorema 3.4.5(1), $\{\gamma_1, \ldots, \gamma_s\}$ se puede extender a bases de W y U respectivamente, las cuales serán denotadas por $\{\gamma_1, \ldots, \gamma_s, \alpha_1, \ldots, \alpha_t\}$ y $\{\gamma_1, \ldots, \gamma_s, \beta_1, \ldots, \beta_m\}$ respectivamente. Afirmamos que $\{\gamma_1, \ldots, \gamma_s, \alpha_1, \ldots, \alpha_t, \beta_1, \ldots, \beta_m\}$ es base de $W + U$.

Si:

$$a_1\gamma_1 + \cdots + a_s\gamma_s + b_1\alpha_1 + \cdots + b_t\alpha_t + c_1\beta_1 + \cdots + c_m\beta_m = 0 \tag{3.19}$$

entonces $a_1\gamma_1 + \cdots + a_s\gamma_s + b_1\alpha_1 + \cdots + b_t\alpha_t = -(c_1\beta_1 + \cdots + c_m\beta_m)$ es un elemento en la intersección (¿por qué?). De esto concluimos que $c_1\beta_1 + \cdots + c_m\beta_m = d_1\gamma_1 + \cdots + d_s\gamma_s$, para algunos escalares $d_1, d_2, \ldots, d_s$. Pasando todos los sumandos al primer miembro se tiene $c_1\beta_1 + \cdots + c_m\beta_m - d_1\gamma_1 - \cdots - d_s\gamma_s = 0$. Como $\{\gamma_1, \ldots, \gamma_s, \beta_1, \ldots, \beta_m\}$ es base, se concluye que $c_1 = c_2 = \cdots = c_m = d_1 = d_2 = \cdots = d_s = 0$. Usando esta condición en (3.19) se tiene $a_1\gamma_1 + \cdots + a_s\gamma_s + b_1\alpha_1 + \cdots + b_t\alpha_t = 0$. Utilizando que $\{\gamma_1, \ldots, \gamma_s, \alpha_1, \ldots, \alpha_t\}$ es base de W se concluye que $a_1 = a_2 = \cdots = a_s = b_1 = b_2 = \cdots = b_t = 0$, probando que el conjunto $\{\gamma_1, \ldots, \gamma_s, \alpha_1, \ldots, \alpha_t, \beta_1, \ldots, \beta_m\}$ es linealmente independiente.

Dejamos como ejercicio demostrar que este conjunto también genera.

3.5. Espacios vectoriales generales

La formulación y solución de muchos problemas que ocurren en matemáticas y ciencias hacen uso de la idea de espacio vectorial abstracto, siendo la razón fundamentalmente el hecho que cuando se estudia un problema, la información se puede cuantificar y una vez hecho esto, es importante poder operar con ella. En cálculo se estudian funciones y sus propiedades, algunas veces una sola función. A veces es conveniente estudiar colecciones de funciones. Por ejemplo, el conjunto de todas las funciones continuas en el intervalo $[0, 2\pi]$ o todas las funciones periódicas en ese intervalo. Algunos elementos importantes de dicho conjunto son las funciones seno y coseno.

Muchas funciones importantes se construyen a partir de funciones sencillas, usando operaciones de suma y producto. Por ejemplo, las funciones polinominales se obtienen al sumar funciones de la forma ax^k y éstas se obtienen multiplicando funciones del tipo cx, las cuales son relativamente sencillas de entender. Estos casos ilustran el comentario respecto a operar con datos.

En esta sección desarrollaremos la idea abstracta de espacio vectorial a partir de lo que hemos hecho para el caso de $\mathbb{R}^n$. Antes de formular la definición de espacio vectorial en general, notemos que la suma y el producto por escalar en $\mathbb{R}^n$ satisfacen las siguientes propiedades.

1. Propiedades de la suma.
   a) Para todos α y $\beta \in \mathbb{R}^n$ se cumple, $\alpha + \beta = \beta + \alpha$ (conmutatividad).
   b) Para todos α, β y γ en $\mathbb{R}^n$ se cumple, $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ (asociatividad).
   c) Existe un elemento en $\mathbb{R}^n$ llamado cero y denotado 0 tal que $0 + \alpha = \alpha$, para todo $\alpha \in \mathbb{R}^n$ (existencia de neutro aditivo).

d) Para todo $\alpha \in \mathbb{R}^n$ existe un elemento α' tal que $\alpha + \alpha' = 0$ (existencia de elementos inversos).

2. Propiedades del producto por escalar.

a) Para todo $\alpha \in \mathbb{R}^n$ se tiene $1\alpha = \alpha$, con $1 \in \mathbb{R}$.

b) Para todo $\alpha \in \mathbb{R}^n$ y para todos λ y $\mu \in \mathbb{R}$ se tiene $\lambda(\mu\alpha) = (\lambda\mu)\alpha$.

c) El producto por escalar es distributivo, es decir, $(\lambda + \mu)\alpha = \lambda\alpha + \mu\alpha$; $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$, para todos $\lambda, \mu \in \mathbb{R}$ y para todos $\alpha, \beta \in \mathbb{R}^n$.

Las propiedades anteriores son fundamentales y son las que deben cumplir la suma y producto por escalar en un espacio vectorial general. En la siguiente definición precisamos esto.

Definición 3.5.1. *Un espacio vectorial sobre los reales es un conjunto no vacío* V *en donde está definida una operación, llamada suma y un producto por escalar que cumplen:*

1. *Propiedades de la suma.*

a) *Para todos* α *y* $\beta \in V$ *se cumple* $\alpha + \beta = \beta + \alpha$ (*conmutatividad*).

b) *Para todos* α, β *y* γ *en* V *se cumple* $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ (*asociatividad*).

c) *Existe un elemento en* V *llamado cero y denotado* 0 *tal que* $0 + \alpha = \alpha$*, para todo* $\alpha \in V$ (*existencia de neutro aditivo*).

d) *Para todo* $\alpha \in V$ *existe un elemento* α' *tal que* $\alpha + \alpha' = 0$ (*existencia de elementos inversos*).

2. *Propiedades del producto por escalar.*

a) *Para todo* $\alpha \in V$ *se tiene* $1\alpha = \alpha$*, con* $1 \in \mathbb{R}$.

b) *Para todo* $\alpha \in V$ *y para todos* $\lambda, \mu \in \mathbb{R}$ *se tiene* $\lambda(\mu\alpha) = (\lambda\mu)\alpha$.

c) *El producto por escalar es distributivo, es decir,* $(\lambda + \mu)\alpha = \lambda\alpha + \mu\alpha$; $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$.

Si V *es un espacio vectorial, a sus elementos les llamaremos* **vectores**,

Antes de presentar ejemplos que usualmente aparecen en los textos queremos presentar uno para ilustrar lo general y abstracto del concepto definido.

Ejemplo 3.5.1. *Considere su objeto favorito, por ejemplo su teléfono celular, y formemos el conjunto* V = {*teléfono celular*}. *Definiremos en* V *una suma. Para facilitar la notación, al único elemento de* V *lo denotaremos por* tc.

Como V *tiene solamente un elemento:* tc, *la única forma de sumar elementos de* V *es tomar a* tc *y sumarlo consigo mismo y el resultado debe ser un elemento de* V, *entonces hay una única forma de hacer esto, es decir, definimos* tc + tc := tc (*note que estamos usando el símbolo usual de suma*).

Si deseamos definir un producto por escalar, al tomar un escalar $x \in \mathbb{R}$ *y multiplicarlo por* **tc** *el resultado debe ser elemento de* V, *entonces la forma de hacerlo debe ser.*

$$\boldsymbol{x(tc) := tc}.$$

¿Es V *un espacio vectorial con la suma y producto por escalar definidos antes?*

1. *Propiedades de la suma.*

a) *Como* tc *es el único elemento de* V, *entonces la conmutatividad de la suma es clara.*

b) *También se cumple la asociatividad, pues para el único elemento en* V *se cumple* (tc + tc) + tc = tc + (tc + tc) = tc (*asociatividad*).

c) *Existe un elemento en* V *llamado cero y denotado* 0 *tal que* 0 + tc = tc. *Tomamos* 0 = ***tc***, *claramente se tiene satisfecha la condición* (*existencia de neutro aditivo*).

d) *Para todo* $\alpha \in V$ *existe un elemento* α' *tal que* $\alpha + \alpha' = 0$. *Como el único elemento en* V *es* tc, *entonces su inverso es el mismo* (*existencia de elementos inversos*).

2. *Propiedades del producto por escalares* (*explique por qué se cumplen*).

a) *Para todo* $\alpha \in V$ *se tiene* $1\alpha = \alpha$, $1 \in \mathbb{R}$.

b) *Para todo* $\alpha \in V$ *y para todos* $\lambda, \mu \in \mathbb{R}$ *se tiene* $\lambda(\mu\alpha) = (\lambda\mu)\alpha$.

c) *El producto por escalar es distributivo, es decir,* $(\lambda + \mu)\alpha = \lambda\alpha + \mu\alpha$; $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$.

Con la discusión del ejemplo anterior, es natural preguntarse si se puede construir un ejemplo de espacio vectorial con dos, tres o más elementos. Explique si se puede o no construir un espacio vectorial con dos elementos.

Ejercicio 3.5.1. *En cada uno de los siguientes casos decida si el conjunto propuesto es espacio vectorial.*

1. *Sea* $V = \mathbb{R}^n$ *con la suma y multiplicación por escalar definidas en la discusión que se hizo antes.*
2. *Proponemos a* $V = \mathcal{M}_{m\times n}(\mathbb{R}) := \{A \mid A$ *es una matriz* $m \times n$ *con entradas reales*$\}$ *con la suma usual de matrices y multiplicación por escalar definida entrada por entrada.*
3. *Considere que* V *es el conjunto de todas las funciones de* $\mathbb{R}$ *en* $\mathbb{R}$ *con la suma usual de funciones y el producto por escalar usual.*
4. *Sea* V *el conjunto de los polinomios con coeficientes en* $\mathbb{R}$, *la suma usual de polinomios y la multiplicación usual de un polinomio por un escalar.*
5. *Dado un entero positivo* n, V *es el conjunto de los polinomios de grado* $\leq n$, *con suma y producto por escalar como en el caso anterior.*
6. *Tomemos a* V *como el conjunto de funciones de* $\mathbb{R}$ *en* $\mathbb{R}$ *que son cero salvo un conjunto finito.*
7. *Consideremos a* $V = \mathbb{Z}^n := \{(x_1, x_2, \ldots, x_n) \mid x_i$ *es entero para todo* $i = 1, 2, \ldots, n\}$ *con la suma entrada por entrada, y multiplicación por escalar entrada por entrada.*

Como lo hicimos en $\mathbb{R}^n$, procedemos a formular algunos de los conceptos fundamentales para espacios vectoriales generales.

Definición 3.5.2. *Sea* V *un espacio vectorial,* $\{\alpha_1, \alpha_2, \ldots, \alpha_n\} \subseteq V$ *se dice linealmente independiente, si los únicos escalares,* $x_1, x_2, \ldots, x_n$ *que satisfacen la ecuación* $x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = 0$ son $x_1 = x_2 = \cdots = x_n = 0$. *En caso contrario el conjunto* $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *se dice linealmente dependiente.*

Definición 3.5.3. *Un espacio vectorial* V *se dice finitamente generado, si existe un subconjunto finito* S *de* V *que lo genera, es decir, si existe* $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\} \subseteq V$ *y cualesquiera* $\alpha \in V$ *es de la forma* $\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$, *para algunos escalares* $x_1, x_2, \ldots, x_n$.

En lo que sigue de la discusión, centraremos la atención en espacios vectoriales finitamente generados, sin embargo, para ilustrar algunas ideas generales sobre espacios vectoriales, de vez en cuando mostraremos ejemplos de espacios vectoriales que no son finitamente generados.

Cuando hicimos la discusión de $\mathbb{R}^n$ mostramos que tiene una base y que todas las bases tienen el mismo número de elementos: n (corolario 3.4.1, página 108). Es importante notar que la demostración del teorema 3.4.2, página 108 y el corolario citado se aplican a cualquier espacio vectorial, por lo que su validez es de tipo general, sin embargo, consideramos importante dar su formulación explícita.

Iniciamos con la siguiente observación.

Observación 3.5.1. *Supongamos que* $\beta = a_1\alpha_1 + \cdots + a_n\alpha_n$. *Si algún* α_i *es combinación lineal de los restantes* $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n$, *entonces,* β *es combinación lineal de* $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n$.

Demostración. Como $\alpha_i = \sum_{j\neq i} x_j\alpha_j$ y $\beta = \sum_{k=1}^{n}\alpha_k\alpha_k$, sustituyendo la primera ecuación en la segunda y agrupando términos se tiene lo afirmado.

Dado que el concepto de base en un espacio vectorial es de suma importancia, formulamos la definición en el caso general.

Definición 3.5.4. *Sea* V *un espacio vectorial. Un subconjunto* $\mathcal{B}$ *de* V *que genera y es linealmente independiente se llama base de* V.

Teorema 3.5.1. *Sea* V *un espacio vectorial finitamente generado distinto de cero,* $S = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ *un conjunto de generadores que no incluye al vector cero. Entonces se puede extraer una base de* S.

Demostración. Si S es linealmente independiente, la demostración termina; en caso contrario, existe un $\alpha_1 \in S$ que es combinación lineal de los restantes elementos de S. Aplicando la observación anterior se tiene que el conjunto $S_1 = S \setminus \{\alpha_i\}$ sigue generando a V. Procediendo con este razonamiento tantas veces como sea necesario se llega a un conjunto $\mathcal{B} \subseteq S$ que genera y es linealmente independiente.

Teorema 3.5.2. *Sea* V *un espacio vectorial finitamente generado,* $S = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ *un conjunto de generadores y* $\{\beta_1, \beta_2, \ldots, \beta_r\}$ *un subconjunto de* V *linealmente independiente, entonces* $r \leq m$.

Demostración. Ver la demostración del teorema 3.4.2, página 108.

Corolario 3.5.1. *En un espacio vectorial finitamente generado, cualquiera dos bases tiene la misma cantidad de elementos.*

Demostración. Ver la demostración del corolario 3.4.1, página 108.

Definición 3.5.5. *Sea* V *un espacio vectorial finitamente generado y* $\mathcal{B}$ *una base. A la cardinalidad de* $\mathcal{B}$ *le llamamos la dimensión de* V *y se denota* $dim(V)$.

Corolario 3.5.2. *Si un espacio vectorial está generado por* m *elementos y* V *tiene dimensión* n, *entonces* $n \leq m$.

Teorema 3.5.3. *Si* V *es un espacio vectorial finitamente generado,* $S = \{\beta_1, \beta_2, \ldots, \beta_r\}$ *un subconjunto de* V *linealmente independiente, entonces existe una base de* V *que contiene a* S.

Demostración. Sea $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ un conjunto de generadores que no incluye al vector cero, entonces el conjunto:

$$S_1 = \{\beta_1, \beta_2, \ldots, \beta_r, \alpha_1, \alpha_2, \ldots, \alpha_m\}$$

también genera a V. Procediendo como en la demostración del teorema 3.4.4, página 85, se puede extraer una base del conjunto S_1 que incluya a los elementos de S.

Definición 3.5.6. *Sea* V *un espacio vectorial,* U *y* W *subespacios. Decimos que* V *es la suma directa de* U *y* W *si* $V = U + W$ *y* $U \cap W = \{0\}$. *En este caso usamos la notación* $V = U \oplus W$.

Por ejemplo, $\mathbb{R}^2$ es la suma directa de cualesquiera dos subespacios diferentes de dimensión uno; $\mathbb{R}^3$ es la suma directa de un subespacio de dimensión uno y uno de dimensión dos.

3.6. Ejercicios

1. Revise todos los teoremas que se enunciaron para el espacio vectorial $\mathbb{R}^n$ y enúncielos para un espacio vectorial V finitamente generado.
2. Proporcione ejemplos de subconjuntos de $\mathbb{R}^2$ que sean cerrados bajo la suma pero no sean subespacios.
3. Sea V un espacio vectorial, S un subconjunto no vacío de V. El subespacio generado por S, denotado $\mathcal{L}(S)$ es: $\mathcal{L}(S) = \left\{ \sum_{j=i}^{n} x_j \alpha_j : x_j \in \mathbb{R}, a_j \in S, n \in \mathbb{N} \right\}$.

   Demuestre que $\mathcal{L}(S) = \cap\, W$, en donde la intersección se toma sobre todos los subespacios que contienen a S. ¿Cuál es el subespacio generado por el conjunto vacío? Con esto en mente, ¿cuál es la dimensión del subespacio cero? ¿Tiene una base cualquier subespacio?
4. Formule la definición de suma directa para $k > 2$ subespacios.
5. Termine la demostración del teorema 3.4.7, página 113.
6. Demuestre que cualesquiera tres elementos de $\mathbb{R}^2$ son linealmente dependientes. ¿Cuál es la condición sobre k para que cualquier conjunto de k elementos sea linealmente dependiente en $\mathbb{R}^n$?
7. Use *Maple* o cualquier otro sistema computacional para decidir si los siguientes conjuntos son linealmente independientes. Primero identifique el espacio que los contiene.

   a) $S = \{(1, 2, 1), (1, 2, -1), (1, 2, 1)\}$

   b) $S = \{(1, 2, 3, 3), (13, 0, 0, 1), (1, -1, 0, l), (1, 2, 3, 30)\}$

   c) $S = \{(1, 2, 3, 3, 1), (1, 2, 0, 0, 1), (1, -1, -1, 0, 1), (1, 2, 3, 3, -1), (1, 1, 1, 1, 10)\}$

   d) $S = \{(1, 2, 3, 3, 0, 0), (1, 2, -1, 0, 2, 1), (1, 3, 3, -1, 0, 1), (1, 2, 3, 3, 1, 0), (1, 2, 3, 3, -1, 1), (1, 2, -1, 0, -2, -1)\}$

   e) $S = \{(1, 2, 3, 3, 0, 0, 1), (1, 2, -1, 3, 0, 0, 1), (1, 2, 3, 3, -1, 0, 1)\}$
8. Demuestre que $\mathbb{R}^n$ es la suma directa de un subespacio de dimensión $n - 1$ y uno de dimensión 1.
9. Sea V un espacio vectorial real de dimensión positiva, S un subconjunto con más de un elemento y finito. Demuestre que S no es un subespacio.
10. Considere la siguiente ecuación, $x(1, -2, 3) + y(2, -4, 0) + z(0, 1, -1) = (1, 0, 0)$. ¿Existen reales x, y, z que la satisfagan?

    Los vectores $(1, -2, 3), (2, -4, 0), (0, 1, -1)$ ¿son linealmente independientes?
11. Encuentre varios conjuntos de tres elementos que generen a $\mathbb{R}^3$. Los elementos de estos conjuntos ¿son linealmente independientes?

12. Sea V el conjunto de los polinomios con coeficientes reales de grado menor o igual que n, con n un entero positivo fijo. Con las operaciones usuales de suma de polinomios y producto por una constante, V es un espacio vectorial. Determine un conjunto de generadores para V.

13. Sean $W_1 = \{(x, y, z) \in \mathbb{R}^3 \mid x + y - z = 0\}$ y $W_2 = \{(x, y, z) \in \mathbb{R}^3 \mid 2x - y + z = 0\}$. Demuestre que W_1 y W_2 son subespacios de $\mathbb{R}^3$ y determine $W_1 \cap W_2$.

14. Sean W_1, W_2 y W_3 subespacios del espacio vectorial V. Muestre con un ejemplo que no necesariamente se cumple $W_1 \cap (W_2 + W_3) = (W_1 \cap W_2) + (W_1 \cap W_3)$.

15. Sea, V un espacio vectorial de dimensión $n > 2$; W_1 y W_2 subespacios de V con $\dim(W_1) = n - 2$. Determine el valor mínimo que debe tener $\dim(W_2)$ para que tenga lugar la condición $W_1 \cap W_2 \neq \{0\}$.

16. Sea V el espacio vectorial de todas las funciones de $\mathbb{R}$ en $\mathbb{R}$, $P = \{f \in V : f$ es par, es decir, $f(x) = f(-x)\}$, $I = \{f \in V : f$ es impar, es decir, $f(-x) = -f(x)\}$. Demuestre que I y P son subespacios y $V = P \oplus I$.

17. ¿Son linealmente independientes las funciones $\operatorname{sen}(x)$ y $\cos(x)$? ¿Son linealmente independientes las funciones e^x y $\operatorname{sen}(x)$? ¿Cuál es la interpretación geométrica, en cuanto a sus gráficas, que refleja el hecho que las funciones f y g sean linealmente dependientes?

18. Sea V un espacio vectorial, $S \subseteq V$. Demuestre que S es linealmente dependiente $\Leftrightarrow$ S contiene un subconjunto propio T tal que S y T generan el mismo subespacio.

19. Sean α_1, α_2, α_3 vectores linealmente independientes en V. Determine cuáles de los siguientes conjuntos son linealmente independientes: *a*) $\{\alpha_1 + \alpha_2, \alpha_2, \alpha_2 - \alpha_3\}$; *b*) $\{\alpha_1 + \alpha_2 + \alpha_3, \alpha_1 - \alpha_3\}$; *c*) $\{2\alpha_1 + \alpha_2, \alpha_2, \alpha_1 - 2\alpha_2\}$.

20. Sean W_1 y W_2 subespacios de V tales que $V = W_1 \oplus W_2$. Demuestre que dado $\alpha \in V$, éste tiene una representación única de la forma $\alpha = \alpha_1 + \alpha_2$, con $\alpha_1 \in W_1$ y $\alpha_2 \in W_2$.

21. Dado $S = \{(1, -3, 0), (-1, 3, -2), (0, 0, -2)\} \subseteq \mathbb{R}^3$, ¿existe un subconjunto de S, linealmente independiente que genere el mismo subespacio que S?

22. Dados $\alpha_1 = (1, 2, 3, 4, 5)$; $\alpha_2 = (1, 0, -1, 2, -2)$; $\alpha_3 = (-1, 0, 1, 0, -1)$; $\alpha_4 = (-2, 0, -1, 2, -1) \in \mathbb{R}^5$.

a) Determine si los α_i son linealmente independientes.

b) Sea W el subespacio generado por $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$. Determine $\dim(W)$.

c) Sea W definido como en el ejercicio anterior. Determine si los siguientes elementos pertenecen a W. *a*) $\alpha = (1, 1, 1, 1, 1)$; *b*) $\alpha = (0, -1, 1, -1, 0)$, *c*) $\alpha = (-1, 1, -1, 1, -1)$.

23. Sean $\alpha_1, \ldots, \alpha_n \in \mathbb{R}^n$, $\alpha_j = (a_{1j}, \ldots, a_{nj})$, $j = 1, \ldots, n$. Sea A la matriz cuyas columnas son los α_j. Demuestre que $\{\alpha_1, \ldots, \alpha_n\}$ es linealmente independiente $\Leftrightarrow$ A es inversible.

24. Sea A una matriz $n \times n$ tal que $A^k = 0$ para algún k. Demuestre que $I_n - A$ es inversible.

25. Se dice que una matriz cuadrada A es nilpotente si $A^r = 0$ para algún entero $r \geq 1$. Sean A y B matrices nilpotentes del mismo orden, y suponga que $AB = BA$. Demuestre que AB y $A + B$ son nilpotentes.

26. Dado el sistema de ecuaciones:

$$\begin{aligned} 2X_1 + 3X_2 + X_3 + 4X_4 - 9X_5 &= 0 \\ X_1 + X_2 + X_3 + X_4 - 3X_5 &= 0 \\ X_1 + X_2 + X_3 + X_4 - 5X_5 &= 0 \\ 2X_1 + 2X_2 + 2X_3 + 3X_4 - 8X_5 &= 0 \end{aligned}$$

determine el espacio solución y calcule su dimensión.

27. Sean W_1 y W_2 subespacios del espacio vectorial V, tales que $\dim W_1 = \dim W_2$. Demuestre que $W_1 = W_2 \Leftrightarrow W_1 \subseteq W_2$ o $W_2 \subseteq W_1$.

28. Los números complejos se pueden considerar como un espacio vectorial sobre los reales. ¿Cuál es su dimensión?

29. Sean, n un entero positivo y $A = (a_{ij})$ una matriz $n \times n$. Supongamos que $a_{i1} + a_{i2} + \cdots + a_{in} = 0$ para todo $i = 1, 2, \ldots, n$. Demuestre que A es singular. Sugerencia: las filas de A satisfacen la ecuación $x_1 + x_2 + \cdots + x_n = 0$.

30. Sea V un espacio vectorial de dimensión $n \geq 2$. Demuestre que V no se puede representar como unión finita de subespacios propios (diferentes de V). Sugerencia: use inducción. Suponga que se tienen $W_1, W_2, \ldots, W_k$ subespacios, ninguno igual a V. Si $k = 1$, es claro que $V \neq W_1$. Suponga $k > 1$ y $V \neq \cup_{i=1}^{k} W_i$, entonces existen $\alpha \in V \setminus \cup_{i=1}^{k} W_i$ y $\beta \in V \setminus W_{k+1}$. Si $V = \cup_{i=1}^{k+1} W_i$, se tiene $t\alpha \in W_{k+1}$ para todo $t \in \mathbb{R}$, por lo que $t\alpha + \beta \notin W_{k+1}$ para todo t. Por el principio de las casillas, existen $t, t_1 \in \mathbb{R}$, diferentes, tales que $t\alpha + \beta, t_1\alpha + \beta \in W_1$ para algún $i \in \{1, 2, \ldots, k\}$, de esto se concluye que $\alpha \in W_i$, contradiciendo lo supuesto sobre α.

31. Considere el espacio de matrices $V = \mathcal{M}_{n\times n}(\mathbb{R})$ (ver ejemplo 3.5.1, página 89). Demuestre que los subconjuntos $S_{n\times n}(\mathbb{R}) := \{A \in \mathcal{M}_{n\times n}(\mathbb{R}) \mid A^t = A\}$ y $A_{n\times n}(\mathbb{R}) := \{A \in \mathcal{M}_{n\times n}(\mathbb{R}) \mid A^t = -A\}$ son subespacios de V y $V = S_{n\times n}(\mathbb{R}) \oplus A_{n\times n}(\mathbb{R})$. ¿Cuál es la dimensión de cada uno de estos subespacios? Sugerencia: Para calcular la dimensión de $S_{n\times n}(\mathbb{R})$, considere el conjunto de matrices simétricas que tienen solamente un 1 en la diagonal y cero en las entradas restantes, junto con las que tienen dos unos simétricos fuera de la diagonal y cero en las entradas restantes. Demuestre que este conjunto es una base y su cardinalidad es $1 + 2 + 3 + \cdots + n = \dfrac{n(n+1)}{2}$. Puede iniciar considerando casos particulares. Por ejemplo, $n = 2, 3, 4, \ldots$

32. Sea $\mathbb{S} := \{f : \mathbb{Z} \to \mathbb{R} \mid f \text{ es función}\}$. A los elementos de S se les denota por $\{x_k\}$, con $k \in \mathbb{Z}$, y $x_k \in \mathbb{R}$. La interpretación es que x_k es el valor de la función $f \in S$ evaluada en k, es decir, $f(k) = x_k$ y la notación $\{x_k\}$ significa que se conoce la "lista" de los valores de f en cada entero. Se puede verificar directamente que S es un espacio vectorial con la suma usual de funciones y el producto de un escalar por una función. Este espacio vectorial es llamado el espacio de *señales en tiempo discreto* y sus elementos aparecen en ingeniería eléctrica o de control al "muestrear" señales. Si V es un espacio vectorial y S es un subconjunto de V, se dice que S es linealmente independiente si todo subconjunto finito de S es linealmente independiente. Sea $S \subseteq \mathbb{S}$ que consiste exactamente de los elementos f_l que satisfacen $f_l(k) = \begin{cases} 1 \text{ si } l = k \\ 0 \text{ en otro caso} \end{cases}$. Demuestre que S es linealmente independiente.

Capítulo 4

Transformaciones lineales y matrices

Uno de los conceptos fundamentales en todas las áreas de matemáticas es el de *función*, pues a través de éste se pueden formular, de manera precisa, una gran variedad de problemas centrales en matemáticas y sus aplicaciones. Por ejemplo, en cálculo, las ideas fundamentales se formulan mediante los conceptos de límite y continuidad de funciones. Veremos, a lo largo de la discusión, en este capítulo, que los dos problemas fundamentales de álgebra lineal (solución de sistemas de ecuaciones lineales y el problema de valores característicos), cuando se formulan usando el concepto de *transformación lineal*, se entienden de manera estructural y a mucha más profundidad.

4.1. Definiciones y resultados básicos

Para la discusión que sigue, recordaremos el ejemplo 2.1.3, página 33.

Ejemplo 4.1.1. *Considere un consorcio con* n *empresas que producen* m *diferentes bienes. Supongamos que todas las empresas producen los mismos bienes. Con estas consideraciones, la producción de cada bien se describe abajo.*

Por cada millón de pesos que se invierte en la empresa 1, la producción es:

$$\begin{gathered} a_{11} \text{ pesos del producto } 1 \\ a_{21} \text{ pesos del producto } 2 \\ \vdots \\ a_{m1} \text{ pesos del producto } m \end{gathered}$$

En general, por cada millón de pesos que se invierte en la empresa j-ésima, ésta produce:

$$\begin{gathered} a_{1j} \text{ pesos del producto } 1 \\ a_{2j} \text{ pesos del producto } 2 \\ \vdots \\ a_{mj} \text{ pesos del producto } m \end{gathered}$$

Si denotamos por x_j la cantidad de dinero que se invierte en la empresa j, entonces la inversión en todas las empresas la podemos representar por un elemento $(x_1, x_2, x_3, \ldots, x_n)$ de $\mathbb{R}^n$. El valor total del bien número i es:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = c_i$$

Con notación de funciones, el proceso anterior se puede representar mediante la expresión:

$$T(x_1, x_2, \ldots, x_n) = (c_1, c_2, \ldots, c_m) \tag{4.1}$$

Al vector $C = (c_1, c_2, \ldots, c_m)$ le llamaremos vector de producción correspondiente al vector de inversión $X = (x_1, x_2, x_3, \ldots, x_n)$.

Supongamos que a otro vector de inversión $Y = (y_1, y_2, \ldots, y_n)$ le corresponde el vector de producción $D = (d_1, d_2, \ldots, d_m)$, entonces se tiene:

$$T(X + Y) = T(X) + T(Y) = C + D \tag{4.2}$$

Si el vector de inversión X se multiplica por un escalar λ, entonces el correspondiente vector de producción también se multiplica por λ, es decir

$$T(\lambda X) = \lambda T(X) \tag{4.3}$$

Hay muchos problemas de aplicaciones que al formularse mediante funciones, éstas satisfacen (4.2) y (4.3). A estas propiedades les llamaremos *propiedades de linealidad* de T. La idea precisa de transformación lineal la establecemos en la siguiente definición.

Definición 4.1.1. *Sean* V *y* W *dos espacios vectoriales. Una transformación lineal es una función* $T : V \to W$ *que satisface las siguientes propiedades.*

1. $T(\alpha + \beta) = T(\alpha) + T(\beta)$, para todos $\alpha, \beta \in V$.
2. $T(r\alpha) = rT(\alpha)$, para todo escalar $r \in \mathbb{R}$ y para todo $\alpha \in V$.

Nota: Para decidir si una función es lineal, es muy importante verificar que se cumplan las dos propiedades de la definición. Puede haber casos en los que se cumpla solamente una o ninguna, en cuyo caso no se trata de una transformación lineal.

Si $V = W = \mathbb{R}$, la función $T(x) = x$ es una transformación lineal, pues $T(x + y) = x + y = T(x) + T(y)$ y $T(rx) = rx = rT(x)$.

Un caso más general que el anterior ocurre al tomar la función $f(x) = ax$, para algún real fijo a. El lector está invitado a verificar que esta función es lineal.

En contraste con los ejemplos anteriores tenemos la función $f(x) = x + 1$. Ésta no es lineal, pues $f(x + y) = x + y + 1 \neq f(x) + f(y) = (x + 1) + (y + 1)$.

Los ejemplos anteriores motivan las siguientes preguntas. Las dos primeras quedan abiertas para que el lector reflexione acerca de respuestas y obtenga conclusiones.

1. ¿Existe una función $f : \mathbb{R} \to \mathbb{R}$ que satisfaga $f(x + y) = f(x) + f(y)$ para todos $x, y \in \mathbb{R}$ y no sea lineal?
2. ¿Existe una función $f : \mathbb{R} \to \mathbb{R}$ que satisfaga $f(rx) = rf(x)$ para todos $r, x \in \mathbb{R}$ y no sea lineal?
3. ¿Cómo son las funciones $T : \mathbb{R} \to \mathbb{R}$ que son transformaciones lineales?

En relación con la última pregunta se tiene, si $T : \mathbb{R} \to \mathbb{R}$ es una transformación lineal, entonces $T(x) = T(x1) = xT(1)$. Sea $a := T(1)$, de esto concluimos que $T(x) = ax$, para todo $x \in \mathbb{R}$. Recíprocamente, si $T(x) = ax$ para algún $a \in \mathbb{R}$ fijo, entonces T es lineal. Resumiendo, una función $T : \mathbb{R} \to \mathbb{R}$ es una transformación lineal $\Leftrightarrow$ existe un real fijo a tal que $T(x) = ax$ para todo $x \in \mathbb{R}$. De esta discusión se llega a que las transformaciones lineales de $\mathbb{R} \to \mathbb{R}$ no son otra cosa que funciones lineales cuya gráfica pasa por el origen; esto traducido al lenguaje algebraico significa: si $T : \mathbb{R} \to \mathbb{R}$ es lineal, entonces $T(0) = 0$. Resulta que esta última formulación es de carácter general, es decir, se tiene:

Observación 4.1.1. *Si* $\mathrm{T} : \mathrm{V} \to \mathrm{W}$ *es una transformación lineal, entonces* $\mathrm{T}(0) = 0$. *En efecto,* $\mathrm{T}(0) = \mathrm{T}(0 + 0) = \mathrm{T}(0) + \mathrm{T}(0)$, *cancelando* $\mathrm{T}(0)$ *en ambos miembros se tiene lo afirmado*.

Grafique las siguientes funciones y determine cuáles son transformaciones lineales.

1. $f(x) = 3x$
2. $f(x) = 2x + 1$
3. $f(x) = x + 10^{-20}$
4. $f(x) = \sqrt{2x}$

¿Se pueden clasificar las transformaciones lineales de $\mathbb{R}^2$ en $\mathbb{R}$? Para contestar a esta pregunta, recordemos que un elemento (x, y) de $\mathbb{R}^2$ se puede representar en términos de la base canónica, es decir, $(x, y) = xe_1 + ye_2$. Si $T : \mathbb{R}^2 \to \mathbb{R}$ es lineal, usando la representación de (x, y) y la definición de linealidad de T, concluimos que $T(x, y) = T(xe_1 + ye_2) = xT(e_1) + yT(e_2)$. Definiendo los reales $a_1 = T(e_1)$ y $a_2 = T(e_2)$, números que solamente dependen de la transformación T, se tiene $T(x, y) = a_1x + a_2y$. Esta última ecuación expresa el hecho que la transformación lineal T *queda completamente determinada por su acción en los elementos básicos* y que el valor de la transformación en (x, y) es una expresión lineal en las variables x, y.

Recíprocamente, si existen constantes $a_1, a_2 \in \mathbb{R}$ tales que el valor de la función T en (x, y) se expresa como $T(x, y) = a_1x + a_2y$, entonces T es lineal (verificarlo).

En la discusión anterior clasificamos a las transformaciones lineales T de $\mathbb{R}$ en $\mathbb{R}$ y a las de $\mathbb{R}^2$ en $\mathbb{R}$. Para ello se utilizó que los elementos de $\mathbb{R}^2$ se expresan como combinación lineal de la base canónica.

Si tenemos una transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$, entonces al evaluar a T en (x, y), el resultado es un elemento de $\mathbb{R}^2$, por lo que $T(x, y)$ también se representa en términos de la base canónica. De manera más precisa, existen números reales a_1 y a_2, que dependen de (x, y), tales que $T(x, y) = a_1e_1 + a_2e_2$. Para enfatizar la dependencia de a_1 y a_2 de (x, y), usaremos la notación $a_1 = T_1(x, y)$ y $a_2 = T_2(x, y)$. Con esto, los escalares a_1 y a_2 son los valores de las funciones T_1 y T_2 en (x, y).

Usando esta terminología y notación, podemos decir que la transformación lineal T está determinada por las funciones $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}$. Esto lo escribimos como $T = (T_1, T_2)$ y el significado de esto es $T(x, y) = (T_1(x, y), T_2(x, y))$. Si T está dada por $T = (T_1, T_2)$, a las funciones T_1 y T_2 se les llama las *funciones coordenadas de T*.

De lo discutido antes se tiene que una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$ está determinada por sus funciones coordenadas, las cuales son funciones de $\mathbb{R}^2$ en $\mathbb{R}$.

Un argumento sencillo demuestra el siguiente:

Teorema 4.1.1. *La función* $\mathrm{T} : \mathbb{R}^2 \to \mathbb{R}^2$ *es lineal si y sólo si sus funciones coordenadas son lineales*.

Con el criterio del teorema anterior y la clasificación de las transformaciones lineales de $\mathbb{R}^2$ en $\mathbb{R}$, es relativamente fácil producir ejemplos de transformaciones lineales

$T : \mathbb{R}^2 \to \mathbb{R}^2$.

Ejercicio 4.1.1. *Decida cuáles de las siguientes funciones son transformaciones lineales.*

1. $T(x, y) = (x, y)$
2. $T(x, y) = (x + 1, y - 1)$

3. $T(x, y) = (x + y, y - x)$
4. $T(x, y) = (x + \cos(y), y)$
5. $T(x, y) = (ax + by, cx + dy)$, *a, b, c* y *d constantes*

En esta parte nos interesa explorar cómo es la imagen de un conjunto bajo la acción de una transformación lineal.

Ejemplo 4.1.2. *Sea* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *dada por* $T(x, y) = (2x + y, x + y)$. *¿En qué transforma* T *a la línea* L *cuya ecuación es* $y = x$*? En general, ¿en qué transforma* T *a una línea de la forma* $y = ax$?

Para contestar a la primera pregunta, consideremos puntos sobre la recta, es decir, puntos (x, y) en donde $y = x$. Estos puntos vistos como parejas son de la forma (x, x) y deseamos saber en qué se transforman cuando les aplicamos T. En otras palabras, deseamos saber qué objeto geométrico es el conjunto de puntos de la forma $T(x, x)$, cuando x varía en $\mathbb{R}$. Por la definición de T se tiene: $T(x, x) = (3x, 2x) = x(3, 2)$. Cuando x toma todos los valores reales, el vector $x(3, 2)$ describe todos los puntos de la recta que pasa por el punto $(3, 2)$ y por el origen (¿por qué?). De esto se tiene que la recta $y = x$ es transformada en la recta que pasa por el origen y tiene pendiente $2/3$. La figura 4.1 ilustra la situación discutida.

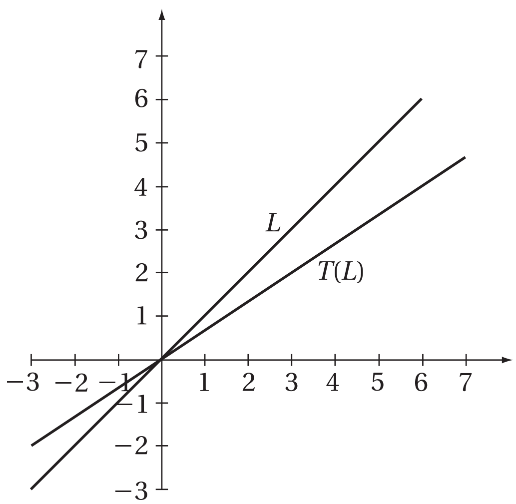


Figura 4.1. Imagen de L bajo $T(x, y) = (2x + y, x + y)$.

La segunda pregunta se aborda de manera análoga, es decir, consideramos puntos de la forma (x, ax) y les aplicamos T, obteniendo, $T(x, ax) = (2x + ax, x + ax) = x(2 + a, 1 + a)$, y de esto se tiene que T transforma a la recta $y = ax$ en la recta que pasa por el origen de coordenadas y por el punto $(2 + a, 1 + a)$, pues el vector $(2 + a, 1 + a)$ nunca es cero, de hecho, observando las coordenadas de este vector se tiene que la recta de pendiente -2 es transformada en el eje y, mientras que la recta de pendiente -1 es transformada en el eje x, (¿por qué?).

Después de esta discusión se puede preguntar: ¿en qué transforma T al cuadrado $\{(x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1\}$? Una forma de abordar esta pregunta es averiguar cuál es la imagen de los lados del cuadrado bajo la acción de T, es decir, deseamos conocer $T(x, 0)$, $T(1, y)$, $T(x, 1)$ y $T(0, y)$ cuando $0 \leq x, y \leq 1$.

Se tiene: $T(x, 0) = (2x, x) = x(2, 1)$, $T(1, y) = (2 + y, 1 + y) = (2, 1) + y(1, 1)$, $T(x, 1) = (2x + 1, x + 1) = (2x, x) + (1, 1) = x(2, 1) + (1, 1)$ y $T(0, y) = (y, y) = y(1, 1)$. Como la x y la

y son números reales entre cero y uno, los resultados anteriores muestran que los lados son transformados en segmentos.

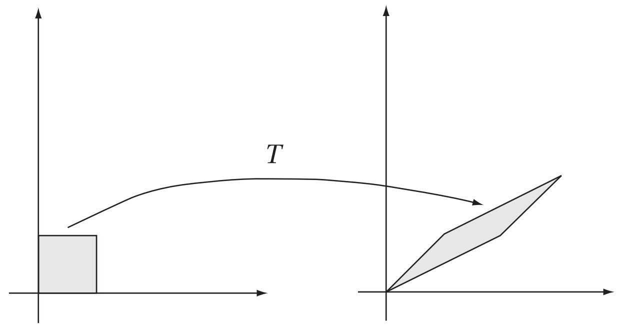


Figura 4.2. Imagen del cuadrado unitario bajo $T(x, y) = (2x + y, x + y)$.

De esto se concluye que el cuadrado es transformado en el paralelogramo que se muestra en la parte derecha de la figura 4.2.

Ejercicio 4.1.2. *Use* Cabri-Geometry *o cualquier otro sistema computacional para contestar las preguntas que se plantean. Después haga un análisis algebraico para corroborar los resultados que obtenga.*

1. *Si* $\mathrm{T} : \mathbb{R}^2 \to \mathbb{R}^2$ *está definida por* $\mathrm{T}(x, y) = (2x - y, x + 2y)$. *¿Cuál es la imagen de la circunferencia de radio* 1 *y centro en el origen bajo* T?
2. *¿Cómo debe estar definida* $\mathrm{T} : \mathbb{R}^2 \to \mathbb{R}^2$ *para que la imagen de una circunferencia bajo* T *sea una circunferencia?*

En la discusión que hicimos para determinar las transformaciones lineales de $\mathbb{R}$ en $\mathbb{R}$ y de $\mathbb{R}^2$ en $\mathbb{R}$ nos dimos cuenta que los valores de la transformación quedan determinados en la base canónica. Esta idea lleva a preguntarse si acaso se tendrá el mismo resultado en general, es decir, ¿será cierto que una transformación lineal queda definida en una base? La respuesta la da el siguiente resultado.

Teorema 4.1.2. *Sea* $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *una base de* V *y* $\beta_1, \beta_2, \ldots, \beta_n$ *cualquier colección de* n *vectores en* W, *entonces existe una única transformación lineal* $\mathrm{T} : \mathrm{V} \to \mathrm{W}$ *tal que* $\mathrm{T}(\alpha_i) = \beta_i$, *para todo* $i = 1, 2, \ldots, n$.

Demostración. Para definir T procedemos como sigue, primero definimos T en cada α_i de tal forma que las conclusiones del teorema se satisfagan, es decir, $T(\alpha_i) := \beta_i$ para cada $i = 1, 2, \ldots, n$. En general, si $\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$ proponemos $T(\alpha) := x_1T(\alpha_1) + x_2T(\alpha_2) + \cdots + x_nT(\alpha_n)$. Hay que notar que la representación de cada $\alpha \in V$ como combinación lineal en términos de una base es única, por lo que la definición de T es consistente. El resto de la demostración lo asignamos como un ejercicio.

4.2. Transformaciones lineales geométricas

En esta sección presentamos ejemplos de transformaciones lineales que surgen de problemas geométricos. Iniciamos discutiendo las:

a) **Reflexiones.** La idea de reflexión, desde el punto de vista geométrico, proviene de considerar la imagen de un objeto reflejada a través de un espejo. En el plano, una reflexión se considera respecto a una recta y, sin perder generalidad se puede suponer que dicha recta pasa por el origen. Con este supuesto, una reflexión en el plano es una transformación que fija a los puntos de la recta W y a un vector $u \neq 0$, perpendicular a W, lo transforma en su negativo. Ver la figura 4.3.

Algebraicamente y con base en las ideas geométricas, esto se puede representar como:

$$T(v) = v - 2\frac{\langle v,u\rangle}{\langle u,u\rangle}u \tag{4.4}$$

pues dado un vector v, la proyección ortogonal de v a lo largo de u está dada por $\frac{\langle v,u\rangle}{\langle u,u\rangle}u$ (definición 3.3.4, página 96), esto y un argumento geométrico muestra la razón de la ecuación anterior.

Ejemplo 4.2.1. *Encuentre una expresión para la reflexión del plano en el plano a través de la recta de ecuación* y = 2x.

Discusión. De acuerdo con la ecuación 4.4, debemos encontrar un vector u perpendicular a la recta de ecuación $y = 2x$. De hecho, para facilitar los cálculos propondremos un vector u de norma uno. Sea $u = \frac{1}{\sqrt{5}}(2,-1)$ y pongamos $v = (x, y)$. Usando la ecuación 4.4 para determinar la reflexión buscada se obtiene:

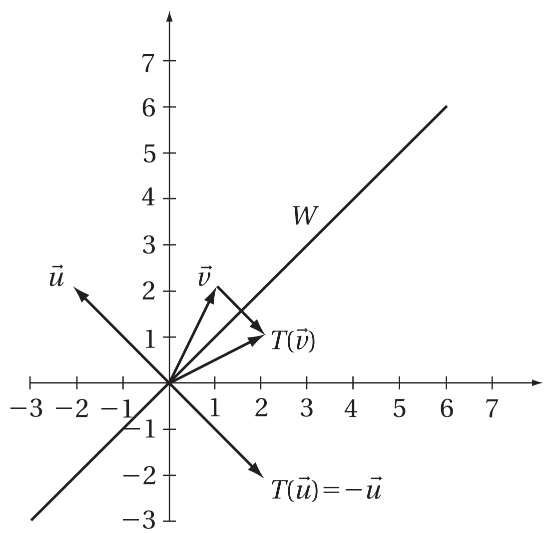


Figura 4.3. Reflexión a través del subespacio *W*.

$$\begin{aligned} T(x, y) &= (x, y) - 2\left\langle (x, y), \frac{1}{\sqrt{5}}(2,-1)\right\rangle \frac{1}{\sqrt{5}}(2,-1) \\ &= (x, y) - \frac{2}{5}(2x - y)(2, -1) \\ &= (x, y) - \left(\frac{4}{5}(2x-y), -\frac{2}{5}(2x-y)\right) \\ &= \frac{1}{5}(-3x + 4y, 4x + 3y) \end{aligned}$$

Una forma de verificar que la función propuesta satisface los requerimientos, es evaluar en puntos de la recta, es decir calcular $T(x, 2x)$. De acuerdo con la descripción para T, se tiene:

$$T(x, 2x) = \frac{1}{5}(-3x + 4(2x), 4x + 3(2x)) = \frac{1}{5}(5x + 10x) = (x + 2x)$$

como se esperaba.

La ecuación 4.4 muestra que una reflexión del plano en el plano queda determinada por un vector no cero u, de manera equivalente, por una recta que pasa por el origen. Esta idea se puede extender a $\mathbb{R}^3$ o en general a $\mathbb{R}^n$, pues para definir una reflexión como la descrita por la ecuación 4.4, tomamos $u = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ no cero. El "plano", W en $\mathbb{R}^n$, es aquel cuyos elementos $v = (x_1, x_2, \ldots, x_n) \in W$ satisfacen:

$$T(v) = v - 2\frac{\langle v,u\rangle}{\langle u,u\rangle}u = v$$

Alternativamente, los elementos de W satisfacen la ecuación homogénea:

$$\langle v, u\rangle = a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0 \tag{4.5}$$

cuyo conjunto solución es precisamente W, el cual resulta ser un subespacio de dimensión $n - 1$; demuéstrelo como ejercicio.

b) Rotaciones. Dado un vector u, se tiene $u = \mathrm{r}(\cos(\theta), \operatorname{sen}(\theta))$, con r la norma de u y θ el ángulo que forma con el eje x. Si u es rotado un ángulo ω y denotamos por $T_\omega(u)$ al resultado de aplicar tal rotación, en notación de funciones se tiene.

$$T_\omega(u) = r(\cos(\theta + \omega), \operatorname{sen}(\theta + \omega)) = u'$$

Haciendo uso de las identidades trigonométricas:

$$\operatorname{sen}(\theta + \omega) = \operatorname{sen}(\theta)\cos(\omega) + \operatorname{sen}(\omega)\cos(\theta)$$
$$\cos(\theta + \omega) = \cos(\theta)\cos(\omega) - \operatorname{sen}(\theta)\cos(\omega)$$

la expresión para $T_\omega(u)$ es:

$$\begin{aligned} T_\omega(u) &= r(\cos(\theta + \omega), \operatorname{sen}(\theta + \omega)) \\ &= r(\cos(\theta)\cos(\omega) - \operatorname{sen}(\theta)\operatorname{sen}(\omega), \operatorname{sen}(\theta)\cos(\omega) + \operatorname{sen}(\omega)\cos(\theta)) \end{aligned}$$

Si representamos al vector $T_\omega(u)$ en forma de columna, la ecuación anterior se expresa como:

$$T_\omega(u) = \begin{bmatrix} \cos(\omega) & -\operatorname{sen}(\omega) \\ \operatorname{sen}(\omega) & \cos(\omega) \end{bmatrix} \begin{bmatrix} r\cos(\theta) \\ r\operatorname{sen}(\theta) \end{bmatrix} \tag{4.6}$$

A la matriz:

$$R_\omega = \begin{bmatrix} \cos(\omega) & -\operatorname{sen}(\omega) \\ \operatorname{sen}(\omega) & \cos(\omega) \end{bmatrix} \tag{4.7}$$

le llamaremos la matriz que representa a la rotación T_ω.

Si u está representado en forma cartesiana, digamos $u = (x, y)$, entonces $T_\omega(x, y)$ se obtiene de:

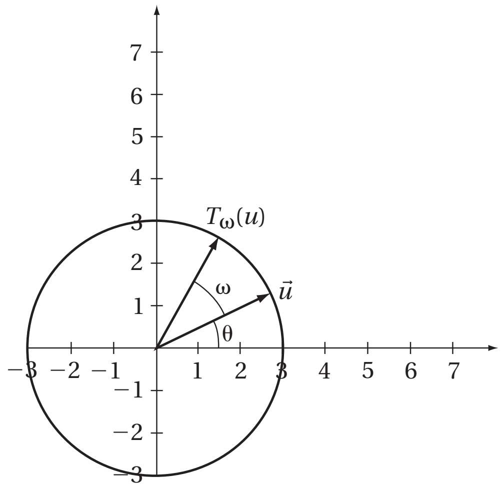


Figura 4.4. Rotación un ángulo ω.

$$T_\omega(x,y)=\begin{bmatrix}\cos(\omega) & -\operatorname{sen}(\omega)\\ \operatorname{sen}(\omega) & \cos(\omega)\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

$$=\begin{bmatrix}x\cos(\omega)-y\operatorname{sen}(\omega)\\ x\operatorname{sen}(\omega)+y\cos(\omega)\end{bmatrix}$$

Este resultado de pareja ordenada toma la forma:

$$T_\omega(x, y) = (x\cos(\omega) - y\operatorname{sen}(\omega), x\operatorname{sen}(\omega) + y\cos(\omega)) \tag{4.8}$$

Ejemplo 4.2.2. *Describa algebraicamente la transformación que rota al plano un ángulo de* $\frac{\pi}{4}$ *radianes.*

Discusión. De acuerdo con la ecuación 4.2, la transformación buscada es:

$$T_{\pi/4}(x, y) = (x\cos(\pi/4) - y\operatorname{sen}(\pi/4), x\operatorname{sen}(\pi/4) + y\cos(\pi/4)) = \frac{1}{\sqrt{2}}(x-y, x+y)$$

Resuelva los siguientes ejercicios.

Ejercicio 4.2.1.

1. *Demuestre que las rotaciones y reflexiones son transformaciones lineales.*
2. *Describa el efecto geométrico de componer una rotación y una reflexión.*
3. *Defina una translación. ¿Es transformación lineal?*
4. *Se define un movimiento rígido del plano en el plano como una función* $f:\mathbb{R}^2\to\mathbb{R}^2$ *que preserva la norma de los vectores, es decir,* $\langle\alpha,\alpha\rangle = \langle f(\alpha), f(\alpha)\rangle$ *para todo* $\alpha\in\mathbb{R}^2$. *Demuestre que las rotaciones, reflexiones y traslaciones son movimientos rígidos.*
5. *Considere un cuadrado* C *de lados paralelos a los ejes y la transformación* $T(x, y) = (2x + y, x + y)$. *Encuentre la imagen de* C *bajo* T.
6. *Encuentre la imagen de una recta bajo la transformación del ejercicio anterior.*

4.3. Rango y núcleo de una transformación lineal

Dada una transformación lineal, hay dos subespacios relacionados con ella que resultan ser de importancia para determinar sus propiedades.

Definición 4.3.1. *Sea* $T : V \to W$ *una transformación lineal. Definimos el núcleo o kernel de* T *como el conjunto* $N_T = \{\alpha \in V \mid T(\alpha) = 0\}$. *La imagen o rango de* T, *denotada* R_T *se define como* $R_T = \{\beta \in W$ *existe un* $\alpha \in V$ *y satisface* $T(\alpha) = \beta\}$.

Ejercicio 4.3.1. *Demuestre que el núcleo y el rango de una transformación lineal son subespacios.*

Ejercicio 4.3.2. *Para cada una de las siguientes transformaciones lineales, encuentre su núcleo y rango.*

1. *Sea* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *definida por* $T(x, y) = (x - y, 3x + 2y)$.
2. *Sea* n *un entero positivo,* V *el espacio de los polinomios de grado* $\leq n$ *y* T *la función definida por* $T(p(x)) = p'(x)$, *con* $p'(x)$ *la derivada de* $p(x)$.
3. *Sea* $T : \mathbb{R}^n \to \mathbb{R}$ *definida por* $T(x_1, x_2, \ldots, x_n) = x_1, + x_2 + x_3 + \cdots + \mathbf{x_n}$.
4. *Sea* V *el espacio de las matrices* 2×2 *y* $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. *Se define* $T_A : V \to V$ *por* $T_A(X) = AX$, *en donde* $X = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$.

Ejercicio 4.3.3. *Revise las definiciones de función inyectiva y suprayectiva. Demuestre lo siguiente acerca de una transformación lineal* T.

1. *La transformación lineal* T *es inyectiva* $\Leftrightarrow N_T = \{0\}$.
2. *La transformación lineal* T *es suprayectiva* $\Leftrightarrow R_T = W$.

Definición 4.3.2. *Sean* V *y* W *espacios vectoriales,* $T : V \to W$ *una transformación lineal.*

1. *Si* T *es inyectiva, se dice que es no singular.*
2. *Si* T *es biyectiva se dice que* T *es un isomorfismo y decimos que* V *es isomorfo a* W, *denotando esto por* $V \cong W$.

Los siguientes son dos de los teoremas más importantes en la teoría de espacios vectoriales de dimensión finita. Las demostraciones estarán a cargo del lector y se darán algunas sugerencias en los ejercicios al final de este capítulo.

Teorema 4.3.1. *Sea* V *un espacio vectorial de dimensión* n, *entonces* $V \cong \mathbb{R}^n$.

Teorema 4.3.2. *Sean* V *y* W *espacios vectoriales con* V *finitamente generado,* $T : V \to W$ *una transformación lineal, entonces la siguiente ecuación se cumple.*

$$dim\,(V) = dim\,(N_T) + dim\,(R_T) \tag{4.9}$$

Ejercicio 4.3.4. *Sean* m *y* n *dos enteros positivos diferentes. Demuestre lo siguiente.*

1. *Si* $m < n$, *entonces no existe una transformación lineal* $T : \mathbb{R}^n \to \mathbb{R}^m$ *inyectiva.*
2. *Si* $n < m$, *entonces no existe una transformación lineal* $T : \mathbb{R}^n \to \mathbb{R}^m$ *suprayectiva.*
3. *Use los ejercicios anteriores para concluir que* $\mathbb{R}^n \cong \mathbb{R}^m$ *si y sólo si* $n = m$. *En el siguiente ejercicio probará este hecho de manera más general.*

Ejercicio 4.3.5. *Sean* V *y* W *dos espacios vectoriales finitamente generados. Demuestre que* $V \cong W \Leftrightarrow dim\,(V) = dim\,(W)$.

Ejercicio 4.3.6. *Si* $T : \mathbb{R}^n \to \mathbb{R}^m$ es *una transformación lineal, de acuerdo con los ejercicios* 13 *y* 14, *página* 146, *encontrar* R_T *se reduce a resolver sistemas de ecuaciones lineales. Explique esto.*

4.4. Matrices y transformaciones lineales

En esta sección estableceremos una relación estrecha entre transformaciones lineales y matrices. También probaremos un resultado importante para matrices: *si **A** es una matriz **m** × **n**, entonces el espacio generado por las columnas y el espacio generado por las filas tienen la misma dimensión*. A este número común le llamaremos el *rango de **A***.

Sabemos que una transformación lineal queda completamente determinada en una base, teorema 4.1.2, página 128. Si $T : V \to W$ es una transformación lineal, $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ y $\{\beta_1, \beta_2, \ldots, \beta_m\}$ son bases de V y W respectivamente, entonces para cada $j = 1, \ldots, n$, $T(\alpha_j)$ se representa como combinación lineal de los elementos de la base $\{\beta_1, \beta_2, \ldots, \beta_m\}$, es decir, existen escalares $a_{1j}, a_{2j}, \ldots, a_{mj}$, únicos, tales que:

$$T(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m \tag{4.10}$$

Los escalares a_{ij} solamente dependen de la transformación lineal y de las bases elegidas, con ellos formamos la matriz:

$$A_{[\alpha,\beta]} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Definición 4.4.1. *La matriz* $A_{[\alpha,\beta]}$ *se llama la matriz asociada a la transformación* T *respecto a las bases* $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *y* $\{\beta_1, \beta_2, \ldots, \beta_m\}$. *Si no hay lugar a confusión en cuanto a las bases, suprimiremos el subíndice de la matriz asociada a una transformación.*

Ejemplo 4.4.1. *Sea* $T : \mathbb{R}^2 \to \mathbb{R}^5$ *dada por* $T(x, y) = (2x + y, x - y, 4x - 3y, x, y)$. *Encuentre la matriz asociada a* T *respecto a las bases canónicas.*

Desarrollo. Para encontrar la matriz asociada a T respecto a las bases canónicas, la evaluamos en $e_1 = (1, 0)$ y en $e_2 = (0, 1)$, encontrándose combinaciones lineales en términos de la base canónica de $\mathbb{R}^5$. Los coeficientes que aparecen en $T(e_1)$ forman la primera columna y los de $T(e_2)$ forman la segunda columna. Se tiene:

$$\begin{aligned} T(1, 0) &= (2, 1, 4, 1, 0) \\ &= 2(1, 0, 0, 0, 0) + (0, 1, 0, 0, 0) + 4(0, 0, 1, 0, 0) + (0, 0, 0, 1, 0) + 0(0, 0, 0, 0, 1) \end{aligned}$$

y

$$\begin{aligned} T(0, 1) &= (1, -1, -3, 0, 1) \\ &= (1, 0, 0, 0, 0) - (0, 1, 0, 0, 0) - 3(0, 0, 1, 0, 0) + 0(0, 0, 0, 1, 0) + (0, 0, 0, 0, 1) \end{aligned}$$

De esto obtenemos que la matriz buscada es $A = \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 4 & -3 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$

Ejemplo 4.4.2. *Sean* $\alpha_1 = (1, 1)$ *y* $\alpha_2 = (3, -2)$. *Justifique que* $\{\alpha_1, \alpha_2\}$ *es una base de* $\mathbb{R}^2$. *De acuerdo con el teorema 4.1.2, existe una única transformación lineal* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *tal que* $T(\alpha_1) = (4, 5)$ y $T(\alpha_2) = (6, -1)$. *Encuentre la matriz asociada a* T *respecto a la base canónica de* $\mathbb{R}^2$.

Desarrollo. Verifique que los vectores $\alpha_1 = (1, 1)$ y $\alpha_2 = (3, -2)$ forman una base de $\mathbb{R}^2$. Para determinar la matriz de T respecto a la base canónica, debemos determinar $T(e_1) = (a, b)$ y $T(e_2) = (c, d)$, usando las ecuaciones $T(1, 1) = T(e_1) + T(e_2)$ y $T(3, -2) = 3T(e_1) - 2T(e_2)$. Estas condiciones se traducen a: $(4, 5) = (a + c, b + d)$ y $(6, -1) = (3a - 2c, 3b - 2d)$, y de esto se tiene el sistema:

$$\begin{aligned} a + c &= 4 \\ b + d &= 5 \\ 3a - 2c &= 6 \\ 3b - 2d &= -1 \end{aligned}$$

Resolviendo este sistema se encuentra $T(e_1) = \left(\frac{14}{5}, \frac{9}{5}\right)$ y $T(e_2) = \left(\frac{6}{5}, \frac{16}{5}\right)$, de lo cual la matriz asociada a T respecto a la base canónica es:

$$A = \begin{bmatrix} \frac{14}{5} & \frac{6}{5} \\ \frac{9}{5} & \frac{16}{5} \end{bmatrix}$$

Con la información anterior se tiene que la expresión que define a $T(x, y)$ se calcula mediante el producto de matrices:

$$AX = \begin{bmatrix} \frac{14}{5} & \frac{6}{5} \\ \frac{9}{5} & \frac{16}{5} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{14}{5}x + \frac{6}{5}y \\ \frac{9}{5}x + \frac{16}{5}y \end{bmatrix}$$

De esto se concluye que $T(x, y) = \left(\frac{14}{5}x + \frac{6}{5}y, \frac{9}{5}x + \frac{16}{5}y\right)$

Ejercicio 4.4.1. *Sea* P_n *el espacio vectorial de los polinomios de grado menor o igual que* n, D *el operador derivación de* P_n *en* P_n, *es decir,* $D(p(x)) = p'(x)$. *Decida si* D *es una transformación lineal y encuentre la matriz asociada a* D *respecto a la base* $\{1, x, x^2, \ldots, x^n\}$. *¿Cuál es el núcleo e imagen de* D*?*

En diferentes problemas, la idea de composición de funciones emerge como una herramienta para formular y entender los procesos involucrados. En situaciones relacionadas con transformaciones lineales la situación es como sigue.

Supongamos que V, W y U son tres espacios vectoriales, $T : V \to W$ y $T_1 : W \to U$ transformaciones lineales, entonces tiene sentido formar la función $T_1 \circ T : \mathrm{V} \to U$ la cual resulta ser lineal.

En efecto, si $\alpha, \beta \in V$ entonces $(T_1 \circ T)\ (\alpha + \beta) = T_1(T(\alpha + \beta)) = T_1(T(\alpha) + T(\beta)) = T_1(T(\alpha)) + T_1(T(\beta)) = (T_1 \circ T)(\alpha) + (T_1 \circ T)(\beta)$.

De manera similar se demuestra que $(T_1 \circ T)(a\alpha) = a(T_1 \circ T)(\alpha)$. Esto prueba lo afirmado.

Sean V, W y U espacios vectoriales con bases $\{\alpha_1, \alpha_2, ..., \alpha_n\}$, $\{\beta_1, \beta_2, ..., \beta_m\}$ y $\{\gamma_1, \gamma_2, ..., \gamma_r\}$ respectivamente. Si $T : V \to W$ y $T_1 : W \to U$ son dos transformaciones lineales, y $A_{[\alpha, \beta]}$ y $B_{[\beta, \gamma]}$ son las matrices asociadas a T y T_1 respecto a las correspondientes bases, se pregunta, ¿cuál es la matriz asociada a $T_1 \circ T$ respecto a las bases $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ y $\{\gamma_1, \gamma_2, ..., \gamma_r\}$?

Para determinar la matriz asociada a $T_1 \circ T$ debemos evaluar a $T_1 \circ T$ en cada uno de los elementos de la base $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ y expresar el resultado como combinación lineal de los elementos de la base $\{\gamma_1, \gamma_2, ..., \gamma_r\}$.

Supongamos que las matrices $A_{[\alpha, \beta]}$ y $B_{[\beta, \gamma]}$ tienen entradas a_{ij} y b_{ts} respectivamente. Llamemos C a la matriz asociada a $T_1 \circ T$ respecto a las bases $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ y $\{\gamma_1, \gamma_2, ..., \gamma_r\}$. Note que la matriz C es de orden $r \times n$ y sus entradas satisfacen:

$$(T_1 \circ T)(\alpha_j) = c_{1j}\gamma_1 + c_{2j}\gamma_2 + \cdots + c_{rj}\gamma_r \tag{4.11}$$

Usando la definición de composición de funciones y de matriz asociada se tiene:

$$\begin{aligned}
(T_1 \circ T)(\alpha_j) &= T_1(T(\alpha_j)) \\
&= T_1(a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m) \\
&= a_{1j}T_1(\beta_1) + a_{2j}T_1(\beta_2) + \cdots + a_{mj}T_1(\beta_m) \\
&= a_{1j}\left(\sum_{t=1}^{r} b_{t1}\gamma_t\right) + a_{2j}\left(\sum_{t=1}^{r} b_{t2}\gamma_t\right) + \cdots + a_{mj}\left(\sum_{t=1}^{r} b_{tm}\gamma_t\right) \\
&= \left(\sum_{k=1}^{m} b_{1k}a_{kj}\right)\gamma_1 + \left(\sum_{k=1}^{m} b_{2k}a_{kj}\right)\gamma_2 + \cdots + \left(\sum_{k=1}^{m} b_{rk}a_{kj}\right)\gamma_r
\end{aligned}$$

Igualando términos con la ecuación 4.11 y usando el hecho que la representación de un vector en una base es única se llega a:

$$c_{ij} = b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{im}a_{mj}, \text{ para todo } i \text{ y para todo } j$$

es decir, la matriz asociada a $T_1 \circ T$ es $B_{[\beta, \gamma]}A_{[\alpha, \beta]}$

La discusión anterior la formalizamos en el siguiente:

Teorema 4.4.1. *Sean* U, V *y* W *espacios vectoriales con bases* $\{\alpha_1, \alpha_2, ..., \alpha_n\}$, $\{\beta_1, \beta_2, ..., \beta_m\}$ y $\{\gamma_1, \gamma_2, ..., \gamma_r\}$ *respectivamente. Si* $\mathrm{T} : \mathrm{V} \to \mathrm{W}$ *y* $\mathrm{T}_1 : \mathrm{W} \to \mathbf{U}$ *son dos transformaciones lineales, y* $\mathrm{A}_{[\alpha, \beta]}$ y $\mathrm{B}_{[\beta, \gamma]}$ *son las matrices asociadas a* T *y* T_1 *respecto a las bases dadas, entonces la matriz asociada a* $\mathrm{T}_1 \circ \mathrm{T}$ *es* $\mathrm{B}_{[\beta, \gamma]}\mathrm{A}_{[\alpha, \beta]}$.

El lector está invitado a explicar los detalles en los argumentos anteriores.

4.4.1. Matrices de cambio de base

Cuando se elige una base en $\mathbb{R}^2$, se ha definido un sistema de referencia. Por ejemplo, cuando tomamos la base canónica y representamos un vector de $\mathbb{R}^2$ en la forma $\alpha = (x, y)$, lo que esto significa es que para llegar al punto que determina al vector α, se debe uno desplazar desde el origen x unidades en la dirección del vector $e_1 = (1, 0)$ y y unidades en la dirección del vector $e_2 = (0, 1)$. Sin embargo, para llegar al punto que determina al vector α, pudimos haber elegido otro sistema de referencia. Por ejemplo los vectores $\alpha_1 = (3, 2)$ y $\alpha_2 = (1, 3/2)$. Entonces para llegar al punto α habrá que desplazarse x_1 unidades en la dirección del vector α_1 y y_1 unidades en la dirección del vector α_2. A la pareja de números x_1 y y_1 les llamamos las *coordenadas de* α *respecto a la base* $\{\alpha_1 = (3, 2), \alpha_2 = (1, 3/2)\}$.

Ejercicio 4.4.2. *Haga un dibujo que ilustre lo discutido en el párrafo anterior.*

Una pregunta interesante es: ¿Cuál es la relación que hay entre las coordenadas de α respecto a la base canónica y respecto a la base $\{\alpha_1 = (3, 2), \alpha_2 = (1, 3/2)\}$?

Para contestar esta pregunta introduciremos la idea de matriz de cambio de base. Supongamos que se tienen un espacio vectorial V y dos bases: $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ y , $\{\alpha_1', \alpha_2', \ldots \alpha_n'\}$ entonces cada uno de los elementos α_j' se puede expresar como combinación lineal de los elementos $\alpha_1, \alpha_2, \ldots, \alpha_n$ es decir, existen escalares $p_{1j}, p_{2j}, \ldots, p_{nj}$, tales que $\alpha_j' = p_{1j}\alpha_1 + p_{2j}\alpha_2 + \cdots + p_{nj}\alpha_n$. Con estos coeficientes se forma la matriz:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

tomando los coeficientes de cada α_j', para obtener la columna j-ésima.

Definición 4.4.2. *Con la notación e hipótesis como antes, a la matriz* P *se le llama la matriz que cambia la base* $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *a la base* $\{\alpha_1', \alpha_2', \ldots \alpha_n'\}$.

Observación 4.4.1. *Dado que* $\{\alpha_1', \alpha_2', \ldots \alpha_n'\}$ *también es base, cada* α_j *se puede expresar como combinación lineal de los* $\{\alpha_1', \alpha_2', \ldots \alpha_n'\}$ *y se obtiene una matriz* Q, *la cual es la inversa de* P.

Ejercicio 4.4.3. *Justificar la observación anterior.*

Para contestar la pregunta sobre la relación que hay entre las coordenadas del vector α en la base canónica de $\mathbb{R}^2$ y en la base $\{\alpha_1 = (3, 2), \alpha_2 = (1, 3/2)\}$, procederemos de forma general y después aplicaremos el método al caso que corresponde. Antes precisaremos lo que entenderemos por coordenadas respecto a una base.

Sea V un espacio vectorial, $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ una base de V. Sabemos que para todo $\alpha \in V$ existen únicos escalares $x_1, x_2, \ldots, x_n$ tales que:

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n \tag{4.12}$$

Definición 4.4.3. *Sean* V, $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *y* $\alpha \in$ V *como en la discusión previa,* A *los escalares* $x_1, x_2, \ldots, x_n$ *que aparecen en (4.12) les llamamos las coordenadas de* α *respecto a la base* $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$.

Observación 4.4.2. *Si* $x_1, x_2, ..., x_n$ *son las coordenadas del vector* α *respecto a la base* $\mathcal{B} = \{\alpha_1, \alpha_2, ..., \alpha_n\}$, *la cual suponemos fija, construimos el vector* $X = (x_1, x_2, ..., x_n)$ *en* $\mathbb{R}^n$ *y le llamamos el vector coordenado de* α *respecto a la base* $\mathcal{B}$. *Al vector* X *también lo podemos considerar como vector columna, y así lo haremos en lo que sigue.*

Sean $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ y $\{\alpha_1', \alpha_2', \dots \alpha_n'\}$ bases de V. Dado $\alpha \in V$, éste se puede representar como $\alpha = \sum_{i=1}^{n} x_i\alpha_i$ y como $\alpha = \sum_{i=1}^{n} y_i\alpha_i$. Si P es la matriz que cambia la base $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ a la base $\{\alpha_1', \alpha_2', \dots \alpha_n'\}$, mostraremos que $X = PY$, en donde X es el vector coordenado de α respecto a la base $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ y Y es el vector coordenado de α respecto a la base $\{\alpha_1', \alpha_2', \dots \alpha_n'\}$.

Por cuestiones de notación, representemos la relación entre los elementos de las bases mediante la ecuación "matricial":

$$\begin{bmatrix} p_{11} & p_{21} & \cdots & p_{n1} \\ p_{21} & p_{22} & \cdots & p_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ p_{n1} & p_{2n} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \alpha_1' \\ \alpha_2' \\ \vdots \\ \alpha_n' \end{bmatrix} \tag{4.13}$$

Note que el primer factor en la ecuación anterior es P^t.

En forma análoga, podemos representar al vector α usando sus vectores coordenados respecto a cada base.

$$\alpha = [x_1 \;\; x_2 \;\; \cdots \;\; x_n] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = [y_1 \;\; y_2 \;\; \cdots \;\; y_n] \begin{bmatrix} \alpha_1' \\ \alpha_2' \\ \vdots \\ \alpha_n' \end{bmatrix}$$

De la ecuación 4.13 se tiene:

$$\alpha = [x_1 \;\; x_2 \;\; \cdots \;\; x_n] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = [y_1 \;\; y_2 \;\; \cdots \;\; y_n] \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{n1} \\ p_{12} & p_{22} & \cdots & p_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ p_{1n} & p_{2n} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

Como la representación de un vector respecto a una base es única, se concluye que $X^t = Y^tP^t$, equivalentemente,

$$X = PY \tag{4.14}$$

Como P es una matriz que tiene inversa, la ecuación $X = PY$ es equivalente a $P^{-1}X = Y$, es decir, las coordenadas de α respecto a la nueva base $\{\alpha_1', \alpha_2', \dots \alpha_n'\}$ se obtienen de la ecuación $P^{-1}X = Y$.

La pregunta que dió origen a la discusión previa la podemos contestar de la siguiente forma. Debemos encontrar la inversa de la matriz que cambia la base canónica a la base $\{(3, 2), (1, 3/2)\}$. Tal matriz es: $P = \begin{pmatrix} 3 & 1 \\ 2 & \frac{3}{2} \end{pmatrix}$. A partir de esto se encuentra que

$P^{-1} = \begin{bmatrix} \frac{3}{5} & -\frac{2}{5} \\ -\frac{4}{5} & \frac{6}{5} \end{bmatrix}$. Si Y es el vector coordinado de α respecto a la base {(3, 2), (1, 3/2} y X es el vector coordenado de α respecto a la base canónica, entonces $Y = P^{-1}X = \begin{pmatrix} \frac{3}{5} & -\frac{2}{5} \\ -\frac{4}{5} & \frac{6}{5} \end{pmatrix} X$, equivalente, $y_1 = \frac{3}{5}x_1 - \frac{2}{5}x_2$ y $y_2 = -\frac{4}{5}x_1 + \frac{6}{5}x_2$.

Ejemplo 4.4.3. *Dada la base canónica* $\{e_1, e_2, e_3\}$ *de* $\mathbb{R}^3$, *encuentre las coordenadas del vector* $\alpha = e_1 - 2e_2 + 5e_3$ *respecto a la base* {(1, 2, 0), (1, −1, 0), (1, 1, 1)}.

Discusión. La matriz de cambio de base es $P = \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$; usando algún método para calcular inversas se encuentra que:

$$P^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix}$$

De esto se tiene que las coordenadas de α respecto a la nueva base están dadas por

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{11}{3} \\ -\frac{1}{3} \\ 5 \end{bmatrix}$$

La ecuación anterior expresa las coordenadas de (1, –2, 5) respecto a la base:

{(1, 2, 0), (1, −1, 0), (1, 1, 1)}, es decir, $(1, -2, 5) = -\frac{11}{3}(1, 2, 0) - \frac{1}{3}(1, -1, 0) + 5(1, 1, 1)$

En muchas aplicaciones es importante encontrar una matriz, lo más simple posible, que represente a una transformación lineal $T : V \to W$. Esto por supuesto depende de poder elegir adecuadamente bases en los espacios V y W. Sabemos que dadas bases en V y W, T tiene asociada una matriz, más precisamente, si $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ y $\{\beta_1, \beta_2, ..., \beta_m\}$ son bases de V y W respectivamente, la matriz asociada a T queda definida. Al cambiar a nuevas bases, digamos $\{\alpha'_1, \alpha'_2, ..., \alpha'_n\}$ y $\{\beta'_1, \beta'_2, ..., \beta'_m\}$, en estas nuevas bases, T tiene asociada una matriz, B. ¿Cuál es la relación entre las matrices A, B y las matrices de cambio de base? La respuesta la proporciona el siguiente teorema, cuya demostración dejamos como ejercicio.

Teorema 4.4.2. (Cambio de base). *Sea* $T : V \to W$ *una transformación lineal. Supongamos que* A *es la matriz asociada a* T *respecto a bases dadas* $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ *en* V *y* $\{\beta_1, \beta_2, ..., \beta_m\}$ *en* W. *Si las bases anteriores se cambian a nuevas bases* $\{\alpha'_1, \alpha'_2, ..., \alpha'_n\}$ *y* $\{\beta'_1, \beta'_2, ..., \beta'_m\}$, *con matrices de cambio de base* P *y* Q *respectivamente y* B *es la matriz asociada a* T *en estas nuevas bases, entonces se tiene:*

$$B = Q^{-1}AP.$$

Demostración. Ejercicio.

Corolario 4.4.1. *Si* $T : V \to V$ *es una transformación lineal,* $\alpha_i = \beta_i$ *y* $\alpha_i' = \beta_i'$ *para todo* $i = 1, 2, \ldots, n$, *entonces la matriz asociada a* T *respecto a la nueva base es* $P^{-1}AP$, P *la matriz de cambio de base.*

Definición 4.4.4. *Sean* A *y* B *dos matrices* $n \times n$. *Se dice que* B *es similar o semejante a* A, *si existe una matriz no singular* P *tal que* $B = P^{-1}AP$.

Ejercicio 4.4.4. *Demuestre lo siguiente.*

1. *Toda matriz es similar a sí misma (reflexividad).*
2. *Si* B *es similar a* A, *entonces* A *es similar a* B *(simetría).*
3. *Si* C *es similar a* B *y* B *es similar a* A, *entonces* C *es similar a* A *(transitividad).*

Ejemplo 4.4.4. *Sea* $T : \mathbb{R}^3 \to \mathbb{R}^3$ *dada por* $T(x, y, z) = (x + y - z, 2x - y + 3z, x - z)$ *encuentre la matriz asociada a* T *respecto a la base* $\{(1, 2, 0), (1, -1, 0), (1, 1, 1)\}$.

Discusión. Primero encontramos la matriz asociada a T respecto de la base canónica, la cual se obtiene evaluando a T en los vectores canónicos.

Se tiene $T(1, 0, 0) = (1, 2, 1)$, $T(0, 1, 0) = (1, -1, 0)$ y $T(0, 0, 1) = (-1, 3, -1)$, por lo que la matriz asociada a T respecto de la base canónica es:

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

La matriz de cambio de base es $P = \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$; aplicando el método de operaciones elementales para calcular la inversa de P se tiene:

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

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 0 & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

De lo anterior se concluye que $P^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 & -2 \\ 2 & -1 & -1 \\ 0 & 0 & 3 \end{bmatrix}$. Aplicando el teorema 4.4.2 inferimos que la matriz asociada a T respecto a la base $\{(1, 2, 0), (1, -1, 0), (1, 1, 1,)\}$ es:

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

En lo que sigue procederemos a demostrar el resultado que mencionamos al inicio de esta sección concerniente al rango de una matriz.

Definición 4.4.5. *Dada una matriz* m × n*, al espacio generado por las columnas de* A *le llamaremos el espacio columna de* A *y a su dimensión, el rango columna de* A*. De manera análoga se define el espacio fila de* A*, así como el rango fila de* A*.*

Teorema 4.4.3. *Sea* A *una matriz* m × n*, entonces el rango fila de* A *es igual al rango columna de* A*.*

Demostración. Si A es la matriz cero, no hay nada que demostrar, pues los espacios fila y columna tienen dimensión cero.

Podemos suponer que A es no cero. Por el teorema 1.3.2, página 29, A se puede llevar mediante operaciones elementales en sus filas a una matriz en forma escalonada reducida R. Se tiene que el espacio fila de A y el de R son el mismo, pues las filas de R son combinaciones lineales adecuadas de las filas de A, por lo que el rango fila de A es igual al número de filas no cero de R, pues las filas no cero de R constituyen una base para su espacio fila (explique esto). Supongamos que R tiene r filas no cero, y que la entrada principal de la fila i de R se encuentra en la columna t_i, entonces dichas columnas de R son linealmente independientes y cualesquiera otras son combinación lineal de esas (explique esto). Es decir, las columnas de R en donde se encuentran las entradas principales de las filas no cero, son una base del espacio columna de R, por lo que R tiene rango columna r, concluyendo que el resultado se cumple en R.

Ahora mostraremos que también se cumple en A. Para esto notamos que los conjuntos solución de las ecuaciones $AX = 0$ y $RX = 0$ son el mismo, esto equivale a decir que si denotamos por $A_1, A_2, \ldots, A_n$ y $R_1, R_2, \ldots, R_n$ a las columnas de A y R respectivamente, entonces para escalares $x_1, x_2, \ldots, x_n$ se tiene $x_1A_1 + x_2A_2 + \cdots + x_nA_n = 0 \Leftrightarrow x_1R_1 + x_2R_2 + \cdots + x_nR_n = 0$. Sea t el rango columna de A. Sin perder generalidad podemos suponer que $A_1, A_2, \ldots, A_t$ son una base del espacio columna de A. Mostraremos que $t = r$.

Sean $x_1, \ldots, x_t$ escalares tales que $x_1R_1 + x_2R_2 + \cdots + x_tR_t = 0$, entonces también se tiene $x_1A_1 + x_2A_2 + \cdots + x_tA_t = 0$, por lo que $x_1 = x_2 = \cdots = x_t = 0$, es decir, $R_1, R_2, \ldots, R_t$ son linealmente independientes, concluyendo que $t \leq r$.

Si $x_1, \ldots, x_r$ satisfacen $x_1A_1 + x_2A_2 + \cdots + x_rA_r = 0$, entonces $x_1R_1 + x_2R_2 + \cdots + x_rR_r = 0$, por lo que $x_1 = x_2 = \cdots = x_r = 0$, es decir, $A_1, A_2, \ldots, A_r$ son linealmente independientes, de lo que se obtiene $r \leq t$, terminando la demostración del teorema.

Definición 4.4.6. *Dada una matriz* A *se define el rango de* A *como el rango fila o el rango columna de* A*, los cuales coinciden gracias al teorema anterior.*

Observación 4.4.3. *De la demostración del teorema 4.4.3 se puede notar que para calcular el rango fila de una matriz* A*, se lleva a forma escalonada reducida. El rango de* A *es el número de filas no cero de su forma escalonada reducida. Si* A *es una matriz* m × n*, su rango es* ≤ k*, en donde* k *es el mínimo entre* m *y* n*.*

Ejemplo 4.4.5. *Sea* $A = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 2 & -1 & 3 & 0 \\ 1 & 0 & -1 & 2 \end{bmatrix}$. *Calculemos el rango de* A*.*

Discusión. Por la observación anterior, es suficiente llevar a A a forma escalonada reducida y el número de filas no cero será el rango buscado.

Aplicando operaciones elementales en las filas de A se tiene:

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

La última matriz es la forma escalonada reducida de A; de esto concluimos que el rango de A es tres.

4.4.2. El espacio de las transformaciones lineales

Sean V y W espacios vectoriales de dimensión n y m respectivamente; el espacio de las transformaciones lineales, denotado $\mathcal{L}(V; W)$, es el conjunto de todas las transformaciones lineales de V en W. Las operaciones que hacen de $\mathcal{L}(V; W)$ un espacio vectorial son:

- **Suma.** Dados $T, T_1 \in \mathcal{L}(V; W)$ se define $(T + T_1)(\alpha) := T(\alpha) + T_1(\alpha_1)$.
- **Producto por escalar.** Dado un escalar λ y un elemento $T \in \mathcal{L}(V; W)$ se define el producto $(\lambda T)(\alpha) : = \lambda T(\alpha)$.

Se deja como ejercicio verificar que con la suma y producto por escalar definidos antes, $\mathcal{L}(V; W)$ es un espacio vectorial.

Recordemos que dadas $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ y $\{\beta_1, \beta_2, ..., \beta_m\}$, bases de V y W, respectivamente, al considerar $T \in \mathcal{L}(V; W)$ se construye la matriz asociada a T, entonces al considerar dos transformaciones lineales $T, T_1 \in \mathcal{L}(V; W)$, hay dos matrices A y B asociadas a T y T_1. ¿Cuál es la matriz asociada a $T + T_1$? De la definición de matriz asociada a una transformación lineal se tiene que la columna j-ésima de A se obtiene de los escalares que aparecen en $T(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m$; análogamente, la columna j-ésima de B se obtiene de los escalares que aparecen en $T_1(\alpha_j) = b_{1j}\beta_1 + b_{2j}\beta_2 + \cdots + b_{mj}\beta_m$. Usando la definición de la suma de transformaciones lineales se tiene:

$$\begin{aligned} (T + T_1)(\alpha_j) &= T(\alpha_j) + T_1(\alpha_j) \\ &= (a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m) + (b_{1j}\beta_1 + b_{2j}\beta_2 + \cdots + b_{mj}\beta_m) \\ &= (a_{1j} + b_{1j})\beta_1 + (a_{2j} + b_{2j})\beta_2 + \cdots + (a_{mj} + b_{mj})\beta_m \end{aligned}$$

De la ecuación anterior concluimos que la transformación $T + T_1$ tiene asociada a la matriz $A + B$. Un cálculo como el anterior muestra que si λ es un escalar, entonces λT tiene asociada a la matriz λA.

De la discusión anterior se tiene definida una función $\varphi : \mathcal{L}(V; W) \to \mathcal{M}_{m\times n}(\mathbb{R})$ por la regla: $\varphi(T) := A$.

Las propiedades de φ se enuncian en el siguiente:

Teorema 4.4.4. *Sean* V *y* W *espacios vectoriales de dimensión* n *y* m *respectivamente. Entonces* φ *es un isomorfismo, consecuentemente, la dimensión de* $\mathcal{L}(V; W)$ *es* mn.

Demostración. El lector está invitado a probar que φ es un isomorfismo, la segunda parte la probaremos directamente, pues consideramos de cierta importancia el presentar una prueba en donde se construya una base de $\mathcal{L}(V; W)$, ya que en este proceso tendremos la oportunidad de ilustrar el uso de varios conceptos y resultados. Sean $\alpha_1, \alpha_2, ..., \alpha_n$ y $\beta_1, \beta_2, ..., \beta_m$, bases de V y W, respectivamente; definiremos transformaciones lineales $T_{ij} : V \rightarrow W$ que formarán una base. Como una transformación lineal queda completamente definida en una base, procedemos a definir los elementos T_{ij} como sigue.

$$T_{ij}(\alpha_k) = \begin{cases} \beta_i \text{ si } j = k \\ 0 \text{ en otro caso} \end{cases}$$

Mostraremos primero que el conjunto de transformaciones lineales $\{T_{ij}\}$ es linealmente independiente, para esto consideremos una combinación lineal $\sum_{i,j} a_{ij} T_{ij} = 0$.

Como esta combinación lineal es la función cero, entonces al evaluarla en cualquier α_k el resultado es cero, es decir, $0 = \sum_{i,j} a_{ij} T_{ij}(\alpha_k) = \sum_{i=1}^{m} a_{ik} \beta_i$, para todo $k = 1, \dots, n$.

Por otro lado, como ésta es una combinación lineal en términos de la base de W, los a_{ik} son cero para todo i, entonces se tiene $a_{ik} = 0$ para todo i y para todo k, probando lo deseado.

Para demostrar que el conjunto de transformaciones $\{T_{ij}\}$ generan, consideremos un elemento $T \in \mathcal{L}(V; W)$ y demostremos que existen escalares b_{ij} tales que $\sum_{i,j} b_{ij} T_{ij} = T$. Como T es dado, esto significa que conocemos su acción en cada α_k. es decir, conocemos los escalares $a_{1k}, a_{2k}, \dots, a_{mk}$ tales que $T(\alpha_k) = \sum_{i=1}^{m} a_{ik} \beta_i$, para cada $k = 1, 2, \dots, n$. Definamos $b_{ij} := a_{ij}$ y formemos la combinación lineal $\sum_{i,j} b_{ij} T_{ij}$; de acuerdo con los cálculos anteriores, se tiene $T(\alpha_k) = \sum_{i,j} b_{ij} T_{ij}(\alpha_k)$, probando con esto lo que deseábamos.

El conjunto $\{T_{ij}\}$ tiene nm elementos, entonces la dimensión de $\mathcal{L}(V; W)$ es nm.

4.5. Ejercicios

1. Enuncie los teoremas y las definiciones que se han discutido en este capítulo.
2. Expanda el conjunto $\{(1, 2, 9)\}$ a una base de $\mathbb{R}^3$.
3. Encuentre dos bases de $\mathbb{R}^4$ que no tengan elementos en común.
4. En el espacio de los polinomios de grado a lo más 3, encuentre una base que incluya al elemento x.
5. Encuentre explícitamente una base para cada uno de los siguientes espacios:
   a) $\mathcal{L}(\mathbb{R}^3; P_4)$.
   b) $\mathcal{L}(\mathbb{R}^4; \mathcal{M}_{2\times 2}(\mathbb{R}))$.
   c) $\mathcal{L}(P_2; \mathcal{M}_{2\times 3}(\mathbb{R}))$.
6. Si W_1 y W_2 son subespacios de dimensión tres en $\mathbb{R}^4$, determine las posibles dimensiones de $W_1 \cap W_2$ y $W_1 + W_2$.
7. ¿Tiene dimensión finita el espacio de las funciones continuas de $\mathbb{R}$ en $\mathbb{R}$?
8. Demuestre que la composición de transformaciones lineales es una transformación lineal.

9. Si $T : \mathbb{R}^2 \to \mathbb{R}^2$ es una transformación lineal tal que $T(1, 0) = (a, b)$ y $T(0, 1) = (c, d)$. Encuentre $T(x, y)$.

10. Sea $T : \mathbb{R}^3 \to \mathbb{R}^4$ dada por $T(x, y, z) = (x - y, y - z, z - x, x)$. Determine N_T y R_T.

11. Encuentre una transformación lineal inyectiva de $\mathbb{R}^3$ en el espacio de los polinomios de grado a lo más tres.

12. Sean V y W dos espacios vectoriales, $T : V \to W$ una función. Demuestre que T es lineal $\Leftrightarrow T(x\alpha + y\beta) = xT(\alpha) + yT(\beta)$, para todos $x, y \in \mathbb{R}$ y para todos $\alpha, \beta \in V$.

13. Sea $T : \mathbb{R}^n \to \mathbb{R}$ una función. Demuestre que T es lineal $\Leftrightarrow$ existen $a_1, a_2, \ldots, a_n \in \mathbb{R}$ tales que $T(x_1, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$.

14. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una función, con funciones coordenadas $T_1, T_2, \ldots, T_m$. Demuestre que T es lineal $\Leftrightarrow$ cada T_i es lineal, $i = 1, 2, \ldots, m$.

15. Sea $T : \mathbb{R} \to \mathbb{R}$ una función tal que $T(x + y) = T(x) + T(y)$ para todo $x, y \in \mathbb{R}$. Demuestre que las siguientes condiciones acerca de T son equivalentes.

a) La función T es lineal.

b) La función T es continua en $\mathbb{R}$.

c) La función T es continua en un punto $x_0 \in \mathbb{R}$.

16. Sea V un espacio vectorial, $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ una base de V y $T : V \to V$ la transformación lineal cuya acción en la base está dada por $T(\alpha_n) = 0$ y para cada $i = 1, 2, \ldots, n - 1$, $T(\alpha_i) = \alpha_{i+1}$.

a) Encuentre la matriz A, asociada a T respecto a la base dada.

b) Demuestre que $T^n = 0$, en donde T^n significa la composición de T consigo misma n veces.

c) Concluya de la parte anterior que $A^n = 0$.

17. Determine cuáles de las siguientes proposiciones son verdaderas o falsas. Explique sus respuestas.

a) Si ninguno de los vectores $\{\alpha, \beta, \gamma\}$ es múltiplo de alguno de los dos restantes, entonces son linealmente independientes.

b) Si los vectores $\{\alpha, \beta, \gamma\}$ son linealmente dependientes, entonces pertenecen a $\mathbb{R}^2$.

c) Las matrices cuadradas $n \times n$ pueden ser consideradas como matrices asociadas a transformaciones lineales de $\mathbb{R}^n \to \mathbb{R}^n$ respecto a alguna base.

d) Si una transformación lineal $T : V \to V$ es sobre y V tiene dimensión finita, entonces T es biyectiva.

e) Si W es un subespacio propio de $\mathbb{R}^n$, entonces $\dim(W) < n$.

f) Una transformación lineal puede transformar un plano en una recta.

g) Una transformación lineal puede transformar a $\mathbb{R}^n$ en todo $\mathbb{R}^m$, con $m > n$.

h) Un movimiento rígido del plano transforma el cero en el cero.

18. ¿Cómo define movimiento rígido en $\mathbb{R}^n$?

19. Demuestre que si un movimiento rígido de $\mathbb{R}^n$ deja fijo al cero, entonces es una transformación lineal.

20. Demuestre que una transformación lineal T es no singular $\Leftrightarrow T$ transforma conjuntos linealmente independientes en conjuntos linealmente independientes.

21. Sea V un espacio vectorial de dimensión finita, $T : V \to V$ una transformación lineal, Demuestre que T es no singular $\Leftrightarrow$ T es biyectiva $\Leftrightarrow$ T es suprayectiva.

22. Complete la demostración del teorema 4.1.2.

23. En este ejercicio se bosqueja la demostración del teorema 4.3.1, página 103. Sea $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ una base de V, entonces dado $\alpha \in V$, existen escalares $a_1, a_2, ..., a_n$, únicos, tales que $\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n$. Definimos $T : V \to \mathbb{R}^n$ por $T(\alpha) := (a_1, a_2, ..., a_n)$. Demuestre que T es un isomorfismo.

24. En este ejercicio se bosqueja la demostración del teorema 4.3.2, página 103. Si T es no singular, de acuerdo con el ejercicio 20, página 114, T transforma conjuntos linealmente independientes en conjuntos linealmente independientes, por lo que la dimensión de V es igual a la dimensión del rango de T y la dimensión del núcleo es cero. Justifique esto y concluya la igualdad establecida en el teorema.

Si el núcleo de T, N_T no es cero, sea $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ una base de N_T. Extienda $\{\alpha_1, \alpha_2, ..., \alpha_r\}$ a una base de V, digamos $\{\alpha_1, \alpha_2, ..., \alpha_r, \alpha_{r+1}, ..., \alpha_n\}$. Demuestre que $\{T(\alpha_{r+1}), ..., \mathrm{T}(\alpha_n)\}$ es una base del rango de T y de esto concluya la ecuación establecida en el teorema.

25. Sea V el espacio vectorial de los polinomios con coeficientes reales. Considere la transformación lineal $T: V \to V$ definida por $T(f(x)) := a_0x + \dfrac{a_1x^2}{2} + \cdots + \dfrac{a_kx^{k+1}}{k+1}$, con $f(x) = a_0 + a_1x + \cdots + a_kx^k$. Sea $D : V \to V$, la transformación derivada, es decir, $D(f(x)) = f'(x)$. Demuestre que $\mathrm{D} \circ T = id$, pero ni D ni T son isomorfismo. Discuta si alguno de D o T es inyectivo, suprayectivo.

26. Sea V un espacio vectorial. Se define el espacio de los operadores en V como $\mathcal{L}(V; V) := \{T: V \to V \mid T \text{ es lineal}\}$. Supongamos que V es de dimensión finita y sea T un operador en V. Demuestre que T es inyectivo $\Leftrightarrow$ T es suprayectivo.

27. Sea $T \in \mathcal{L}(V; V)$. Demuestre que $\dim(R_T)$ = número de filas no cero de la matriz escalonada reducida de A, con A la matriz asociada a T respecto a alguna base.

28. Sea $T \in \mathcal{L}(V; V)$. Demuestre que $T^2 = 0 \Leftrightarrow T(V) \subseteq N_T$.

29. Encuentre un ejemplo de un operador $T: V \to V$ para mostrar que $T(V) \cap N_T \neq \{0\}$ puede ocurrir.

30. Sea $T \in \mathcal{L}(V, V)$. Demuestre que existe un operador no cero $S \in \mathcal{L}(V; V)$ tal que $T \circ S = 0 \Leftrightarrow N_T \neq \{0\}$.

31. Sea V un espacio vectorial de dimensión finita, $S, T \in \mathcal{L}(V; V)$ tales que $S \circ T = id$. Demuestre que $T \circ S = id$. (El ejercicio 25 muestra que la conclusión no es cierta si V no tiene dimensión finita.)

32. Explique si existe una transformación lineal $T : \mathbb{R}^4 \to \mathbb{R}^2$ tal que $T(0, 1, 1, 1) = (2, 0)$, $T(1, 2, 1, 1) = (1, 2)$, $T(1, l, 1, 2) = (3, 1)$, y $T(2, 1, 0, 1) = (2, 3)$.

33. Sea V un espacio de dimensión n. Denotemos por V^* al espacio $\mathcal{L}(V; \mathbb{R})$. Sabemos que este espacio tiene dimensión n. Demuestre este hecho de la siguiente forma. Dada una base $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ de V, existe una base $\{f_1, f_2 ..., f_n\}$ de V^* tal que $f_i(\alpha_j) = 1$, si $i = j$ y $f_i(\alpha_j) = 0$ en otro caso. A la base $\{f_1, f_2 ..., f_n\}$ se le llama base dual a $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ y al espacio V^* se le llama el espacio dual de V.

34. Sea $A = (a_{ij})$ una matriz $n \times n$ con entradas reales que satisface:

i) $a_{ij} < 0$ para $i \neq j$.

ii) $a_{i1} + a_{i2} + \cdots + a_{in} > 0$ para todo $i = 1, \ldots, n$.

Demuestre que A tiene inversa. Sugerencia: Si A no tiene inversa, el sistema $AX = 0$

tiene una solución diferente de cero. Sea $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ tal solución.

Si x_k es la componente de X que satisface $|x_k| \geq x_i$ para todo i, multiplicando por -1 podemos suponer que x_k es positivo. La hipótesis sobre las entradas de la matriz implica $a_{kj}x_j \geq a_{kj}x_k$ para todo $j \neq k$. Use esto, ii) y $a_{k1}x_1 + \cdots + a_{kk}x_k + \cdots + a_{kn}x_n = 0$ para llegar a una contradicción.

35. Supongamos que A es la matriz aumentada de un sistema de n ecuaciones en n variables. Demuestre que el sistema representado por A tiene solución única $\Leftrightarrow$ el rango de A es n.

Capítulo 5

Determinantes

La teoría de los determinantes es una de las herramientas más importantes en el estudio de sistemas de ecuaciones lineales. También tiene aplicaciones en el estudio de transformaciones lineales, el cálculo de áreas y volúmenes, entre otras muchas. Dicho sea de paso, la presentación que haremos de la teoría de determinantes tiene como punto de partida las propiedades de la función que prescribe el área de un paralelogramo determinado por un par de vectores en el plano. Ver la ilustración de la figura 5.1. Esta misma idea se extiende al caso del volumen de un paralelepípedo, es decir, la función que mide el volumen de un paralelepípedo determinado por tres vectores en el espacio, satisface las mismas propiedades que la función que mide el área de un paralelogramo.

5.1. Determinantes y volúmenes de paralelepípedos

En el capítulo 3, ecuación 3.15, encontramos una forma de calcular el área de un paralelogramo cuyos lados se determinan por los vectores $\vec{u}_1$ y $\vec{u}_2$. Ahora estamos interesados en examinar la función que a cada par de vectores $\vec{u}_1$ y $\vec{u}_2$ le asigna el valor de esa área. Denotemos esa función por $A(\vec{u}_1, \vec{u}_2)$ y hagamos algunas observaciones

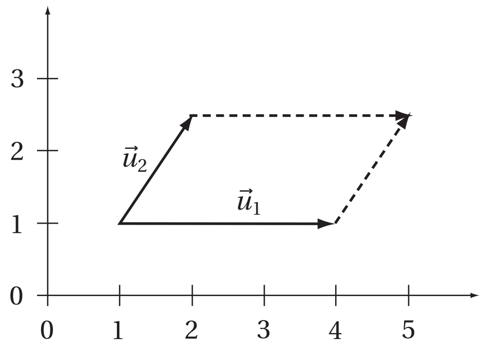


Figura 5.1. Paralelogramo determinado por dos vectores.

de las propiedades de la función *A*. Un argumento, motivado por una gráfica, muestra que la función *A* cumple las siguientes propiedades.

1. $A(\vec{e}_1, \vec{e}_2) = 1$, (área del cuadrado unitario).
2. Si λ es un real positivo, entonces $A(\lambda\vec{u}_1, \vec{u}_2) = A(\vec{u}_1, \lambda\vec{u}_2) = \lambda A(\vec{u}_1, \vec{u}_2)$.
3. $A(\vec{u}_1 + \vec{u}_2) = A(\vec{u}_1, \vec{u}_1 + \vec{u}_2) = A(\vec{u}_1, \vec{u}_2)$.
4. $A(\vec{u}_1, \vec{u}_2) \neq 0 \Leftrightarrow \vec{u}_1$ y $\vec{u}_2$ son linealmente independientes.

La propiedad 1 es evidente, pues los vectores canónicos determinan el cuadrado unitario (cuadrado de lado uno).

La propiedad 2 se obtiene notando que al multiplicar uno de los vectores por λ, el paralelogramo resultante tiene la misma altura[1] y la base es λ veces la base del original.

Ejercicio 5.1.1. *Justifique las propiedades 3 y 4.*

De acuerdo con la observación 3.3.2, página 78, el área de un paralelogramo determinado por un par de vectores, $\vec{u} = (a, b)$ y $\vec{v} = (a_1, b_1)$ está dada por:

$$\|\vec{u} \times \vec{v}\| = \left\| \begin{pmatrix} \vec{i} & \vec{j} & \vec{k} \\ a & b & 0 \\ a_1 & b_1 & 0 \end{pmatrix} \right\| := \|(ab_1 - a_1b)\,\vec{k}\,\| = |ab_1 - a_1b| \tag{5.1}$$

Usando la ecuación 5.1, se tiene una fórmula que expresa el área de un paralelogramo en términos de las coordenadas de los vectores que lo definen, es decir, con la notación de la función A que mide el área de un paralelogramo determinado por los vectores $\vec{u}$ y $\vec{v}$ se tiene:

$$A(\vec{u}, \vec{v}) = |ab_1 - a_1b|$$

A la expresión $ab_1 - a_1b$ le llamaremos el **determinante** de los vectores $\vec{u}$ y $\vec{v}$ o el determinante de la matriz $\begin{pmatrix} a & b \\ a_1 & b_1 \end{pmatrix}$, formada por las coordenadas de los vectores.

Ejercicio 5.1.2. *Calcule el área del paralelogramo determinado por cada par de vectores y haga una representación gráfica en cada caso.*

1. $\vec{u} = (2, 3)$, $\vec{v} = (4, 5)$.
2. $\vec{u} = (0, -2)$, $\vec{v} = (2, \pi)$.
3. $\vec{u} = (-1, 3)$, $\vec{v} = (0, \sqrt{2})$.
4. $\vec{u} = (2, 3)$ $\vec{v} = (2, 6)$.

Después de haber discutido cómo definir una función que a un par de vectores en el plano le asigna el área del paralelogramo determinado por éstos; formularemos la definición de una función, llamada **determinante**, que a cada n-ada de vectores, $\vec{u}_1$, $\vec{u}_2$, ..., $\vec{u}_n$, en $\mathbb{R}^n$ le asigne un número real. Las propiedades que debe satisfacer la función que definiremos, son las correspondientes que satisface la función área definida antes. En el caso de $\mathbb{R}^3$, la función se interpretará como aquella que mide el volumen del paralelepípedo determinado por tres vectores.

Definición 5.1.1. *Un determinante es una función* D *que asigna un número real a cada* n*-ada de vectores* $(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$ *de* $\mathbb{R}^n$*, y satisface las siguientes propiedades:*

1. $D(\vec{u}_1, \ldots, \vec{u}_{i-1}, \vec{u}_i + \vec{u}_j, \vec{u}_{i+1}, \ldots, \vec{u}_n) = D(\vec{u}_1, \ldots, \vec{u}_{i-1}, \vec{u}_i, \vec{u}_{i+1}, \ldots, \vec{u}_n)$, para todo $i = 1, \ldots, n$ y para todo $j \neq i$.

[1] La altura de un paralelogramo es la de uno de los triángulos, determinado por cualesquiera tres de sus vértices; el área de un paralelogramo se obtiene multiplicando la altura por la base.

2. $D(\vec{u}_1, \ldots., \vec{u}_{i-1}, \lambda \vec{u}_i, \vec{u}_{i+1}, \ldots, \vec{u}_n) = \lambda D(\vec{u}_1, \ldots., \vec{u}_{i-1}, \lambda \vec{u}_i, \vec{u}_{i+1}, \ldots, \vec{u}_n)$, para todo $\lambda \in \mathbb{R}$ y para todo $i = 1, \ldots, n$.
3. $D(\vec{e}_1, \ldots, \vec{e}_n) = 1$, en donde $\vec{e}_1, \ldots, \vec{e}_n$ son los vectores canónicos.

Observación 5.1.1. *Nótese, como en el caso del plano, que si los vectores* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$, *tienen coordenadas* $\vec{u}_1 = (a_{i1}, a_{i2}, \ldots, a_{in})$, *entonces se puede formar la matriz*

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

y abusando de terminología diremos que la función D *expresa el determinante de la matriz* A.

Observación 5.1.2. *Si para algún* $i = 1, 2, \ldots, n$ *se tiene* $\vec{u}_1 = 0$, *entonces de la parte 2 de la definición 5.1.1 se concluye que* $D(\vec{u}_1, \ldots, \vec{u}_n) = 0$.

En el siguiente teorema se establecen algunas de las propiedades básicas de la función determinante que permitirán calcular su valor. Esto se logrará haciendo uso de operaciones elementales en las filas de la matriz determinada por las coordenadas de los vectores.

Teorema 5.1.1. *Sea* D *como en la definición anterior, entonces satisface las siguientes propiedades.*

1. $D(\vec{u}_1, \ldots., \vec{u}_{i-1}, \vec{u}_i, \vec{u}_{i+1}, \ldots, \vec{u}_j, \ldots, \vec{u}_n) = -D(\vec{u}_1, \ldots., \vec{u}_{i-1}, \vec{u}_j, \vec{u}_{i+1}, \ldots, \vec{u}_i, \ldots, \vec{u}_n)$ es decir, D cambia de signo al intercambiar cualesquiera dos elementos.
2. $D(\vec{u}_1, \ldots., \vec{u}_i, \vec{u}_n) = D(\vec{u}_1, \ldots., \vec{u}_i + \sum_{j \neq i} \lambda_j \vec{u}_j, \vec{u}_{i+1}, \ldots, \vec{u}_n)$, para todo $i = 1, \ldots, n$.
3. D es multilineal, es decir,
   a) $D(\vec{u}_1, \ldots., \lambda \vec{u}_i, \ldots, \vec{u}_n) = \lambda D(\vec{u}_1, \ldots., \vec{u}_i, \vec{u}_{i+1}, \ldots, \vec{u}_n)$, para todo i y para todo $\lambda \in \mathbb{R}$ y.
   b) $D(\vec{u}_1, \ldots., \vec{u}_i + \vec{w}, \ldots, \vec{u}_n) = D(\vec{u}_1, \ldots., \vec{u}_i, \ldots, \vec{u}_n) + D(\vec{u}_1, \ldots., \vec{w}_i, \ldots, \vec{u}_n)$ para todo i y para cualquier par de vectores $\vec{u}, \vec{w} \in \mathbb{R}^n$.
4. $D(\vec{u}_1, \ldots, \vec{u}_n) = 0 \Leftrightarrow \{\vec{u}_1, \ldots, \vec{u}_n\}$ es linealmente dependiente.
5. Si para todo $i = 1, 2, \ldots, n$, $u_i = (0, \ldots, 0, a_{ii}, \ldots, a_{in})$ para todo $i = 1, 2, \ldots, n$, $u_i = (a_{i1}, \ldots, a_{ii}, 0, \ldots, 0)$, entonces $D(\vec{u}_1, \ldots., \vec{u}_i, \ldots, \vec{u}_n) = a_{11} a_{22} \cdots a_{nn}$.

Demostración. 1) Aplicando sucesivamente las propiedades (1) y (2) de la definición 5.1.1 se tienen las siguientes igualdades (identifique cuándo se usó cada propiedad):

$$\begin{aligned} D(\vec{u}_1, \ldots., \vec{u}_i, \ldots, \vec{u}_j, \ldots, \vec{u}_n) &= -D(\vec{u}_1, \ldots., \vec{u}_{i-1}, -\vec{u}_i, \vec{u}_{i+1}, \ldots, \vec{u}_j, \ldots, \vec{u}_n) \\ &= -D(\vec{u}_1, \ldots., \vec{u}_{i-1}, -\vec{u}_i, \vec{u}_{i+1}, \ldots, \vec{u}_j - \vec{u}_i, \ldots, \vec{u}_n) \\ &= D(\vec{u}_1, \ldots., \vec{u}_{i-1}, -\vec{u}_i, \vec{u}_{i+1}, \ldots, \vec{u}_i - \vec{u}_j, \ldots, \vec{u}_n) \end{aligned}$$

$$= D(\vec{u}_1, ..., \vec{u}_{i-1}, -\vec{u}_j, \vec{u}_{i+1}, ..., \vec{u}_i - \vec{u}_j, ..., \vec{u}_n)$$
$$= -D(\vec{u}_1, ..., \vec{u}_{i-1}, \vec{u}_j, \vec{u}_{i+1}, ..., \vec{u}_i - \vec{u}_j, ..., \vec{u}_n)$$
$$= -D(\vec{u}_1, ..., \vec{u}_{i-1}, \vec{u}_j, \vec{u}_{i+1}, ..., \vec{u}_i, ..., \vec{u}_n)$$

2) Procediendo por inducción, basta probar que

$$D(\vec{u}_1, ..., \vec{u}_{i-1}, \vec{u}_i + \lambda\vec{u}_j, \vec{u}_{i+1}, ..., \vec{u}_j, ..., \vec{u}_n) = D(\vec{u}_1, ..., \vec{u}_{i-1}, \vec{u}_i, \vec{u}_{i+1}, ..., \vec{u}_j, ..., \vec{u}_n)$$

para todo λ y para todo $j \neq i$. Si $\lambda = 0$, la igualdad anterior es obvia. Supongamos que $\lambda \neq 0$, entonces:

$$D(\vec{u}_1, ..., \vec{u}_i, ..., \vec{u}_j, \vec{u}_n) = \frac{1}{\lambda}D(\vec{u}_1, ..., \vec{u}_i, ..., \lambda\vec{u}_j, \vec{u}_n)$$
$$= \frac{1}{\lambda}D(\vec{u}_1, ..., \vec{u}_i + \lambda\vec{u}_j, ..., \lambda\vec{u}_j, \vec{u}_n)$$
$$= D(\vec{u}_1, ..., \vec{u}_i + \lambda\vec{u}_j, \vec{u}_j, ..., \vec{u}_n)$$

3) La parte a) es la propiedad 2 de la definición de la función D. Para demostrar la parte b), es suficiente demostrar que se cumple la igualdad:

$$D(\vec{u}_1 + \vec{v}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n) = D(\vec{u}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n) + D(\vec{v}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n)$$

Primer caso. Si los vectores $\{\vec{u}_2, ..., \vec{u}_n\}$ son linealmente dependientes, podemos suponer que $\vec{u}_2 = \sum_{i=3}^{n}\lambda_i\vec{u}_i$, entonces por la propiedad 2, ya probada, se tiene que:

$$D(\vec{u}_1 + \vec{v}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n) = D(\vec{u}_1 + \vec{v}_1, \vec{u}_2 - \sum_{i=3}^{n}\lambda_i\vec{u}_i, \vec{u}_3, ..., \vec{u}_i, ..., \vec{u}_n)$$
$$= D(\vec{u}_1 + \vec{v}_1, 0, ..., \vec{u}_i, ..., \vec{u}_n)$$
$$= 0$$

De igual forma se tiene que cada uno de los sumandos $D(\vec{u}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n)$ y $D(\vec{v}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n)$ es cero, obteniéndose la igualdad en este caso.

Segundo caso. Los vectores $\{\vec{u}_2, ..., \vec{u}_n\}$ son linealmente independientes, entonces existe $w \in \mathbb{R}^n$ tal que el conjunto $\{w, \vec{u}_2, ..., \vec{u}_n\}$ es una base de $\mathbb{R}^n$.

De esto se tiene $\vec{u}_1 = aw + \sum_{i=2}^{n}\lambda_i\vec{u}_i$ y $\vec{v}_1 = bw + \sum_{i=2}^{n}\mu_i\vec{u}_i$, para algunos escalares a y b.

$$D(\vec{u}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n) = D(a\vec{w} + \sum_{i=2}^{n}\lambda_i\vec{u}_i, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n)$$
$$= D(aw, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n)$$
$$= aD(\vec{w}, \vec{u}_2, ..., \vec{u}_n)$$

Un cálculo análogo muestra que $D(\vec{v}_1, \vec{u}_2, ..., \vec{u}_n) = bD = (\vec{w}, \vec{u}_2, ..., \vec{u}_n)$.

Por otro lado, $\vec{u}_i + \vec{v}_1 = (a+b)\ \vec{w} + \sum_{i=2}^{n}(\lambda_i + \mu_i)\vec{u}_i$. Procediendo como en los cálculos anteriores se obtiene que:

$$\begin{aligned} D(\vec{u}_1 + \vec{v}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n) &= (a+b)D(w, \vec{u}_2, ..., \vec{u}_n) \\ &= aD(w, \vec{u}_2, ..., \vec{u}_n) + bD(w, \vec{u}_2, ..., \vec{u}_n) \\ &= D(\vec{u}_1, \vec{u}_2, ..., \vec{u}_n) + D(\vec{v}_1, \vec{u}_2, ..., \vec{u}_n) \end{aligned}$$

probando lo que se deseaba.

4) Demostraremos primero que si dos vectores son iguales, digamos el i-ésimo y el j-ésimo, entonces $D(\vec{u}_1, ..., \vec{u}_n) = 0$.

Por la parte 1 del teorema y la hipótesis $\vec{u}_i = \vec{u}_j$ tenemos que:

$$\begin{aligned} D(\vec{u}_1, ..., \vec{u}_{i-1}, \vec{u}_i, \vec{u}_{i+1}, ..., \vec{u}_j, ..., \vec{u}_n) &= -D(\vec{u}_1, ..., \vec{u}_{i-1}, \vec{u}_j, \vec{u}_{i+1}, ..., \vec{u}_i, ..., \vec{u}_n) \\ &= -D(\vec{u}_1, ..., \vec{u}_{i-1}, \vec{u}_i, \vec{u}_{i+1}, ..., \vec{u}_j, ..., \vec{u}_n) \end{aligned}$$

Agrupando en un miembro se tiene $2D(\vec{u}_1, ..., \vec{u}_n) = 0$, y de esto $D(\vec{u}_1,..., \vec{u}_n) = 0$.

Supongamos que los vectores son linealmente dependientes. Sin perder generalidad, podemos suponer que $\vec{u}_1 = \sum_{i=2}^{n}\lambda_i\vec{u}_i$, entonces usando la parte 2, ya probada, y la observación 5.1.2 se tiene:

$$\begin{aligned} D(\vec{u}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n) &= D(\vec{u}_1 - \sum_{i=2}^{n}\lambda_i\vec{u}_i,\ \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n) \\ &= D(0, \vec{u}_2, ..., \vec{u}_n) \\ &= 0 \end{aligned}$$

Para demostrar la implicación recíproca, procederemos por contradicción, es decir, supongamos que los vectores son linealmente independientes y $D(\vec{u}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n) = 0$. Como los vectores $\vec{u}_1, ..., \vec{u}_n$ son linealmente independientes, entonces son una base de $\mathbb{R}^n$, por lo que los vectores canónicos se pueden expresar como combinación lineal de $\vec{u}_1, ..., \vec{u}_n$. De manera precisa, para cada $i = 1, 2, ..., n$, se tiene $\vec{e}_i = \sum_{j=1}^{n}a_{ji}\vec{u}_j$.

Usando esta representación de los vectores canónicos y la condición 3 en la definición del determinante se tiene que:

$$\begin{aligned} 1 &= D(\vec{e}_1, \vec{e}_2, ..., \vec{e}_n) \\ &= D\left(\sum_{j=1}^{n}a_{j1}\vec{u}_j, \sum_{j=1}^{n}a_{j2}\vec{u}_j, ..., \sum_{j=1}^{n}a_{jn}\vec{u}_j\right) \end{aligned}$$

Al desarrollar la última expresión en la ecuación anterior, usando que la función D es multilineal, propiedad 3 ya probada, se tiene que cada sumando contiene un factor de la forma $D(\vec{u}_{i1}, \vec{u}_{i2}, ..., \vec{u}_{in})$ el cual es cero (¿por qué?). De esto se concluye que $1 = 0$, contradicción que termina la prueba.

5) Abordaremos el caso en que $\vec{u}_i = (0, 0, ..., a_{ii}, ..., a_{in})$, el otro se trata de manera análoga.

Note que si $a_{ii} = 0$ para algún i, entonces los vectores son linealmente dependientes (verificar esto como ejercicio) y aplicando la propiedad 4 se tiene $D(\vec{u}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n) = 0 = a_{11} \cdots a_{ii} \cdots a_{nn}$.

Si para todo $i = 1, 2, ..., n$, $a_{ii} = 1$, aplicaciones sucesivas de la propiedad 2, ya probada, llevan a la ecuación $D(\vec{u}_1, \vec{u}_2, ..., \vec{u}_i, ..., \vec{u}_n) = D(\vec{e}_1, \vec{e}_2, ..., \vec{e}_n) = 1$, que es el resultado que se desea probar.

Supongamos que $a_{ii} \neq 0$ para todo $i = 1, 2, ..., n$. Definamos $\vec{u}_i = \dfrac{\vec{u}_i}{a_{ii}}$ para cada i.

Note que los vectores $\vec{u}_i'$ satisfacen que todas sus entradas, antes de la i-ésima, sean cero y la i-ésima uno, entonces por lo observado antes se tiene $D(\vec{u}_i', \vec{u}_2', ..., \vec{u}_n') = 1$. Por otro lado, $D(\vec{u}_1, \vec{u}_2,..., \vec{u}_n) = D(a_{11}\vec{u}_i', a_{22}\vec{u}_2', ..., a_{nn}\vec{u}_n')$. Aplicando la propiedad 2 de la definición 5.1.1 y lo comentado antes se concluye que $D(\vec{u}_1, \vec{u}_2,..., \vec{u}_n) = a_{11} \cdots a_{ii} \cdots a_{nn} D(\vec{u}_i', \vec{u}_2', ..., \vec{u}_n') = a_{11} \cdots a_{ii} \cdots a_{nn}$, como se afirmó.

Antes de proceder a calcular determinantes usando las propiedades del teorema anterior, es necesaria algo de terminología.

Definición 5.1.2. *Sea* $A = (a_{ij})$ *una matriz* $n \times n$. *Se dice que* A es:

1. Triangular superior, si $a_{ij} = 0$ para todo $i > j$.
2. Triangular inferior, si $a_{ij} = 0$ para todo $i < j$.
3. Triangular, si es triangular superior o inferior.
4. Diagonal si $a_{ij} = 0$ para todo $i \neq j$.

Recordemos también la relación que hay entre matrices elementales y operaciones elementales. Las matrices elementales se han denotado por E_{ij}, $E_i(r)$ y $E_{ij}(r)$. Las definiciones correspondientes se recuerdan abajo.

1. E_{ij} se obtiene al intercambiar las filas i y j de la matriz identidad.
2. $E_i(r)$ se obtiene de la identidad al multiplicar la fila i por $r \neq 0$.
3. $E_{ij}(r)$ se obtiene al multiplicar la fila i por r y sumarla a la fila j.

Introduciremos algo de notación para calcular determinantes. Si A es una matriz cuadrada cuyas filas se obtienen con las coordenadas de los vectores $\vec{u}_i, \vec{u}_2, ..., \vec{u}_n$, entonces usaremos la notación:

$$D(\vec{u}_i, \vec{u}_2, ..., \vec{u}_n) := |A| := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

Con esta notación, las propiedades del teorema 5.1.1 se interpretan abajo y serán de gran utilidad para calcular determinantes.

1. Intercambiar dos filas de una matriz, cambia el signo del determinante, propiedad 1.
2. Si la matriz B se obtiene de la matriz A, multiplicando una fila de A por un escalar r y sumándola a otra fila, entonces A y B tienen el mismo determinante, propiedad 2.
3. Si una matriz B se obtiene de la matriz A multiplicando una fila de A por un escalar r, se tiene $|B| = r|A|$, propiedad 3(a).
4. Las filas de A son linealmente dependientes si y sólo si su determinante es cero, propiedad 4.
5. El determinante de una matriz triangular es igual al producto de los elementos de su diagonal, propiedad 5.

Ejemplo 5.1.1. *Calcule el determinante de la matriz* $A = \begin{pmatrix} 5 & 2 \\ 3 & 3 \end{pmatrix}$

Discusión.

$$\begin{aligned}
\begin{vmatrix} 5 & 2 \\ 3 & 3 \end{vmatrix} &= \begin{vmatrix} 2 & -2 \\ 3 & 3 \end{vmatrix} \text{ Restando la fila dos a la uno.} \\
&= \begin{vmatrix} 2 & -1 \\ 1 & 4 \end{vmatrix} \text{ Restando la fila uno a la dos.} \\
&= \begin{vmatrix} 0 & -9 \\ 1 & 4 \end{vmatrix} \text{ Multiplicando la fila dos por } -2 \text{ y sumándola a la fila uno.} \\
&= \begin{vmatrix} 1 & 4 \\ 0 & -9 \end{vmatrix} \text{ Intercambiando filas.} \\
&= 9 \quad \text{Aplicando la propiedad del determinante a matrices triangulares.}
\end{aligned}$$

Ejemplo 5.1.2. *Calcule el determinante de la matriz* $A = \begin{bmatrix} 5 & 2 & 3 \\ 3 & 3 & -1 \\ 4 & 5 & 7 \end{bmatrix}$

Discusión.

$$\begin{aligned}
\begin{vmatrix} 5 & 2 & 3 \\ 3 & 3 & -1 \\ 4 & 5 & 7 \end{vmatrix} &= \begin{vmatrix} 1 & -3 & -4 \\ 3 & 3 & -1 \\ 4 & 5 & 7 \end{vmatrix} \text{ Restando la fila tres a la uno.} \\
&= \begin{vmatrix} 1 & -3 & -4 \\ 0 & 12 & 11 \\ 0 & 17 & 23 \end{vmatrix} \text{ ¿Qué operaciones se efectuaron?} \\
&= \begin{vmatrix} 1 & -3 & -4 \\ 0 & 12 & 11 \\ 0 & 5 & 12 \end{vmatrix} \text{ Restando la fila dos a la tres.} \\
&= \begin{vmatrix} 1 & -3 & -4 \\ 0 & 2 & -13 \\ 0 & 5 & 12 \end{vmatrix} \text{ ¿Cómo pasó del anterior a éste?}
\end{aligned}$$

$$= \begin{vmatrix} 1 & -3 & -4 \\ 0 & 2 & -13 \\ 0 & 1 & 38 \end{vmatrix} \text{ ¿Cómo se obtuvo este determinante?}$$

$$= \begin{vmatrix} 1 & -3 & -4 \\ 0 & 0 & -89 \\ 0 & 1 & 38 \end{vmatrix} \text{ ¿Por qué?}$$

$$= -\begin{vmatrix} 1 & -3 & -4 \\ 0 & 1 & 38 \\ 0 & 0 & -89 \end{vmatrix} \text{ Intercambiando la fila tres y la dos.}$$

$$= 89 \qquad \text{Por la propiedad en matrices triangulares.}$$

Ejercicio 5.1.3. *Calcule el determinante de la matriz* $A = \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}$, *suponiendo que* a, b *y* c *son diferentes.*

5.1.1. Propiedades del determinante

A continuación presentamos algunas de las propiedades fundamentales del determinante, en su demostración se usan los resultados contenidos en los teoremas 2.2.1 y 3.4.4, por lo que urgimos al lector a revisarlos.

Teorema 5.1.2. *Sean* A, B *y* E *matrices cuadradas del mismo orden, con* E *elemental. Entonces se cumple lo siguiente.*

1. *El determinante de* E *es no cero y* $|EA| = |E|\,|A|$.
2. *La matriz* A *tiene inversa* $\Leftrightarrow |A| \neq 0$ *y en este caso* $|A^{-1}| = \dfrac{1}{|A|}$.
3. *El determinante es multiplicativo, es decir,* $|AB| = |A|\,|B|$.
4. *Si* A^t *denota a la transpuesta de* A, *entonces* $|A| = |A^t|$*; es decir, la matriz* A *y su transpuesta tienen el mismo determinante.*
5. *Las filas o columnas de* A *son linealmente independientes* $\Leftrightarrow |A| \neq 0$.

Demostración.

1. La primera parte se obtiene directamente de la definición de matriz elemental y de la propiedad 3 de la definición 5.1.1; la segunda se obtiene de la interpretación que hicimos del teorema 5.1.1 y de la definición de matriz elemental.
2. Sabemos, teorema 3.4.4, que A tiene inversa $\Leftrightarrow$ A es equivalente por filas a la identidad; $\Leftrightarrow$ A es producto de matrices elementales. Como el determinante de una matriz elemental no es cero y si E y F son matrices elementales se tiene $|EF| = |E|\,|F|$, entonces: $|A| \neq 0 \Leftrightarrow A$ tiene inversa. Si A tiene inversa, entonces existen matrices elementales $F_1, F_2 \ldots, F_s$, tales que $F_1F_2 \cdots F_sA = I_n$. Usando la propiedad 1, que ya hemos probado, obtenemos:

$$|F_1F_2 \cdots F_sA| = |F_1|\,|F_2| \cdots |F_s|\,|A| = |I_n| = 1 \tag{5.2}$$

También sabemos que $A^{-1} = F_1F_2 \cdots F_s$. Aplicando el determinante a esta última ecuación y usando nuevamente la propiedad 1 y la ecuación 5.2, se concluye $|A^{-1}| = \dfrac{1}{|A|}$

3. Para demostrar esta propiedad, notemos que el producto de matrices tiene inversa $\Leftrightarrow$ cada factor la tiene. Por esta observación y la parte 2 se tiene que si alguna de A o B no tiene inversa, entonces tampoco AB, por lo que $|AB| = 0$ y como alguna de A o B no tiene inversa, entonces uno de $|A|$ o $|B|$ es cero, por lo que $|A|\,|B| = 0$, y de esto $|AB| = |A|\,|B|$. Si A y B son inversibles, entonces existen matrices elementales $F_1, \ldots, F_k$ y $H_1, \ldots, H_r$ tales que $A = F_1 \cdots F_k$ y $B = H_1 \cdots H_r$.

   De estas dos ecuaciones y la parte 1 se tiene:

$$|AB| = |F_1 \cdots F_k H_1 \cdots H_r| = |F_1| \cdots |F_k|\,|H_1| \cdots |H_r| = |A|\,|B|.$$

4. Esta parte la dejamos como ejercicio al lector y le sugerirnos un camino.
   - *a*) Defina operaciones elementales en las columnas de A.
   - *b*) Examine la relación que hay entre operaciones elementales en filas y operaciones elementales en columnas.
   - *c*) Observe qué ocurre al aplicar reducción en las columnas de A iniciando con la primer columna.
   - *d*) Si las operaciones elementales en las filas de A, que la llevan a forma escalonada reducida R, están representadas por las matrices elementales $F_1, \ldots, F_k$, es decir, $F_1 \cdots F_k A = R$, entonces $A^t F_k^t \cdots F_1^t = R^t$.
   - *e*) Una matriz triangular y su transpuesta tienen el mismo determinante, propiedad 5, teorema 5.1.1.
5. Para este punto le pedimos al lector que aplique la parte 4 del teorema 5.1.1, y como ya demostró la parte 4 comentada arriba, se obtiene la conclusión.

5.1.2. Existencia y unicidad del determinante

Una vez hecha la discusión de las propiedades del determinante, es muy importante justificar la existencia de una función que satisfaga dichas propiedades. La justificación de la existencia la proporciona el siguiente:

Teorema 5.1.3. *Existe una única función* $D : \underbrace{\mathrm{R}^n \times \cdots \times \mathrm{R}^n}_{n \text{ veces}} \to \mathbb{R}$ *que satisface las propiedades de la definición 5.1.1.*

Demostración. Si $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ son los vectores canónicos, definamos la función D que satisfaga las siguientes dos condiciones.

1. $D(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) := 1$
2. $D(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_i, \ldots, \vec{e}_j, \ldots, \vec{e}_n) := -D(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_j, \ldots, \vec{e}_i, \ldots, \vec{e}_n)$ para cualesquiera dos índices diferentes i, j.

Las condiciones anteriores implican directamente $D(\vec{e}_{j1}, \vec{e}_{j2}, \ldots, \vec{e}_{jn}) = \pm 1$ para cualquier posible elección de los índices $\{j_1, j_2, \ldots, j_n\}$. El signo se puede determinar notando que cualquier elección de los índices $\{j_1, j_2, \ldots, j_n\}$ se puede llevar al orden natural intercambiando dos cada vez.

Como D debe satisfacer la propiedad 2 de la definición 5.1.1 y cada vector $\vec{u}$ es combinación lineal de la base canónica, entonces proponemos:

$$D(\vec{u}_i, \vec{u}_2, ..., \vec{u}_n) := \sum_{j_1,...,j_n} a_{1j1} a_{2\,j2} \cdots a_{n\,jn} D(e_{j1}, e_{j2}, ..., e_{jn}) \qquad (5.3)$$

en donde la suma se toma sobre las $n!$ posibles elecciones de los índices $\{j_1, j_2, ..., j_n\}$ y $\vec{u}_i = (a_{i1}, a_{i2}, ..., a_{in})$.

Dejamos sin demostración las propiedades 1 y 3 de la definición 5.1.1 y la unicidad de la función. Le pedimos al lector que proporcione los argumentos para terminar la prueba del teorema.

Observación 5.1.3. *Note que la definición del determinante expresada en la ecuación 5.3, tiene poca utilidad desde el punto de vista práctico, pues es difícil calcular todas las elecciones de los índices* $\{j_1, j_2, ..., j_n\}$ *y los signos que aparecen; sin embargo, para el caso de matrices* 2×2, *se tiene*:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

5.2. Regla de Cramer, menores y cofactores

Uno de los problemas centrales en álgebra lineal es resolver un sistema de ecuaciones lineales, el cual se puede representar en la forma $AX = B$. Para tal efecto existen varios métodos, uno de los cuales ha sido discutido, nos referimos al método de reducción de Gauss-Jordan. Este método ofrece ciertas ventajas, sobre todo si tratamos de encontrar soluciones numéricas. Por otro lado, cuando los coeficientes del sistema son parámetros que dependen del tiempo, como es el caso de algunos modelos económicos lineales, y las soluciones se han de analizar tomando en cuenta su variación en el tiempo, el método de Gauss-Jordan puede no ofrecer la mejor solución. En este caso, un método que permita expresar las soluciones por medio de una ecuación pudiese ser más adecuado; éste es el caso del método conocido como *Regla de Cramer*. Para desarrollarlo, requeriremos calcular determinantes, por lo que se llama método de *menores y cofactores*. Requerimos un poco de terminología. Dada una matriz cuadrada $A = (a_{ij})$ y dos enteros $1 \leq i, j \leq n$, se define el menor M_{ij} como el determinante de la matriz que se obtiene de A al omitir la fila i-ésima y la columna j-ésima. Así mismo, el cofactor C_{ij} del elemento a_{ij} se expresa mediante $C_{ij} := (-1)^{i+j} M_{ij}$. También se define la matriz *adjunta clásica* de A, denotada Adj(A) como:

$$\text{Adj}(A) := \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

Por ejemplo, si $A = \begin{pmatrix} a_{11} & b_{22} \\ a_{21} & b_{22} \end{pmatrix}$, entonces $M_{11} = a_{22}$, $M_{12} = a_{21}$, $M_{21} = a_{12}$, $M_{22} = a_{11}$, $C_{11} = a_{22}$, $C_{12} = -M_{12} = -a_{21}$, $C_{21} = -M_{21} = -a_{12}$ y $C_{22} = M_{22} = a_{11}$. De esto se tiene que la adjunta clásica de A es:

$$\text{Adj}(A) = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

De la expresión que define a la adjunta de A y aplicando la observación 5.1.3 se obtiene:

$$A(\text{Adj}(A)) = \begin{pmatrix} a_{11} & b_{12} \\ a_{21} & b_{22} \end{pmatrix}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = \begin{pmatrix} |A| & 0 \\ 0 & |A| \end{pmatrix} = |A|I_2$$

Con la notación y terminología establecidas anteriormente, se tienen los resultados que siguen, los cuales se presentan sin demostración. El primero proporciona un método para evaluar determinantes, llamado desarrollo de determinantes por *menores y cofactores*, mientras que el segundo es la base del método conocido como *Regla de Cramer* para resolver sistemas de ecuaciones lineales.

Teorema 5.2.1. *Sea* A *una matriz* $n \times n$, C_{ij} *como se definió antes, entonces*

$$|A| = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \tag{5.4}$$

para cualesquiera i, j *entre* 1 *y* n.

Teorema 5.2.2. (Regla de Cramer). *Dada una matriz cuadrada* A, *se tiene* $\text{Adj(A)}\,\text{A} = \text{A} \cdot \text{Adj(A)} = |\text{A}|\text{I}_\text{n}$. *En particular, si* A *admite inversa, ésta se representa como* $\text{A}^{-1} = \dfrac{\text{Adj(A)}}{|\text{A}|}$.

Recordemos que si A es una matriz cuadrada que tiene inversa y B es un vector columna con entradas b_i, entonces el sistema $AX = B$ tiene solución única. En efecto, multiplicando en la ecuación anterior por A^{-1}, se tiene $A^{-1}AX = A^{-1}B$; de esto y la definición de Adj(A) obtenemos:

$$X = A^{-1}B = \frac{\text{Adj}(A)}{|A|}\, B = \frac{1}{|A|}\text{Adj}(A)B = \frac{1}{|A|}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} B$$

Desarrollando el producto anterior y escribiendo las componentes de X se llega a:

$$x_i = \frac{b_1C_{1i} + b_2C_{2i} + \cdots + b_nC_{ni}}{|A|}, \text{ para todo } i = 1, 2, \ldots, n \tag{5.5}$$

De acuerdo con (5.4), la ecuación anterior se puede expresar como:

$$x_i = \frac{|A_i|}{|A|} \tag{5.6}$$

en donde la matriz A_i se obtiene de A sustituyendo la columna i-ésima de A por el vector columna B.

Observación 5.2.1. *A la ecuación 5.6 se le conoce usualmente como Regla de Cramer para resolver sistemas de ecuaciones. Nosotros le hemos llamado Regla de Cramer al teorema que sirve de base para obtener este resultado.*

5.3. Determinantes y ecuaciones diferenciales

En esta sección presentamos una demostración del criterio que usa el wronskiano para determinar si un conjunto de funciones, n-veces diferenciables en un intervalo, es linealmente independiente. Estamos dando por hecho que el conjunto de las funciones con valores reales, definidas en un intervalo I, son n-veces diferenciables es un espacio vectorial.

Consideremos la ecuación diferencial:

$$y' + p(x)y = 0 \tag{5.7}$$

en donde la función $p(x)$ es continua en un intervalo I. Sea $P(x)$ una primitiva de $p(x)$ en I, es decir, $P'(x) = p(x)$ para todo $x \in I$. Multiplicando ambos miembros de (5.7) por $e^{P(x)}$ y aplicando la regla para derivar un producto de funciones se tiene:

$$(e^{P(x)}y)' = e^{P(x)}y' + p(x)e^{P(x)}y = e^{P(x)}(y' + p(x)y) = 0 \tag{5.8}$$

concluyendo que $e^{P(x)}y = k$, para alguna constante k. La última ecuación equivale a:

$$y = ke^{-P(x)} \tag{5.9}$$

La función $P(x)$ puede ser definida por $P(x) := \int_{x_0}^{x} p(t)dt$, para algún $x_0 \in I$. Con esta elección de $P(x)$, la constante k se obtiene al evaluar y en x_0, es decir, de la ecuación 5.9 se llega a: $y(x_0) = ke^{-P(x_0)} = ke^0 = k$.

Lema 5.3.1. *Sea* $\{f_{ij}\}$ *un conjunto con* n^2 *funciones diferenciables en el intervalo* I. *Para cada* $x \in I$ *se define:*

$$G(x) := \begin{vmatrix} f_{11}(x) & \cdots & f_{1n}(x) \\ f_{21}(x) & \cdots & f_{1n}(x) \\ \vdots & & \vdots \\ f_{n1}(x) & \cdots & f_{nn}(x) \end{vmatrix}$$

Entonces G(x) *es diferenciable y*

$$G'(x) := \begin{vmatrix} f'_{11}(x) & \cdots & f'_{1n}(x) \\ f_{21}(x) & \cdots & f_{2n}(x) \\ \vdots & & \vdots \\ f_{n1}(x) & \cdots & f_{nn}(x) \end{vmatrix} + \begin{vmatrix} f_{11}(x) & \cdots & f_{1n}(x) \\ f'_{21}(x) & \cdots & f'_{2n}(x) \\ \vdots & & \vdots \\ f_{n1}(x) & \cdots & f_{nn}(x) \end{vmatrix} + \cdots + \begin{vmatrix} f_{11}(x) & \cdots & f_{1n}(x) \\ f_{21}(x) & \cdots & f_{2n}(x) \\ \vdots & & \vdots \\ f'_{n1}(x) & \cdots & f'_{nn}(x) \end{vmatrix}$$

Demostración. Recordemos que el determinante de una matriz se expresa mediante la ecuación:

$$G(x) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) f_{1\sigma(1)} \cdots f_{n\sigma(n)} \tag{5.10}$$

(5.3), página 126. Derivando ambos miembros de esta ecuación y usando la regla de Leibniz para calcular la derivada del producto de n funciones se llega a:

$$G'(x) = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)(f_{1\sigma(1)} \cdots f_{n\sigma(n)})'$$

$$= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)\left(\sum_{i=1}^{n} f'_{i\sigma(i)}(f_{1\sigma(1)} \cdots \hat{f}_{i\sigma(i)} \cdots f_{n\sigma(n)})\right)$$

$$= \sum_{i=1}^{n}\left(\sum_{\sigma=S_n} \operatorname{sign}(\sigma) f'_{i\sigma(i)} f_{1\sigma(1)} \cdots \hat{f}_{i\sigma(i)} \cdots f_{n\sigma(n)}\right)$$

en donde la notación $\hat{f}_{i\sigma(i)}$ significa que ese factor se ha omitido. Observemos que cada suma:

$$\sum_{\sigma\in S_n} \operatorname{sign}(\sigma) f'_{i\sigma(i)} f_{1\sigma(1)} \cdots \hat{f}_{i\sigma(i)} \cdots f_{n\sigma(n)}$$

es el determinante de la matriz A_i que se ha obtenido de $G(x)$ sustituyendo la fila i-ésima por $(f'_{i1}, ..., f'_{in})$, es decir, $G'(x) = \sum_{i=1}^{n} |A_i|$, como se afirmó en el lema.

Sean $g_1(x)$, $g_2(x)$, ..., $g_n(x)$ funciones n-veces diferenciables en un intervalo I. La j-ésima derivada de g_i la denotaremos por $g_i^{(j)}$ y convenimos, como es costumbre, que $g_i^{(0)} = g_i$. Haciendo $f_{ij}(x) := g_i^{(j)}$ para $i = 1, 2, ..., n$ y $j = 0, 1, ..., n - 1$, la función $G(x)$ se expresa como:

$$G(x) := \begin{vmatrix} g_1(x) & \cdots & g_n(x) \\ g'_1(x) & \cdots & g'_n(x) \\ \vdots & & \vdots \\ g_1^{(n-1)}(x) & \cdots & g_n^{(n-1)}(x) \end{vmatrix}$$

comúnmente llamada el wronskiano de $g_1(x)$, $g_2(x)$, ..., $gn(x)$ y denotada por $W(g_1, g_2, ..., g_n)(x)$.

Lema 5.3.2. *Si* $g_1(x)$, $g_2(x)$, ..., $g_n(x)$ *son funciones* n*-veces diferenciables en un intervalo* I, *entonces la función* $W(g_1, g_2, ..., g_n)(x)$ *es diferenciable y se tiene*:

$$W'(g_1, g_2, ..., g_n)(x) = \begin{vmatrix} g_1(x) & \cdots & g_n(x) \\ g'_1(x) & \cdots & g'_n(x) \\ \vdots & & \vdots \\ g_1^{(n)}(x) & \cdots & g_n^{(n)}(x) \end{vmatrix}$$

Demostración. Por el lema 5.3.1 se tiene que $W'(g_1, g_2, ..., g_n)(x) = G'(x) = \sum_{i=1}^{n} |A_i|$, con A_i la matriz que se obtiene de $W(g_1, g_2, ..., g_n)$ sustituyendo la fila i-ésima por su derivada. Entonces la matriz A_i tiene una fila repetida para $i = 1, 2, ..., (n - 1)$, por lo que su determinante es cero. De esto se tiene que $W'(g_1, g_2, ..., g_n)(x) = G'(x) = |A_n|$, que es la expresión buscada.

Teorema 5.3.1 *Sean* $a_n(x)$, $a_{n-1}(x)$, ..., $a_0(x)$ *funciones continuas en un intervalo cerrado* I. *Adicionalmente, supongamos que* $a_n(x) \neq 0$ *para todo* $x \in$ I. *Consideremos la ecuación diferencial lineal homogénea de orden* n:

$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x) = 0. \tag{5.11}$$

Si $g_1, g_2, ..., g_n$ *son soluciones de (5.11), entonces* $\{g_1, g_2, ..., g_n\}$ *es linealmente independiente* $\Leftrightarrow$ *existe* $x_0 \in$ I *tal que* $W(g_1, g_2, ..., g_n)(x_0) \neq 0$.

Demostración. Sean $c_1, c_2, ..., c_n$ escalares tales que $h = c_1g_1 + c_2g_2 + \cdots + c_ng_n$ es la función cero en I. De esto se tiene que $h^{(j)} = c_1g_1^{(j)} + c_2g_2^{(j)} + \cdots + c_ng_n^{(j)}$ también es la función cero en I para todo $j = 0, 1, 2 ..., (n-1)$. Si A tiene por determinante a: $W(g_1, g_2, ..., g_n)$, entonces $(c_1, c_2, ..., c_n)$ es solución no trivial de $AX = 0 \Leftrightarrow W(g_1, g_2 ..., g_n)(x) = 0$ para todo $x \in I$. Para terminar la demostración probaremos que $W(g_1, g_2 ..., g_n)(x) \neq 0$ para todo $x \in I \Leftrightarrow$ existe $x_0 \in I$ tal que $W(g_1, g_2 ..., g_n)(x_0) \neq 0$. Esto último será consecuencia de que $W(g_1, g_2 ..., g_n)(x)$ es solución de (5.7), con $p(x) = \dfrac{a_{n-1}(x)}{a_n(x)}$.

Del lema 5.3.2, se tiene:

$$W'(g_1, g_2, ..., g_n)(x) = \begin{vmatrix} g_1(x) & \cdots & g_n(x) \\ g_1'(x) & \cdots & g_n'(x) \\ \vdots & & \vdots \\ g_1^{(n)}(x) & \cdots & g_n^{(n)}(x) \end{vmatrix} \tag{5.12}$$

Por hipótesis, cada g_i es solución de (5.11), entonces:

$$g_i^{(n)}(x) = \frac{1}{a_n(x)}(a_{n-1}g_i^{(n-1)} + \cdots + a_0g_i)(x).$$

Sustituyendo estas ecuaciones en (5.12), usando la multilinealidad del determinante y la propiedad de que cuando hay dos filas iguales el determinante vale cero, obtenemos:

$$W'(g_1, g_2, ..., g_n)(x) = -\frac{a_{n-1}}{a_n}\begin{vmatrix} g_1(x) & \cdots & g_n(x) \\ g_1'(x) & \cdots & g_n'(x) \\ \vdots & & \vdots \\ g_1^{(n-1)}(x) & \cdots & g_n^{(n-1)}(x) \end{vmatrix}$$

probando que $W(g_1, g_2 ..., g_n)(x)$ es solución de $y' + \dfrac{a_{n-1}}{a_n}y = 0$. De manera más precisa, por la discusión al inicio de esta sección sabemos que:

$$W(g_1, g_2, ..., g_n)(x) = W(g_1, g_2, ..., g_n)(x_0)e^{\int_{x_0}^{x} p(t)\,dt} \tag{5.13}$$

para algún $x_0 \in$ I y $p(t) = \dfrac{a_{n-1}(t)}{a_n(t)}$. Como la función exponencial no es cero en todo punto, entonces $W(g_1, g_2, ..., g_n)(x) = 0 \Leftrightarrow W(g_1, g_2, ..., g_n)(x_0) = 0$.

Observación 5.3.1. *Note que en la demostración del teorema solamente se usaron las condiciones de continuidad para* a_n y a_{n-1}. *No se usó condición alguna sobre los restantes coeficientes.*

5.4. Ejercicios

1. Calcule los siguientes determinantes usando menores y cofactores.

a) $\begin{vmatrix} 2 & -2 & 2 \\ 0 & 1 & 3 \\ -1 & 2 & 2 \end{vmatrix}$, $\begin{vmatrix} a & 0 & b \\ 0 & a & 1 \\ -1 & 0 & 1 \end{vmatrix}$, $\begin{vmatrix} 1 & -2 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \end{vmatrix}$

2. Sean (a, b) y (c, d) dos puntos diferentes en $\mathbb{R}^2$. Demuestre que la ecuación de la recta que los contiene se puede expresar mediante la ecuación:

$$\begin{vmatrix} x & y & 1 \\ a & b & 1 \\ c & d & 1 \end{vmatrix} = 0$$

3. Demuestre que el área de un triángulo que tiene por vértices a los puntos (a, b), (c, d) y (e, f) es el valor absoluto de:

$$\frac{1}{2}\begin{vmatrix} a & b & 1 \\ c & d & 1 \\ e & f & 1 \end{vmatrix}$$

4. Demuestre que el volumen de un tetraedro[2] con vértices en los puntos (a_1, b_1, c_1), (a_2, b_2, c_2), (a_3, b_3, c_3) y (a_4, b_4, c_4) está dado por el valor absoluto de:

$$\frac{1}{6}\begin{vmatrix} a_1 & b_1 & c_1 & 1 \\ a_2 & b_2 & c_2 & 1 \\ a_3 & b_3 & c_3 & 1 \\ a_4 & b_4 & c_4 & 1 \end{vmatrix}$$

5. Calcule el volumen de los paralelepípedos determinados por los vectores que se indican en cada caso. Ilustre con figuras.

a) $(1, 2, 3)$, $(3, -4, 6)$ y $(1, 1, 1)$

b) $(1, 0, 3)$, $(0, 1, 6)$ y $(1, 1, 1)$

c) $(1, \pi, 3)$, $(\sqrt{2}, -4, 6)$ y $(l, 1, 1)$

6. Calcule el área de los paralelogramos determinados por los vectores que se especifican. Ilustre con figuras.

a) $(1, 2)$, $(3, -4)$

b) $(1, 0)$, $(0, 1)$

c) $(\pi, 6)$, $(\sqrt{2}, -4)$

[2] Un tetraedro es un poliedro de cuatro caras, cada una de las cuales es un triángulo equilátero.

7. Sean U, V y W vectores en $\mathbb{R}^3$. Demuestre que el volumen del paralelepípedo determinado por estos tres vectores, ver la figura 5.2, es $|\langle U \times V, W\rangle|$, en donde $U \times V$ denota al producto cruz de U y V. De esta consideración geométrica debiera ser claro que $\langle U \times V, W\rangle = \langle U, V \times W\rangle = D(U, V, W)$, con D la función determinante.

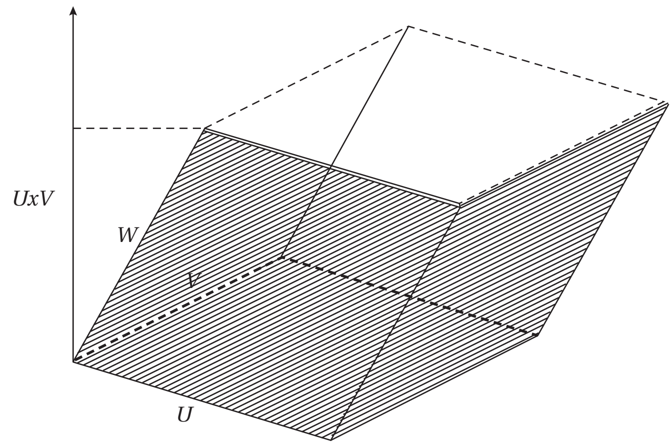


Figura 5.2. Volumen de un paralelepípedo.

8. Resuelva los siguientes sistemas de ecuaciones en términos de t, usando la Regla de Cramer. No olvide considerar el caso en que la matriz de coeficientes es singular.

   a) $2(t + 1)x + 3y = 1$
   $3x + 4(t^2 + 1)y = 3$

   b) $2tx + 3y + z = 1$
   $2x + \mathrm{t}y + z = 0$
   $-x + 3y + t^2 z = -1$

   c) $2tx + 3y = t^2$
   $3x + 4t^2 y = 3t$

   d) $2x + 3y + z + w = 1$
   $2x - y + z - 3w = 0$
   $-x + 3y + z - w = -1$
   $x - y - z - w = 2$

9. Enuncie todos los teoremas y definiciones que se han discutido en este capítulo.

10. Si A y B son matrices similares, demuestre que tienen el mismo determinante.

11. Sea A una matriz $n \times n$, demuestre que hay a lo más n valores de $c \in \mathbb{R}$ tales que la matriz $B = A - cI_n$ es singular. Encuentre un ejemplo para el cual la matriz B es singular exactamente para n valores de c.

12. Sean (a_1, a_2, a_3), (b_1, b_2, b_3) y (c_1, c_2, c_3) tres puntos no colineales en $\mathbb{R}^3$. Demuestre que la ecuación del plano que los contiene está dada por:

$$\begin{vmatrix} x & y & z & 1 \\ a_1 & a_2 & a_3 & 1 \\ b_1 & b_2 & b_3 & 1 \\ c_1 & c_2 & c_3 & 1 \end{vmatrix} = 0$$

13. Sea A una matriz $n \times n$. Supongamos que sus columnas, consideradas como elementos de $\mathbb{R}^n$, son ortogonales. ¿Cuál es el valor de $|A|$?

14. Sea A una matriz cuadrada diagonal por bloques, ¿cómo calcular el determinante de A en términos de los bloques de A?

15. Sean A, B, C y D matrices cuadradas del mismo orden que conmutan. Adicionalmente, suponga que A y C tienen inversa. Pongamos $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$. Demuestre que $|M| = |AD - BC|$.

16. Sean $x_1, ..., x_n \in \mathbb{R}$, $A = (a_{ij})$ matriz $n \times n$, con $a_{ij} = x_i^{j-1}$. Demuestre que:

$$|A| = (-1)^{\frac{n(n-1)}{2}} \prod_{1 \leq i < j \leq n} (x_i - x_j)$$

A este determinante se le llama *determinante de Vandermonde.*

17. Use el resultado del ejercicio 16 para demostrar que:

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2^2 & \cdots & 2^{n-1} \\ 1 & 3 & 3^2 & \cdots & 3^{n-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & n & n^2 & \cdots & n^{n-1} \end{vmatrix} = 1!2!3! \cdots (n-1)!$$

18. Recordemos que una matriz es antisimétrica si $A^t = -A$. Sea A una matriz $n \times n$, antisimétrica con n impar. Demuestre que A es singular.

19. Demuestre que:

$$\begin{vmatrix} 1+x_1 & x_2 & x_3 & \cdots & x_n \\ x_1 & 1+x_2 & x_3 & \cdots & x_n \\ x_1 & x_2 & 1+x_3 & \cdots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & x_3 & \cdots & 1+x_n \end{vmatrix} = 1 + x_1 + x_2 + \cdots + x_n$$

20. Una matriz cuadrada A, se dice **idempotente** si $A^2 = A$. ¿Cuáles son los posibles valores de $|A|$ para matrices idempotentes?

21. Sea A una matriz 2×2 y suponga que $A^2 = 0$. Demuestre que para todo escalar c, $|cI - A| = c^2$. ¿Se cumple el mismo resultado si A es $n \times n$ y $A^k = 0$ para algún entero positivo k?

22. Calcule el siguiente determinante:

$$\begin{vmatrix} a & a+h & a+h_2 & \cdots & \cdots & a+(n-1)h \\ -a & a & 0 & \cdots & \cdots & 0 \\ 0 & -a & a & \cdots & \cdots & 0 \\ 0 & 0 & -a & a & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & 0 \\ 0 & 0 & 0 & \cdots & -a & a \end{vmatrix}$$

23. Sin efectuar los cálculos, demuestre que las siguientes igualdades tienen lugar.

a) $\begin{vmatrix} 1 & a & bc \\ 1 & b & ca \\ 1 & c & ab \end{vmatrix} = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}$

b) $\begin{vmatrix} a-b & b-c & c-a \\ 1 & 1 & 1 \\ a & b & c \end{vmatrix} = \begin{vmatrix} a & b & c \\ 1 & 1 & 1 \\ b & c & a \end{vmatrix}$

24. Suponga que la matriz siguiente es de orden $n \times n$, $n > 2$. ¿Para qué valores de n es singular?

$$\begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

25. Sea A una matriz $n \times n$. Demuestre que $|\text{Adj}(A)| = |A|^{n-1}$. Si A es no singular use la igualdad anterior para demostrar que $\text{Adj}(\text{Adj}(A)) = |A|^{n-2}A$, si $n \geq 2$.

26. Sea A una matriz $n \times n$ tal que sus columnas $A_1, A_2, \ldots, A_n$ vistas como elementos de $\mathbb{R}^n$ tienen norma uno y son ortogonales. Demuestre que $|A| = \pm 1$.

27. Sea A como en el ejercicio anterior y considere vectores $Y_1, Y_2, \ldots, Y_n$ en $\mathbb{R}^n$. Si B es la matriz que tiene a los elementos $AY_1, AY_2, \ldots, AY_n$ como columnas, ¿cuál es el determinante de B en términos de los elementos $Y_1, Y_2, \ldots, Y_n$? ¿Cuál es la interpretación geométrica de este resultado?

28. Sea A una matriz $n \times n$ tal que sus columnas $A_1, A_2, \ldots, A_n$ se pueden considerar elementos de $\mathbb{R}^n$. Demuestre que el valor absoluto del determinante de A es $\leq$ que $\|A_1\|\,\|A_2\| \cdots \|A_n\|$, en donde $\|A_i\|$ denota la norma de A_i. Sugerencia: reduzca el problema al caso en que las columnas son linealmente independientes y tienen norma uno. La desigualdad anterior se llama *Desigualdad de Hadamard*.

Capítulo 6

Eigenteoría: estructura de operadores

En diversos problemas de economía, física, ingeniería, entre otras áreas; la formulación matemática se hace por medio de un sistema de funciones lineales, y esto se puede representar por medio de un producto de matrices AX, en donde A es $n \times n$ y X es un vector columna. Las mismas condiciones del problema imponen hipótesis sobre X, una de ésas es que no sea el vector cero. Por ejemplo, X puede representar las cantidades que se invierten para producir bienes de consumo. Con estas suposiciones es natural preguntarse, ¿existe una X de tal forma que la producción sea proporcional a la inversión? En lenguaje matemático esto se traduce a ¿existe una $X \neq 0$ tal que:

$$AX = \lambda X? \tag{6.1}$$

La forma usual de abordar la pregunta anterior es notar que la ecuación 6.1 tiene una solución, con $X \neq 0$, $\Leftrightarrow$ la matriz $A - \lambda I$ es singular, y esta última condición se puede formular a través del determinante de la matriz $A - \lambda I$, dando lugar al concepto de polinomio característico. El usar determinantes para discutir el problema planteado tiene algunas desventajas, entre éstas se encuentra la necesidad de hacer una presentación más o menos exhaustiva de las propiedades del determinante. También se tiene que el polinomio característico no proporciona condiciones necesarias y suficientes para que una matriz sea diagonalizable o triangulable.

Con el enfoque abordado en este capítulo, iniciamos definiendo el polinomio mínimo de una matriz u operador, con lo cual estamos en posibilidades de establecer propiedades relacionadas con la estructura de la matriz u operador. Por ejemplo, uno de los primeros resultados que se demuestran es el teorema de la Descomposición Primaria, teorema 6.1.3, página 140, el cual juega un papel central en la discusión posterior. Otro aspecto relevante de este enfoque es que el polinomio mínimo se obtiene usando un algoritmo que solamente requiere de operaciones elementales en las filas y columnas de la matriz $A - \lambda I$, lo cual, desde el punto de vista computacional, resulta más eficiente que calcular determinantes. Asimismo, a partir del polinomio mínimo, podemos definir el polinomio característico y dar una prueba muy corta del importante teorema de Cayley-Hamilton.

6.1. Definiciones y resultados básicos

En esta sección discutimos la existencia del polinomio mínimo, así como sus propiedades fundamentales.

6.1.1. El polinomio mínimo

Iniciamos la discusión presentando algunos conceptos relacionados con la existencia de valores característicos.

El problema de valores característicos se aborda estudiando las soluciones de la ecuación 6.1. Esto también se puede formular de la siguiente manera. ¿Para qué valores de λ es singular la matriz $A - \lambda I$?

Tomando la ruta vía determinantes, la condición para que $A - \lambda I$ sea singular se formula en términos de su determinante, esto es, se debe cumplir que $f(\lambda) := |A - \lambda I| = 0$. Entonces, la existencia de valores característicos de A equivale a la existencia de raíces de $f(x)$, y esto a la vez equivale a que $f(x)$ admita factores lineales, lo cual no siempre queda garantizado.

En la ruta que tomamos, partimos de la misma condición: la matriz $A - \lambda I$ es singular. La existencia de valores característicos toma otro camino, en donde se hace uso sistemático de aspectos relacionados con la matriz u operador y el espacio en donde actúa.

Si X es un vector no cero en $\mathbb{R}^n$, el conjunto $\{X, AX, A^2X, \ldots, A^nX\}$ es linealmente dependiente, por tener $n + 1$ elementos, por lo que existen escalares $a_0, a_1, \ldots, a_n$, no todos cero, tales que:

$$a_0X + a_1AX + a_2A^2X + \cdots + a_nA^nX = 0 \tag{6.2}$$

Definiendo el polinomio $h(x) := a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$, la ecuación 6.2 se puede escribir como $h(A)X = 0$, en donde hemos definido $h(A) := a_0I + a_1A + a_2A^2 + \cdots + a_nA^n$. Notemos que el polinomio $h(x)$ depende de X.

Si para algún $X \neq 0$, $h(x)$ se factoriza como producto de polinomios lineales, digamos:

$$h(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n) \tag{6.3}$$

entonces, al ser $h(A) = (A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_{nI})$ singular ($h(A)X = 0$), necesariamente uno de los factores de $h(A)$ debe ser singular, es decir, A tiene un valor propio.

Como señalamos antes, el polinomio $h(x)$ depende del vector X; lo que haremos en seguida es mostrar que existe un polinomio de mínimo grado $m(x)$, que solamente depende de A y $m(A) = 0$.

La discusión que sigue muestra dos formas de probar la existencia de $m(x)$, con las mismas condiciones de antes, salvo que usaremos cierta notación para $h(x)$. Si $\{X_1, X_2, \ldots, X_n\}$ es una base de $\mathbb{R}^n$, al polinomio $h(x)$ asociado a cada X_i lo denotaremos por $h_i(x)$, entonces el polinomio:

$$g(x) = h_1(x)h_2(x) \cdots h_n(x) \tag{6.4}$$

tiene grado $\leq n^2$ y satisface que $g(A)$ es la matriz cero (justifíquelo, ejercicio 17, página 168). Una vez garantizada la existencia de un polinomio de grado positivo que tiene por raíz a A, se toma el de menor grado.

Lo desarrollado antes se hizo para matrices, sugerimos al lector hacer la traducción para operadores. En lo que sigue mostramos la existencia de $m(x)$ para operadores siguiendo otra idea.

Sea V un espacio vectorial de dimensión finita, sabemos que el espacio de los operadores, $\mathcal{L}(V; V)$ también tiene dimensión finita, más precisamente, si V tiene dimensión

n, entonces dim $(\mathcal{L}(V; V)) = n^2$. Sea T un operador en V, entonces el conjunto $\{I, T, T^2, \ldots, T^{n^2}\}$ es linealmente dependiente, por lo que existen escalares, $a_0, a_1 \ldots, a_{n^2}$ no todos cero tales que $a_0 I + a_1 T + \cdots + a_{n^2} T^{n^2} = 0$. Esto se traduce diciendo que existe un polinomio no cero, $f(x)$ tal que $f(T) = 0$. En este caso decimos que T *es un cero del polinomio* f(x). De todos los polinomios que tienen a T como cero, existe uno de mínimo grado y mónico.

En el siguiente teorema se discuten algunas de las propiedades importantes de este polinomio.

Teorema 6.1.1. *Sea* V *un espacio vectorial de dimensión* n, $\mathrm{T} : \mathrm{V} \to \mathrm{V}$ *un operador, entonces existe un único polinomio mónico* m(x) *tal que*:

1. $\mathrm{m(T)} = 0$.
2. *Si* f(x) *es otro polinomio tal que* $\mathrm{f(T)} = 0$, *entonces* m(x) *divide a* f(x).

Demostración. La existencia de $m(x)$ ha sido discutida anteriormente; la unicidad será consecuencia de la parte dos del teorema, por lo que probaremos eso. Sea $f(x)$ un polinomio tal que $f(T) = 0$. Por el algoritmo de la división, existen polinomios $q(x)$ y $r(x)$ tales que:

$$f(x) = m(x)q(x) + r(x), \text{ con } r(x) = 0 \text{ o } \deg(r(x)) < \deg(m(x)) \tag{6.5}$$

Evaluando la ecuación 6.5 en T y tomando en cuenta que T es raíz de $f(x)$ y de $m(x)$ se tiene $r(T) = 0$. Como $m(x)$ es el polinomio de mínimo grado que tiene a T por raíz, necesariamente se debe tener que $r(x) = 0$, probando lo afirmado. La unicidad de $m(x)$ se obtiene del siguiente argumento. Si $m(x)$ y $g(x)$ son polinomios mónicos de mínimo grado que tienen a T por raíz, entonces, por lo probado antes se dividen mutuamente; ahora el hecho de ser mónicos implica que son iguales.

Definición 6.1.1. *Al polinomio del teorema 6.7.1 le llamaremos el polinomio mínimo de* T *y lo denotaremos* $\mathrm{m_T(x)}$.

Antes de presentar propiedades generales e importantes del polinomio mínimo de un operador, discutimos el caso en que el espacio tiene dimensión dos. La generalización del ejemplo siguiente es el teorema 6.1.2.

Ejemplo 6.1.1. *Sea* V *un espacio vectorial de dimensión* 2, T *un operador en* V. *Entonces el polinomio mínimo de* T *tiene grado a lo más dos*.

Caso I. Para todo $\alpha \in V$ los vectores α y $T(\alpha)$ son linealmente dependientes. La condición equivale a que para cada α existe un escalar c_α tal que $T(\alpha) = c_\alpha \alpha$. Mostraremos que el escalar c_α no depende de α. Para esto es suficiente demostrar que si α y β son linealmente independientes, entonces $c_\alpha = c_\beta$.

Se tiene $T(\alpha + \beta) = c(\alpha + \beta)$. Por otro lado, $T(\alpha + \beta) = T(\alpha) + T(\beta) = c_\alpha \alpha + c_\beta \beta$. De estas ecuaciones se concluye que $c\alpha + c\beta = c_\alpha \alpha + c_\beta \beta$. Como α y β son linealmente independientes, entonces $c = c_\alpha = c_\beta$.

Resumiendo, hemos probado que existe una constante c tal que $T(\alpha) = c\alpha$ para todo $\alpha \in V$; esto equivale a $(T - cI)(\alpha) = 0$ para todo $\alpha \in V$, es decir, $T - cI$ es el operador cero. Entonces el polinomio mínimo de T es $x - c$.

Caso II. Existe un $\alpha \in V$ tal que α y $T(\alpha)$ son linealmente independientes. Como V tiene dimensión dos, α, $T(\alpha)$, $T^2(\alpha)$ son linealmente dependientes, entonces existen escalares a y b tales que $T^2(\alpha) = aT(\alpha) + b\alpha$.

Afirmamos que el polinomio mínimo de T es $x^2 - ax - b$. Esto equivale a probar que $S = T^2 - aT - bI$ es el operador cero, para lo cual basta probar que $S(\alpha) = S(T(\alpha)) = 0$.

De la ecuación $T^2(\alpha) = aT(\alpha) + b\alpha$ se obtiene $S(\alpha) = T^2(\alpha) - aT(\alpha) - b\alpha - 0$.

También tenemos que $S(T(\alpha)) = (T^2 - aT - bI)(T(\alpha)) = T(T^2(\alpha) - aT(\alpha) - b\alpha) = T(0) = 0$, probando lo afirmado. De los casos discutidos se concluye que $m_T(x)$ tiene grado a lo más dos, como afirmamos.

Ejemplo 6.1.2. *Encuentre el polinomio mínimo del operador* $\mathrm{T} : \mathbb{R}^2 \to \mathbb{R}^2$ *dado por* $\mathrm{T(x, y) = (2x - 3x + y)}$.

Se tiene $T(1, 0) = (2, 3)$ por lo que $(1, 0)$ y $T(1, 0) = (2, 3)$ son linealmente independientes. También tenemos $T^2(1, 0) = T(T(1, 0)) = T(2, 3) = (1, 9)$. Como $(1, 0)$ y $T(1, 0) = (2, 3)$ son linealmente independientes, deseamos encontrar escalares a y b tales que $T^2(1, 0) = aT(1, 0) + b(1, 0)$. Resolviendo esta ecuación se encuentra que $a = 3$ y $b = -5$, por lo que el polinomio mínimo de T es $m_T(x) = \mathrm{x}^2 - 3x + 5$.

El método del ejemplo anterior funciona para operadores en espacios de dimensión mayor, sin embargo puede fallar en algunos casos. Veamos un ejemplo en dimensión tres.

Ejemplo 6.1.3. *Encuentre el polinomio mínimo del operador* $\mathrm{T} : \mathbb{R}^3 \to \mathbb{R}^3$ *dado por* $\mathrm{T(x, y, z) = (2x + y, 2y + z, x + y + 2z)}$.

Se tiene $T(1, 0, 0) = (2, 0, 1)$, $T^2(1, 0, 0) = T(2, 0, 1) = (4, 1, 4)$ y los vectores $(1, 0, 0)$, $(2, 0, 1)$, $(4, 1, 4)$ son linealmente independientes, pues la matriz cuyas columnas son los vectores dados es equivalente por filas a la identidad. También tenemos $T^3(1, 0, 0) = (9, 6, 13)$. Este vector es combinación lineal de $(1, 0, 0)$, $T(1, 0, 0) = (2, 0, 1)$, $T^2(1, 0, 0) = (4, 1, 4)$, más precisamente,

$$T^3(1, 0, 0) = (9, 6, 13) = 7(1, 0, 0) - 11(2, 0, 1) + 6(4, 1, 4)$$

De esta ecuación se tiene que el polinomio mínimo de T es $m_T(x) = x^3 - 6x^2 + 11x - 7$, pues $m_T(T)$ se anula en la base $\{(1, 0, 0), T(1, 0, 0), T^2(1, 0, 0)\}$.

Como se mencionó al inicio del capítulo, el concepto de valor y vector característico de un operador o matriz es fundamental en varios problemas de álgebra lineal y teoría de operadores. Estos conceptos se pueden formular de la siguiente forma.

Definición 6.1.2. *Sea* V *un espacio de dimensión finita,* $\mathrm{T : V \to V}$ *un operador y* λ *un escalar. Se dice que* λ *es un valor característico o valor propio de* T, *si* $\mathrm{T} - \lambda\mathrm{I}$ *es singular. A los elementos del núcleo de* $\mathrm{T} - \lambda\mathrm{I}$, *excluyendo el cero, se les llama vectores característicos de* T *asociados a* λ.

Más adelante, en este mismi capítulo, aparece otra definición de valor característico, invitamos al lector a verificar que aquélla y la anterior son equivalentes.

Un concepto que generaliza al de vector característico y es de gran utilidad es el de *subespacio invariante*.

Definición 6.1.3. *Sea* V *un espacio vectorial,* $\mathrm{T : V \to V}$ *un operador. El subespacio* U *de* V *se llama* T-*invariante, si* $\mathrm{T(U) \subseteq U}$.

Si U es un subespacio propio de V e invariante bajo el operador T, entonces la restricción de T a U es un operador actuando en un espacio de dimensión menor, lo cual en muchos casos resulta de utilidad. Una ventaja aún mayor resulta cuando U tiene un complemento que también es invariante, es decir, existe un subespacio U_1 que es T-invariante, tal que $V = U \oplus U_1$. Estas ideas las extenderemos en lo que sigue.

De la argumentación en la demostración del teorema 6.1.1 se tiene que el grado de $m_T(x)$ es $\leq n^2$; el teorema que sigue, después del lema, mejora esa cota de manera significativa.

Lema 6.1.1. *Sea* V *un espacio de dimensión* $n \geq 1$, T *un operador en* V y W *un subespacio* T-*invariante de dimensión* $k \geq 1$. *Entonces existe un polinomio* $g(x)$ *de grado a lo más* $n - k$ *tal que* $g(T)(V) \subseteq W$.

Demostración. Pongamos $\dim(W) = n - r$, con $0 \leq r < n$. Aplicaremos inducción sobre r. Si $r = 0$, entonces $W = V$ y cualquier polinomio constante funciona. Si $r = 1$, entonces W tiene dimensión $n - 1$. Sean $\{\alpha_1, \alpha_2, ..., \alpha_{n-1}\}$ una base de W y $\beta \in V$ tal que $\{\alpha_1, \alpha_2, ..., \alpha_{n-1}, \beta\}$ es base de V. Entonces $T(\beta) = \gamma + c\beta$, con $\gamma \in W$ y $c \in \mathbb{R}$. Como W es T-invariante, entonces $(T - cI)(\alpha_i) \in W$ para todo $i = 1, 2, ..., n - 1$. De esto se tiene $(T - cI)(\alpha) \in W$ para todo $\alpha \in V$, es decir, el polinomio $x - c$ satisface lo requerido.

Supongamos $r > 1$ y sea $\alpha \in V \setminus W$. Se tiene que existe un polinomio $h(x)$ de grado positivo tal que $h(T)(\alpha) \in W$, por ejemplo el polinomio mínimo de T cumple esto. Sea $g(x)$ el polinomio de mínimo grado y mónico que satisface $g(T)(\alpha) \in W$. Si el grado de $g(x)$ es l, consideremos los elementos: α, $T(\alpha)$, $T^2(\alpha)$, ..., $T^{l-1}(\alpha)$. Por la minimalidad del grado de $g(x)$ se tiene que estos elementos son linealmente independientes, de hecho se puede decir algo más; el subespacio $W_1 = \mathcal{L}\{\alpha, T(\alpha), T^2(\alpha), ..., T^{l-1}(\alpha)\}$ forma una suma directa con W y el subespacio $W + W_1$ es T-invariante de dimensión mayor que la de W. Por hipótesis de inducción, existe un polinomio $g_1(x)$ de grado $\leq n - \dim(W + W_1) = n - \dim(W) - \dim(W_1)$ tal que $g_1(T)(V) \subseteq W + W_1$.

Por otro lado, se tiene $g(T)(W_1) \subseteq W$; esto y la hipótesis sobre la invariancia de W bajo T implican:

$$g(T)g_1(T)(V) \subseteq g(T)(W) + g(T)(W_1) \subseteq W$$

Como W_1 tiene dimensión l y $\deg(g_1(x)) \leq n - \dim(W + W_1) = n - \dim(W) - \dim(W_1)$, concluimos que $\deg(g(x)g_1(x)) \leq n - \dim(W)$, terminando la demostración.

Teorema 6.1.2. *Sea* V *un espacio vectorial de dimensión* n, $T : V \to V$ *un operador y* $m_T(x)$ *su polinomio mínimo, entonces* $\deg(m_T(x)) \leq n$.

Demostración. Aplicaremos inducción sobre n. Si $n = 1$, el resultado es claro, pues si α es un generador de V y T es un operador, entonces $T(\alpha) = a\alpha$, por lo que $T - aI$ es el operador cero en V, es decir, $x - a$ es el polinomio mínimo de T. Otra forma de argumentar este caso es notando que $n = n^2$ y aplicando que en general, el polinomio mínimo tiene grado a lo más n^2.

Supongamos que $n > 1$ y el resultado cierto para todos los espacios vectoriales de dimensión menor que n. Sea $\alpha \in V \setminus \{0\}$, entonces el conjunto $\{\alpha, T(\alpha), ..., T^n(\alpha)\}$ es linealmente dependiente, por lo que existen escalares $a_0, a_1 \ldots, a_n$, no todos cero, tales que $a_0\alpha + a_1T(\alpha) + \cdots + a_nT^n(\alpha) = 0$. Sea $T_1 := a_0I + a_1T + \cdots + a_nT^n$ y $U = N_{T_1}$ núcleo de T_1.

Si $U = V$, entonces T_1 es el operador cero, por lo que el polinomio mínimo de T tiene grado $\leq n$.

Supongamos $U \neq V$, como $\alpha \in U$, se tiene que U es un subespacio propio de V, es decir, no cero y tampoco V. También se tiene que U es un subespacio invariante, pues si $\beta \in U$ entonces $T(\beta) \in U$.

Sea $g(x)$ el polinomio mínimo de T restringido a U, es decir, $g(T)(U)$ es el subespacio cero. Por hipótesis de inducción, $\deg(g(x)) \leq \dim(U)$. Aplicando el lema anterior se tiene que existe un polinomio $h(x)$ tal que $h(T)(V) \subseteq U$ y $\deg(h(x)) \leq n - \dim(U)$. De todo esto se concluye que el operador $g(T)h(T)$ es el operador cero y $\deg(g(x)h(x)) \leq n$, por lo que el polinomio mínimo de T tiene grado a lo más n, probando el teorema.

El siguiente teorema jugará un papel crucial en lo que sigue de la discusión. Antes de continuar recordaremos lo que significa la suma directa de más de dos subespacios.

Dado un espacio vectorial V y una colección $W_1, W_2, \ldots, W_k$ de subespacios de V se define la suma de estos subespacios como:

$$W_1 + W_2 + \cdots + W_k = \{\alpha_1 + \alpha_2 + \cdots + \alpha_k \mid \alpha_i \in W_i, \text{ para } i = 1, 2, \ldots, k\}$$

Si se cumple la condición: $\alpha_1 + \alpha_2 + \cdots + \alpha_k = 0$, implica que cada $\alpha_i = 0$, diremos que los subespacios forman una suma directa y en este caso usaremos la notación:

$$W_1 \oplus W_2 \oplus \cdots \oplus W_k := \{\alpha_1 + \alpha_2 + \cdots + \alpha_k \mid \alpha_i \in W_i, \text{ para } i = 1, 2, \ldots, k\}$$

Por ejemplo, en $\mathbb{R}^3$ los subespacios $W_1 = \{(t, t, -t) : t \in \mathbb{R}\}$, $W_2 = \{(u, u/2, u) : u \in \mathbb{R}\}$ y $W_3 = \{(s, -s, s) : s \in \mathbb{R}\}$ forman una suma directa, pues dados $(t, t, -t) \in W_1$, $(u, u/2, u) \in W_2$ y $(s, -s, s) \in W_3$ tales que $(t, t, -t) + (u, u/2, u) + (s, -s, s) = (0, 0, 0)$, entonces se tiene el sistema de ecuaciones:

$$\begin{aligned} t + u + s &= 0 \\ t + u/2 - s &= 0 \\ -t + u + s &= 0 \end{aligned}$$

y claramente la única solución de este sistema es $t = u = s = 0$. Note que cada uno de los subespacios del ejemplo tiene dimensión uno. Discuta la siguiente pregunta, dado $n > 2$, ¿cómo deben ser los subespacios de dimensión uno en $\mathbb{R}^n$ para que formen suma directa?

Una manera alternativa de formular el concepto de suma directa es: los subespacios $W_1, W_2, \ldots, W_k$ forman una suma directa $\Leftrightarrow$ para cada $i = 1, 2, \ldots, k$ se cumple:

$$W_i \cap \left(\sum_{j \neq i} W_j \right) = \{0\}$$

Ejercicio 6.1.1. *Demuestre que las dos formulaciones de suma directa son equivalentes.*

Teorema 6.1.3. (Descomposición primaria). *Sea* V *un espacio vectorial de dimensión* n, T *un operador en* V *y* $\mathrm{m_T(x)} = \mathrm{p_1^{r_1}(x)\, p_2^{r_2}(x) \cdots p_k^{r_k}(x)}$ *la factorización del polinomio mínimo de* T *en factores irreducibles. Si* $\mathrm{W_i}$ *es el núcleo del operador* $\mathrm{S_i} = \mathrm{p}_i^{r_i}(\mathrm{T})$, *entonces*:

1. $\mathrm{V} = \mathrm{W_1} \oplus \mathrm{W_2} \oplus \cdots \oplus \mathrm{W_k}$.
2. *Cada* $\mathrm{W_1}$ *es* T-*invariante*.
3. *El operador* $\mathrm{S_i}$ *tiene por polinomio mínimo a* $p_i^{r_i}(\mathrm{x})$.

Demostración. Para cada i, sea $f_i(x) = \prod_{j \neq i} p_j^{r_j}(x)$. De la definición de los polinomios $f_1(x), f_2(x), \ldots, f_k(x)$ se tiene que son primos relativos, por lo que existen polinomios $g_1(x)$, $g_2(x), \ldots, g_k(x)$ tales que:

$$f_1(x)g_1(x) + \cdots + f_k(x)g_k(x) = 1 \qquad 1(6.6)$$

Definamos $T_i := f_i(T)g_i(T)$. Se verifica directamente lo siguiente:

1. $T_1 + T_2 + \cdots + T_k = I$, el operador identidad.
2. Si $i \neq j$, entonces $T_iT_j = 0$.
3. Para todo $i = 1, 2, \ldots, k$ se tiene $T_i^2 = T_i$.

De la parte 1 mencionada antes se tiene $T_1(V) + T_2(V) + \cdots + T_k(V) = V$. La primera conclusión del teorema se tendrá si probamos que $T_i(V) = W_i$ y que los subespacios $T_i(V)$ forman una suma directa.

Sea $\beta \in T_i(V)$, es decir, $\beta = T_i(\alpha) = f_i(T)g_i(T)(\alpha)$ para algún $\alpha \in V$, entonces $S_i(\beta) = p_i^{r_i}(T)(\beta) = p_i^{r_i}(T)f_i(T)g_i(T)(\alpha) = m(T)g_i(T)(\alpha) = 0$, probando que $\beta \in W_i$.

Recíprocamente, sea $\alpha \in W_i$; como $T_1 + T_2 + \cdots + T_k = I$, entonces $\alpha = T_1(\alpha) + T_2(\alpha) + \cdots + T_k(a)$. También se tiene que $p_i^{r_i}(x)$ divide a $f_j(x)$ para todo $j \neq i$, entonces $T_j(\alpha) = f_j(T)g_j(T)(\alpha) = 0$, de lo que se tiene $\alpha = T_i(\alpha) \in T_i(V)$.

Ahora mostraremos que $T_i(V) \cap \left| \sum_{j\neq i} T_j(V) \right| = \{0\}$. Sea $\alpha \in T_i(V) \cap \left(\sum_{j\neq i} T_j(V) \right)$, entonces, $\alpha = T_i(\gamma) = \left(\sum_{j\neq i} T_j(\alpha_j) \right)$. Aplicando T_i a esta ecuación y usando la propiedad 2 enunciada antes se tiene $T_i(\alpha) = T_i^2(\gamma) = \sum_{j\neq i} T_i T_j(\alpha_j) = 0$, ahora usando la propiedad 3 se tiene $0 = T_i(\alpha) = T_i^2(\gamma) = T_i(\gamma) = \alpha$, como se afirmó, finalizando la prueba de la primera parte del teorema.

La segunda parte del teorema se tiene de manera directa usando que T conmuta con $p_i^{r_i}(T)$.

Para probar la parte tres, note que $p_i^{r_i}(T)$ es el operador cero en W_i por lo que el polinomio mínimo del operador inducido por T en W_i, divide a $p_i^{r_i}(x)$. Por otro lado, si $h(x)$ es cualquier otro polinomio tal que $h(E_i) = 0$, con E_i la restricción de T a W_i, entonces $h(T)f_i(T)$ es el operador cero, por lo que $m(x) = p_i^{r_i}(x)f_i(x)$ divide a $h(x)f_i(x)$, es decir $p_i^{r_i}(x)$ divide a $h(x)$, probando el teorema.

Corolario 6.1.1. *Sea* V *un espacio vectorial de dimensión finita sobre cualquier campo,*[1] T *un operador en* V. *Entonces* T *es singular* $\Leftrightarrow$ *el cero es raíz del polinomio mínimo de* T.

Demostración. Supongamos que el cero es raíz del polinomio mínimo de T, entonces x^e exactamente divide al polinomio mínimo de T, para algún $e \geq 1$. Por el teorema anterior, el núcleo de T^e es un sumando no cero de V, por lo que el núcleo de T es no cero.

Recíprocamente, sea W el núcleo de T; claramente W es T-invariante de dimensión positiva, entonces el polinomio mínimo de T restringido a W es x y divide al polinomio mínimo de T, es decir, el cero es raíz del polinomio mínimo de T.

Corolario 6.1.2. *Sea* V *un espacio vectorial de dimensión finita sobre cualquier campo*, T *un operador en* V. *Entonces* T *tiene un subespacio invariante de dimensión uno* $\Leftrightarrow$ *el polinomio mínimo de* T *tiene un factor lineal.*

Demostración. Sea W un subespacio invariante de dimensión uno, entonces $T - cI$ es el operador cero en W, para algún c; de esto se tiene que $x - c$ es el polinomio mínimo de T restringido a W, por lo que $x - c$ divide al polinomio mínimo de T.

Recíprocamente, si $m_T(x)$ tiene un factor lineal, digamos $x - c$, sea $(x - c)^e$ la potencia exacta de $x - c$ que divide al polinomio mínimo. Por el teorema anterior, el núcleo de $(T - cI)^e$ es un subespacio invariante no cero de V, entonces el núcleo de $T - cI$ es no cero y cualquier elemento no cero de éste genera un subespacio invariante de dimensión uno.

[1] Un campo es un conjunto no vacío en el que están definidas una suma y un producto que satisfacen las mismas propiedades que la suma y producto de los números reales.

Lema 6.1.2. *Sean* V *un espacio de dimensión finita,* $T : V \to V$ *un operador y* W *un subespacio* T-*invariante. Dado* $\alpha \in V$, *existe un único polinomio de mínimo grado y mónico,* g(x), *tal que* $g(T)\alpha \in W$. *Si* f(x) *es otro polinomio tal que* $f(T)\alpha \in W$, *entonces* g(x) *divide a* f(x).

Demostración. Como el polinomio mínimo de T satisface $m_T(T)\alpha = 0 \in W$, entonces podemos tomar un polinomio mónico de mínimo grado y mónico que satisface la condición requerida. Denotemos por $g(x)$ a este polinomio y sea $f(x)$ otro polinomio que satisface $f(T)\alpha \in W$. Por el algoritmo de la división, existen polinomios $q(x)$ y $r(x)$ tales que:

$$f(x) = q(x)g(x) + r(x), \quad \text{con} \quad r(x) = 0 \quad \text{o} \quad \deg(r(x)) < \deg(g(x))$$

De la ecuación anterior se tiene $f(T)\alpha = q(T)g(T)\alpha + r(T)\alpha$. Como $g(T)\alpha \in W$ y W es T-invariante, entonces $q(T)g(T)\alpha \in W$, por lo que $r(T)\alpha \in W$. La elección de $g(x)$ y la condición $\deg(r(x)) < \deg(g(x))$ implican que $r(x) = 0$, probando que $g(x)$ divide a $f(x)$. Si $g(x)$ y $g_1(x)$ son dos polinomios mónicos de mínimo grado tales que $g(T)\alpha \in W$ y $g_1(T)\alpha \in W$, por lo probado antes, estos polinomios se dividen mutuamente, entonces son iguales.

Ejercicio 6.1.2. *¿Qué relación hay entre los lemas 6.1.1 y 6.1.2?*

Definición 6.1.4. *Sean* V, T, W *y* α *como en el lema anterior, al polinomio mónico de menor grado del que se habla allí se le llama el* T-*anulador de* α *respecto a* W. *En caso que* $W = \{0\}$, *simplemente le llamaremos el* T-*anulador de* α *y lo denotaremos por* ann (α, x).

Lema 6.1.3. *Sean* V *un espacio de dimensión finita,* $T : V \to V$ *un operador,* $\beta \in V \setminus \{0\}$. *Entonces la dimensión del subespacio generado por* $\{\beta, T(\beta), T^2(\beta), \ldots,\}$ *es igual al grado de* ann (β, x).

Demostración. Sea ann $(\beta, x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1x + a_0$; demostraremos que una base del espacio generado por $\{\beta, T(\beta), T^2(\beta), \ldots,\}$ es $\{\beta, T(\beta), T^2(\beta), \ldots, T^{k-1}(\alpha)\}$. La condición ann (β, T) $(\beta) = 0$, implica que $T^k(\beta)$ pertenece al subespacio generado por $\{\beta, T(\beta), T^2(\beta), \ldots, T^{k-1}(\alpha)\}$. Por inducción se demuestra que para cada $i = 1, 2, 3\ldots$ el elemento $T^{k+i}(\beta)$ pertenece al espacio generado por $\{\beta, T(\beta), T^2(\beta), \ldots, T^{k-1}(\alpha)\}$ mostrando con esto que $\{\beta, T(\beta), T^2(\beta), \ldots, T^{k-1}(\alpha)\}$ genera el mismo subespacio que $\{\beta, T(\beta), T^2(\beta), \ldots,\}$.

Falta mostrar que $\{\beta, T(\beta), T^2(\beta), \ldots, T^{k-1}(\alpha)\}$ es linealmente independiente. Esto se tiene de la minimalidad del grado de ann (β, x).

Definición 6.1.5. *Sean* V *un espacio de dimensión finita,* $T : V \to V$ *un operador y* $\beta \in V$. *Al subespacio generado por* $\{\beta, T(\beta), T^2(\beta), \ldots, \}$ *le llamaremos el subespacio* T-*cíclico generado por* β *y le denotaremos por* $C(\beta, T)$.

Teorema 6.1.4. *Sea* V *un espacio de dimensión* n *sobre cualquier campo* (*los reales, los complejos o cualquier otro*), T *un operador en* V *con polinomio mínimo* $m_T(x) = p^l(x)$, p(x) *irreducible de grado* r. *Entonces* r *divide a* n.

Demostración. Aplicaremos inducción sobre la dimensión de V. Si V tiene dimensión uno, entonces el polinomio mínimo de T es irreducible de grado uno y la conclusión del teorema se tiene. Podemos suponer que V tiene dimensión mayor que uno y que el resultado es cierto para todos los espacios vectoriales de dimensión menor que la de V.

Sea $\alpha \in V \setminus \{0\}$; por el lema 6.1.2, ann (α, x) divide a $p^l(x)$, por lo que ann $(\alpha, x) = p^{l_\alpha}(x)$, con $l_\alpha \leq l$. Si $l_\alpha < l$ para cada α, entonces en particular $l_{\alpha_i} < l$ para cada $\alpha_i \in \{\alpha_1, \alpha_2, \ldots, \alpha_n\} = B$, base de V. Sea $m = \max\{l_{\alpha_1}, l_{\alpha_2}, \ldots, l_{\alpha_n}\} < l$, entonces se tiene $p^m(T)(\alpha_i) = 0$, para todo $i = 1, 2, \dots n$. De aquí obtenemos que $p^m(T)$ anula a cualquier combinación lineal de los elementos $\alpha_1, \alpha_2, \dots, a_n$; de esto se concluye que $p^m(T) = 0$, contradiciendo que el polinomio mínimo de T es $p^l(x)$. Del argumento anterior se tiene que existe un $\alpha \in V$ tal que ann $(\alpha, x) = p^l(x)$. Para este α, sea $W = C(\alpha, T)$. Por el lema 6.1.3, la dimensión de W es rl.

Si $W = V$, hemos concluido la demostración, en caso contrario sea U un subespacio T-invariante maximal que contiene a W y diferente de V. Tenemos que el polinomio mínimo de T restringido U es una potencia de $p(x)$. Aplicando la hipótesis de inducción a U se tiene que r divide a su dimensión.

Sea $\alpha \in V \setminus U$, entonces el polinomio de mínimo grado, $h(x)$ que satisface $h(T)(\alpha) \in U$, divide a $p^l(x)$, lema 6.1.2, por lo que $h(x) = p^k(x)$, para algún k. La condición sobre el grado de $h(x)$ implica que $W_1 := \mathcal{L}\,\{\alpha, T(\alpha), \dots, T^{kr-1}(\alpha)\}$ tiene dimensión kr; esta misma condición implica que U y W_1 forman suma directa; como ambos son T-invariantes, entonces $U + W_1$ también es T-invariante y la maximalidad de U implica que $V = U \oplus W_1$. De esta ecuación se tiene $\dim(V) = \dim(U) + \dim(W_1)$. Como r es factor de cada uno de estos sumandos, entonces r es factor de $\dim(V)$, probando el teorema.

Corolario 6.1.3. *Sea* V *un espacio vectorial de dimensión finita sobre cualquier campo,* T *un operador en* V, $\mathrm{m_T(x)} = \mathrm{p_1^{r_1}(x)\, p_2^{r_2}(x) \cdots p_k^{r_k}(x)}$ *la factorización del polinomio mínimo de* T *y* $\mathrm{V} = \mathrm{W_1 \oplus W_2 \oplus \cdots \oplus W_k}$ *la descomposición garantizada en el teorema 6.1.3. Entonces el grado de* $\mathrm{p_i(x)}$ *divide a dim* $\mathrm{W_i}$.

Demostración. Cada W_i es T-invariante y el polinomio mínimo de T restringido a W_i es $p_i^{r_i}(x)$. La conclusión se tiene del teorema anterior.

Definición 6.1.6. *Sean* V *un espacio de dimensión* n *sobre cualquier campo,* $\mathrm{T : V \to V}$ *un operador*, $\mathrm{m_T(x)} = \mathrm{p_1^{r_1}(x)\, p_2^{r_2}(x) \cdots p_k^{r_k}(x)}$ *la factorización del polinomio mínimo de* T *como producto de irreducibles. Definimos el polinomio característico de* T *como* $\mathrm{f_T(x)} :=$ $(-1)^n \mathrm{p_1^{d_1}(x)\, p_2^{d_2}(x) \cdots p_k^{d_k}(x)}$, *en donde* $\mathrm{d_i} = \dfrac{\dim \mathrm{W_i}}{\deg \mathrm{p_i(x)}}$.

Observación 6.1.1. *El grado de* $\mathrm{f_T(x)}$ *es igual a la dimensión de* V.

Teorema 6.1.5 (Cayley-Hamilton). *Sea* V *un espacio de dimensión* n, T *un operador en* V. *Entonces* T *es un cero de su polinomio característico.*

Demostración. De los teoremas 6.1.2 y 6.1.3 se tiene que el polinomio mínimo de T divide al polinomio característico. La conclusión sigue inmediatamente.

Teorema 6.1.6. *Sea* V *un espacio vectorial sobre los reales de dimensión* $\mathrm{n} \geq 1$, T *un operador en* V. *Entonces,* T *tiene un subespacio invariante de dimensión uno o dos*.

Demostración. Sea $\alpha \in V \setminus \{0\}$, entonces el conjunto $\{\alpha, T(\alpha), \dots, T^n(\alpha)\}$ es linealmente dependiente, es decir, existen escalares $a_0, a_1 \dots, a_n$, no todos cero, tales que $a_0\alpha + a_1T(\alpha) + \cdots + a_nT^n(\alpha) = 0$. Sea $p(x) = a_0 + a_1x + \cdots + a_nx^n$. Note que la hipótesis sobre α garantiza que $p(x)$ tiene grado positivo. Por el teorema fundamental del álgebra,[2] $p(x)$ se factoriza como producto de factores lineales y cuadráticos, es decir, $p(x) = p_1(x)p_2(x) \cdots p_t(x)$, con cada $p_i(x)$ lineal o cuadrático.

Como $p(T)$ es singular, entonces al menos uno de los $p_i(T)$ también es singular. Si $p_i(x) = ax + b$, con $a \neq 0$, entonces existe un $\beta \neq 0$ tal que $aT(\beta) + b\beta = 0$, de esto concluimos que el subespacio generado por β es invariante. Si $p_i(x) = ax^2 + bx + c$, con

[2] El teorema fundamental del álgebra establece que todo polinomio de grado positivo con coeficientes en los números complejos admite por lo menos una raíz. Un corolario de este teorema establece que los polinomios irreducibles con coeficientes reales son lineales o cuadráticos.

$a \neq 0$, entonces existe un $\beta \neq 0$ tal que $aT^2(\beta) + bT(\beta) + c\beta = 0$. De esto se tiene que el subespacio generado por $\{\beta, T(\beta)\}$ es invariante y tiene dimensión uno o dos, terminando la prueba del teorema. En las siguientes líneas probaremos algo más preciso: el subespacio generado por $\{\beta, T(\beta)\}$ tiene dimensión dos, pues si este subespacio tuviese dimensión uno, entonces el polinomio mínimo de T restringido a éste sería lineal y dividiría a $p_i(x) = ax^2 + bx + c$, lo cual es imposible, pues $p_i(x)$ es irreducible.

Teorema 6.1.7. *Sea* V *un espacio de dimensión impar sobre los reales,* T *un operador en* V. *Entonces* T *tiene un subespacio invariante de dimensión uno*.

Demostración. Sea $m(x) = p_1^{r_1}(x)\, p_2^{r_2}(x) \cdots p_k^{r_k}(x)$ la factorización del polinomio mínimo de T en factores irreducibles. Por el Teorema Fundamental del Álgebra, cada $p_i(x)$ tiene grado uno o dos.

Si $k = 1$, aplicando el teorema 6.1.4 concluimos que $m(x)$ es potencia de un polinomio lineal, digamos $m(x) = (x - c)^l$, entonces el núcleo de $T - cI$ es no cero, concluyéndose la prueba en este caso.

Si $k > 2$, sea $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$, la descomposición que garantiza el teorema de la descomposición primaria, teorema 6.1.3. Este teorema también garantiza que cada W_i es T-invariante. Como V tiene dimensión impar, entonces al menos uno de $W_1, W_2, \ldots, W_k$ tiene dimensión impar, menor que la de V; haciendo inducción sobre el subespacio apropiado, se tiene la conclusión del teorema.

6.2. Valores y vectores característicos

De los resultados anteriores sabemos que los factores irreducibles del polinomio mínimo y característico de un operador son los mismos, por lo que tienen las mismas raíces. Por lo común, se formula la definición de valor característico en términos del polinomio característico, nosotros la formulamos usando al polinomio mínimo.

Definición 6.2.1. *Si* T *es un operador en* V, *a las raíces del polinomio mínimo de* T *les llamaremos valores característicos de* T.

Con este lenguaje el corolario 6.1.2 establece que un valor característico c de T tiene asociado un subespacio invariante de dimensión uno, y a cualquier generador de este subespacio le llamaremos un vector característico asociado a c. Lo anterior se puede formular diciendo que c es un *valor característico de T* si existe $\alpha \neq 0$ tal que $T(\alpha) = c\alpha$. Recíprocamente, cada subespacio invariante de dimensión uno tiene asociado un valor característico.

Definición 6.2.2. *Si* T *es un operador en* V, *diremos que* T *es diagonalizable si existe una base de* V *que consiste de vectores característicos de* T.

Una forma equivalente de enunciar la definición anterior es: T es diagonalizable $\Leftrightarrow$ existe una base de V en la que T se representa por medio de una matriz diagonal.

El polinomio mínimo de una matriz se define de forma análoga a como se hace para un operador.

Ejercicio 6.2.1. *Demuestre que los polinomios mínimos de matrices que representan al mismo operador son iguales.*

Teorema 6.2.1. *Sea* V *un espacio de dimensión* n, T *un operador en* V. *Entonces* T *es diagonalizable* $\Leftrightarrow$ *el polinomio mínimo de* T *es producto de factores lineales diferentes*.

Demostración. Supongamos que T es diagonalizable, entonces existe una base de V que consiste de vectores característicos de T. Respecto a esta base, T se representa por medio de una matriz diagonal. Mostraremos que el polinomio mínimo de una matriz diagonal es producto de factores lineales diferentes. En efecto, si los elementos de la diagonal de la matriz A son $a_1, a_2, \ldots, a_n$, de los cuales los únicos diferentes son $a_1, a_2 \ldots, a_r$, entonces $m_T(x) = (x - a_1)(x - a_2) \cdots (x - a_r)$ es el polinomio mínimo de A. Esto se debe a que $m_T(A)$ es producto de matrices diagonales, por lo cual también es diagonal; cada elemento de la diagonal de $m_T(A)$ es de la forma $(a_i - a_1)(a_i - a_2) \cdots (a_i - a_r) = 0$. Como el polinomio mínimo divide a cualquier otro que tiene a A por raíz, entonces $m_T(x)$ debe ser el polinomio mínimo.

Recíprocamente, supongamos que el polinomio mínimo de T es $m_T(x) = (x - c_1)(x - c_2) \cdots (x - c_k)$, con $c_i \neq c_j$ para $i \neq j$. Por el teorema 6.1.3, $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$, con W_i el núcleo del operador $T - c_iI$, es decir, W_i consiste de vectores característicos de T asociados a c_i. Sea $\mathcal{B}_i$ una base de W_i, entonces $\mathcal{B} = \bigcup_{i=1}^{k} \mathcal{B}_i$ es una base de V que consiste de vectores característicos, es decir, T es diagonalizable.

Lema 6.2.1. *Sean* V *un espacio de dimensión finita,* $T : V \to V$ *un operador y* $W \neq V$ *un subespacio* T-*invariante. Supongamos que el polinomio mínimo de* T, $m_T(x)$ *tiene todas sus raíces en* $\mathbb{R}$, *entonces existe* $\alpha \in V \setminus W$ *tal que* $T(\alpha) - c\alpha \in W$, *con* c *raíz del polinomio mínimo de* T.

Demostración. Sea $\beta \in V \setminus W$ y $g(x)$ el T-anulador de β respecto a W, entonces $g(x)$ divide a $m_T(x)$. Como $m_T(x) = (x - c_1)^{r_1}(x - c_2)^{r_2} \cdots (x - c_k)^{r_k}$, con todos los $c_j \in \mathbb{R}$, y $g(x)$ lo divide, entonces $g(x) = (x - c_1)^{s_1}(x - c_2)^{s_2} \cdots (x - ck)^{s_k}$, con $s_j \leq r_j$ para todo j. También se tiene que $g(x)$ tiene grado positivo, por lo cual podemos suponer que $s_1 > 0$, así que $g(x) = (x - c_1)h(x)$; de esto último se tiene $g(T)(\beta) = (T - c_1I)h(T)(\beta) \in W$. Como $g(x)$ es de mínimo grado tal que $g(T)(\beta) \in W$, entonces $h(T)(\beta) = \alpha \notin W$ es el elemento buscado.

Definición 6.2.3. *Sean* V *un espacio de dimensión finita,* $T : V \to V$ *un operador. Se dice que* T *es triangulable si existe una base de* V *en la cual* T *se representa por una matriz triangular*.

Teorema 6.2.2. *Sea* V *un espacio de dimensión* n, T *un operador en* V. *Entonces* T *es triangulable* $\Leftrightarrow$ *el polinomio mínimo de* T *es producto de factores lineales.*

Demostración. Supongamos que el polinomio mínimo se factoriza como $m_T(x) = (x - c_1)^{r_1}(x - c_2)^{r_2} \cdots (x - c_k)^{r_k}$. Aplicando el lema anterior con $W_0 = \{0\}$ se tiene que existe $\alpha_1 \neq 0$ tal que $T(\alpha_1) = c\alpha_1$, con c raíz de $m_T(x)$. Si $n > 1$, entonces continúa el proceso con W_1, el espacio generado por α_1 que es claramente T-invariante. Continuando este proceso se obtiene una base de V, en la cual T se representa por una matriz triangular. (Explique esto último.)

Recíprocamente, supongamos que T es triangulable, entonces existe una base de V, en la cual T se representa por una matriz triangular, con entradas en la diagonal $a_{11}, a_{22}, \ldots, a_{nn}$. Supongamos que de estos elementos solamente r de ellos son diferentes, sin perder generalidad podemos suponer que éstos son $a_{11}, a_{22}, \ldots, a_{ll}$, entonces (justifique esta afirmación) el polinomio mínimo de T es $m_T(x) = (x - a_{11})^{r_1}(x - a_{22})^{r_2} \cdots (x - a_{ll})^{r_l}$, para algunos enteros positivos $r_1, r_2, \ldots, r_l$.

Lema 6.2.2. *Si* V *es un espacio de dimensión finita,* $T : V \to V$ *un operador,* α *y* β *elementos de* V *tales que* ann (α, x) *y* ann (β, x) *son primos relativos, entonces* ann $(\alpha + \beta, x) =$ ann (α, x)ann (β, x).

Demostración. Con las hipótesis sobre α y β mostraremos que los subespacios cíclicos generados por α y β forman una suma directa, es decir, se cumple: $C(\alpha, T) \cap C(\beta, T) = \{0\}$. Si $\gamma \in C(\alpha, T) \cap C(\beta, T) \setminus \{0\}$, ann (γ, x) tiene grado positivo y $\gamma = f_1(T)(\alpha) = f_2(T)(\beta)$, para algunos polinomios $f_1(x)$ y $f_2(x)$. De esta ecuación se concluye que ann $(\alpha, T)(\gamma) =$ ann $(\beta, T)(\gamma) = 0$, por lo que ann (γ, x) divide a ann (α, x) y a ann (β, x), contradiciendo que son primos relativos.

Por evaluación se verifica que ann (α, T)ann $(\beta, T)(\alpha + \beta) = 0$. Por lo que ann $(\alpha + \beta, x)$ divide a ann (α, x)ann (β, x). También tenemos ann $(\alpha + \beta, T)(\alpha + \beta) =$ ann $(\alpha + \beta, T)(\alpha) +$ ann $(\alpha + \beta, T)(\beta) = 0$, es decir, ann $(\alpha + \beta, T)(\alpha) = -$ann $(\alpha + \beta, T)(\beta) \in C(\alpha, T) \cap C(\beta, T) = \{0\}$. De esto concluimos que ann (α, x) y ann (β, x) dividen a ann $(\alpha + \beta, x)$. Como ann (α, x) y ann (β, x) son primos relativos, su producto también divide a ann $(\alpha + \beta, x)$. De lo probado antes se concluye que ann $(\alpha + \beta, x) =$ ann (α, x)ann (β, x).

Sea V un espacio de dimensión finita, $\alpha \in V$ y $T : V \to V$ un operador, se tiene que ann (α, x) divide al polinomio mínimo de T. De esto se concluye que existe un $\alpha_0 \in V$ tal que deg(ann $(\alpha_0, x)) \geq$ deg(ann $(\beta, x))$ para todo $\beta \in V$. El siguiente lema establece una relación más precisa entre ann (α_0, x) y ann (β, x).

Lema 6.2.3. *Con las hipótesis y notación anteriores se tiene: ann* (β, x) *divide a ann* (α_0, x) *para todo* $\beta \in V$.

Demostración. Podemos representar a los anuladores de α y β como producto de polinomios irreducibles, de manera precisa. Sean ann $(a_0, x) = p_1^{e_1}(x)p_2^{e_2}(x) \cdots p_r^{e_r}(x)$ y ann $(\beta, x) = p_1^{a_1}(x)p_2^{a_2}(x) \cdots p_r^{a_r}(x)$. Si la conclusión del lema fuese falsa, entonces para algún i, $a_i > e_i$. Sean, $\gamma_1 = p_1^{e_1}(T)(\alpha)$ y $\gamma_2 = \dfrac{\text{ann }(\beta)}{p_i^{a_i}}(T)(\beta)$. Es claro que el T-anulador de γ_1 es $\dfrac{\text{ann }(\alpha_0, x)}{p_i^{e_i}(x)}$ y el T-anulador de γ_2 es $p_i^{a_i}(x)$. Estos polinomios son primos relativos. Aplicando el lema 6.2.2 se tiene que el T-anulador de $\gamma_1 + \gamma_2$ es $\dfrac{\text{ann }(\alpha_0, x)}{p_i^{e_i}(x)}p_i^{a_i}(x)$, que tiene grado mayor que ann (α_0, x), contradiciendo la elección de α_0.

Lema 6.2.4. *Sean,* V *un espacio de dimensión finita,* $T : V \to V$ *un operador. Entonces* V *tiene un subespacio* T-*cíclico de dimensión igual al grado del polinomio mínimo de* T.

Demostración. Demostraremos algo más; existe $\alpha_0 \in V$ cuyo T-anulador es igual al polinomio mínimo de T; por tanto, la conclusión se tendrá del lema 6.1.3.

Sea $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ una base de V y α_0 como en la discusión anterior. Aplicando el lema 6.2.3 se tiene que ann (α_i, x) divide a ann (α_0, x), lo cual implica ann $(\alpha_0, T)(\alpha_i) = 0$ para todo $i = 1, 2, \ldots, n$. De esto se concluye que ann $(\alpha_0, T) = 0$, por lo que necesariamente, $m_T(x) =$ ann (α_0, x), probando el lema.

Teorema 6.2.3. *Sea* V *un espacio de dimensión finita,* T *un operador en* V, *y* $\beta \in V$ *tal que* ann $(\beta, x) = m_T(x)$. Sea $C(\beta, T)$ *el subespacio* T-*cíclico generado por* β, U *un subespacio* T-*invariante y* $g : U \to C(\beta, T)$ *una transformación lineal tal que* T *conmuta con* g. *Entonces* g *se puede extender a una transformación lineal* $g_1 : V \to C(\beta, T)$ *que conmuta con* T.

Demostración. Si $U = V$ no hay nada que demostrar, por lo que podemos suponer la existencia de $\alpha_1 \in V \setminus U$. Definamos $U_1 = U + C(\alpha_1, T)$. Sea $s(x)$ el T-anulador de α_1 respecto a U, es decir, $s(T)(\alpha_1) \in U$. Como $m(T)(\alpha_1) = 0 \in U$, entonces $s(x)$ divide a $m_T(x)$, por lo que $m_T(x) = s(x)h(x)$. De la definición de $s(x)$ se tiene que $g(s(T)(\alpha_1)) = u(T)\beta$ para algún $u(x)$. La hipótesis sobre la conmutatividad de g y T implica $h(T)g(s(T)\alpha_1) = g(h(T)s(T)(\alpha_1)) = g(0) = 0$, de lo cual se tiene que el T-anulador de $g(s(T)\alpha_1)$ en $C(\beta, T)$ divide a $h(x)$, por lo que $h(T)(g(s(T)(\alpha_1))) = h(T)(u(T)\beta) = 0$, y de esto se tiene que $m_T(x)$ divide a $h(x)u(x)$, es decir, $h(x)u(x) = m_T(x)q(x) = s(x)h(x)q(x)$, de lo que se tiene $u(x) = s(x)q(x)$, y de esto se concluye que $g(s(T)(\alpha_1)) = s(T)q(T)\beta$. Sea $\beta_0 = q(T)\beta$. Para todo $\alpha \in U$, definamos $g_1(\alpha + r(T)\alpha_1) := g(\alpha) + r(T)\beta_0$. Demostraremos que la definición de g_1 es consistente en el sentido que si $\alpha + r(T)\alpha_1 = \alpha' + r_1(T)\alpha_1$ implica $g(\alpha) + r(T)\beta_0 = g(\alpha') + r_1(T)\beta_0$. La condición $\alpha + r(T)\alpha_1 = \alpha' + r_1(T)\alpha_1$ lleva a que $\alpha' - \alpha =$

$[r(T) - r_1(T)]\alpha_1 \in U$, por lo que $s(x)$ divide a $r(x) - r_1(x)$, es decir, $r(x) - r_1(x) = s(x)l(x)$, para algún $l(x)$. De lo anterior se tiene:

$$\begin{aligned} g([r(T) - r_1(T)]\alpha_1) &= g(s(T)l(T))\alpha_1 \\ &= l(T)g(s(T)\alpha_1) \\ &= l(T)s(T)q(T)\beta \\ &= l(T)s(T)\beta_0 \\ &= (r(T) - r_1(T))\beta_0 \end{aligned}$$

Por otro lado, $g([r(T) - r_1(T)]\alpha_1) = g(\alpha - \alpha') = g(\alpha) - g(\alpha')$, de esto junto con lo anterior se obtiene $g(\alpha) + r(T)\beta_0 = g(\alpha') + r_1(T)\beta_0$, como se afirmó. Se verifica de manera directa que g_1 conmuta con T. La demostración termina en a lo más tantos pasos como la dimensión de V.

Teorema 6.2.4. (Descomposición cíclica). *Sea* V *un espacio de dimensión finita,* T *un operador en* V. *Entonces* $\mathrm{V} = \mathrm{W}_1 \oplus \mathrm{W}_2 \oplus \cdots \oplus \mathrm{W}_k$, *en donde* W_i *es* T-*cíclico con* T-*anulador*[3] $\mathrm{m}_i(\mathrm{x})$ *y éstos satisfacen* $\mathrm{m}_i(\mathrm{x})$ *divide a* $\mathrm{m}_{i-1}(\mathrm{x})$ *para todo* $\mathrm{i} \in \{2, 3, \ldots, \mathrm{k}\}$ *y* $\mathrm{m}_1(\mathrm{x})$ *es el polinomio mínimo de* T.

Demostración. Aplicando el lema 6.2.4 se tiene que existe $\alpha_1 \in V$, cuyo anulador es el polinomio mínimo de T. Sea W_1 el espacio T-cíclico generado por α_1. Aplicando el teorema anterior con $U = W_1$ y $g = I$, identidad, se tiene que existe una transformación lineal $g_1 : V \to W_1$ que conmuta con T. Esta condición garantiza que el núcleo de g_1 es un subespacio T invariante. Sea N_2 tal núcleo, mostraremos que V es la suma directa de W_1 y N_2.

Dado $\alpha \in V$ se tiene $g_1(\alpha - g_1(\alpha)) = g_1(\alpha) - g_1(g_1(\alpha))$. Como $g_1(\alpha) \in W_1$ y g_1 restringido a W_1 es la identidad, entonces $g_1(g_1(\alpha)) = g_1(\alpha)$, es decir $\alpha - g_1(\alpha) \in N_2$ y de esto se concluye que $V = W_1 + N_2$. Si $\alpha \in W_1 \cap N_2$, entonces $\alpha = g_1(\alpha) = 0$, probando con esto que V es la suma directa de W_1 y N_2. Ahora aplicamos el mismo proceso al subespacio N_2, probando primeramente que existe un $\alpha_2 \in N_2$ cuyo anulador es el polinomio mínimo de T restringido a N_2. Si este polinomio lo denotamos por $m_2(x)$, claramente se tiene que $m_2(x)$ divide a $m_1(x)$ y $N_2 = W_2 \oplus N_3$ con W_2 el subespacio T-cíclico generado por α_2, entonces $V = W_1 \oplus W_2 \oplus N_3$. Este proceso termina y produce la descomposición deseada.

El siguiente teorema concluye la discusión sobre la descomposición cíclica de un espacio vectorial de dimensión finita, estableciendo la unicidad. Este resultado es un caso particular del teorema que se tiene para módulos finitamente generados sobre dominios de ideales principales, lo enunciamos sin demostración.

Teorema 6.2.5. (Unicidad de la descomposición cíclica). *Sea* V *un espacio de dimensión finita,* T *un operador en* V. *Supongamos que* $\mathrm{V} = \mathrm{W}_1 \oplus \mathrm{W}_2 \oplus \cdots \oplus \mathrm{W}_k = \mathrm{U}_1 \oplus \mathrm{U}_2 \oplus \cdots \oplus \mathrm{U}_l$, *en donde* W_i y U_j *son* T-*cíclicos para cada* $\mathrm{i} = 1, 2, \ldots, \mathrm{k}$ *y para cada* $\mathrm{j} = 1, 2, \ldots, l$; *con anuladores* $\mathrm{m}_i(\mathrm{x})$ *y* $\mathrm{n}_j(\mathrm{x})$, *respectivamente. Los anuladores satisfacen*: $\mathrm{m}_i(\mathrm{x})$ *divide a* $\mathrm{m}_{i-1}(\mathrm{x})$ *para todo* $\mathrm{i} \in \{2, 3, \ldots \mathrm{k}\}$; $\mathrm{n}_j(\mathrm{x})$ *divide a* $\mathrm{n}_{j-1}(\mathrm{x})$ *para todo* $\mathrm{j} \in \{2, 3, \ldots, \mathrm{l}\}$. *Entonces* $\mathrm{k} = \mathrm{l}$ *y* $\mathrm{m}_i(\mathrm{x}) = \mathrm{n}_i(\mathrm{x})$ *para todo* $\mathrm{i} = 1, 2, \ldots, \mathrm{k}$.

Definición 6.2.4. *Sea* $\mathrm{p}(\mathrm{x}) = \mathrm{c}_0 + \mathrm{c}_1\mathrm{x} + \mathrm{c}_2\mathrm{x}^2 + \cdots + \mathrm{c}_{l-1}\mathrm{x}^{l-1} + \mathrm{x}^l$ *un polinomio. Se define la matriz:*

[3] Como W_i es T-cíclico, al T-anulador de un generador le llamamos el T-anulador de W_i.

$$A(p) := \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{l-1} \end{pmatrix}$$

y le llamamos la matriz asociada a p(x).

Observación 6.2.1. *La matriz* A(p) *puede ser considerada en la siguiente forma. Sea* $\{\alpha_1, \alpha_2, \cdots, \alpha_l\}$ *una base de* $\mathbb{R}^l$ *y* $T : \mathbb{R}^l - \mathbb{R}^l$ *el operador definido por* $T(\alpha_i) = \alpha_{i+1}$, $i = 1, 2, \ldots, (l-1)$ *y* $T(\alpha_l) = -c_0\alpha_1 - c_1\alpha_2 - \cdots - c_{l-1}\alpha_{l-1}$. *Entonces* A(p) *es la matriz que representa a* T *respecto de la base dada.*

Ejercicio 6.2.2. *Demuestre que el polinomio mínimo de la matriz* A(p) *es* p(x).

Corolario 6.2.1. *Sea* A *una matriz* $n \times n$. *Entonces existe* P *no singular tal que:*

$$P^{-1}AP = \begin{pmatrix} A_1 & 0 & \cdots & 0 & 0 \\ 0 & A_2 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & A_k \end{pmatrix} \tag{6.8}$$

en donde cada una de las matrices A_i *es de la forma (6.7). Más aún, esta representación de* A *es única. Si* $m_i(x)$ *es el polinomio mínimo de* A_i, *entonces* $m_1(x)$ *es el polinomio mínimo de* A *y* $m_{i+1}(x)$ *divide a* $m_i(x)$ *para todo* $i = 1, 2, \ldots, k-1$.

Definición 6.2.5. *Sea* A *una matriz cuadrada. La representación de* A *dada por (6.8) se le llama la forma racional de* A. *Si* $m_1(x), \ldots, m_k(x)$ *son los polinomios mínimos de las matrices* $A_1, A_2, \ldots, A_k$ *respectivamente, a ésos se les llaman los factores invariantes de* A.

Corolario 6.2.2. *Dos matrices cuadradas son similares* $\Leftrightarrow$ *tienen la misma forma racional.*

Definición 6.2.6. *Si* T *es un operador en un espacio* V, *a los factores invariantes de cualquier matriz que represente a* T *se les llama los factores invariantes de* T.

6.2.1. Calculando el polinomio mínimo

En esta parte presentamos un algoritmo que determina el polinomio mínimo de una matriz.

Sea A una matriz $n \times n$. Por el corolario 6.2.1, existe una matriz P no singular tal que:

$$P^{-1}AP = \begin{pmatrix} A_1 & 0 & \cdots & 0 & 0 \\ 0 & A_2 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & A_k \end{pmatrix}$$

en donde cada una de las matrices de la diagonal son de la forma (6.7).

De esto se tiene:

$$P^{-1}AP - xI = \begin{pmatrix} A_1 - xI & 0 & \cdots & 0 & 0 \\ 0 & A_1 - xI & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & A_1 - xI \end{pmatrix} \tag{6.9}$$

en donde I representa a la matriz identidad del orden adecuado en cada caso. *Para cada bloque* $\mathrm{A_i - xI}$, *mostraremos que por medio de operaciones elementales en las filas y columnas, éste se puede llevar a una matriz diagonal que en la entrada* (1,1) *tiene a su polinomio mínimo,* $\mathrm{m_i(x)}$ *y en las restantes entradas de la diagonal tiene unos*.

Para facilitar la notación podemos suponer que:

$$A_i = \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{l-1} \end{pmatrix}$$

en donde $m_i(x) = x^l + c_{l-1}x^{l-1} + \cdots + c_1x + c_0$ representa a su polinomio mínimo.

Escribiendo explícitamente a la matriz $A_i - xI$ se tiene:

$$A_i - xI = \begin{pmatrix} -x & 0 & \cdots & 0 & -c_0 \\ 1 & -x & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{l-1} - x \end{pmatrix}$$

Multiplicando la segunda fila por x y sumándola a la primera se obtiene la matriz:

$$B_1 = \begin{pmatrix} 0 & -x^2 & \cdots & 0 & -c_0 - c_1x \\ 1 & -x & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{l-1} - x \end{pmatrix}$$

Multiplicando la tercera fila de B_1 por x^2 y sumándola a la primera resulta la matriz B_2, que en la primera fila tiene ceros excepto en la tercera y última posiciones, cuyas entradas allí son $-x^3$ y $-c_0 - c_1x - c_2x^2$, respectivamente. Las restantes coinciden con las de B_1.

Multiplicando ahora la cuarta fila de B_2 por x^3 y sumándola a la primera, se obtiene la matriz B_3, que en la primera fila tiene ceros, excepto en la cuarta y última posiciones, cuyas entradas son $-x^4$ y $-c_0 - c_1x - c_2x^2 - c_3x^3$, respectivamente. Las filas restantes coinciden con las de B_1. Continuando de esta forma $l - 1$ veces se obtiene una matriz B_{l-1} que en la entrada $(1, l)$ tiene a $-p(x) = -c_0 - c_1x + \cdots + c_{l-1}x^{l-1} - x^l$, en las restantes entradas de la primera fila tiene ceros y las entradas de las filas restantes coinciden con las de B_1. Explícitamente:

$$B_{l-1} = \begin{pmatrix} 0 & 0 & \cdots & 0 & -p(x) \\ 1 & -x & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{l-1}-x \end{pmatrix}$$

Aplicando operaciones elementales adecuadas (¿cuáles?) en las columnas de B_{l-1}, se obtiene:

$$B = \begin{pmatrix} 0 & 0 & \cdots & 0 & -p(x) \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

Multiplicando la última columna de B por -1 y trasladándola a la primera, la matriz B se transforma en:

$$C = \begin{pmatrix} p(x) & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

como se afirmó.

Note que las operaciones elementales que se aplicaron en el bloque $A_i - xI$ de (6.9) no afectan a ninguno de los bloques restantes y en cada uno se puede aplicar el mismo procedimiento. En resumen, hemos probado que la matriz $P^{-1}(A - xI)P$, y por ende $A - xI$, es equivalente a una matriz diagonal. De forma más precisa, existen matrices inversibles Q y R con entradas en los polinomios tales que:

$$Q(A - xI)R = \begin{pmatrix} m_1(x) & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & m_2(x) & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & m_k(x) & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

en donde $m_{i+1}(x)$ divide a $m_i(x)$ para todo $i = 1, 2, ..., k - 1$ y $m_1(x)$ es el polinomio mínimo de A.

Para describir un algoritmo que permite calcular los factores invariantes de A hace falta una definición y un teorema.

Definición 6.2.7. *Sea* N *una matriz cuyas entradas son polinomios. Diremos que se encuentra en forma normal de Smith si satisface:*

1. *Todo elemento fuera de la diagonal principal es cero.*
2. *En la diagonal principal de* N *aparecen en orden los polinomios* $f_1, f_2 \ldots, f_l$ *y satisfacen que* f_j *divide a* f_{j+1} *para todos* $j = 1, 2, ..., l - 1$.

El siguiente teorema lo presentamos sin demostración, éste y lo probado antes proveen de un algoritmo para determinar los factores invariantes de cualquier matriz, en particular permiten calcular el polinomio mínimo.

Teorema 6.2.6. *Sea* B *una matriz con entradas en los polinomios con coeficientes en un campo. Entonces,* B *es equivalente a una única matriz en forma normal de Smith.*

En esta última parte de la sección, las matrices tienen entradas en el conjunto de los polinomios con coeficientes en los reales o los complejos. Por esta razón, precisaremos el tipo de operaciones elementales permitidas cuando las entradas de las matrices son polinomios.

1. Intercambiar filas o columnas.
2. Multiplicar una fila (columna) por un polinomio y sumarla a otra fila (columna).
3. Multiplicar una fila por un polinomio constante no cero (éstos son los únicos polinomios que tienen inverso multiplicativo).

Algoritmo 6.2.1[4]

Requiere: *matriz cuadrada* A.

1. *Construya la matriz* $A_1 = A - xI$.
2. *Con operaciones elementales en las filas y columnas de* A_1 *se obtiene* $B = \text{diag}\{p(x), C\}$, *en donde* p(x) *es el máximo común divisor de los elementos de la primera fila y la primera columna de* A_1, *y* C *es cuadrada.*

[4] El interesado en obtener el código del algoritmo implementado en Maple, favor de solicitarlo al autor por correo electrónico a la dirección: barrera@uaeh.edu.mx.

3. ***mientras*** p(x) *no divida a todas la entradas de* C ***haga***
4. *Encuentre la primera columna, de izquierda a derecha, que contiene un elemento no divisible por* p(x) *y sume esta columna a la primera de* B.
5. *Aplique el paso* 2 *a* B.
6. ***fin mientras***
7. *Haga* $A_1 = C$ *y aplique este mismo procedimiento.*

En un número finito de iteraciones llegará a una matriz de la forma

$$\mathrm{diag}\{m_1(x), m_2(x), \ldots, m_k(x), 1, \ldots, 1\},$$

en donde $m_i(x)$ *divide a* $m_{i+1}(x)$. *El polinomio mínimo de* A *es* $m_k(x)$.

Ejemplo 6.2.1. *Encuentre el polinomio mínimo de* $A = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$

Sea $A_1 = \begin{pmatrix} 1-x & 0 & -1 \\ -1 & 1-x & 0 \\ 0 & -1 & 1-x \end{pmatrix}$. Multiplicando la segunda fila de A_1 por $1-x$ y sumándola a la primera se obtiene:

$A_2 = \begin{pmatrix} 0 & (1-x)^2 & -1 \\ -1 & 1-x & 0 \\ 0 & -1 & 1-x \end{pmatrix}$. Multiplicando la segunda fila de A_2 por -1 e intercambiándola con la uno se tiene: $A_3 = \begin{pmatrix} 1 & x-1 & 0 \\ 0 & (1-x)^2 & -1 \\ 0 & -1 & 1-x \end{pmatrix}$

Multiplicando la primera columna de A_3 por $-(x-1)$ y sumándola a la segunda se tiene: $A_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & (1-x)^2 & -1 \\ 0 & -1 & 1-x \end{pmatrix}$. Multiplicando la tercera fila de A_4 por $(1-x)^2$ y sumándola a la segunda se obtiene: $A_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1+(1-x)^3 \\ 0 & -1 & 1-x \end{pmatrix}$

Multiplicando la segunda columna de A_5 por $(1 - x)$ y sumándola a la tercera se obtiene:

$$A_6 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1+(1-x)^3 \\ 0 & -1 & 0 \end{pmatrix}$$

Multiplicando las columnas dos y tres de A_6 por –1 e intercambiando las filas dos y tres se obtiene:

$A_7 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (x-1)^3+1 \end{pmatrix}$. Con este proceso hemos obtenido la siguiente información:

1. El polinomio mínimo de A es $(x - 1)^3 + 1 = x(x^2 - 3x + 3) = x^3 - 3x^2 + 3x$.
2. La matriz A es diagonalizable, pues su polinomio mínimo tiene raíces diferentes.
3. La forma racional de A es: $\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{pmatrix}$.
4. La matriz A es singular.

Ejercicio 6.2.3. *Si* $m_1(x), \ldots, m_k(x)$ *son los factores invariantes de* A, *¿cómo se obtiene el polinomio característico de* A*?*

6.3. Forma canónica de Jordan

Sea V un espacio de dimensión finita, T un operador en V y $m_T(x) = p_1^{e_1}(x)\, p_2^{e_2}(x) \cdots p_r^{e_r}(x)$, la representación del polinomio mínimo de T como producto de irreducibles. Por el teorema 6.1.3 (teorema de la descomposición primaria) se tiene:

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k, \quad \text{con} \quad W_i = V_{p_i^{e_i}}(T)$$

También sabemos que el polinomio mínimo de T restringido a W_j es $p_j^{e_j}(x)$. Aplicando el teorema 6.2.4 (teorema de la descomposición cíclica) a cada W_j se tiene que $W_j = W_{1j} \oplus W_{2j} \oplus \cdots \oplus W_{i_j j}$, en donde cada W_{ij} es T-cíclico con anulador $p_j^{e_{ij}}(x)$; los exponentes satisfacen $e_j = e_{1j} \geq e_{2j} \geq \cdots \geq e_{i_j j}$.

Definición 6.3.1. *Los polinomios* $p_y^{e_{ij}}(x)$ *se llaman divisores elementales de* T.

Supongamos que algún $p_j(x)$ es lineal y que el anulador de W_{ij} es $(x - cj)^{e_{ij}}$. Si v es un vector cíclico de W_{ij} entonces:

$$\{v, (T - c_j I)v, \ldots, (T - c_j I)^{e_{ij}-1}\}$$

es una base.

En efecto, como la dimensión de W_{ij} es e_{ij}, basta demostrar que el conjunto propuesto es linealmente independiente. Sea:

$$a_0 v + a_1(T - c_j I)v + \cdots + a_l(\mathrm{T} - c_j I)^l v = 0 \tag{6.10}$$

con $l = e_{ij} - 1$. Como $(T - c_j I)^{e_{ij}} v = 0$, entonces aplicando $(T - c_j I)^l$ en la ecuación 6.10 se tiene $a_0(T - c_j I)^l v = 0$ y como $(T - c_j I)^l v \neq 0$, entonces $a_0 = 0$. Continuando en esta forma se concluye que todos los coeficientes en la citada ecuación son cero.

La matriz de T restringida a W_{ij} respecto a la base $\{v, (T - c_j I)v, \ldots, (T - c_j I)^{e_{ij}-1}\}$ se obtiene aplicando T a cada elemento.

$$\begin{aligned}
T(v) &= c_j v + (T - c_j I)v \\
T(T - c_j I)v &= (T - c_j I)(T(v)) \\
&= (T - c_j I)(c_j v + (T - c_j I)v) \\
&= c_j(T - c_j I)v + (T - c_j I)^2 v \\
&\vdots \\
T(T - c_j I)^m v &= (T - c_j I)^m(c_j v + (T - c_j I)v) \\
&= c_j(T - c_j I)^m v + (T - c_j I)^{m+1} v \\
&\vdots \\
T((T - c_j I)^{e_{ij}-1}(v)) &= c_j(T - c_j)^{e_{ij}-1}(v)
\end{aligned}$$

De estas ecuaciones se tiene que la matriz asociada a la restricción de T en W_{ij} es:

$$\begin{pmatrix}
c_j & 0 & \cdots & 0 & 0 & 0 \\
1 & c_j & \cdots & 0 & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & c_j & 0 \\
0 & 0 & \cdots & 0 & 0 & c_j
\end{pmatrix} \tag{6.11}$$

consecuentemente, existe una base de W_j respecto de la cual la matriz asociada a T restringida a W_j es diagonal por bloques con cada bloque de la forma (6.11), llamado *bloque de Jordan*. Si el polinomio mínimo se expresa como producto de factores lineales, entonces el anulador en cada W_{ij} es de la forma $(x - c_j)^{e_{ij}}$ y procediendo como en el caso anterior se tiene que la restricción de T a cada W_j es diagonal por bloques con cada bloque de la forma (6.11). Resumiendo, hemos demostrado el siguiente:

Teorema 6.3.1. (Forma canónica de Jordan). *Sea* V *un espacio de dimensión finita,* T *un operador en* V*. Supongamos que el polinomio mínimo de* T *se expresa como* $\mathrm{m(x)} = (\mathrm{x} - \mathrm{c}_1)^{e_1} (\mathrm{x} - \mathrm{c}_2)^{e_2} \cdots (\mathrm{x} - \mathrm{c}_k)^{e_k}$*. Entonces existe una base de* V *respecto de la cual* T *se representa por una matriz de la forma* $\mathrm{J} = \mathrm{diag}\{\mathrm{J}_1, \ldots, \mathrm{J}_k\}$*, con cada* J_m *a la vez diagonal por bloques*: $\mathrm{J}_m = \mathrm{diag}\{\mathrm{J}_{1m}, \ldots, \mathrm{J}_{i_m m}\}$ *y cada* J_{rm} *un bloque de Jordan de orden* e_{rm}*, los cuales satisfacen* $\mathrm{e}_m = \mathrm{e}_{1m} \geq \mathrm{e}_{2m} \geq \cdots \geq \mathrm{e}_{r_m m}$.

La formulación del teorema anterior para matrices es:

Corolario 6.3.1. *Sea* A *una matriz cuadrada con entradas en cualquier campo. Supongamos que su polinomio mínimo se factoriza como* $m(x) = (x - c_1)^{e_1}(x - c_2)^{e_2}\cdots(x - c_k)^{e_k}$, *entonces* A *es similar a una matriz de la forma* $J = \mathrm{diag}\{J_1, \ldots, J_k\}$, *en donde* $J_m = \mathrm{diag}\{J_{1m}, \ldots, J_{r_m m}\}$ *y cada* J_{lm} *un bloque de Jordan de orden* e_{lm}, *los cuales satisfacen* $e_m = e_{1m} \geq e_{2m} \geq e_{r_m m}$.

Definición 6.3.2. *A la matriz* J *del corolario anterior se le llama la* forma canónica de Jordan *de* A.

Observación 6.3.1. *La forma canónica de Jordan de una matriz* A *queda completamente descrita por sus valores característicos y por los órdenes de los bloques de Jordan. Más precisamente, si los valores característicos de* A *son* $\{c_1, c_2, \ldots, c_k\}$ *y los órdenes de los bloques de Jordan son* $\{(e_{11}, e_{21}, \ldots, e_{i_1 1}), (e_{12}, e_{22}, \ldots, e_{i_2 2}), \ldots, (e_{1k}, e_{2k}, \ldots, e_{i_k k})\}$, *entonces la forma canónica de Jordan de* A *está únicamente determinada por el corolario anterior.*

Ejemplo 6.3.1. *Supongamos que una matriz* A *tiene valores característicos* $\{2, 4, -1\}$ y órdenes de bloques de Jordan $\{(3, 2, 1), (2, 2), (2)\}$. *Entonces la forma canónica de Jordan de* A *es* $J = \mathrm{diag}\{J_{11}, J_{21}, J_{31}, J_{12}, J_{22}, J_{13}\}$, *en donde*:

$$J_{11} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}, J_{21} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}, J_{31} = (2), J_{12} = J_{22} = \begin{pmatrix} 4 & 0 \\ 1 & 4 \end{pmatrix}$$

$$J_{13} = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix}, \text{ de forma explícita:}$$

$$J = \begin{pmatrix}
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1
\end{pmatrix}$$

Ejercicio 6.3.1. *Escriba los divisores elementales de la matriz del ejemplo anterior.*

Ejercicio 6.3.2. *Sea* A *una matriz* 2×2. *Describa todas las posibles formas canónicas de Jordan de* A.

Ejercicio 6.3.3. *Sea* A $= \begin{pmatrix} 2 & 0 & 0 \\ a & -1 & 0 \\ c & b & -1 \end{pmatrix}$

1. *Determine condiciones necesarias y suficientes para que* A *sea diagonalizable y la forma canónica de Jordan de* A.
2. *Si* A *no es diagonalizable, encuentre su forma canónica de Jordan*.

6.4. Matrices reales con valores característicos no reales

Bajo la fuerte hipótesis que el polinomio mínimo de un operador se factorice completamente, en la sección anterior presentamos una discusión general sobre la representación de un operador mediante una matriz diagonal por bloques. Gracias al teorema fundamental del álgebra, esta hipótesis se tiene si el espacio vectorial es complejo. Entonces surge una pregunta, ¿qué es lo mejor que se tiene en el caso de espacios vectoriales reales? Sabemos, por el teorema 6.2.1, que si el polinomio mínimo tiene raíces diferentes, el operador es diagonalizable, sin embargo, en el caso complejo la matriz que diagonaliza, así como la matriz diagonal, no son reales. En esta sección demostraremos que el operador puede ser representado por una matriz diagonal por bloques reales de órdenes uno o dos. Los bloques de orden dos son matrices similares a una rotación seguida de una homotecia. El caso general lleva a lo que en algunos textos le llaman la *Forma Canónica de Jordan Real*.

6.4.1. Matrices 2×2

Iniciamos con el caso de matrices 2×2 que no tienen valores característicos reales.

Sea:

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

una matriz cuyo polinomio mínimo no tiene raíces reales; mostraremos que B es similar a una de la forma:

$$A = \begin{pmatrix} m & -n \\ n & m \end{pmatrix}, \text{ con } n \neq 0$$

Por lo observado antes, iniciamos la discusión considerando este tipo de matrices.

Dada la matriz A, pongamos $\lambda = \sqrt{m^2+n^2}$ λ: y

$$R = \begin{pmatrix} \frac{m}{\lambda} & \frac{-n}{\lambda} \\ \frac{n}{\lambda} & \frac{m}{\lambda} \end{pmatrix}$$

Haciendo sen $(\theta) = \frac{n}{\lambda}$ y $\cos(\theta) = \frac{m}{\lambda}$, se tiene que R es la matriz de una rotación, ecuación 4.7, página 101.

De esto se concluye que A es el producto de la homotecia λI y la rotación R, como lo muestra la ecuación:

$$A = \begin{bmatrix} m & -n \\ n & m \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{pmatrix} \dfrac{m}{\lambda} & \dfrac{-n}{\lambda} \\ \dfrac{n}{\lambda} & \dfrac{m}{\lambda} \end{pmatrix}$$

En términos geométricos, el efecto de aplicar A a un vector X se ilustra en la figura 6.1.

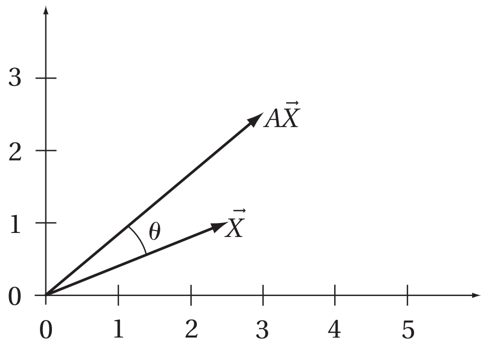


Figura 6.1. Rotación y homotecia del vector $\vec{x}$ bajo la acción de A.

Resumiendo, dada la matriz:

$$A = \begin{bmatrix} m & -n \\ n & m \end{bmatrix}, \text{ con } n \neq 0,$$

la acción de A en un vector X es rotarlo un ángulo $\theta = \operatorname{tg}^{-1}\left(\dfrac{n}{m}\right)$ y magnificarlo un factor $\lambda = \sqrt{m^2 + n^2}$.

Otro aspecto, que es de importancia notar acerca de la matriz A, es la existencia de su inversa, condición que se obtiene de la hipótesis sobre n.

Con A como en la discusión anterior, pongamos $P = \begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix}$, cuya inversa es $P^{-1} = \dfrac{1}{n}\begin{bmatrix} n & -m \\ 0 & 1 \end{bmatrix}$. De esto se obtiene:

$$P^{-1} AP = \frac{1}{n}\begin{bmatrix} n & -m \\ 0 & 1 \end{bmatrix}\begin{bmatrix} m & -n \\ n & m \end{bmatrix}\begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix} = \begin{bmatrix} 0 & -n^2 - m^2 \\ 1 & 2m \end{bmatrix} \tag{6.12}$$

Sea:

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

tal que su polinomio mínimo no tiene raíces reales. Esta hipótesis y el resultado $B\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix}$ muestran que los vectores $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} a \\ c \end{bmatrix}$ son linealmente independientes, por lo que la matriz:

$$P_1 = \begin{bmatrix} 1 & a \\ 0 & c \end{bmatrix}$$

es no singular con inversa:

$$P_1^{-1} = \frac{1}{c}\begin{bmatrix} c & -a \\ 0 & 1 \end{bmatrix}$$

Haciendo los cálculos indicados se tiene:

$$P_1^{-1}BP_1 = \frac{1}{c}\begin{bmatrix} c & -a \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & a \\ 0 & c \end{bmatrix} = \begin{bmatrix} 0 & bc-ad \\ 1 & a+d \end{bmatrix}$$

Notemos que la matriz de la derecha es de la forma dada en la ecuación 6.12. Tomado esto como referencia concluimos:

$$\begin{bmatrix} 1 & -n^2-m^2 \\ 0 & 2m \end{bmatrix} = \begin{bmatrix} 1 & bc-ad \\ 0 & a+d \end{bmatrix} \Leftrightarrow m = \frac{a+d}{2} \text{ y } -n^2 - m^2 = bc - ad$$

sustituyendo en esta última ecuación el valor de m y simplificando obtenemos $n^2 = \frac{-(a-d)^2-4bc}{4}$.

Por otro lado, la hipótesis sobre B es equivalente a que el discriminante de su polinomio mínimo sea negativo. Se encuentra que el polinomio mínimo de B es $m(x) = x^2 - (a+d)x + ad - bc$, por lo que B no tiene valores característicos reales $\Leftrightarrow$

$$(a-d)^2 + 4bc < 0 \tag{6.13}$$

Esta última condición garantiza que $n = \frac{\sqrt{-(a-d)^2-4bc}}{2}$ es real.

Definiendo $P = \begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix}$ se tiene:

$$PP_1^{-1}BP_1P^{-1} = \begin{bmatrix} m & -n \\ n & m \end{bmatrix} \tag{6.14}$$

probando que B es similar a la matriz A tratada al inicio y ésta a su vez es similar a una rotación seguida de una homotecia, en resumen, hemos probado el siguiente teorema.

Teorema 6.4 1. *Toda matriz real* 2×2 *con valores característicos no reales es similar a una de la forma* $\begin{bmatrix} m & -n \\ n & m \end{bmatrix}$, *con* $n \neq 0$.

6.4.2. Matrices reales con valores característicos diferentes

Sea A una matriz real $n \times n$, supongamos que $\lambda \in \mathbb{C}$ es un valor característico de A, entonces existe un $X \in \mathbb{C}^n \setminus \{0\}$, tal que $AX = \lambda X$. Definiendo el conjugado de una matriz, entrada por entrada y, denotando el conjugado de A por $\overline{A}$, y tomando en cuenta que A tiene entradas reales, se tiene $\overline{AX} = A\overline{X}$, por otro lado, $\overline{AX} = \overline{\lambda X} = \overline{\lambda}\,\overline{X}$, concluyendo que $A\overline{X} = \overline{\lambda}\,\overline{X}$, es decir, λ tiene asociado al vector característico $\overline{X}$.

Con esto hemos preparado el siguiente teorema:

Teorema 6.4.2. *Sea* $T : \mathbb{R}^n \to \mathbb{R}^n$ *un operador. Supongamos que el polinomio característico de* T, f(x), *tiene raíces de multiplicidad uno. Entonces existe una base de* $\mathbb{R}^n$ *respecto de la que* T *se representa por una matriz de la forma*:

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_l & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & D_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & D_r \end{pmatrix} \tag{6.15}$$

en donde $\lambda_1, \ldots, \lambda_l$ *son las raíces reales de* f(x) *y* $D_j = \begin{bmatrix} m_j & -n_j \\ n_j & m_j \end{bmatrix}$, *con* $m_j, n_j \in \mathbb{R}$ *para todo* $j = 1, 2, \ldots, r$.

Demostración. Aplicando el teorema fundamental del álgebra, podemos suponer que $f(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_l)(x - \mu_1)(x - \bar{\mu}_1)(x - \mu_2)(x - \bar{\mu}_2) \cdots (x - \mu_r)(x - \bar{\mu}_r)$, con $\lambda_j \in \mathbb{R}$ y $\mu_j \in \mathbb{C} \setminus \mathbb{R}$. Sean $X_1, \ldots, X_l, Y_1 \overline{Y_1}, \ldots, Y_r, \overline{Y_r}$, vectores característicos correspondientes a los valores característicos. Note que $X_k \in \mathbb{R}^n$ para todo $k = 1, 2, \ldots, l$ y $Y_t \in \mathbb{C}^n$ para todo $t = 1, 2, \ldots, r$.

Para $j = 1, 2, \ldots, r$ definamos $Z_{2j-1} = \frac{1}{2}(Y_j - \overline{Y_j})$ y $Z_{2j} = \frac{i}{2}(Y_j - \overline{Y_j})$.

Se tiene que $\{X_1, \ldots, X_l, Z_1, Z_2, \ldots, Z_{2r-1}, Z_{2r}\}$ es una base de $\mathbb{R}^n$ (demuéstrelo); incidentalmente note que $n = 2r + l$.

También se tiene $T(X_k) = \lambda_k X_k$ para $k = 1, 2, \ldots, l$,

$$\begin{aligned} T(Z_{2j-1}) &= T(\tfrac{1}{2}(Y_j - \overline{Y_j})) \\ &= \frac{1}{2}T(Y_j + T\overline{Y_j})) \\ &= \frac{1}{2}(\mu_j Y_j + \bar{\mu}\overline{Y_j}) \\ &= \frac{1}{2}[(m_j + in_j)Y_j + (m_j - in_j)\overline{Y_j}] \\ &= \frac{1}{2}[m_j(Y_j + \overline{Y_j}) + in_j(Y_j + \overline{Y_j})] \\ &= m_j Z_{2j-1} - n_j Z_{2j} \end{aligned}$$

para $j = 1, 2, \ldots, r$. Un cálculo como el anterior muestra que $T(Z_{2j}) = n_j Z_{2j-1} + m_j Z_{2j}$.

Poniendo $D_j = \begin{bmatrix} m_j & -n_j \\ n_j & m_j \end{bmatrix}$ y usando la definición de matriz asociada a una transformación en una base se tiene la conclusión del teorema.

6.5. Aplicaciones

En esta sección presentamos algunas aplicaciones de la teoría de valores y vectores característicos. Posiblemente no sean de los más importantes e interesantes; su objetivo es sólo ilustrar la forma de usar la herramienta que hemos desarrollado. El lector tendrá oportunidad de aplicar la herramienta a mejores e interesantes ejemplos, y por qué no, desarrollar los propios.

6.5.1. Especies que interactúan

Las ideas centrales de lo que se presenta aquí fueron tomadas de [3]. Nuestra labor ha consistido en adaptar el caso continuo al discreto, con la finalidad de ilustrar lo que se llama un sistema dinámico lineal discreto y darle un sentido más "real" al caso continuo.

Cuando se analiza la interacción de las diferentes especies vivientes, es claro que el crecimiento de población de una, depende de la de otras. Las relaciones entre las diferentes especies pueden ser de cooperación, competencia y depredación. Para entender estas complejas relaciones se puede iniciar considerando sólo dos especies.

Supongamos que se harán observaciones a intervalos de tiempo determinados para estimar las poblaciones de las dos especies bajo estudio. Los intervalos de tiempo serán numerados: $k = 0, 1, 2, 3$, etcétera. Si denotamos por x_k y y_k a las poblaciones en el tiempo k, nos interesa describir la relación que hay entre las cantidades x_k y y_k, dependiente del tipo de interacción que mantienen las especies.

El modelo que analizaremos es el llamado *depredador y presa*, para una discusión más completa de otras interacciones ver ([3]).

Bajo aislamiento de las especies y con las condiciones apropiadas de supervivencia y crecimiento es razonable suponer que la población en el tiempo $k + 1$ sea proporcional a la que había en el tiempo k. En forma algebraica esto se expresa mediante el sistema de ecuaciones:

$$x_{k+1} = R_1 x_k \tag{6.16}$$

$$y_{k+1} = R_2 y_k \tag{6.17}$$

Suponiendo que las constantes de proporcionalidad, R_1 y R_2, no cambian con el tiempo y agregando que se conocen las poblaciones al momento de iniciar la experimentación, las ecuaciones anteriores se resuelven directamente por sustitución sucesiva, es decir, si x_0 y y_0 son las poblaciones iniciales, entonces $x_1 = R_1 x_0$; $x_2 = R_1 x_1 = R_1 R_1 x_0 = R_1^2 x_0$; en general $x_k = R_1^k x_0$. De manera análoga se obtiene $y_k = R_2^k y_0$.

En general, R_1 y R_2 dependen tanto del tiempo como de las cantidades x_k y y_k. Para modelar la relación depredador presa supondremos que la población x_k corresponde a los depredadores y la población y_k a las presas; $R_1 = -a + by_k$ y $R_2 = c - dx_k$, con a, b, c, y d constantes positivas. Sustituyendo estas expresiones en las ecuaciones 6.16 y 6.17 se tiene:

$$x_{k+1} = (-a + by_k)x_k \tag{6.18}$$

$$y_{k+1} = (c - dx_k)y_k \tag{6.19}$$

y la interpretación de estas nuevas ecuaciones es como sigue:

En la ecuación 6.18 se tiene que sin presas, la población de depredadores decrece, mientras que con presas, el cambio neto en la población es el decrecimiento "natural" más lo que aumenta gracias a las presas. Este aumento está representado por by_kx_k.[5]

En la ecuación 6.19 se tiene que sin depredadores, la población aumenta a tasa constante, mientras que con la presencia de depredadores el cambio neto es lo que aumenta sin depredadores, menos dx_ky_k que se debe a la interacción de los depredadores sobre las presas.

Note que el sistema representado por las ecuaciones 6.18 y 6.19 no es lineal. Para aplicar los métodos lineales introduciremos una hipótesis adicional, supondremos que las ecuaciones se transforman a:

$$x_{k+1} = -ax_k + by_k \tag{6.20}$$

$$y_{k+1} = cy_k - dx_k \tag{6.21}$$

con interpretaciones como sigue:

La ecuación 6.20 se interpreta diciendo que el cambio neto en la población de depredadores se debe a los que mueren naturalmente, más una cantidad proporcional a la cantidad de presas que hacen aumentar la población.

De manera análoga, la ecuación 6.21 significa que la cantidad neta de presas es la diferencia de los que aumentan naturalmente, menos los que son devorados por los depredadores. Esto último se interpreta diciendo que cada depredador consume una cantidad constante de presas por unidad de tiempo.

Con estas consideraciones, las ecuaciones 6.20 y 6.21 se pueden representar en forma matricial.

$$AX_k = X_{k+1} \tag{6.22}$$

En donde $A = \begin{bmatrix} -a & b \\ -d & c \end{bmatrix}$ y $X_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix}$. Suponiendo conocidas las poblaciones al momento de iniciar la experimentación y llamándolas x_0 y y_0 se tiene:

$$X_1 = \begin{bmatrix} -a & b \\ -d & c \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

$$X_2 = \begin{bmatrix} -a & b \\ -d & c \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$= \begin{bmatrix} -a & b \\ -d & c \end{bmatrix} \begin{bmatrix} -a & b \\ -d & c \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

$$= \begin{bmatrix} -a & b \\ -d & c \end{bmatrix}^2 \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

[5] *Ley de acción de masas para una población*. La tasa de cambio de una población debida a la interacción con otra es proporcional al producto de las dos poblaciones.

De manera inductiva se obtiene:

$$X_k = \begin{bmatrix} -a & b \\ -d & -c \end{bmatrix}^k \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \tag{6.23}$$

Para entender la matriz A^k y su efecto en $X_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$, utilizaremos la forma canónica de Jordan. Para tal fin, encontraremos sus valores característicos. El polinomio característico de A es:

$$f(x) = \begin{vmatrix} -a-x & b \\ -d & c-x \end{vmatrix} = x^2 + (a-c)x - ac + bd \tag{6.24}$$

de esto se tiene que los valores característicos de A son:

$$\begin{aligned} x &= \frac{c-a \pm \sqrt{(a-c)^2 - 4(bd-ac)}}{2} \\ &= \frac{c-a \pm \sqrt{(a+c)^2 - 4bd}}{2} \end{aligned}$$

Sean $\lambda_1 = \dfrac{c-a-\sqrt{(a+c)^2-4bd}}{2}$ y $\lambda_2 = \dfrac{c-a+\sqrt{(a+c)^2-4bd}}{2}$ los valores característicos de A. Distinguimos los siguientes casos.

1. Supongamos que $\lambda_1 = \lambda_2 = \dfrac{c-a}{2}$. Si además A fuese diagonalizable, entonces existe P no singular tal que $P^{-1}AP = \lambda_1 I$. De esto se concluye que $A = \lambda_1 I$, lo cual es imposible debido a que las entradas de A no son cero. De lo anterior se tiene que $\lambda_1 = \lambda_2$ implica que A no es diagonalizable, por lo que su forma canónica de Jordan es $J = \begin{bmatrix} \lambda_1 & 0 \\ 1 & \lambda_1 \end{bmatrix}$ y existe una matriz no singular P tal que $P^{-1}AP = J$, en forma equivalente, $A = PJP^{-1}$. De esta última ecuación se tiene que $A^k = PJ^kP^{-1}$. Por inducción se demuestra que $J^k = \begin{bmatrix} \lambda_1^k & 0 \\ k\lambda_1^{k-1} & \lambda_1^k \end{bmatrix}$.

   Sustituyendo estas expresiones en la ecuación 6.23 se tiene:

$$X_k = P \begin{bmatrix} \lambda_1^k & 0 \\ k\lambda_1^{k-1} & \lambda_1^k \end{bmatrix} P \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

   Como $\lambda_1 = \dfrac{c-a}{2}$ las conclusiones sobre el crecimiento de las poblaciones de presas y depredadores dependen sólo de las tasas de crecimiento de cada una de las poblaciones, de acuerdo con las ecuaciones 6.20 y 6.21.

2. Supongamos que $\lambda_1 \neq \lambda_2 \in \mathbb{R}$. En este caso A es diagonalizable, es decir, los vectores característicos de A forman una base de $\mathbb{R}^2$. Denotemos a los vectores característicos de A por V_1 y V_2, entonces existen escalares reales μ_1 y μ_2, tales que $X_0 = \mu_1 V_1 + \mu_2 V_2$. De esta ecuación se tiene:

$$X_k = A^k X_0 = \mu_1 V_1 + \mu_2 V_2 \tag{6.25}$$

Si $|\lambda_1|$, $|\lambda_2| < 1$, las dos especies se extinguirán, pues λ_1^K y λ_2^K se aproximan a cero cuando k se aproxima a infinito. ¿Qué ocurre si $|\lambda_1| = |\lambda_2|$?

3. Si los valores característicos son complejos, entonces uno es conjugado del otro, por lo que sus módulos son iguales. También se tiene que la matriz A es diagonalizable sobre los complejos, es decir, existe una matriz compleja P no singular, cuyas columnas son precisamente los vectores característicos de A cuando se le considera como matriz compleja. De esto se tiene una ecuación análoga a (6.25). De manera más precisa, se tiene la ecuación: $X_0 = \mu_1 V_1 + \mu_2 V_2$, con μ_1, μ_2 como números complejos y V_1 y V_2 vectores característicos complejos asociados a los valores característicos. A partir de esto se tiene:

$$X_k = A^k X_0 = \mu_1 \lambda_1^K V_1 + \mu_2 \lambda_2^K V_2 \tag{6.26}$$

como en el caso real. Ahora, el análisis se hace en términos del módulo de los valores característicos.

- Si $|\lambda_1| = 1$, los vectores X_k giran en una órbita elíptica, por lo que las especies coexisten sin extinción de alguna de ellas. (Explique esto.)
- Si $|\lambda_1| < 1$, los vectores X_k se aproximan a cero, es decir, las dos especies se extinguen.
- Si $|\lambda_1| > 1$, los vectores X_k se "alejan" del origen, es decir, las poblaciones de las especies crecen sin cota. Por supuesto, esto es imposible desde el punto de vista biológico.

6.5.2. Sistemas dinámicos lineales discretos

Hay muchas situaciones en las que el comportamiento de un "estado" soló depende del previo; esto ocurre en la naturaleza, en la sociedad y en muchas otras situaciones.

El siguiente es un caso hipotético de movimiento en las categorías de una asociación deportiva.

La asociación deportiva "Venados de Pisaflores" tiene tres categorías: novatos, profesionales y masters. Los miembros de la asociación deben permanecer exactamente dos años en cada categoría antes de pasar a la siguiente. No hay penalización por abandonar antes de concluir el periodo correspondiente. Los masters deben abandonar la asociación al concluir su periodo. La asociación tiene las siguientes reglas: los novatos no tienen obligación alguna; mientras que cada profesional está obligado a traer una cantidad a de novatos durante los dos años que dura la categoría profesional; asimismo, cada master está obligado a traer b novatos antes de que emigre. Se estima que un porcentaje c de novatos pasará a profesionales, mientras que un porcentaje d de profesionales pasará a masters. Si al iniciar la asociación se tienen x_0, y_0

y z_0 novatos, profesionales y masters, respectivamente, ¿cuántos miembros habrá en cada categoría al cabo de 2, 4, 6, etcétera, años?

Discusión. Tomaremos periodos bianuales y en el i-ésimo, llamemos x_i, y_i y z_i a las cantidades de miembros de la asociación en las categorías novato, profesional y master, respectivamente. La cantidad de novatos al concluir el primer periodo es la suma de los que incorporaron los profesionales y masters, pues ningún novato permanece más de dos años. En símbolos, $x_1 = ay_0 + bz_0$. De igual forma, el número de profesionales será $y_1 = cx_0$ y el de masters será $z_1 = dy_0$. Estas tres ecuaciones se pueden formular en forma matricial.

$$X_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 0 & a & b \\ c & 0 & 0 \\ 0 & d & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \tag{6.27}$$

De forma análoga, para el siguiente periodo de dos años, los miembros en cada categoría están determinados por las cantidades x_1, y_1 y z_1, pues las reglas no han cambiado. En forma simbólica:

$$X_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 & a & b \\ c & 0 & 0 \\ 0 & d & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 0 & a & b \\ c & 0 & 0 \\ 0 & d & 0 \end{bmatrix}^2 \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \tag{6.28}$$

En general, después de $2k$ años y aplicando inducción se tiene:

$$X_k = \begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} = \begin{bmatrix} 0 & a & b \\ c & 0 & 0 \\ 0 & d & 0 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ z_{k-1} \end{bmatrix} = \begin{bmatrix} 0 & a & b \\ c & 0 & 0 \\ 0 & d & 0 \end{bmatrix}^k \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \tag{6.29}$$

Para estudiar el comportamiento de X_k estudiaremos los valores característicos de $A = \begin{bmatrix} 0 & a & b \\ c & 0 & 0 \\ 0 & d & 0 \end{bmatrix}$. Por medio del algoritmo 6.2.1, página 151, se encuentra que el polinomio mínimo de A es $m(x) = x^3 - acx - bdc$.

Para estudiar la naturaleza de las raíces de $m(x)$ se utilizan las fórmulas de Cardano. Una condición necesaria y suficiente para que $m(x)$ tenga raíces diferentes es que $4a^3c - 27b^2d^2 \neq 0$. Sugerimos al lector, a modo de ejercicio, que use esta condición para decidir si A es diagonalizable y partiendo de esto que haga un análisis del comportamiento de X_k.

Ejercicio 6.5.1. *Generalice la situación del ejemplo anterior a un club con* n *diferentes categorías, adaptando las reglas de manera que el proceso se pueda describir mediante una matriz que generalice a la matriz* A *del ejemplo.*

6.5.3. Sistema de masas acopladas con resortes

Consideremos un sistema de n masas interconectadas mediante resortes, colocadas sobre una mesa como se ilustra en la figura 6.2. Supondremos que no hay fuerza de fricción y que la masa de los resortes es despreciable en comparación con las masas. También supondremos que las fuerzas con dirección hacia la izquierda son negativas y aquellas cuya dirección es hacia la derecha son positivas.

A manera de ilustración, iniciamos discutiendo el caso $n = 2$.

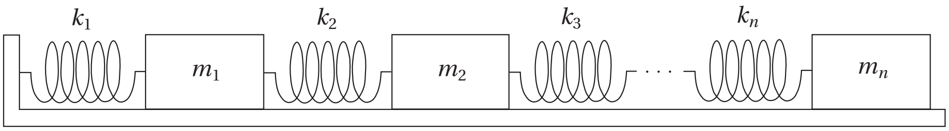


Figura 6.2. Masas acopladas con resortes.

Antes de iniciar la discusión, acordamos que la segunda derivada de y respecto del tiempo la denotaremos por $\ddot{y}$.

A partir de la segunda ley de Newton y de la ley de Hooke tenemos:

$$\begin{aligned} m_1\ddot{y}_1 &= -k_1y_1 + k_2(y_2 - y_1) \\ m_2\ddot{y}_2 &= -k_2(y_2 - y_1) \end{aligned} \tag{6.30}$$

pues en m_1 actúan dos fuerzas, mientras que en m_2 sólo actúa una.

Desde luego estamos asumiendo que m_1 y m_2 no son cero, así que el sistema 6.30 se transforma en:

$$\begin{aligned} \ddot{y}_1 &= -\frac{k_1+k_2}{m_1}\, y_1 + \frac{k_2}{m_1} y_2 \\ \ddot{y}_2 &= -\frac{k_2}{m_2} y_2 + \frac{k_2}{m_2} y_1 \end{aligned} \tag{6.31}$$

Este sistema se puede representar en forma matricial:

$$\ddot{Y} = \begin{bmatrix} -\dfrac{k_1+k_2}{m_1} & \dfrac{k_2}{m_1} \\ -\dfrac{k_2}{m_2} & \dfrac{k_2}{m_2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{6.32}$$

Para resolver el sistema anterior, procederemos a indicar cómo se resuelve el caso general.

Si l_i es la distancia entre m_i y m_{i+1}, en la posición inicial, y l'_i es la distancia al instante t, entonces la deformación del i-ésimo resorte es $l'_i - l_i$. Tenemos las siguientes ecuaciones:

$$\begin{aligned} l_1 &= y_1 + x_1 & l'_1 &= x_1 + y_2 \\ l_2 &= y_2 + x_2 & l'_2 &= x_2 + y_3 \\ &\vdots & &\vdots \\ l_{n-1} &= y_{n-1} + x_{n-1} & l'_{n-1} &= x_{n-1} + y_n \end{aligned} \tag{6.33}$$

de las cuales se obtiene:

$$l'_i = y_{i+1} + l_i - y_{\text{i}} \quad \text{o} \quad l'_i - l_i = y_{i+1} - y_i \tag{6.34}$$

La masa m_1 está bajo la acción de los resortes 1 y 2, por tanto $m_1\ddot{y}_1 = -k_1y_1 + k_2(y_2 - y_1)$; la masa m_2 está bajo la acción de los resorte 2 y 3; la fuerza proveniente del resorte 2 es la opuesta a la que actúa sobre m_1, así que $m_2\ddot{y}_2 = -k_2(y_2 - y_1) + k_3(y_3 - y_2)$, en general, para $2 \leq i \leq n - 1$ tenemos:

$$\begin{aligned} m_i\ddot{y}_i &= -k_i(y_i - y_{i-1}) + k_{i+1}(y_{i+1} - y_i) \\ &= k_iy_{i-1} - (k_i + k_{i+1})\, y_i + k_{i+1}y_{i+1} \end{aligned}$$

Para $i = n$ se tiene $m_n\ddot{y}_n = -k_n(y_n - y_{n-1})$.

Este sistema puede ser escrito en forma matricial: $\ddot{Y} = AY$, en donde A es la matriz:

$$A = \begin{vmatrix} -\dfrac{k_1 + k_2}{m_1} & \dfrac{k_1}{m_1} & 0 & 0 & \cdots & 0 & 0 \\ \dfrac{k_2}{m_2} & -\dfrac{k_2 + k_3}{m_2} & \dfrac{k_3}{m_2} & 0 & \cdots & 0 & 0 \\ 0 & \dfrac{k_3}{m_3} & -\dfrac{k_3 + k_2}{m_3} & \dfrac{k_4}{m_3} & \cdots & 0 & 0 \\ \vdots & & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \dfrac{k_n}{m_n} & -\dfrac{k_n}{m_n} \end{vmatrix}$$

Por el corolario 6.3.1, sabemos que existe una matriz compleja no singular P tal que $P^{-1}AP = \operatorname{diag}\{J_1, \ldots J_k\}$ es la forma canónica de Jordan. Estamos suponiendo que el vector:

$$Y = \begin{bmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{bmatrix}$$

está en $\mathbb{R}^n$. Dado que P es no singular para cada t, existe $Z(t)$ tal que $Y(t) = PZ(t)$, $Z(t) \in \mathbb{C}^n$. Como $Y(t)$ tiene segunda derivada, también $Z(t)$ admite segunda derivada y $\ddot{Y}(t) = \mathrm{P}\ddot{\mathrm{Z}}(t)$. De esto obtenemos $AY = \ddot{\mathrm{Y}}(t) = P\ddot{Z}(t)$, por lo que $APZ = P\ddot{Z}(t)$, concluyendo de lo anterior $P^{-1}APZ = JZ = \ddot{\mathrm{Z}}$.

Como cada matriz J_i es de la forma $J_i = \operatorname{diag}\{J_{i1}, \ldots, J_{i_{s(i)}}\}$ entonces el problema ha sido reducido a resolver el sistema:

$$\begin{pmatrix} c & 0 & 0 & \cdots & 0 & 0 \\ 1 & c & 0 & \cdots & 0 & 0 \\ 0 & 1 & c & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & c & 0 \\ 0 & 0 & 0 & \cdots & 1 & c \end{pmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix} = \begin{bmatrix} \ddot{z}_1 \\ \ddot{z}_2 \\ \ddot{z}_3 \\ \vdots \\ \ddot{z}_{n-1} \\ \ddot{z}_n \end{bmatrix} \tag{6.35}$$

el cual es equivalente a

$$\ddot{z}_1 = cz_1 \tag{6.36}$$

$$\ddot{z}_2 = z_1 + cz_2 \tag{6.37}$$

$$\ddot{z}_3 = z_2 + cz_3 \tag{6.38}$$

$$\vdots$$

$$\ddot{z}_n = z_{n-1} + cz_n$$

y este sistema puede ser resuelto fácilmente resolviendo 6.36, luego 6.37 y así sucesivamente.

Observación 6.5.1. *Si* A *es diagonalizable, el problema es más sencillo; se reduce a resolver un sistema en donde cada ecuación es de la forma* $\ddot{z} = cz$. *El nuevo sistema se llama no acoplado*.

Observación 6.3.2. *La matriz del sistema original se llama tridiagonal y existe toda una teoría para estudiar este tipo de matrices.*

6.6. Ejercicios

1. Calcule el polinomio mínimo de las siguientes matrices:

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

b) $A = (a_{ij})$, en donde $a_{ij} = \begin{cases} 1 & \text{si } j = i + 1,\ i = 1, 2, \ldots, n - 1 \\ 0 & \text{en otro caso} \end{cases}$

2. ¿Cómo deben ser las raíces del polinomio mínimo de un operador para que sea inversible?

3. Sea A una matriz $n \times n$, λ un escalar tal que $A - \lambda I$ es no singular. Pongamos $B = (A - \lambda I)^{-1}$. Demuestre que B es diagonalizable $\Leftrightarrow$ A lo es. Más aún, A y B se diagonalizan simultáneamente. Sugerencia: si $\{X_1, X_2, \ldots, X_n\}$ es una base de vectores propios de A, también será una base de vectores propios de B. Para probar la implicación recíproca, note que los valores característicos de B son no cero. Si X es un vector característico de B con valor característico μ, entonces X es vector característico de A con valor característico $\frac{1 + \lambda\mu}{\mu}$.

4. Sea A una matriz $n \times n$, cuyas entradas son todas iguales a uno. Demuestre que los únicos valores característicos de A son n y 0.
5. Sea A una matriz inversible con polinomio mínimo $m(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1x + a_0$. ¿Cuál es el polinomio mínimo de A^{-1}? ¿Cómo representa a A^{-1} en términos de A? Si A es diagonalizable, ¿qué puede decir al respecto de A^{-1}?
6. Suponga que la matriz A es diagonalizable. ¿Es A^k diagonalizable para todo entero positivo k? Si A^k es diagonalizable para algún entero positivo k, ¿es A diagonalizable?
7. Sean, A una matriz diagonalizable con valores característicos $\lambda_1, \lambda_2, \ldots, \lambda_n$ y $q(x)$ un polinomio no constante. Definimos la matriz $B := q(A)$. ¿Es diagonalizable la matriz B? Si su respuesta es afirmativa, ¿a cuál matriz diagonal es similar B?
8. Sea A matriz 6×6 con polinomio mínimo $x^2(x - 3)^3$. Determine las posibles formas canónicas de Jordan de A.
9. Sean T y T_1 dos operadores en V que conmutan y son diagonalizables. Demuestre que existe una base de V en la cual ambos se representan por matrices diagonales.
10. Demuestre que la matriz del ejercicio 1*b* es similar a su transpuesta. Sugerencia: A puede ser considerada la matriz de un operador T respecto a la base $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, entonces $T(\alpha_i) = \alpha_{i+1}$ para $i = 1, 2, \ldots, n - 1$ y $T(\alpha_n) = 0$.
11. Use el ejercicio anterior para demostrar que toda matriz es similar a su transpuesta.
12. Sea A una matriz diagonalizable. Demuestre que todos los valores característicos de A son iguales $\Leftrightarrow A = cI$, para algún c.
13. Sea $p(x)$ un polinomio mónico. ¿Existen matrices no similares que tengan a $p(x)$ por polinomio mínimo?
14. ¿Existen matrices simétricas reales que tengan por polinomio mínimo a $x^4 - 1$?
15. Sean A y B matrices nilpotentes de orden 3. Demuestre que son similares $\Leftrightarrow$ tienen el mismo polinomio mínimo.
16. Si los polinomios mínimo y característico de A son $(x - 3)^2$ y $(x - 3)^3$ respectivamente. Determine las posibles formas canónicas de Jordan de A.
17. Sea $g(x)$ el polinomio en la ecuación 6.4, página 136. Demuestre que $g(x)$ tiene grado $\leq n^2$ y que $g(A) = 0$, en donde A es la matriz que se usa para construir a $g(x)$.
18. Sea $p(x)$ un polinomio con coeficientes enteros, irreducible y mónico. Si A es una matriz $n \times n$ cuyo polinomio mínimo es $p(x)$, conteste las siguientes preguntas.

    a) ¿Puede ocurrir que A no tenga entradas enteras?

    b) ¿Es A diagonalizable?

    c) ¿Cuál debe ser el grado de $p(x)$ para que A pueda ser simétrica?

    d) ¿Puede ocurrir que $p(x)$ tenga grado 13?

19. Sea A una matriz diagonalizable, $\{X_1, X_2, ..., X_n\}$ una base de $\mathbb{R}^n$ que consiste de vectores característicos de A y P la matriz cuyas columnas son los vectores $X_1, X_2, ..., X_n$. Demuestre que $P^{-1}AP = \text{diag}\{\lambda_1, \lambda_2, ..., \lambda_n\}$, en donde $AX_i = \lambda_i X_i$.

20. Sea A una matriz $n \times n$, λ un escalar no cero y $B = \lambda A$. Suponga que el polinomio mínimo de A es $m(x) = x^r + a_{r-1}x^{r-1} + \cdots + a_1 x + a_0$. Demuestre que el polinomio mínimo de B es $m_1(x) = x^r + \lambda a_{r-1}x^{r-1} + \cdots + \lambda^{r-1}a_1 + \lambda^r a_0$. ¿Qué relación existe entre las raíces de $m(x)$ y las de $m_1(x)$?

21. Sea D la matriz del ejercicio 22, página 76 y $A = \begin{pmatrix} 9 & 7 & 4 & 1 \\ 7 & \frac{19}{2} & 7 & 4 \\ 4 & 7 & \frac{19}{2} & 7 \\ 1 & 4 & 7 & 9 \end{pmatrix}$. Note que $D = \frac{1}{10^3}A$. Use *Maple* en el ejercicio anterior para calcular los valores característicos de D, conociendo los de A.

22. En relación con el ejercicio 6.5.1 haga una discusión completa del caso $n = 2$, es decir, la asociación deportiva tiene solamente dos categorías.

23. Sea $A = \begin{bmatrix} .25 & .95 \\ .75 & .05 \end{bmatrix}$ la matriz que describe la movilidad de población en el ejemplo 1.1.3, página 15. Demuestre que el polinomio mínimo de A es $m(x) = (x - 1)\left(x + \frac{7}{10}\right)$. Encuentre una matriz P tal que $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{7}{10} \end{bmatrix}$. De esta ecuación se tiene: $A^k = P\begin{bmatrix} 1 & 0 \\ 0 & -\left(\frac{7}{10}\right)^k \end{bmatrix}P^{-1}$. Use esta representación de A^k para hacer un análisis de la movilidad de población en el citado ejemplo.

24. En relación con el ejemplo anterior note que la suma de las entradas de cada columna en la matriz A es uno; también se tiene que un valor característico de A es $\lambda = 1$. Suponga que A es una matriz $n \times n$ con entradas ≥ 0 tales que la suma de los elementos de cada columna es uno. Demuestre que uno de los valores característicos de A es $\lambda = 1$. Matrices del tipo anterior se les llama *matrices estocásticas*, por representar procesos estocásticos discretos con un número finito de estados. Algunos autores llaman matriz estocástica a la transpuesta de A. Con este lenguaje, la interpretación del valor característico igual a uno es que hay un estado estacionario, es decir, si los estados de un proceso estocástico se representan por X, entonces se tiene $AX = X$, para algún X.

25. Demuestre que el producto de matrices estocásticas es estocástica. En particular, si A es una matriz estocástica, entonces A^k es estocástica para todo entero $k \geq 1$.

26. Suponga que A es una matriz estocástica, no singular 2×2. ¿Cuánto suman las entradas de cada columna de A^{-1}? ¿Es estocástica A^{-1}?

27. Suponga que las características laborales de una persona pueden ser: profesional, obrero calificado u obrero no calificado. Suponga que de los hijos de un profesional, 80% son profesionales, 10% son obreros calificados y 10% son no calificados. Para el caso de hijos de un obrero calificado, 60% son profesionales, 20% son obreros calificados y 20% son no calificados. Finalmente, 50% de los hijos de obreros no calificados son no calificados, solamente 25% son profesionales y el otro 25% son obreros calificados. Suponga que cada persona tiene al menos un hijo. Use una matriz estocástica para seguir la profesión de un hijo de familia a través de varias generaciones. Encuentre la probabilidad de que, eligiendo al azar un nieto de un obrero no calificado sea un profesional.

28. ¿Puede usar las ideas del ejercicio anterior para modelar la propagación de una enfermedad contagiosa en una población dada? Explique.

Capítulo 7

Espacios con producto interno

Los espacios euclidianos $\mathbb{R}^2$ y $\mathbb{R}^3$ son los que por lo común utilizamos para representar problemas que surgen de situaciones del mundo real *macro*. Una de las características de éstos es lo concerniente a sus propiedades geométricas, las cuales se precisan por medio de los conceptos fundamentales de ángulo y distancia. Usando la estructura de espacio vectorial, más precisamente el concepto de producto interno en estos espacios, se pueden precisar dichos conceptos, los cuales resultan de gran utilidad en muchos problemas que admiten una representación geométrica. Dado que $\mathbb{R}^n$ y los espacios vectoriales abstractos son una generalización de $\mathbb{R}^2$ y $\mathbb{R}^3$, es natural preguntarse: ¿qué se requiere para introducir en estos espacios los conceptos que permitan definir distancia y ángulo? La respuesta se obtiene a través de la formulación del concepto de producto interno.

7.1. Aspectos geométricos de un espacio vectorial

En esta sección definimos el concepto de producto interno y demostramos algunos de los resultados de gran importancia, tales como el teorema de Pitágoras, la desigualdad de Cauchy-Schwarz, entre otros.

Definición 7.1.1. *Diremos que el espacio vectorial real* V *es euclidiano si hay definida una función* $\langle\ ,\ \rangle : \mathrm{V} \times \mathrm{V} \to \mathbb{R}$ *que satisfaga*:

1. *Para todos* $\alpha, \beta, \in \mathrm{V}, \langle\alpha, \beta\rangle = \langle\beta, \alpha\rangle$.
2. *Para todos* $\alpha, \beta, \gamma \in \mathrm{V}$ *y para todos* $\mathrm{a}, \mathrm{b} \in \mathbb{R}, \langle \mathrm{a}\alpha + \mathrm{b}\beta, \gamma\rangle = \mathrm{a}\langle\alpha, \gamma\rangle + \mathrm{b}\langle\beta, \gamma\rangle$.
3. *Para todo* $\alpha \in \mathrm{V}, \langle\alpha, \alpha\rangle \geq 0$, y $\langle\alpha, \alpha\rangle = 0$ *solamente si* $\alpha = 0$.

Ejemplo 7.1.1. *Sea* V *el espacio de las matrices* $\mathrm{n} \times \mathrm{n}$. *Definimos en* $\mathrm{V} \times \mathrm{V}$ *la función* $\langle\ ,\ \rangle$ *determinada por la ecuación*: $\langle \mathrm{A}, \mathrm{B}\rangle := \mathrm{tr}(\mathrm{AB}^t)$. *Mostraremos que esta función hace de* V *un espacio euclidiano*.

Usando la definición de transpuesta de una matriz, es directo demostrar que la función propuesta satisface las dos primeras propiedades de la definición 7.1.1, por lo que justificaremos sólo la tercera. Si A tiene entradas a_{ij} y C denota la matriz AA^t, entonces la entrada c_{ii} de C se obtiene mu1tip1icando la fila i-ésima de A por la columna i-ésima de A^t, es decir, $c_{ii} = a_{i1}^2 + a_{i2}^2 + \cdots + a_{in}^2$, entonces:

$$\langle A, A\rangle = \mathrm{tr}(AA^t) = \sum_{i,j=1}^{n} a_{ij}^2 \geq 0$$

por lo que $\langle A, A\rangle = 0 \Leftrightarrow$ cada $a_{ij} = 0$.

Ejemplo 7.1.2. *Sea* V *el espacio de las funciones continuas en el intervalo* [0, 1]. *Definiendo en* V × V *la función,* ⟨ , ⟩ *dada por* $\langle f, g \rangle = \int_0^1 fg$, *mostraremos que* V *es un espacio euclidiano*.

Como en el ejemplo anterior, la propiedad que requiere ser verificada es la tercera, es decir, se debe justificar que $\langle f, f \rangle = \int_0^1 f^2 \geq 0$, con igualdad $\Leftrightarrow f = 0$. Un resultado de cálculo garantiza que si una función es continua en x_0 y $f(x_0) > 0$, entonces $f(x) > 0$ para todo x en un intervalo con centro en x_0. Usando esto y el hecho que si c es un real entre a y b entonces $\int_a^b f = \int_a^c f + \int_c^b f$, se concluye que $\langle f, f \rangle = \int_0^1 f^2 = 0 \Leftrightarrow f = 0$.

Definición 7.1.2. *Sea* V *un espacio euclidiano*:

1. *Dados* $\alpha, \beta \in$ V, *diremos que son ortogonales si* $\langle \alpha, \beta \rangle = 0$.
2. *Dado* $\alpha \in$ V, *definimos* $\|\alpha\| := \sqrt{\langle \alpha, \alpha \rangle}$ *y le llamamos la norma de* α.

Observación 7.1.1. *Si* a *es un escalar, entonces* $\|a\alpha\| = |a|\,\|\alpha\|$.

Demostración. Por definición, $\|a\alpha\| := \sqrt{\langle a\alpha, a\alpha \rangle} = \sqrt{a^2 \langle \alpha, \alpha \rangle} = |\alpha| \sqrt{\langle \alpha, \alpha \rangle}$.

Teorema 7.1.1 (Teorema de Pitágoras). *Sea* V *un espacio euclidiano*, $\alpha, \beta \in$ V. *Entonces* α *es ortogonal a* $\beta \Leftrightarrow \|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2$.

El aspecto geométrico de este teorema se ilustra en la figura 7.1.

Demostración. De la definición de norma de un vector se tiene: $\|\alpha + \beta\|^2 = \langle \alpha + \beta, \alpha + \beta \rangle = \|\alpha\|^2 + 2\langle \alpha, \beta \rangle + \|\beta\|^2$. Por definición, α es ortogonal a $\beta \Leftrightarrow \langle \alpha, \beta \rangle = 0$. La conclusión se tiene de esto.

Teorema 7.1.2. (Desigualdad de Cauchy-Schwarz). *Si* V *es un espacio euclidiano, y* $\alpha, \beta \in$ V, *entonces* $|\langle \alpha, \beta \rangle| \leq \|\alpha\|\,\|\beta\|$. Con igualdad $\Leftrightarrow \alpha$ *y* β *son linealmente dependientes.*

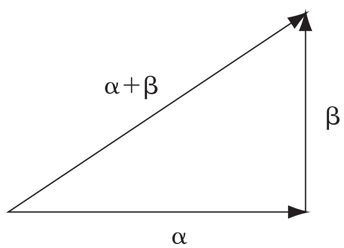


Figura 7.1. Teorema de Pitágoras.

Demostración. Dados $\alpha, \beta \in V$ fijos y cualquier $x \in \mathbb{R}$ se tiene que:

$$f(x) := \langle \alpha + x\beta, \alpha + x\beta \rangle = \|\alpha\|^2 + 2\|\langle \alpha, \beta \rangle x + \|\beta\|^2 x^2 \qquad (7.1)$$

define una función cuadrática no negativa para todo $x \in \mathbb{R}$, por lo que su discriminante d, satisface $d := 4\langle \alpha, \beta \rangle^2 - 4\|\alpha\|^2\,\|\beta\|^2 \leq 0$, y tiene un cero doble $\Leftrightarrow d = 0$. En forma equivalente, esto último ocurre $\Leftrightarrow 4\langle \alpha, \beta \rangle^2 - 4\|\alpha\|^2\|\beta\|^2 = 0$ y esto a su vez ocurre $\Leftrightarrow |\langle \alpha, \beta \rangle| = \|\alpha\|\,\|\beta\|$. Entonces $f(x) = \langle \alpha + x\beta, \alpha + x\beta \rangle = 0 \Leftrightarrow \alpha + x\beta = 0 \Leftrightarrow |\langle \alpha, \beta \rangle| = \|\alpha\|\,\|\beta\|$. De un argumento como el anterior se tiene $f(x) = \langle \alpha + x\beta, \alpha + x\beta \rangle > 0 \Leftrightarrow 4\langle \alpha, \beta \rangle^2 - 4\|\alpha\|^2\|\beta\|^2 < 0$, es decir, α y β son linealmente independientes $\Leftrightarrow |\langle \alpha, \beta \rangle| < \|\alpha\|\,\|\beta\|$.

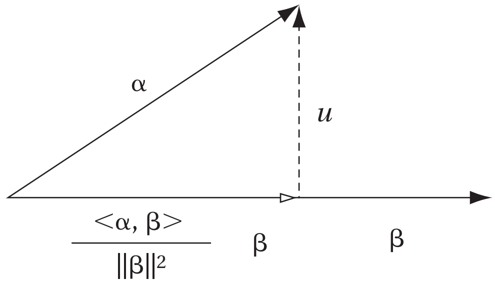


Figura 7.2. Proyección ortogonal de un vector sobre otro.

Otra demostración. La idea de esta demostración se ilustra en la figura 7.2. Por la definición 3.3.4, página 96, sabemos que $\alpha = u + \frac{\langle \alpha, \beta \rangle}{\|\beta\|^2}\beta$, con u y $\frac{\langle \alpha, \beta \rangle}{\|\beta\|^2}\beta$ ortogonales.

Aplicando el teorema de Pitágoras y la observación 7.1.1 se tiene:

$$\|\alpha\|^2 = \|u\|^2 + \left\|\frac{\langle \alpha, \beta \rangle}{\|\beta\|^2}\beta\right\|^2 \geq \left\|\frac{\langle \alpha, \beta \rangle}{\|\beta\|^2}\beta\right\|^2 = \frac{|\langle \alpha, \beta \rangle|^2}{\|\beta\|^4}\|\beta\|^2 = \frac{|\langle \alpha, \beta \rangle|^2}{\|\beta\|^2}$$

Por lo que $\|\alpha\|^2 \geq \frac{|\langle \alpha, \beta \rangle|^2}{\|\beta\|^2}$, y de esto la conclusión se obtiene tomando raíz cuadrada y pasando el denominador al primer miembro de la desigualdad.

Nótese que hay igualdad $\Leftrightarrow$ $u = 0$, y esto ocurre $\Leftrightarrow$ α coincide con su proyección sobre β, es decir, α y β son linealmente dependientes.

La ventaja de la segunda demostración es que no utiliza las propiedades de orden en los números reales, por lo que se puede aplicar a espacios vectoriales complejos con producto interno.

Teorema 7.1.3. (Desigualdad del triángulo). *Sea* V *un espacio euclidiano*, $\alpha, \beta \in$ V. *Entonces* $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$.

Demostración. Se tiene $\|\alpha + \beta\|^2 = \langle \alpha + \beta, \alpha + \beta \rangle = \|\alpha\|^2 + 2\langle \alpha, \beta \rangle + \|\beta\|^2 \leq \|\alpha\|^2 + 2|\langle \alpha, \beta \rangle| + \|\beta\|^2$. Usando la desigualdad de Cauchy-Schwarz en la expresión de la derecha se tiene:

$$\|\alpha + \beta\|^2 \leq \|\alpha\|^2 + 2\|\alpha\|\,\|\beta\| + \|\beta\|^2 = (\|\alpha\| + \|\beta\|)^2$$

La conclusión se obtiene tomando raíz cuadrada en esta desigualdad.

Definición 7.1.3. *Sea* V *un espacio euclidiano. Un subconjunto* S *de* V *se dice ortonormal si* $\langle \alpha, \beta \rangle = 0$ *para todos* $\alpha, \beta \in$ S, *con* $\alpha \neq \beta$ y $\langle \alpha, \alpha \rangle = 1$ *para todo* $\alpha \in$ S. *Cuando* S *es base de* V, *se dice que* S *es base ortonormal*.

Teorema 7.1.4. (Proceso de Gram-Schmidt). *Todo espacio euclidiano de dimensión finita admite una base ortonormal*.

Demostración. Sea V un espacio euclidiano de dimensión finita; procederemos a demostrar el teorema por inducción sobre la dimensión de V. Denotemos a la dimensión de V por n. Si $n = 1$ y α es un generador de V, el elemento $\frac{\alpha}{\|\alpha\|}$ tiene norma 1.

Para ilustrar la construcción de la base, discutiremos con detalle el caso $n = 2$.

Supongamos que $\{\alpha_1, \alpha_2\}$ es una base de V, definamos $\beta_1 = \frac{\alpha_1}{\|\alpha_1\|}$ de esto se tiene que $\{\beta_1, \alpha_2\}$ es una base. Deseamos construir $\beta_2 \neq 0$ de forma que $\langle \beta_1, \beta_2 \rangle = 0$, para esto iniciamos construyendo un elemento $\gamma_2 \neq 0$, que sea ortogonal a β_1, normalizándolo obtendremos el elemento buscado. Proponemos $\gamma_2 = b_2\alpha_2 + b_1\beta_1$; claramente podemos suponer que $b_2 = 1$. De la condición $\langle \gamma_2, \beta_1 \rangle = 0$, se tiene $0 = \langle \alpha_2 + b_1\beta_1, \beta_1 \rangle = \langle \alpha_2, \beta_1 \rangle + b_1\langle \beta_1, \beta_1 \rangle$. Como β_1 tiene norma 1, entonces $b_1 = -\langle \alpha_2, \beta_1 \rangle$. Por construcción, $\gamma_2 = \alpha_2 - \langle \alpha_2, \beta_1 \rangle \beta_1$ es ortogonal a β_1. Definiendo $\beta_2 = \frac{\gamma_2}{\|\gamma_2\|}$ se tiene que $\{\beta_1, \beta_2\}$ es la base deseada.

Supongamos que $n > 2$ y que se han construido $\beta_1, \beta_2, \ldots, \beta_r$ elementos ortonormales que generan el mismo subespacio que $\alpha_1, \alpha_2, \ldots, \alpha_r$, entonces $\{\beta_1, \beta_2, \ldots, \beta_r, \alpha_{r+1}\}$ genera el mismo subespacio que $\{\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha_{r+1}\}$. Construiremos un elemento β_{r+1} tal que $\langle \beta_{r+1}, \beta_i \rangle = 0$, para todo $i = 1, 2, \ldots, r$ y de forma que el conjunto $\{\beta_1, \beta_2, \ldots, \beta_r, \beta_{r+1}\}$ genere el mismo subespacio que $\{\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha_{r+1}\}$.

Si γ_{r+1} es un elemento en el subespacio generado por $\{\beta_1, \beta_2, \ldots, \beta_r, \alpha_{r+1}\}$, entonces $\gamma_{r+1} = b_{r+1}\alpha_{r+1} + b_r\beta_r + \cdots + b_1\beta_1$. Podemos suponer que $b_{r+1} = 1$, adicionalmente, si $\langle \gamma_{r+1}, \beta_i \rangle = 0$ para todo i, entonces $0 = \langle \alpha_{r+1}, \beta_i \rangle + b_i\langle \beta_i, \beta_i \rangle$. Como cada β_i tiene norma 1, entonces $b_i = -\langle \alpha_{r+1}, \beta_i \rangle$, es decir, se tiene:

$$\gamma_{r+1} = \alpha_{r+1} - \langle \alpha_{r+1}, \beta_r \rangle \beta_r - \cdots - \langle \alpha_{r+1}, \beta_1 \rangle \beta_1 \tag{7.2}$$

Definimos $\beta_{r+1} = \frac{\gamma_{r+1}}{\|\gamma_{r+1}\|}$, con esto se tiene que el conjunto $\{\beta_1, \beta_2, \ldots, \beta_r, \beta_{r+1}\}$ es ortonormal y genera el mismo subespacio que $\{\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha_{r+1}\}$.

Definición 7.1.4. *Sea* V *un espacio euclidiano,* W *un subespacio. Definimos el complemento ortogonal de* W *como* $W^\perp := \{\alpha \in V : \langle \alpha, \beta \rangle = 0$ *para todo* $\beta \in W\}$.

Ejercicio 7.1.1. *Demuestre que* $W^\perp$ es un subespacio.

Teorema 7.1.5. *Sea* V *un espacio euclidiano de dimensión finita,* W *un subespacio, entonces* $V = W \oplus W^\perp$.

Demostración. Si $W = \{0\}$ o $W = V$ el resultado es inmediato pues se tiene $W^\perp = V$, o $W^\perp = \{0\}$ respectivamente. Supongamos que W tiene dimensión positiva y menor que la de V. Sea $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ una base ortonormal de W. Por el proceso de Gram-Schmidt podemos completar a una base ortonormal de V, sea esta base $\{\alpha_1, \alpha_2, \alpha_r, \ldots, \alpha_n\}$. Afirmamos que $W^\perp$ tiene por base a $\{\alpha_{r+1}, \ldots, \alpha_n\}$. Es claro que $\alpha_i \in W^\perp$ para todo $i = r + 1, \ldots, n$. Todo $\gamma \in W^\perp$ se puede expresar como $\gamma = c_1\alpha_1 + \cdots + c_r\alpha_r + c_{r+1}\alpha_{r+1} + \cdots + c_n a_n$.

Para cada $i = 1, 2, \ldots, r$ se tiene $0 = \langle \alpha_i, \gamma \rangle = c_i$, entonces γ pertenece al subespacio generado por $\{\alpha_{r+1}, \ldots, \alpha_n\}$, probando con esto lo afirmado y también el teorema.

Si V es un espacio euclidiano de dimensión finita y W un subespacio de V, entonces para todo $\alpha \in V$ se tiene $\alpha = \hat{\alpha} + \beta$, con $\hat{\alpha} \in W$ y $\beta \in W^\perp$. Al vector $\hat{\alpha}$ le llamaremos la proyección ortogonal de α sobre W. La figura 7.3 ilustra esta situación.

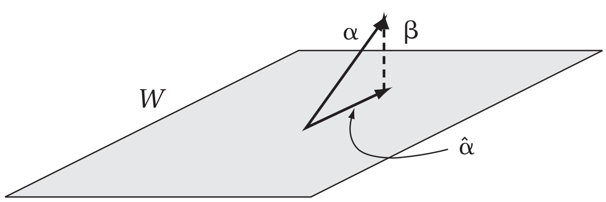


Figura 7.3. Proyección ortogonal de un vector sobre un subespacio.

Si $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ es una base ortonormal de W, entonces:

$$\hat{\alpha} = \langle \alpha, \alpha_1\rangle\alpha_1 + \langle \alpha, \alpha_2\rangle\alpha_2 + \cdots + \langle \alpha, \alpha_r\rangle\alpha_r$$

En efecto, pues si $\hat{\alpha} = c_1\alpha_1 + \cdots + c_r\alpha_r$, tomando el producto interno de $\hat{\alpha}$ con α_i se tiene $\langle \hat{\alpha}, \alpha_i\rangle = c_i$.

Teorema 7.1.6. *Sea* V *un espacio euclidiano de dimensión finita,* W *un subespacio de* V, $\alpha \in$ V *y* $\hat{\alpha}$ *la proyección de* α *sobre* W, *entonces:*

$$\|\alpha - \hat{\alpha}\| < \|\alpha - \beta\|$$

para todo $\beta \in W\setminus\{\hat{\alpha}\}$.

Demostración. Sea $\beta \in W\setminus\{\hat{\alpha}\}$. Se tiene $\alpha - \beta = \alpha - \beta + \hat{\alpha} - \hat{\alpha} = (\alpha - \hat{\alpha}) + (\hat{\alpha} - \beta)$ y $(\alpha - \hat{\alpha}) \in W^{\perp}$. Aplicando el teorema de Pitágoras y tomando en cuenta que $\beta \neq \hat{\alpha}$ obtenemos:

$$\|\alpha - \beta\|^2 = \|\alpha - \hat{\alpha}\|^2 + \|\hat{\alpha} - \beta\|^2 > \|\alpha - \hat{\alpha}\|^2$$

La conclusión final se logra tomando raíz cuadrada en la desigualdad anterior, por supuesto usando que la función raíz cuadrada es monótona.

7.1.1. Método de mínimos cuadrados

Cuando se tiene un sistema de ecuaciones $AX = B$ que no tiene solución, una forma de abordar el problema puede consistir en encontrar la mejor aproximación, es decir, encontrar un vector $\hat{X}$ tal que $\|A\hat{X} - B\|$ sea lo más pequeño posible. ¿Qué significa esto en términos más precisos?

Definición 7.1.5. *Sea* A *una matriz* $m \times n$, B *un vector columna en* $\mathbb{R}^m$. *Una solución por mínimos cuadrados de* AX = B *es un* $\hat{X}$ *en* $\mathbb{R}^n$ tal que:

$$\|A\hat{X} - B\| \leq \|AX - B\|$$

para todo X en $\mathbb{R}^n$.

Recordemos que el sistema $AX = B$ tiene solución $\Leftrightarrow$ B se encuentra en el espacio columna de A. De manera más precisa, si $A_1, A_2, \ldots, A_n$ representan a las columnas de A, entonces el sistema $AX = B$ es equivalente a la ecuación $x_1A_1 + x_2A_2 + \cdots + x_nA_n = B$, en donde $x_1, x_2, \ldots, x_n$ son las entradas de X y esta última ecuación es la que establece que B pertenece al espacio columna de A siempre y cuando el sistema tenga solución. Entonces encontrar un $\hat{X}$ que haga mínimo $\|A\hat{X} - B\|$ equivale a encontrar un vector $\hat{B}$ en el espacio columna de A que satisfaga $\|\hat{B} - B\|$ es mínimo. Si llamamos W al espacio columna de A, entonces aplicando el teorema 7.1.6 se tiene que el vector que satisface la condición de minimalidad es precisamente la proyección ortogonal de B en W.

¿Cómo encontrar $\hat{B}$? Dado que $\hat{B}$ pertenece al espacio columna de A, el sistema $AX = \hat{B}$ tiene solución, llamemos $\hat{X}$ a esta solución.

Por otro lado, el vector $B - \hat{B}$ se encuentra en el complemento ortogonal del espacio columna de A, teorema 7.1.5. Entonces, $B - \hat{B} = B - A\hat{X}$ es ortogonal a todas las columnas de A, lo cual equivale a $A^t(B - A\hat{X}) = 0$, o en forma alternativa:

$$A^tA\hat{X} = A^tB \tag{7.3}$$

Nótese que las soluciones de (7.3) corresponden exactamente a las soluciones del problema de mínimos cuadrados formulado en la definición 7.1.5, el cual siempre tiene solución.

Se tiene que la matriz A^tA es una matriz $n \times n$ y por lo observado antes, el sistema (7.3) tiene solución única $\Leftrightarrow$ la matriz A^tA es inversible. El siguiente teorema establece condiciones necesarias y suficientes para que esto ocurra.

Teorema 7.1.7. *Sea* A *una matriz* m $\times$ n. *Entonces la matriz* $\mathrm{A^tA}$ *es inversible* $\Leftrightarrow$ *las columnas de* A *son linealmente independientes.*

Demostración. Recordemos que el sistema $AX = 0$ tiene sólo la solución cero las columnas de A son linealmente independientes. Para concluir la prueba del teorema, demostraremos que las soluciones de $A^tAX = 0$ coinciden con las de $AX = 0$.

Es claro que si X es solución de $AX = 0$, también lo es de $A^tAX = 0$. Si $A^tAX = 0$ entonces $X^tA^tAX = X^t0 = 0$. La parte de la izquierda en esta última ecuación es el producto punto de AX consigo mismo, por lo que $AX = 0$, terminando la prueba.

Otro resultado que ayudará a resolver el problema de mínimos cuadrados es el siguiente.

Teorema 7.1.8. *Sea* A *una matriz* m $\times$ n *cuyas columnas, como elementos de* $\mathbb{R}^m$, *son linealmente independientes. Entonces* A = QR, *donde* Q *es una matriz cuyas columnas son ortonormales y* R *es una matriz triangular superior, cuyos elementos de la diagonal principal son positivos.*

Demostración. Sea V el espacio generado por las columnas de A. Por el proceso de Gram-Schmidt, existe una base ortonormal $\{B_1, B_2, ..., B_n\}$ de V, más precisamente, si las columnas de A son denotadas $A_1, A_2, ..., A_n$ y haciendo las identificaciones correspondientes, de la ecuación 7.2 se tiene:

$$C_i = A_i - \langle A_i, B_1\rangle B_1 - \cdots - \langle A_i, B_{i-1}\rangle B_{i-1} \tag{7.4}$$

Con la notación adecuada, $B_i = \dfrac{C_i}{\|C_i\|}$; de esto se concluye que $C_i = \|C_i\|B_i$. Sustituyendo esta expresión en (7.4) y despejando A_i tenemos:

$$A_i = c_{ii}B_i + c_{i-1i}B_{i-1} + \cdots + c_{1i}B_1 \tag{7.5}$$

con c_{ii} positivo. En términos matriciales la ecuación 7.5 dice que la columna i de A es el producto de la matriz que tiene por columnas a $B_1, B_2, ..., B_n$ y la transpuesta de $[c_{1i}\ \ c_{2i} \cdots c_{ii}\ \ 0 \cdots 0]$; en resumen $A = QR$, con Q la matriz cuyas columnas son los vectores $B_1, B_2, ..., B_n$ y R la matriz triangular cuya i-ésima columna es la transpuesta de $[c_{1i}\ \ c_{2i} \cdots c_{ii}\ \ 0 \cdots 0]$, terminando la demostración.

Supongamos que las columnas de la matriz A son linealmente independientes y sea $A = QR$ la descomposición del teorema 7.1.8, entonces el sistema $A^tA\hat{X} = A^tB$ equivale a $R^tQ^tQR\hat{X} = R^tQ^tB$. Como las entradas de la diagonal de R son positivas, entonces R y R^t son inversibles. La condición sobre Q implica que $Q^tQ = \mathrm{I}$, por lo que la ecuación anterior se reduce a $R\hat{X} = Q^tB$ o a $\hat{X} = R^{-1}Q^tB$. Desde el punto de vista computacional, la ecuación $R\hat{X} = Q^tB$ tiene ventajas, dado que se trata de un sistema triangular y resolverlo se logra mediante sustitución regresiva.

7.2. Espacios vectoriales complejos

Al abordar varios problemas de ciencias e ingeniería es de gran utilidad considerar espacios vectoriales en donde los escalares son números complejos, razón por la cual hace falta precisar la terminología y resultados básicos que se utilizan en este caso.

Recordemos que el conjunto de los números complejos se denota usualmente por $\mathbb{C}$ y a sus elementos los representamos en la forma $z = a + bi$, en donde i es una raíz fija de la ecuación $x^2 + 1 = 0$, es decir, i satisface $i^2 = -1$. Dado un número complejo $z = a + bi$, se definen las partes real e imaginaria de z como $\operatorname{Re}(z) := a$ e $\operatorname{Im}(z) := b$ respectivamente. También se define el conjugado de z, denotado $\bar{z} := a - bi$.

Una de las propiedades fundamentales de los números complejos es el *teorema fundamental del álgebra*.

Otro aspecto que es importante señalar es que los números reales se consideran como subconjunto de $\mathbb{C}$, haciendo la parte imaginaria de un complejo igual a cero.

La definición de espacio vectorial complejo es exactamente la misma que la de espacio vectorial real, definición 3.5.1, página 91, salvo que los escalares son números complejos. El ejemplo típico de espacio vectorial real es $\mathbb{R}^n$; en el caso complejo lo análogo a $\mathbb{R}^n$ es $\mathbb{C}^n$.

Todos los conceptos fundamentales y resultados de los espacios vectoriales reales se establecen para el caso de espacios vectoriales complejos, haciendo notar sólo que los escalares son números complejos.

Es importante notar que si V es un espacio vectorial complejo, también es un espacio vectorial real; el recíproco no es cierto. Por ejemplo, $\mathbb{R}^n$ no es un espacio vectorial complejo. Otro aspecto que se debe enfatizar es el cambio de dimensión al considerar escalares reales y complejos. Por ejemplo, en $\mathbb{C}^n$ el conjunto:

$$A = \{(1, 0, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, 0, 0, \dots, 0, 1)\}$$

es una base cuando los escalares son complejos, pues cualquier elemento de $\mathbb{C}^n$ es combinación lineal, de manera única, de los elementos de A. Sin embargo, cuando se consideran escalares reales el conjunto A sigue siendo linealmente independiente pero no genera, pues el elemento $(i, 0, \dots, 0)$ no se puede obtener como combinación lineal, con escalares reales, de los elementos de A. De hecho ninguno de los elementos $(0, \dots, 0, i, 0, \dots, 0)$ se puede expresar como combinación lineal de los elementos de A, con escalares reales. Esta observación sugiere proponer el conjunto:

$$B = \{(i, 0, 0, \dots, 0), (0, i, 0, \dots, 0), \dots, (0, 0, 0, \dots, 0, i)\}$$

para tratar de completar a una base de $\mathbb{C}^n$ sobre los reales.

Invitamos al lector a demostrar que el conjunto $A \cup B$ es una base de $\mathbb{C}^n$ sobre los reales. Con esto se tiene: dimensión de $\mathbb{C}^n = 2n$, cuando se considera sobre los reales.

Si A es una matriz con entradas complejas, definimos la conjugada de A, denotada $\bar{A}$, como la matriz que se obtiene de A conjugando los elementos de sus entradas.

Por ejemplo, si:

$$A = \begin{bmatrix} 1+i & 3-i \\ 2-i & 1+i \end{bmatrix}, \text{ entonces } \bar{A} = \begin{bmatrix} 1-i & 3+i \\ 2+i & 1-i \end{bmatrix}$$

En particular, se define el conjugado de un vector columna Y de $\mathbb{C}^n$ como el vector que se obtiene de Y conjugando sus entradas.

Otro aspecto que se debe hacer notar es la forma en que se define un producto interno en un espacio vectorial complejo. Esto se precisa en la siguiente definición.

Definición 7.2.1. *Diremos que el espacio vectorial complejo* V *tiene producto interno si hay una función* $\langle\, , \rangle : V \times V \to \mathbb{C}$ *que satisfaga*:

1. $\langle \alpha, \beta \rangle = \overline{\langle \beta, \alpha \rangle}$ *para todos* $\alpha, \beta \in \mathrm{V}$.
2. $\langle a\alpha + b\beta, \gamma \rangle = a\langle \alpha, \gamma \rangle + b\langle \beta, \gamma \rangle$, *para todos* $a, b \in \mathbb{C}$ *y para todos* $\alpha, \beta, \gamma \in \mathrm{V}$.
3. $\langle \alpha, \alpha \rangle, \geq 0$, *para todo* $\alpha \in \mathrm{V}$ *y* $\langle \alpha, \alpha \rangle = 0$ *solamente si* $\alpha = 0$.

Observación 7.2.1. *Note que las dos primeras propiedades en la definición anterior implican* $\langle \alpha, \lambda\beta \rangle = \lambda\langle \alpha, \beta \rangle$, *para todos* $\alpha, \beta \in \mathrm{V}$ *y para todo* $\lambda \in \mathbb{C}$. *Esta propiedad coincide con la del caso real, pues el conjugado de un real es el mismo.*

Ejemplo 7.2.1. *Sea* $\mathrm{V} = \mathbb{C}^n$; *si* $\mathrm{X}, \mathrm{Y} \in \mathbb{C}^n$ *se consideran como vectores columna, entonces la función* $\langle\ ,\ \rangle : \mathrm{V} \times \mathrm{V} \to \mathbb{C}^n$ *definida por* $\langle \mathrm{X}, \mathrm{Y} \rangle := \overline{Y}^t\ \mathrm{X}$ *representa un producto interno en* V.

Si A es una matriz $n \times n$ entonces $\langle AX, Y \rangle := \overline{Y}^t AX$. Usando una propiedad de la transpuesta y de la operación de conjugación de complejos se tiene:

$$\langle AX, Y \rangle = \langle X, \overline{A}^t\ Y \rangle \tag{7.6}$$

Los conceptos de ortogonalidad y norma se establecen igual que en el caso real.

7.3. Formas cuadráticas y bilineales

Por el ejercicio 13, del capítulo 4, página 114, sabemos que una transformación lineal f de $\mathbb{R}^n$ a los reales es de la forma $f(x_1, x_2, ..., x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$, en donde a_1, a_2, ..., a_n son números reales fijos. Notemos que f se puede representar en términos del producto interno en $\mathbb{R}^n$, es decir, si hacemos $X = (x_1, x_2, ..., x_n)$ y $A = (a_1, a_2, ..., a_n)$ entonces $f(x_1, x_2, ..., x_n) = f(X) = \langle X, A \rangle = a_1x_1 + a_2x_2 + \cdots + a_nx_n$. Esta interpretación de f muestra la relación que existe entre las funciones lineales de $\mathbb{R}^n$ a los reales y el producto interno. De manera más general, si V es un espacio con producto interno y $\beta \in V$ es un vector fijo, se puede definir una transformación lineal $T_\beta : V \to \mathbb{R}$ como sigue:

$$T_\beta(\alpha) = \langle \alpha, \beta \rangle$$

Las propiedades de linealidad de T_β equivalen a las propiedades de linealidad en la primera entrada de la función producto interno. Para definir el concepto de función bilineal tomaremos como punto de partida las propiedades del producto interno.

Definición 7.3.1. *Una función bilineal en el espacio vectorial real* V, *es una función* $\mathcal{B}: \mathrm{V} \times \mathrm{V} \to \mathbb{R}$ *que satisface las propiedades:*

1. *Para todos* $a, b \in \mathbb{R}$ *y para todos* α, β *y* γ *elementos de* V, $\mathcal{B}(a\alpha + b\beta, \gamma) = a\mathcal{B}(\alpha, \gamma) + b\mathcal{B}(\beta, \gamma)$.
2. *Para todos* $b, c \in \mathbb{R}$ *y para todos* α, β *y* γ *elementos de* V, $\mathcal{B}(\alpha, b\beta + c\gamma) = b\mathcal{B}(\alpha, \beta) + c\mathcal{B}(\alpha, \gamma)$.

Una forma alternativa de expresar que $\mathcal{B}$ es bilineal es considerándola como función de dos variables, lineal en cada una.

Cuando $\mathcal{B}$ satisface $\mathcal{B}(\alpha, \beta) = \mathcal{B}(\beta, \alpha)$ para todos α y β en V, se dice que $\mathcal{B}$ es una función bilineal simétrica.

Observación 7.3.1. *Note que todo producto interno en un espacio vectorial real es una función bilineal, pues un producto interno requiere las condiciones de simetría y positividad además de las de linealidad.*

Ejemplo 7.3.1. *Si* V *es el espacio de las funciones Riemann integrables en el intervalo* $[a, b]$, *entonces la función* $\mathcal{B} : \mathrm{V} \times \mathrm{V} \to \mathbb{R}$ *determinada por* $\mathcal{B}(f, g) := \int_a^b fg$ *es bilineal simétrica.*

Discusión. Sean $f, g, h \in V$ y $c, d \in \mathbb{R}$. Por las propiedades de la integral se tiene $\mathcal{B}(af + bg, h) := \int_a^b (cf + dg)h = c\int_a^b f\,h + d\int_a^b gh$, probando que es lineal en la primera entrada. La linealidad en la segunda se obtiene de la conmutatividad del producto de funciones y de la igualdad anterior. Es importante notar que $\mathcal{B}$, no es un producto interno, pues es posible que $\mathcal{B}(f, f) = 0$ sin que f sea la función cero. Proporcione un ejemplo de tal situación.

Sea V un espacio de dimensión finita, $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ una base de V y $\mathcal{B}$ una función bilineal. Dados $\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$ y $\beta = y_1\alpha_1 + \cdots + y_n\beta_n$ y usando las propiedades de bilinealidad de $\mathcal{B}$, se tiene:

$$
\begin{aligned}
\mathcal{B}(\alpha,\beta) &= \mathcal{B}(x_1\alpha_1 + \cdots + x_n\alpha_n, y_1\alpha_1 + \cdots + y_n\alpha_n)\\
&= x_1\mathcal{B}(\alpha_1, y_1\alpha_1 + \cdots + y_n\alpha_n) + \cdots + x_n\mathcal{B}(\alpha_n, y_1\alpha_1 + \cdots + y_n\alpha_n)\\
&= x_1y_1\mathcal{B}(\alpha_1, \alpha_1) + \cdots + x_1y_n\mathcal{B}(\alpha_1, \alpha_n) + \cdots +\\
&\quad + x_ny_1\mathcal{B}(\alpha_n, \alpha_1) + \cdots + x_ny_n\mathcal{B}(\alpha_n, \alpha_n)\\
&= [\, x_1 \cdots x_n \,] \begin{bmatrix} B(\alpha_1,\alpha_1) & \cdots & B(\alpha_1,\alpha_n)\\ B(\alpha_2,\alpha_1) & \cdots & B(\alpha_2,\alpha_n)\\ \vdots & \vdots & \vdots\\ B(\alpha_n,\alpha_1) & \cdots & B(\alpha_n,\alpha_n) \end{bmatrix} \begin{bmatrix} y_1\\ \vdots\\ y_n \end{bmatrix}
\end{aligned}
$$

Definiendo la matriz A, cuyas entradas son los elementos $\mathcal{B}(\alpha_i, \alpha_j)$, y acordando que X y Y son matrices con una columna y n filas con entradas x_i y y_i, respectivamente; la expresión para $\mathcal{B}(\alpha, \beta)$ se representa en la forma:

$$\mathcal{B}(\alpha, \beta) = X^tAY \tag{7.7}$$

Note que la expresión de la derecha en la ecuación 7.7 depende de la base elegida en V y le llamaremos la forma bilineal de $\mathcal{B}$ respecto a la base $\{\alpha_1, \alpha_2 ..., \alpha n\}$.

Sea $\{\alpha_1', \alpha_2', ..., \alpha_n'\}$ otra base de V, con P la matriz de cambio de base y sean X' y Y' los vectores coordenados de α y β respecto a esta nueva base, entonces de acuerdo con la ecuación 4.14, página 108 se tiene $X = PX'$ y $Y = PY'$. Si C denota la matriz asociada a $\mathcal{B}$ respecto a la base $\{\alpha_1', \alpha_2', ..., \alpha_n'\}$, se tiene:

$$\mathcal{B}(\alpha, \beta) = X'^t CY'^t = X^t(P^{-1})^tCP^{-1}Y = X^tAY \tag{7.8}$$

por lo que necesariamente, $A = (P^{-1})^tCP^{-1}$. Esto último equivale a:

$$C = P^tAP. \tag{7.9}$$

Resumiendo la discusión anterior se ha demostrado el siguiente teorema:

Teorema 7.3.1. *Sea* V *un espacio vectorial de dimensión finita y* $\mathcal{B}$ *una función bilineal en* V. *Si* A *y* C *son matrices representando a* $\mathcal{B}$ *en las bases* $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ *y* $\{\alpha_1', \alpha_2', ..., \alpha_n'\}$, *respectivamente, y la matriz de cambio de base está representada por* P, *entonces* $C = P^tAP$.

Definición 7.3.2. *Sean* A *y* C *matrices* $n \times n$. *Se dice que* C *es congruente con* A, *si existe una matriz inversible* P *tal que* $C = P^tAP$.

Observación 7.3.2. *Dos matrices* $n \times n$ *representan a la misma función bilineal* $\Leftrightarrow$ *son congruentes.*

Ejemplo 7.3.2. *Sea* $\mathcal{B} : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ *expresada por:*

$$\mathcal{B}((x_1, x_2, x_3), (y_1, y_2, y_3)) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_3 + x_3y_1 + x_3y_3$$

Encuentre la expresión para $\mathcal{B}$ respecto a la base $\{(1, 1, 0), (1, 1, 1), (1, 0, 0)\}$

Discusión. La representación dada para $\mathcal{B}$ es respecto a la base canónica, y la matriz de $\mathcal{B}$ respecto a esta base se obtiene usando la ecuación 7.7, notando que sus entradas son los coeficientes de los términos $x_i y_j$, en donde el índice i representa el número de fila mientras que j representa el número de columna. De acuerdo con esto obtenemos que la matriz de $\mathcal{B}$ respecto de la base canónica es: $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

Por otro lado, la matriz de cambio de base es $P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Aplicando el teorema 7.3.1 se tiene que la matriz de $\mathcal{B}$ respecto a la nueva base es:

$$P^t AP = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 3 \\ 5 & 7 & 4 \\ 3 & 3 & 2 \end{bmatrix}$$

Si las coordenadas de un vector respecto a la base $\{(1, 1, 0), (1, 1, 1), (1, 0, 0)\}$ las denotamos por (x_1', x_2', x_3'), entonces se tiene que $\mathcal{B}$ está dada por:

$$\begin{aligned} \mathcal{B}((x_1', x_2', x_3'), (y_1', y_2', y_3')) &= [x_1' \ x_2' \ x_3'] \begin{bmatrix} 4 & 5 & 3 \\ 5 & 7 & 4 \\ 3 & 3 & 2 \end{bmatrix} \begin{bmatrix} y_1' \\ y_2' \\ y_3' \end{bmatrix} \\ &= 4\,x_1' y_1' + 5\,x_1' y_2' + 3\,x_1' y_3' + 5\,x_2' y_1' + 7\,x_2' y_2' + 4\,x_2' y_3' \\ &\quad + 3\,x_3' y_1' + 3\,x_3' y_2' + 2\,x_3' y_3' \end{aligned}$$

El ejemplo anterior muestra que la representación de $\mathcal{B}$ puede ser más sencilla en una base que en otra. Con esto en mente surge una pregunta, ¿existe una base de V de tal forma que la representación de una función bilineal dada sea suficientemente sencilla? Más precisamente, ¿se puede representar a una función bilineal por medio de una matriz diagonal? Note que una respuesta afirmativa equivale a decir que existe una base $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ tal que:

$$\mathcal{B}\ (\alpha_i, \alpha_j) = \begin{cases} a_i & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \tag{7.10}$$

De existir tal base, se tendría una especie de análogo del proceso de ortogonalización de Gram-Schmidt. Algo adicional a notar es que las condiciones de la ecuación 7.10 implican que $\mathcal{B}$ es simétrica, pues si C es la matriz asociada a $\mathcal{B}$ respecto a otra base, entonces $\text{diag}\{a_1, a_2, ..., a_n\} = P^tCP$, con P la matriz de cambio de base. Tomando transpuesta en esta ecuación y considerando que la transpuesta de una matriz diagonal es ella misma, se tiene $\text{diag}\{a_1, a_2, ..., a_n\} = P^tC^tP$, luego $P^tC^tP = P^tCP$. Ahora, utilizando que P tiene inversa se concluye que $C^t = C$.

Las observaciones del párrafo anterior demuestran que para la existencia de la base con las condiciones de la ecuación 7.10 es necesario que $\mathcal{B}$ sea simétrica. ¿Será suficiente la condición de simetría en $\mathcal{B}$ para que exista tal base?

Primero, si $\mathcal{B}$ es la función idénticamente cero, es decir, si $\mathcal{B}(\alpha, \alpha) = 0$ para todo $\alpha \in V$, entonces $\mathcal{B}$ se representa por medio de una matriz diagonal.

Por la consideración hecha, podemos suponer que existe $\alpha_1 \in V \setminus \{0\}$ tal que $\mathcal{B}(\alpha_1, \alpha_1) \neq 0$. Como estamos buscando elementos "ortogonales", es natural considerar a los elementos $\beta \in V$ tales que $\mathcal{B}(\alpha, \beta) = 0$ y de allí escoger algunos que formen parte de la base buscada.

Note que el conjunto de los $\beta \in V$ que satisfacen $\mathcal{B}(\alpha, \beta) = 0$ es el núcleo de la función lineal $f_{\alpha_1} : V \to \mathbb{R}$ definida por $f_{\alpha_1}(\beta) = \mathcal{B}(\alpha_1, \beta)$. Como f_{α_1} no es la función cero, entonces es suprayectiva, por lo que su imagen tiene dimensión uno. Aplicando el teorema 4.3.2, página 103, se concluye que la dimensión del núcleo de f_{α_1} es $\dim (V) - 1$.

Llamemos W al núcleo de f_{α_1}. Es claro que $\mathcal{B}$ puede ser considerada una función bilineal en W. Un argumento inductivo sobre la dimensión del espacio sobre el que está definida $\mathcal{B}$ terminará la demostración de la existencia de la base buscada. De manera más precisa. Si V tiene dimensión uno, claramente cualquier matriz asociada a $\mathcal{B}$ es diagonal. Supongamos que V tiene dimensión mayor que uno y supongamos que cuando $\mathcal{B}$ está definida en subespacios de dimensión menor que $\dim (V)$ existe una base en la cual se representa por medio de una matriz diagonal.

Por lo demostrado antes, existe una base de W, $\{\alpha_2, \alpha_3, ..., \alpha_n\}$, tal que $\mathcal{B}(\alpha_i, \alpha_j) = \mathcal{B}(\alpha_j, \alpha_i) = 0$, si $i \neq j$ y para todos $i = 2, 3, ..., n$ y $j = 2, 3, ..., n$.

Como α_1 satisface $\mathcal{B}(\alpha_1, \alpha_1) \neq 0$, entonces $\alpha_1 \notin W$. De esto concluimos que $\{\alpha_1, \alpha_2, \alpha_3, ..., \alpha_n\}$, es una base de V.

Por construcción de W se tiene $\mathcal{B}(\alpha_1, \alpha_j) = 0$ para todo $j = 2, 3, ..., n$. Como $\mathcal{B}$ es simétrica también se tiene $\mathcal{B}(\alpha_j, \alpha_1) = 0$ para todo $j = 2, 3, ..., n$, concluyendo con esto que $\{\alpha_1, \alpha_2, \alpha_3, ..., \alpha_n\}$ es la base buscada.

Resumiendo la discusión anterior, hemos demostrado el siguiente teorema:

Teorema 7.3.2. *Sea* V *un espacio vectorial de dimensión finita,* $\mathcal{B}$ *una función bilineal. Se tiene lo siguiente: existe una representación de* $\mathcal{B}$ *por medio de una matriz diagonal* $\Leftrightarrow$ $\mathcal{B}$ *es simétrica.*

Observación 7.3.3. *Si* V *es un espacio vectorial y* $\mathcal{B}$ *es una función bilineal simétrica, entonces existe una base* $\{\alpha_1, \alpha_2, \alpha_3, ..., \alpha_n\}$ *de* V *tal que* $\mathcal{B}(\alpha, \beta) = a_1x_1y_1 + \cdots + a_nx_ny_n$, *en donde* $\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$ y $\beta = y_1\alpha_1 + y_2\alpha_2 + \cdots + y_n\alpha_n$.

7.3.1. Formas cuadráticas

Si $\mathcal{B}$ es una forma bilineal en el espacio V representada por la matriz A respecto a una base, entonces $\mathcal{B}(\alpha, \alpha) = X^tAX$, con X el vector coordenado de α respecto a la base dada. La expresión X^tAX es lo que usualmente se llama una *forma cuadrática*. Nosotros tomaremos un enfoque un poco diferente en donde la formulación no hace uso de base alguna; esto para ir acorde con la definición de función bilineal formulada al inicio de la sección.

Definición 7.3.3. *Sea* V *un espacio vectorial, una función* $Q : V \to \mathbb{R}$ *es una función cuadrática si cumple las siguientes condiciones.*

1. *Para todo* $\alpha \in V$ *y para todo* $a \in \mathbb{R}$, $Q(a\alpha) = a^2Q(\alpha)$.
2. *La función* $\mathcal{B}(\alpha, \beta) = \dfrac{Q(\alpha+\beta)-Q(\alpha-\beta)}{4}$ es bilineal.

Observación 7.3.4. *Note que de la segunda parte de la definición 7.3.3 y de la discusión anterior se tiene que una función cuadrática se puede representar mediante una matriz simétrica, pues la función bilineal expresada allí es simétrica.*

Como consecuencia del teorema 7.3.2 y de la observación anterior tenemos:

Teorema 7.3.3. *Sea* V *un espacio de dimensión* n. *Si* Q *es una función cuadrática en* V, *entonces existe una base de* V *respecto de la cual se tiene*: $Q(\alpha) = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_n x_n^2$, *con* $x_1, x_2, \ldots, x_n$, *las coordenadas de* α *respecto a dicha base*.

Cuando una función cuadrática Q se ha representado por medio de una matriz A respecto a una base y el vector coordenado de α respecto de esta base es X, a la expresión $Q(X) = X^tAX$ se le suele llamar *forma cuadrática*.

Notemos que los signos de los valores de una forma cuadrática Q representada por $Q(\alpha) = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_n x_n^2$ se pueden examinar considerando los signos de los coeficientes $d_1, d_2, \ldots, d_n$. Por ejemplo, si todos los d_i son positivos, entonces $Q(\alpha) > 0$ para todo $\alpha \neq 0$; si todos los d_i son negativos, entonces $Q(\alpha) < 0$ para todo $\alpha \neq 0$. Estas observaciones llevan a la siguiente definición:

Definición 7.3.4. *Una forma cuadrática* $Q(X)$ *es positiva* (*negativa*) *definida si* $Q(X) > 0$ ($Q(X) < 0$) *para todo* $X \neq 0$.

Notemos que la definición de forma cuadrática positiva definida se puede formular en términos de una matriz que la represente. Esta matriz es necesariamente simétrica. Más adelante serán consideradas con mayor amplitud las matrices positivas definidas.

7.3.2. Teorema de los ejes principales

Lema 7.3.1. *Sea* A *una matriz simétrica con entradas reales. Entonces las raíces de su polinomio característico son reales.*

Demostración. Sea λ un valor característico de A y X un vector característico asociado a λ. Como A es real y simétrica, aplicando la ecuación 7.6, página 178, se tiene $\langle AX, X\rangle = \langle X, AX\rangle$. De esta ecuación y la propiedad del producto interno complejo, se tiene $\lambda\langle X, X\rangle = \bar{\lambda}\langle X, X\rangle$; o en forma equivalentemente, $(\lambda - \bar{\lambda})\langle X, X\rangle = 0$. Como X no es cero, entonces $\langle X, X\rangle \neq 0$, por lo que $(\lambda - \bar{\lambda}) = 0$, es decir, λ es real.

Lema 7.3.2. *Sean* A *una matriz simétrica con entradas reales;* μ *y* λ *valores característicos de* A, *diferentes. Si* X_1 *y* X_2 *son vectores característicos correspondientes a* μ *y* λ *respectivamente, entonces* X_1 *es ortogonal a* X_2.

Demostración. De la ecuación 7.6 y la hipótesis sobre A se tiene $\langle AX_1, X_2\rangle = \langle X_1, AX_2\rangle$. Usando la definición de vector característico, la linealidad del producto interno y agrupando, obtenemos $(\mu - \lambda)\langle X_1, X_2\rangle = 0$. La conclusión se tiene de la hipótesis sobre μ y λ.

Lema 7.3.3. *Sea* B *una matriz simétrica con entradas reales. Supongamos que* $B^mX = 0$ *para algún entero positivo* m, *entonces* $BX = 0$.

Demostración. Sea m el mínimo entero positivo tal que $B^mX = 0$. Si $m > 1$ entonces $B^{m-1}X = Y \neq 0$ y $BY = 0$. De esto se tiene $0 = \langle BY, B^{m-2}X\rangle = \langle Y, B^{m-1}X\rangle = \langle Y, Y\rangle$, contradiciendo lo supuesto sobre Y.

Definición 7.3.5. *Sea* P *una matriz real no singular, diremos que* P *es ortogonal si* $P^{-1} = P^t$.

Observación 7.3.5. *Una matriz* P *es ortogonal* $\Leftrightarrow$ *sus columnas son ortonormales.*

Teorema 7.3.4. (Teorema de los ejes principales). *Sea* A *una matriz simétrica. Entonces existe una matriz ortogonal* P *tal que* $P^{-1}AP = P^tAP = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$, *con* $\lambda_1, \ldots, \lambda_n$ *los valores característicos de* A.

Demostración. Primero demostraremos que A es diagonalizable, lo cual equivale a demostrar que el polinomio mínimo de A tiene factores lineales diferentes, teorema 6.2.1. Sea $m(x) = p(x)^l q(x)$ el polinomio mínimo de A, con $p(x)$ irreducible y primo relativo a $q(x)$. Por el lema 7.3.1, $p(x)$ es lineal. Para concluir esta parte de la demostración, basta demostrar que $l = 1$. Sea $X \in \mathbb{R}^n$, entonces $m(A)X = p(A)^l q(A)X = 0$. Es directo verificar que $p(A)$ es una matriz simétrica. Aplicando el lema 7.3.3 a $p(A)$ se tiene que $p(A)q(A)X = 0$, es decir, $p(A)q(A)$ anula a cualquier X de $\mathbb{R}^n$, lo cual implica que $p(A)q(A)$ es la matriz cero. Entonces $m(x)$ divide a $p(x)q(x)$, concluyéndose que $l = 1$.

Sea $m(x) = (x - \lambda_1) \cdots (x - \lambda_r)$ la factorización en irreducibles diferentes y W_1, W_2, ..., W_r los núcleos de los operadores $A - \lambda_1 I$, $A - \lambda_2 I$, ..., $A - \lambda_r I$. Aplicando el teorema de la descomposición primaria, teorema 6.1.3, página 140, se tiene que:

$$\mathbb{R}^n = W_1 \oplus W_2 \oplus \cdots \oplus W_r$$

Sea $\mathcal{B}_i$ una base ortonormal de W_i. Por el lema 7.3.2, la unión de $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_r$ es una base ortonormal de $\mathbb{R}^n$. Defina P como la matriz cuyas columnas son los elementos de esta base y note que P^tAP puede tener elementos repetidos en la diagonal, de hecho cada λ_i aparece tantas veces como $\dim(W_i)$.

Corolario 7.3.1. *La única matriz nilpotente y simétrica es la matriz cero.*

Ejercicio 7.3.1. *Sea* A *una matriz real* $n \times n$. *Demuestre que todos los valores característicos de* A *son reales* $\Leftrightarrow$ *existen:* Q *matriz ortogonal y* R *matriz triangular superior tales que* $A = QRQ^t$.

Sugerencia. Si $A = QRQ^t$ entonces $Q^tAQ = R$, es decir, A es similar a R y los valores característicos de R son los elementos de su diagonal. Recíprocamente, suponga que los valores característicos de A son todos reales. Sea U_1 un vector unitario asociado al valor característico λ_1, es decir, $A\,U_1 = \lambda_1\,U_1$. Expandiendo U_1 a una base ortonormal de $\mathbb{R}^n$ se construye una matriz ortogonal Q_1 cuyas columnas son esta base, entonces $Q_1^t AQ$ es una matriz que en la primer entrada de la primer columna tiene a λ_1 y cero en las restantes. Los valores característicos de la submatriz que se obtiene omitiendo la primera fila y la primera columna de $Q_1^t AQ$ tiene por valores característicos a λ_2, ..., λ_n. Haga construcciones adecuadas y proceda inductivamente.

Nota. Un corolario de este ejercicio es el teorema de los ejes principales.

7.3.3. Matrices positivas definidas

Uno de los problemas importantes en cálculo de una variable es la clasificación de los puntos críticos de una función. Esto se hace de manera eficiente si la función tiene suficientes derivadas, más precisamente, si $f(x)$ es una función que tiene segunda derivada en un punto x_0 de su dominio y $f''(x_0) \neq 0$, entonces el signo de $f''(x_0)$ determina si se trata de un máximo o un mínimo. Para funciones de $\mathbb{R}^n$ a $\mathbb{R}$ hay un resultado similar, salvo que en este caso, la segunda diferencial es una función bilineal simétrica determinada por una matriz simétrica y el signo de los valores característicos de esa matriz determinan si la función tiene un máximo o un mínimo. En esta parte discutiremos algunas propiedades de las matrices simétricas que tienen valores característicos positivos.

Definición 7.3.6. *Sea* A *una matriz simétrica; diremos que* A *es positiva definida si* $X^tAX > 0$ *para todo* $X \neq 0$.

Teorema 7.3.5. *Sea* A *una matriz simétrica, entonces* A *es positiva definida* $\Leftrightarrow$ *sus valores característicos son positivos.*

Demostración. Supongamos que todos los valores característicos, λ_1, λ_2, ..., λ_n de A son positivos. Por el teorema de los ejes principales, teorema 7.3.4, existe P matriz ortogonal tal que $P^tAP = \operatorname{diag}\{\lambda_1, \lambda_2, ..., \lambda_n\} = D$.

Dado $X \neq 0$, como P es inversible existe $Y \neq 0$ tal que $X = PY$. De esto se tiene:

$$\begin{aligned} X^tAX &= (PY)^tAPY \\ &= Y^tP^tAPY = Y^tDY \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 > 0 \end{aligned}$$

probando que A es positiva definida.

Recíprocamente, si $\lambda i < 0$ para algún i, tomando E_i igual al vector columna que tiene uno en la entrada i y cero en las restantes se tiene $PE_i = Y_i \neq 0$, pues P es inversible. De esto se tiene $Y_i^t AY_i = E_i^t P^tAPE_i = E_i^t DE_i = \lambda_i < 0$, contradiciendo que A es positiva definida; esto prueba que los valores característicos de A deben ser positivos.

El teorema que se ha probado tiene una restricción, para decidir si una matriz simétrica es positiva definida es indispensable conocer sus valores característicos, lo cual puede ser un problema en sí mismo. El siguiente teorema puede funcionar como un método práctico para producir matrices positivas, además de que evita el conocer los valores característicos.

Teorema 7.3.6. *Sea* A *una matriz, entonces* A *es positiva definida* $\Leftrightarrow$ *existe una matriz no singular* B *tal que* $A = B^tB$.

Demostración. Supongamos que A es positiva definida, por el teorema 7.3.5 los valores característicos de A son positivos; por el teorema de los ejes principales existe P ortogonal tal que $P^tAP = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, con $\lambda_i > 0$ para todo $i = 1, 2, \ldots, n$. Pongamos $D = \text{diag}\left\{\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}\right\}$, entonces $P^tAP = D^2$ o $A = PDDP^t = (DP^t)^t(DP^t)$. Definiendo $B = DP^t$ se tienen las condiciones requeridas, es decir, B es no singular y $A = B^tB$.

Recíprocamente, si $A = B^tB$, con B no singular, entonces $BX \neq 0$ para todo $X \neq 0$. De esto se tiene $X^tAX = X^tB^tBX = (BX, BX) > 0$, probando que A es positiva definida.

Si se conoce el polinomio característico de una matriz simétrica, el siguiente criterio es muy útil para decidir si es positiva definida. En la demostración usaremos la regla de los signos de Descartes.

Teorema 7.3.7. (Descartes). *Sea* $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x_1 + a_0$ *un polinomio con coeficientes reales. Si* m *denota el número de cambios de signo en los coeficientes de* f(x) *y* r *el número de raíces positivas de* f(x). *Entonces* $m = r + 2h$, *con* h *entero no negativo*.

Teorema 7.3.8. *Sea* A *una matriz simétrica con polinomio característico* $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x_1 + a_0$. *Entonces* A *es positiva definida* $\Leftrightarrow a_ia_{i-1} < 0$ *para todo* $i = 1, 2, \ldots, n$.

Demostración. Como A es simétrica, entonces todas las raíces de $f(x)$ son reales; de esto, $f(x)$ tiene solamente raíces positivas $\Leftrightarrow a_0 \neq 0$ y $f(-x)$ no tiene raíces positivas. Por la regla de los signos de Descartes, $f(-x)$ no tiene raíces positivas $\Leftrightarrow a_0 \neq 0$ y sus coeficientes tienen el mismo signo que a_0; esto último $\Leftrightarrow a_0 \neq 0$ y $(-1)^ia_i(-1)^{i-1}a_{i-1} > 0$, equivalentemente, $f(-x)$ no tiene raíces positivas o cero $\Leftrightarrow a_0 \neq 0$ y $-a_ia_{i-1} > 0 \Leftrightarrow a_0 \neq 0$ y $a_ia_{i-1} < 0$.

Ejemplo 7.3.3. *Sea* $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, *con* a, b *y* c *reales. El polinomio característico de* A *es* $f(x) = x^2 - (a + c)x + ac - b^2$. *Aplicando el teorema anterior tenemos que* f(x) *tiene raíces positivas* $\Leftrightarrow a + c > 0$ *y* $ac - b^2 > 0$; $\Leftrightarrow a > 0$ *y* $ac - b^2 > 0$. *Este criterio de positividad para matrices* 2×2 *es útil cuando* A *representa a la segunda diferencial de funciones de* $\mathbb{R}^2$ a $\mathbb{R}$.

Clasificación de cónicas

La ecuación general de segundo grado en las variables x y y se puede representar como la suma de una forma cuadrática más una lineal, más precisamente, la ecuación:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{7.11}$$

se puede representar en la forma:

$$[\,x\;\;y\,]\begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + dx + ey + f = 0 \tag{7.12}$$

Haciendo $A = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}$ y aplicando el teorema de los ejes principales, teorema 7.3.4, existe una matriz ortogonal P tal que:

$$P^tAP = \text{diag}\{\lambda_1, \lambda_2\} \tag{7.13}$$

con λ_1 y λ_2 los valores característicos de A.

Poniendo $\begin{bmatrix} x \\ y \end{bmatrix} = P\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} p_{11}x' + p_{12}y' \\ p_{21}x' + p_{22}y' \end{bmatrix}$, y de lo anterior se tiene:

$$\begin{aligned} ax^2 + bxy + cy^2 + dx + ey + f &= [\,x\;\;y\,]\begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + dx + ey + f \\ &= \lambda_1 x'^2 + \lambda_2 y'^2 + d(p_{11}x' + p_{12}y') + e(p_{21}x' + p_{22}y') + f \\ &= \lambda_1 x'^2 + (p_{11}d + p_{21}e)x' + \lambda_2 y'^2 + (p_{12}d + p_{22}e)y' + f \end{aligned}$$

De la representación anterior y de lo que sabemos de geometría analítica se tiene que la ecuación 7.11 representa:

1. Una parábola, si exactamente uno de los valores característicos es cero.
2. Una circunferencia si, $\lambda_1 = \lambda_2$.
3. Una elipse si, $\lambda_1\lambda_2 > 0$.
4. Una hipérbola si, $\lambda_1\lambda_2 < 0$.

Para determinar cada uno de estos casos procedemos a encontrar los valores de λ_1 y λ_2. El polinomio característico de A es: $p(x) = \begin{vmatrix} a-x & \frac{b}{2} \\ \frac{b}{2} & c-x \end{vmatrix} = x^2 - (a+c)x + ac - \frac{b^2}{4}$,

por lo que sus valores característicos son:

$$\lambda = \frac{a+c \pm \sqrt{(a+c)^2 - 4ac + b^2}}{2} = \frac{a+c \pm \sqrt{(a-c)^2 + b^2}}{2} \tag{7.14}$$

Notemos que si los dos valores característicos de A son cero, entonces de la ecuación 7.13 se tiene que $A = 0$ y por tanto no tenemos una ecuación cuadrática. Podemos suponer que $A \neq 0$.

De esto se tiene que exactamente uno de los valores característicos es cero $\Leftrightarrow$ exactamente una de las ecuaciones $a+c+\sqrt{(a-c)^2+b^2}=0$, $a+c-\sqrt{(a-c)^2+b^2}=0$ se cumple, y esto ocurre exactamente cuando $4ac - b^2 = 0$, es decir, la ecuación 7.13 representa una parábola si $4ac - b^2 = 0$.

La ecuación 7.13 representa una circunferencia $\Leftrightarrow a + c - \sqrt{(a-c)^2+b^2} = a+c+\sqrt{(a-c)^2+b^2}$ y esto ocurre $\Leftrightarrow (a - c)^2 + b^2 = 0$ y esto último $\Leftrightarrow a = c$ y $b = 0$.

La ecuación 7.13 representa una elipse $\Leftrightarrow (a+c-\sqrt{(a-c)^2+b^2})(a+c+\sqrt{(a-c)^2+b^2}) > 0$ y esto ocurre $\Leftrightarrow (a + c)^2 - (a - c)^2 - b^2 > 0 \Leftrightarrow 4ac - b^2 > 0$.

La ecuación 7.13 representa una hipérbola $\Leftrightarrow (a+c-\sqrt{(a-c)^2+b^2})\ (a+c+\sqrt{(a-c)^2+b^2}) < 0$ y esto ocurre $\Leftrightarrow (a + c)^2 - (a - c)^2 - b^2 < 0 \Leftrightarrow 4ac - b^2 < 0$.

Lo discutido antes lo podemos resumir en el siguiente:

Teorema 7.3.9. *La ecuación:*

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

representa una circunferencia, elipse, hipérbola o parábola de acuerdo con: $\mathrm{a} = \mathrm{c}$ *y* $\mathrm{b} = 0$; $4\mathrm{ac} - \mathrm{b}^2 > 0$; $4\mathrm{ac} - \mathrm{b}^2 < 0$; $4\mathrm{ac} - \mathrm{b}^2 = 0$, *respectivamente. Los ejes respecto de los que se tienen las nuevas coordenadas* $(\mathrm{x}', \mathrm{y}')$ *están determinados por los vectores característicos de* $\mathrm{A} = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}$

7.4. Operadores adjuntos y normales

En esta parte discutiremos algunos conceptos fundamentales de la teoría de operadores. Iniciamos demostrando que dado un operador T en un espacio de dimensión finita con producto interno, existe otro operador llamado el adjunto de T con propiedades que describe el siguiente teorema.

Teorema 7.4.1. *Sea* V *un espacio de dimensión finita con producto interno*, T *un operador en* V, *entonces existe un operador* T^* *que satisface lo siguiente.*

$$\langle \mathrm{T}(\alpha), \beta\rangle = \langle \alpha, \mathrm{T}^*(\beta)\rangle \quad \textit{para todos } \alpha, \beta \in \mathrm{V} \tag{7.15}$$

Demostración. Presentamos la demostración para el caso real, el caso complejo se obtiene conjugando los coeficientes en la definición de T^*. Por el proceso de Gram-Schmidt, podemos suponer que $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ es una base ortonormal de V. Dado $\beta \in V$ definimos:

$$T^*(\beta) := \langle T(\alpha_1), \beta\rangle\alpha_1 + \langle T(\alpha_2), \beta\rangle\alpha_2 + \cdots + \langle T(\alpha_n), \beta\rangle\alpha_n$$

Utilizando las propiedades del producto interno se tiene directamente la linealidad de T^*.

Si $\alpha, \beta \in V$ y $\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n$, entonces $\langle T(\alpha), \beta\rangle = a_1\langle T(\alpha_1), \beta\rangle + a_2\langle T(\alpha_2), \beta\rangle + \cdots + a_n\langle T(\alpha_n), \beta\rangle$. Por otro lado, usando la bilinealidad del producto interno, la definición de T^* y que la base es ortonormal se obtiene:

$$
\begin{aligned}
\langle \alpha, T^*(\beta)\rangle &= \langle a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n, T^*(\beta)\rangle \\
&= \langle a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n, \langle T(\alpha_1), \beta\rangle\alpha_1 + \langle T(\alpha_2), \beta\rangle\alpha_2 + \cdots \\
&\quad + \langle T(\alpha_n), \beta\rangle\alpha_n\rangle \\
&= a_1\langle T(\alpha_1), \beta\rangle + a_2\langle T(\alpha_2), \beta\rangle + \cdots + a_n\langle T(\alpha_n), \beta\rangle \\
&= \langle T(\alpha), \beta\rangle
\end{aligned}
$$

terminando la prueba del teorema.

Definición 7.4.1. *Sea* A *una matriz con entradas complejas. Se define la adjunta de* A *denotada* A*, *como* $A^* = \overline{A}^t$.

Teorema 7.4.2. *Sea* V *un espacio de dimensión finita con producto interno y* T *un operador en* V*. Si* A *es la matriz asociada a* T *respecto a una base ortonormal, entonces* A* *es la matriz asociada a* T* *respecto a esa misma base.*

Demostración. Sea $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ una base ortonormal. Si a_{ij} denotan las entradas de A, entonces para cada i y cada j se tiene:

$$
\begin{aligned}
\langle T(\alpha_j), \alpha_i\rangle &= \langle a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n, \alpha_i\rangle \\
&= a_{1j}\langle \alpha_1, \alpha_i\rangle + a_{2j}\langle \alpha_2, \alpha_i\rangle + \cdots + a_{nj}\langle \alpha_n, \alpha_i\rangle \\
&= + a_{ij}
\end{aligned}
$$

Por otro lado, denotando por B a la matriz asociada a T^* respecto a la misma base tenemos:

$$
\begin{aligned}
\langle \alpha_j, T^*(\alpha_i)\rangle &= \langle \alpha_j, b_{1i}\alpha_1 + b_{2i}\alpha_2 + \cdots + b_{ni}\alpha_n\rangle \\
&= \overline{b_{1i}}\langle \alpha_j, \alpha_1\rangle + \overline{b_{2i}}\langle \alpha_j, \alpha_2\rangle + \cdots + \overline{b_{ni}}\langle \alpha_j, \alpha_n\rangle \\
&= \overline{b_{ji}}
\end{aligned}
$$

Usando la condición que define al operador adjunto y lo calculado antes tenemos: $a_{ij} = \langle T(\alpha_j), \alpha_i\rangle = \langle \alpha_j, T^*(\alpha_i)\rangle = \overline{b_{ji}}$, probando lo afirmado.

Definición 7.4.2. *Sea* V *un espacio con producto interno,* T *un operador en* V *que admite adjunto.*

1. *Se dice que* T *es normal si* TT* = T*T.
2. *Se dice que* T *es autoadjunto si* T = T*.

Teorema 7.4.3. *Sea* V *un espacio vectorial complejo de dimensión finita con producto interno y* T *un operador en* V*. Entonces* T *es normal* ⇔ V *tiene una base ortonormal de vectores característicos de* T.

Demostración. Si T es normal, la demostración sigue exactamente las mismas líneas que el teorema de los ejes principales, teorema 7.3.4, cambiando matriz simétrica por operador normal.

Recíprocamente, si V tiene una base ortonormal de vectores característicos de T, entonces T se representa por una matriz diagonal D. Por el teorema 7.4.2, la matriz asociada a T^* es $\overline{D}^t$. Claramente D y $\overline{D}^t$ conmutan, entonces T y T^* también conmutan, probando que T es normal.

Corolario 7.4.1. *Si* A *es una matriz tal que* AA* = A*A *entonces existe una matriz* C *cuyas columnas son ortonormales tal que* C*AC *es diagonal.*

7.5. Ejercicios

1. Sea A una matriz positiva definida y k un entero positivo. Demuestre que existe B tal que $A = B^k$.

2. Sean $X_1, X_2, ..., X_m$ elementos de $\mathbb{R}^n$ linealmente independientes con entradas enteras. Demuestre que existen $Y_1, Y_2, ..., Y_m$, tales que:

a) $\langle Y_i, Y_j \rangle = 0$, si $i \neq j$.

b) Cada Y_i tiene entradas enteras y $\{Y_1, Y_2, ..., Y_m\}$ genera el mismo subespacio que $\{X_1, X_2, ..., X_m\}$

3. Sea V un espacio con producto interno, $\alpha_1, \alpha_2, ..., \alpha_r$ elementos de V. Definimos la matriz $A = (\langle \alpha_i, \alpha_j \rangle)$. Demuestre que $\alpha_1, \alpha_2, ..., \alpha_r$ son linealmente dependientes $\Leftrightarrow$ A es singular.

4. Sea A una matriz simétrica con entradas en los reales y $k \geq 1$ entero tal que $A^k = I$. Demuestre que $A^2 = I$.

5. Sea V un espacio con producto interno, T un operador en V. Se dice que T es ortogonal si $\langle T(\alpha), T(\alpha) \rangle = \langle \alpha, \alpha \rangle$ para todo $\alpha \in V$. Demuestre lo siguiente:

a) El operador T es ortogonal $\Leftrightarrow \langle T(\alpha), T(\beta) \rangle = \langle \beta, \beta \rangle$ para todos $\alpha, \beta \in$ V.

b) El operador T es ortogonal $\Leftrightarrow$ transforma bases ortonormales en bases ortonormales.

c) Sea $\mathcal{B} = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ una base ortonormal de V, T un operador ortogonal en V y A la matriz que representa a T en la base $\mathcal{B}$. Demuestre que las columnas y las filas de A son ortonormales en el espacio apropiado. ¿Cuál es el valor del determinante de A?

6. Clasifique los operadores ortogonales en $\mathbb{R}^2$.

7. Sea V un espacio con producto interno, $T : V \to V$ una función que satisface la condición $\langle T(\alpha), T(\alpha) \rangle = \langle \alpha, \alpha \rangle$ para todo $\alpha \in V$. Demuestre que T es un operador, es decir, demuestre que T es una transformación lineal.

8. Sea V el espacio de las funciones continuas en el intervalo $[0, 2\pi]$ con el producto interno dado por $\langle f, g \rangle = \int_0^{2\pi} fg$. Demuestre que el subconjunto de V,

$$\mathcal{B}_{\mathrm{n}} = \{1, \operatorname{sen}(x), \operatorname{sen}(2x), ..., \operatorname{sen}(nx), \cos(x), \cos(2x), ..., \cos(nx)\}$$

es ortogonal. Si W_n es el subespacio generado por $\mathcal{B}_{\mathrm{n}}$ y $f \in V$, determine la proyección ortogonal de f sobre W. Tome $n = 5$ y $f(x) = x^2$ para ilustrar el caso mencionado.

9. Sean, V el espacio vectorial de las funciones continuas en el intervalo $[a, b]$ y $S = \{f_1, f_2, ..., f_n\} \subseteq V$. Para cada $i = 1, 2, ..., n$ y para cada $j = 1, 2, ..., n$ se define $a_{ij} = \ a_{ij} = \int_a^b f_i f_j$. Sea A la matriz que tiene por entradas a los elementos a_{ij}. Demuestre que $f_1, f_2, ..., f_n$ son linealmente dependientes $\Leftrightarrow$ A es singular.

10. Clasifique las cónicas dadas por las ecuaciones siguientes, grafíquelas y encuentre los ejes respecto de los cuales no aparece el término xy.

a) $4x^2 - 24xy + 11y^2 + 56x - 58y + 95 = 0$

b) $12x^2 + 8xy - y^2 + x + y - 2 = 0$

c) $8x^2 + 4y^2 + x - y + 3 = 0$

d) $2x^2 + 4xy + 2y^2 + x + y - 8 = 0$

Bibliografía

[1] S. Axler. Down with determinants. *Am. Math. Monthly*, pp. **102**, 139-154 (1995).

[2] S. Axler. *Linear Algebra Done Right*. Springer-Verlag, New York USA, second edition, 1997.

[3] R. Borrelli and C. S. Colleman. *Ecuaciones Diferenciales: Una perspectiva de modelación*. Oxford University Press, México, 2002.

[4] R. Bru, J-J. Climent, J. Mas, and A. Urbano. *Álgebra Lineal*. Alfaomega, Universidad Politécnica de Valencia, México, 2004.

[5] Ch. W. Curtis. *Linear Algebra: An Introductory approach*. Springer-Verlag, New York, third edition, 1991.

[6] J-L. Dorier(Editor). *On the Teaching of Linear Algebra*. Kluwer Academic Publisher, Dordrecht, Boston, London, 2000.

[7] D. C. Lay. *Álgebra lineal y sus aplicaciones*. Pearson, México, 1999.

[8] A. Mas-Colell, M. Whinston, and J. Green. *Microeconomic Theory*. Oxford University Press, New York USA, first edition, 1995.

[9] J. T. Moore. *Elements of Linear Algebra and Matrix Theory*. McGraw-Hill, New York, first edition, 1968.

[10] M. Newman. *Integral Matrices: Pure and applied mathematics*. Academic Press, Orlando, Florida, first edition, 1972.

[11] A. Tucker. *A Unified Introduction to Linear Algebra: Models, Methods, and Theory*. Macmillan Publishing Company, New York, first edition, 1988.

Índice alfabético

Álgebra
Teorema Fundamental del, 143, 177
ángulo entre vectores, 71
ángulos directores, 80
adjunta clásica, 126
antisimétrica
matriz, 133
anulador
polinomio T, 137
base de un espacio, 84
base dual, 115
Bezout, 31
bilineal
forma, 178
cíclica
teorema de la descomposición, 147
cíclico
subespacio T, 142
cónicas, 184-185
cambio de base
teorema de, 109
característico
polinomio, ix
Cardano, 31, 164
casillas
principio de las, 21, 94
Cauchy, 31
Cauchy-Schwarz
desigualdad de, 72, 172
Cayley, 31
Cayley-Hamilton
teorema de, x, 135, 143
cofactores, 126
combinación lineal, 64, 81
complejo
espacio con producto interno, 177
espacio vectorial, 176-177
conjunto
generador, 83
coordenadas de $\propto$ respecto a una base, 107
cosenos
ley de los, 70-71
Cramer, 31
regla de, 126, 127, 128
cuadrática
forma, 178
dependientes
linealmente, 65
Descartes
regla de los signos de, 184
descomposición QR, 176
descomposición primaria
teorema de, 140
determinante, 117, 119
diagonalizable
operador, 144
Diofanto, 54
divisores elementales, 153
ecuación de calor, 49
ecuación lineal, 12
ecuaciones diofantinas, 54
ecuaciones lineales
sistema de, 12
ejes principales
teorema de los, 182
equivalentes por filas
matrices, 45
espacio columna, 82, 111
espacio dual, 115
espacio fila, 111
espacio vectorial
definición, 89
dimensión de, 85, 92
finitamente generado, 90
estocástica
matriz, 169
euclidiano
espacio, 171
Fermat
ecuación de, 00
forma cuadrática, 181
positiva definida, 182
forma escalonada reducida, 17
Frobenius, 31
funciones coordenadas, 97
Gauss, 31
Gauss-Jordan
método de, 12, 15
Gram-Schmidt
proceso de, 173
Hadamard, 134
homogéneo
sistema, 00
idempotente
matriz, 133
independencia lineal, 89
independientes
linealmente, 66
inversible
matriz, 41
Jordan, 31
bloque de, 154
forma canónica (real) de, 156
forma canónica de, ix, 154, 155
Kirchhoff
leyes de, 47

Lagrange, 31
identidad de, 72
Leibniz, 129
Leontief, 1
modelo de, 4, 57
linealmente independientes
vectores, 82
mínimo
polinomio, ix
mínimos cuadrados
método de, 175
Maple
uso de, 25
matrices
producto de, 33, 36, 37
suma de, 32
matrices elementales, 39
matriz adjunta, 187
matriz antisimétrica, 45
matriz asociada a un polinomio, 147
matriz asociada a una transformación, 104
matriz aumentada, 13
matriz de cambio de base, 107
matriz de coeficientes, 13
matriz de flexibilidad, 46
matriz entera, 53, 55
matriz identidad, 39
matriz inversa, 42
matriz positiva definida, 183
matriz simétrica, 45
matriz transpuesta, 45
menores y cofactores, 126
núcleo de una transformación, 103
norma de un vector, 62, 70, 71, 172
norma de una matriz, 57
Ohm
ley de, 47
operaciones elementales, 14
ortogonal
complemento, 174
matriz, 182
operador, 188
proyección, 73, 74
proyección de un vector, 174
ortogonales
vectores, 172
ortonormal
base, 173
conjunto, 174
paralelogramo
ley del, 62
paramétrica
representación, 68-69
parte imaginaria de un complejo, 177
parte real de un complejo, 177
pirámide, 68
Pitágoras
teorema de, 70, 171
polinomio característico, 143
polinomio mínimo, 135, 136
producto cruz, 77
producto interno, 70, 71
rango columna, 111
rango de una matriz, 104, 111
rango de una transformación, 103
rango fila, 111
simétrica
función bilineal, 178
similares
matrices, 110
sistema consistente, 13
sistemas de ecuaciones lineales, 1
sistemas equivalentes, 14
Smith
forma normal de, 55
solución de un sistema, 13
subespacio T-invariante, 138
subespacio vectorial, 82
suma de subespacios, 87
suma directa, 92, 139
Sylvester, 31
tetraedro, 131
transformación lineal, 95, 96
transformaciones lineales geométricas, 99
traza de una matriz, 45
triángulo
desigualdad del, 173
triangulable
operador, 145
triple producto escalar, 80
triple producto vectorial, 80
valor característico, 138, 144
Vandermonde, 31
determinante de, 133
variables libres, 20
vector característico, 138, 144
vector coordenado, 108
vectores en el plano, 61
wronskiano, 128