Álgebra lineal elemental y aplicaciones

Álgebra lineal elemental y aplicaciones

Clara E. Mejía Laverde

Rector de la Universidad de Antioquia
Alberto Uribe Correa

Vicerrector de Docencia
Óscar Sierra Rodríguez

Coordinador del Programa de Educación Ude@
Guillermo León Ospina Gómez

Autora
Clara E. Mejía Laverde

Corrector de estilo
Daniel Aldana Estrada

Primera edición, 2006
Segunda edición, 2009

Imagen de la portada

Fotografía de la escultura *Fuente ceremonial*

La *Fuente ceremonial* fue la obra ganadora del concurso promovido por el municipio de Medellín para la celebración de los 200 años de la Universidad. Esta pieza en ferroconcreto, enchapada en granito, está inspirada en los sitios rituales y hace parte de una serie que su autor, el maestro Germán Botero Giraldo, viene trabajando desde 1990 con el tema de lo prehispánico. La forma circular invita a la reunión y a la comunión cósmica presente en las culturas indígenas. El elemento vertical es un hacha de la cual fluye agua, convirtiéndola en fuente. «El agua está presente como elemento refrescante, de sonido, brillo y vida, así como los árboles dentro de la escultura, que crecerán y crearán sombra; más que una escultura, es un espacio escultórico, un sitio para que la gente lo habite», dice su autor.

La *Fuente ceremonial*, situada junto a la entrada occidental de la calle Barranquilla, forma parte de las grandes obras que embellecen el campus universitario y refuerza el ideal de hacer de éste un espacio abierto a la cultura, la ciencia y el conocimiento.

ISBN 978-958-8748-27-6

Impreso en Medellín (Colombia).

Acerca de la autora

Clara E. Mejía Laverde

Ingeniera industrial (1977) de la Universidad Nacional y magíster (1996) en Educación (Pensamiento Lógico-Matemático) de la Universidad de Antioquia. Actualmente es profesora jubilada del Departamento de Matemáticas de la Universidad de Antioquia.

Cómo usar este texto

Como estudiante del programa de educación no presencial de la Universidad de Antioquia, Ude@, usted es el centro del modelo educativo y puede controlar el proceso de aprendizaje mediante la organización del tiempo alrededor de sus intereses. La autonomía, la disciplina, la creatividad y el trabajo en equipo son características que le ayudarán en su formación para solucionar problemas reales de la sociedad, recurriendo al método de la ingeniería.

Los cursos Ude@ permiten fortalecer estas características mediante el desarrollo de diferentes actividades[1].

- Estudio individual, apoyado en diferentes medios (impresos, audiovisuales, multimedia).
- Estudio en grupo y acompañamiento del profesor a través del aula virtual.
- Tutorías presenciales, cuya finalidad es apoyar el aprendizaje y afianzar los temas estudiados.

El texto Ude@

En el modelo Ude@ los contenidos educativos son aportados por cada medio, teniendo en cuenta las fortalezas propias de cada uno de ellos. Desde el punto de vista pedagógico, el texto impreso es por tradición un medio idóneo para los procesos de educativos, ya que facilita el aprendizaje de hechos, la comprensión de principios generalizados o abstractos y el desarrollo del razonamiento lógico. En estos aspectos, el texto Ude@ es un medio eficaz para desarrollar y adquirir tales destrezas.

Estructura del texto

El texto *Álgebra lineal elemental y aplicaciones* ha sido desarrollado como parte del material educativo de los estudiantes del programa; sin embargo, su contenido puede ser de gran utilidad para cualquier persona que desee estudiar este tema.

La estructura del texto es lineal, con una progresión gradual de cada tema, lo cual hace más fácil la transmisión del contenido de una manera lógica.

La división del texto está dada por capítulos que, a su vez, agrupan módulos o temas. Al empezar cada capítulo se encuentra un «Contenido breve» que muestra el número y el título de los módulos que componen el capítulo. Por su parte, cada módulo contiene, en su primera página, una introducción, los objetivos de aprendizaje, unas preguntas básicas (relacionadas con los conocimientos previos requeridos) y el índice temático del contenido, que le guiarán en el proceso de aprendizaje sobre el tema en particular de cada sesión de clase.

[1] Los cursos tienen un cronograma semanal de actividades que lo orientará en su proceso de aprendizaje.

Los iconos y la interrelación de medios

El material Ude@ ha sido producido de manera integral, teniendo como objetivo primordial el autoestudio; por lo tanto, la producción de los contenidos se desarrolla en los diferentes formatos (audiovisuales, web, multimedia, videoconferencias), con enlaces entre estos. La esencia de estos enlaces está dada por los iconos Ude@.

Los iconos, como representaciones gráficas de la realidad, serán los elementos gráficos que le ayudarán a guiarse en su navegación por los diferentes medios.

Sugerencias para los estudiantes

En la lectura del libro:

- Antes de iniciar el estudio de un capítulo, lea el contenido breve y la presentación.

- Trate de resolver las preguntas básicas de cada módulo; estas preguntas están diseñadas para ayudarle a comprender los conceptos o temas presentados.

- Lea los ejemplos intercalados en los bloques de texto y trate de resolver los ejercicios con el fin de mejorar sus habilidades en la solución de problemas reales.

- Complemente la lectura del libro con las herramientas de comunicación que posee en el aula virtual y en su correo electrónico.

- Recuerde que sobre el tema que está estudiando en el módulo impreso también existe material disponible en otros medios, y que ese material representa valor agregado, puesto que el contenido de los diferentes formatos no se repite, sino que se complementa.

En el aula virtual:

- Aprenda cómo funcionan las herramientas indispensables para participar en un curso por red: sistema de correo electrónico, sistema de chat, grupos de discusión, búsquedas en Internet, consulta en bases de datos especializadas, entre otras.

- Revise el correo electrónico todos los días.

- Visite con relativa frecuencia el sitio Ude@ y la plataforma donde se publica el curso en Internet para enterarse de cualquier nueva información. Apóyese en la red como un sistema de consulta y establezca criterios para seleccionar la información requerida.

- Introduzca sus datos personales en el aula virtual para que sus tutores y compañeros tengan acceso a ellos.

- Desarrolle, en la primera semana, las actividades preparativas para el curso indicadas en el aula virtual.

- Dedique por lo menos tres horas semanales por cada crédito asignado a un curso para leer los módulos, realizar trabajos, participar en los foros de discusión y presentar evaluaciones, de acuerdo con lo establecido en el cronograma.

- Planee su agenda personal para participar activamente en cada curso y entregar oportunamente sus tareas. En caso de algún imprevisto, debe comunicarse inmediatamente con el tutor.

- Participe de las actividades propuestas para realizar en forma individual y en grupos de trabajo. Haga parte de grupos de trabajo conformados con sus compañeros de curso y en ningún caso pretenda realizar todas las actividades sin ayuda de los demás.

- Manifieste oportunamente a sus compañeros y al profesor las dificultades que se le presentan con las actividades propuestas.

- Elabore su propio horario de trabajo independiente para el curso y cumpla con el cronograma del curso.

- Realice con honradez las actividades de evaluación, autoevaluación y coevaluación que encuentre programadas en el curso.

- Durante su proceso de aprendizaje trate de adquirir autonomía con el conocimiento, es decir, intente construir nuevos conocimientos recurriendo a fuentes de información bibliográfica y a sus habilidades de comparación, análisis, síntesis y experimentación.

- Mantenga una actitud de colaboración con compañeros, tutores y monitores, y esté siempre dispuesto a realizar las actividades de aprendizaje.

- Relaciónese de manera respetuosa y cordial con los demás estudiantes, con el tutor y con los monitores.

Tabla de contenido

Prólogo

El objetivo al redactar este texto ha sido desarrollar las ideas básicas del álgebra lineal y mostrar algunas aplicaciones interesantes, haciendo un balance para no recargar demasiado el desarrollo teórico y mantener un equilibrio entre la abstracción y la aplicación. Así, algunas demostraciones de teoremas que resultan fácilmente accesibles a los estudiantes se dejan propuestas como ejercicios, y otras, demasiado dificiles para el nivel de un curso introductorio, no se presentan.

El hilo conductor en la presentación de los temas ha sido la solución de los problemas básicos del álgebra lineal, enunciados por el profesor William Perry, así:

1. El problema de la solución de un sistema de ecuaciones lineales.
2. El problema de la construcción de una base para un espacio vectorial.
3. El problema de la construcción de una matriz que represente una transformación lineal.
4. El problema de los valores y los vectores característicos.
5. El problema de la diagonalización.

De estos, el problema de la «solución de un sistema de ecuaciones lineales» ya fue tratado en el texto de *Geometría vectorial*, en el cual se hace una introducción al álgebra lineal. Es importante entonces hacer notar que en este texto no hay un desarrollo completo del álgebra lineal básica, sino que continúa la exposición de los temas ya comenzada en el curso anterior y, por este motivo, no se tratan los temas de sistemas de ecuaciones lineales y determinantes.

El problema de la «construcción de una base para un espacio vectorial» se trata en el capítulo 1 sobre espacios vectoriales, y en el capítulo 2, ortogonalidad, donde se introduce un algoritmo para la construcción de una base ortonormal en un espacio vectorial.

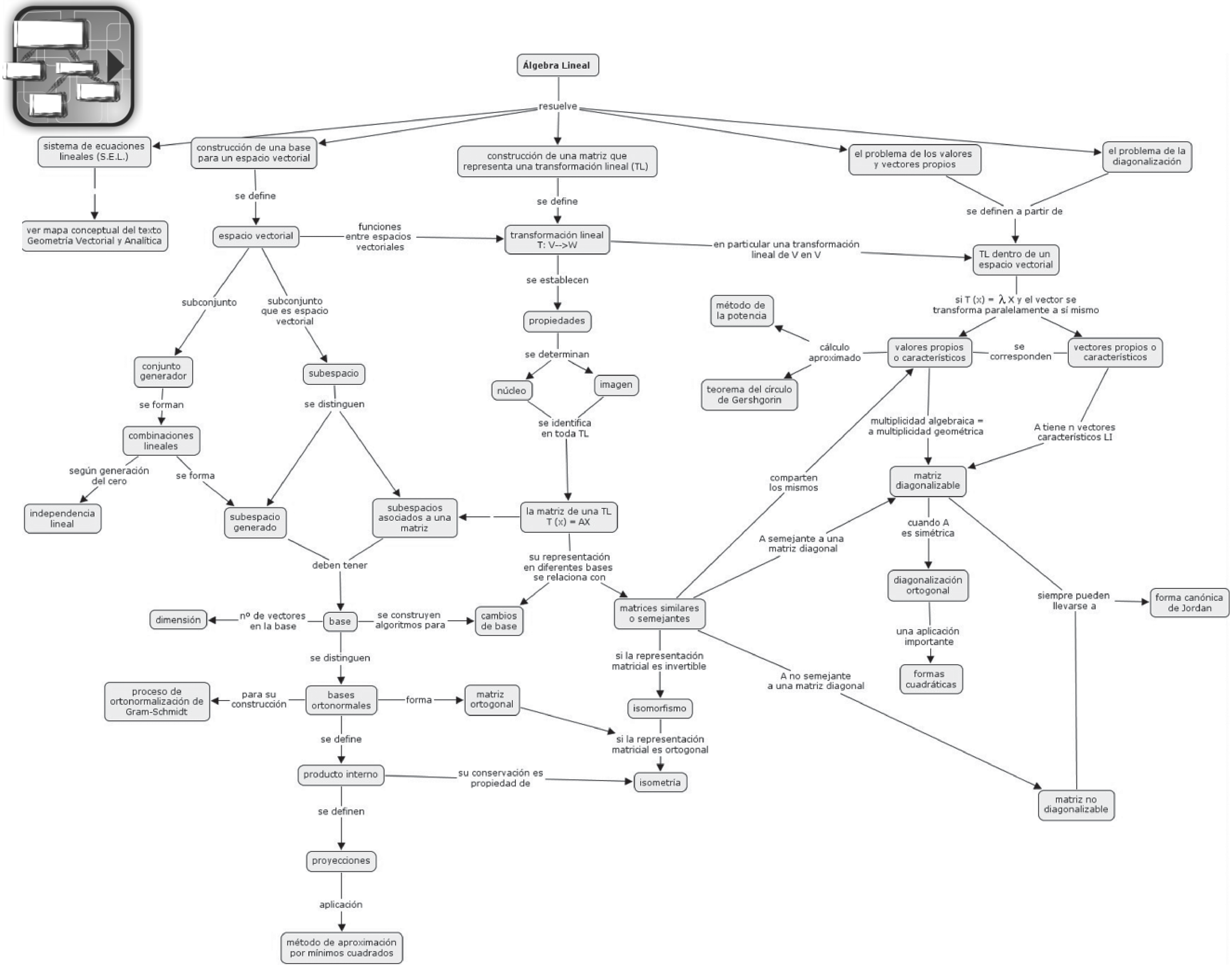
El tercer problema, la «construcción de una matriz que represente una transformación lineal», se resuelve en el capítulo 3 sobre transformaciones lineales, y los problemas de «valores y vectores característicos» y «diagonalización» son tratados en los capítulos 4 y 5.

En la actualidad existen muchos programas de computador que permiten desarrollar una amplia variedad de tópicos y resolver gran cantidad de problemas que de otra forma consumirían mucho tiempo y probablemente inducirían a cometer errores, dada la magnitud de los cálculos que se tienen que realizar. El lenguaje de computación MATLAB es un excelente programa, fácil de usar, y calificado para trabajar problemas de álgebra lineal. Este libro no enseña el manejo de ese programa; sin embargo, en la multimedia disponible para el curso se plantean algunos ejercicios utilizando esta herramienta.

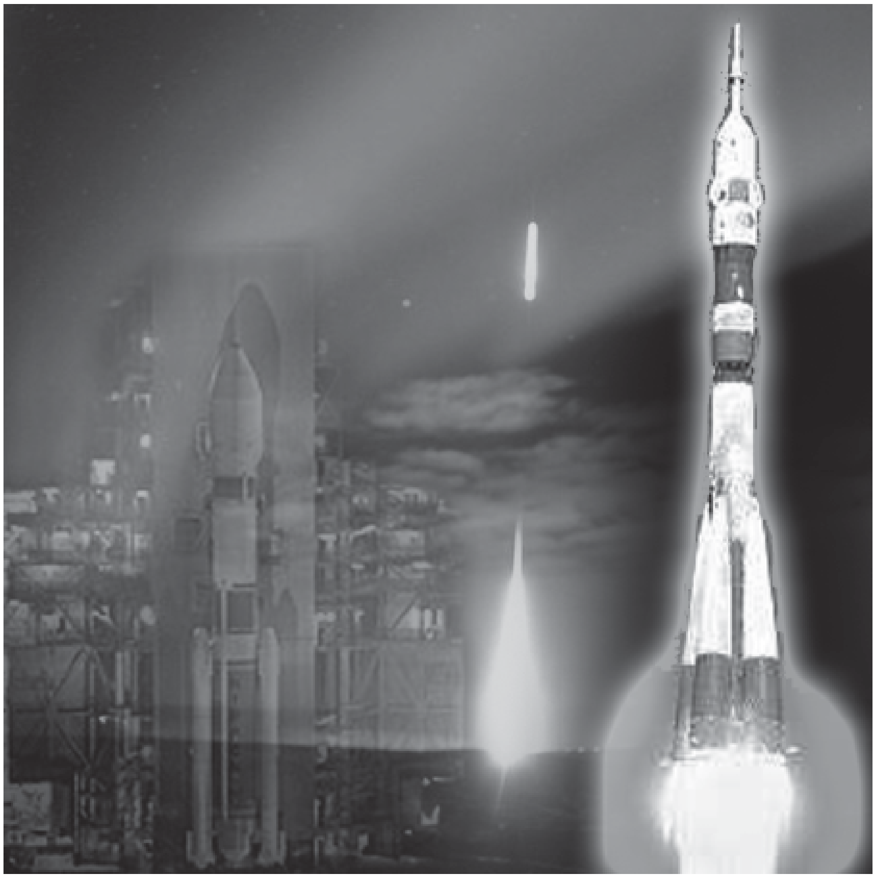
Agradezco cualquier sugerencia que pueda contribuir a mejorar posteriores ediciones de este texto. Todas ellas pueden enviarse a:

clame@matematicas.udea.edu.co

La autora



Álgebra Lineal
resuelve
sistema de ecuaciones lineales (S.E.L.)
ver mapa conceptual del texto Geometría Vectorial y Analítica
construcción de una base para un espacio vectorial
se define
espacio vectorial
funciones entre espacios vectoriales
subconjunto
conjunto generador
se forman
combinaciones lineales
según generación del cero
independencia lineal
se forma
subespacio generado
subconjunto que es espacio vectorial
subespacio
se distinguen
subespacios asociados a una matriz
deben tener
base
nº de vectores en la base
dimensión
se construyen algoritmos para
cambios de base
se distinguen
bases ortonormales
para su construcción
proceso de ortonormalización de Gram-Schmidt
forma
matriz ortogonal
se define
producto interno
su conservación es propiedad de
se definen
proyecciones
aplicación
método de aproximación por mínimos cuadrados
construcción de una matriz que representa una transformación lineal (TL)
se define
transformación lineal T: V-->W
se establecen
propiedades
se determinan
núcleo
imagen
se identifica en toda TL
la matriz de una TL T (x) = AX
su representación en diferentes bases se relaciona con
matrices similares o semejantes
si la representación matricial es invertible
isomorfismo
si la representación matricial es ortogonal
isometría
en particular una transformación lineal de V en V
el problema de los valores y vectores propios
el problema de la diagonalización
se definen a partir de
TL dentro de un espacio vectorial
si T (x) = λ X y el vector se transforma paralelamente a sí mismo
valores propios o característicos
se corresponden
vectores propios o característicos
cálculo aproximado
método de la potencia
teorema del círculo de Gershgorin
comparten los mismos
multiplicidad algebraica = a multiplicidad geométrica
A tiene n vectores característicos LI
A semejante a una matriz diagonal
matriz diagonalizable
cuando A es simétrica
diagonalización ortogonal
una aplicación importante
formas cuadráticas
siempre pueden llevarse a
forma canónica de Jordan
A no semejante a una matriz diagonal
matriz no diagonalizable



Lanzar al espacio un transbordador es el fruto de años de trabajo y un triunfo de la ingeniería de sistemas de control. Matemáticamente, las señales de entrada y de salida de un sistema de control son funciones y es importante que estas señales puedan sumarse y multiplicarse por escalares. Estas dos operaciones con funciones tienen propiedades algebraicas análogas a las operaciones de sumar vectores en $\mathbb{R}^n$ y multiplicar un vector por un escalar; por esta razón, el conjunto de todas las posibles entradas (funciones) se llama espacio vectorial.

Capítulo 1

Espacios vectoriales

Contenido breve

Presentación

En el desarrollo de la geometría vectorial hemos identificado los vectores con los elementos de $\mathbb{R}^2$ y $\mathbb{R}^3$ escribiéndolos como parejas o ternas ordenadas de números reales, respectivamente; una vez hecho esto, surge el problema de generalizar este concepto y tener vectores como *n*-tuplas, conjuntos ordenados de *n* componentes en un espacio de *n* dimensiones. Ahora, ¿por qué querríamos hacer esto? La respuesta puede ser tan simple como decir que la mayoría de los problemas de la vida real involucran más variables que sólo dos o tres.

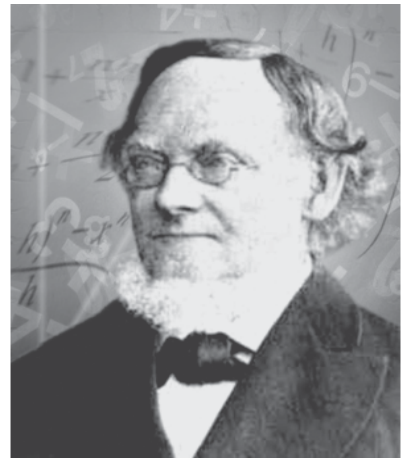
Es claro que cuando generalizamos el concepto de vector a *n* dimensiones, ya no tenemos la intuición geométrica que desarrollamos para el plano $(\mathbb{R}^2)$ y el espacio $(\mathbb{R}^3)$, pero podemos conservar las propiedades algebraicas.

Nos proponemos definir una estructura algebraica llamada *espacio vectorial*, donde se pueden reunir los conjuntos $\mathbb{R}^2$, $\mathbb{R}^3$ y, en general, $\mathbb{R}^n$, además de otros conjuntos tales como las matrices y los polinomios. Con esta abstracción podemos identificar los vectores como los elementos de un espacio vectorial y, en esta forma, tratar las matrices o los polinomios como vectores.

Dentro de los espacios vectoriales podemos reconocer conjuntos que guardan en su interior la estructura de espacio vectorial, llamados subespacios. Una vez establecida la estructura desarrollamos los conceptos de combinación lineal, independencia lineal, base y dimensión, los cuales constituyen los «ladrillos» para la construcción de espacios vectoriales. Ubicamos como problema central en este capítulo la determinación de una base para un espacio vectorial. Otros elementos que aportan a la solución de este problema básico se presentan en el estudio de los subespacios asociados con una matriz y en el establecimiento de coordenadas y cambio de base.

Módulo 1

Definición y propiedades del espacio vectorial



El matemático y lingüista alemán Hermann Grassmann (1809-1877) fue el primero en definir un espacio vectorial *n* dimensional y la independencia lineal.

Introducción

Nos proponemos construir e identificar la estructura de espacio vectorial con diferentes objetos matemáticos: *n*-tuplas, funciones, polinomios, matrices, etc., alcanzando una asimilación algebraica del concepto de vector, con el fin de unificar los procedimientos de trabajo para problemas, en apariencia, diversos. Esto permite, además, tener una visión de las posibles aplicaciones del concepto.

Objetivos

1. Construir la estructura de espacio vectorial.
2. Identificar como vectores todos aquellos elementos de conjuntos que con las operaciones dadas verifiquen los axiomas de la estructura.
3. Aprender el manejo algebraico dentro de un espacio vectorial.

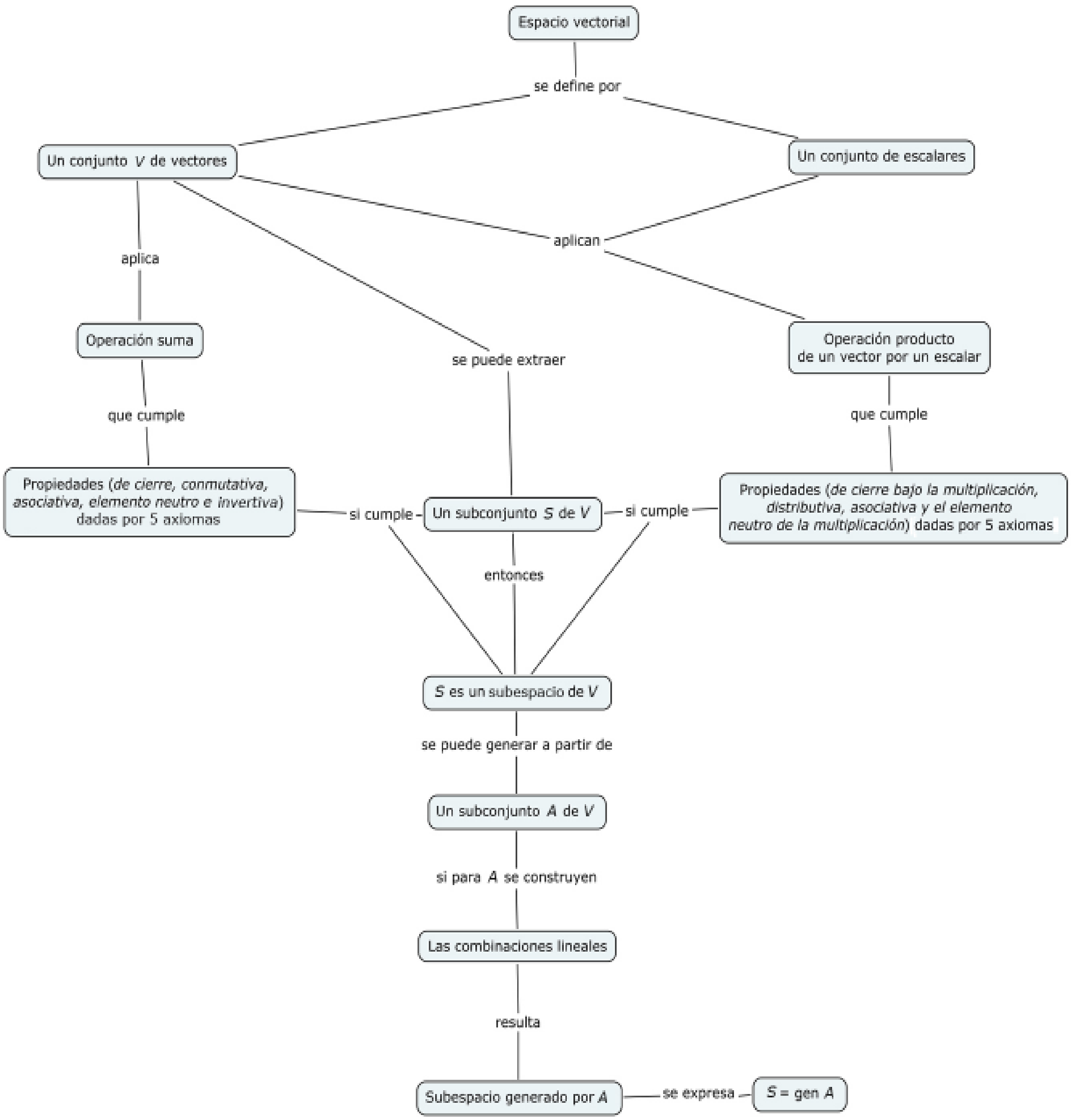
Preguntas básicas

1. ¿Qué es un espacio vectorial?
2. ¿Qué propiedades deben cumplir las operaciones definidas de suma y producto por un escalar?
3. ¿Cuáles son los espacios vectoriales más usuales?
4. ¿Qué propiedades tienen el módulo del espacio y el inverso aditivo de cada vector?
5. ¿Qué otras propiedades algebraicas se desprenden de la definición de espacio vectorial?

Contenidos

1.1 Ejemplos introductorios
 1.1.1 Las fuerzas que pueden actuar sobre un punto
 1.1.2 Las parejas de números reales $(\mathbb{R}^2)$
1.2 Definición
1.3 Ejemplos
1.4 Propiedades

Vea el módulo 1 del programa de televisión *Álgebra lineal*.



Mapa 1: módulos 1, 2 y 3.

1.1 Ejemplos introductorios

1.1.1 Las fuerzas que pueden actuar sobre un punto

Consideremos un punto material P y las fuerzas que pueden actuar sobre él. Si F_1 y F_2 son dos fuerzas que actúan sobre P, se puede encontrar una fuerza F resultante de F_1 y F_2 con la propiedad de ejercer sobre P la misma acción que F_1 y F_2 actuando simultáneamente. Así tenemos una fuerza F, que resulta de sumar F_1 y F_2 $(F = F_1 + F_2)$.

Veamos qué propiedades tiene esta operación.

Representemos F_1 y F_2 por vectores cuya dirección y sentido corresponden a los de las fuerzas dadas, y su longitud (en una escala determinada) mide la magnitud; entonces, la resultante queda determinada por la diagonal del paralelogramo, dos de cuyos lados son los segmentos correspondientes a F_1 y F_2 (figura 1.1).

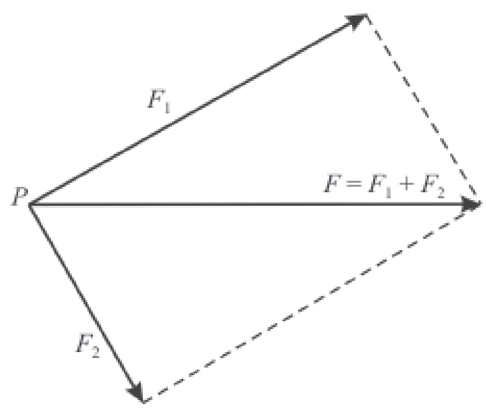


Figura 1.1

Sea T el conjunto de todas las fuerzas que actúan sobre P; entonces, podemos decir que:

i. Si F_1 y F_2 pertenecen a T, $F_1 + F_2$ pertenece a T.

Si sobre P tenemos tres fuerzas distintas F_1, F_2 y F_3, podemos calcular su resultante de dos formas distintas: $(F_1 + F_2) + F_3$ o bien $F_1 + (F_2 + F_3)$. Es decir, calculando primero $F = F_1 + F_2$ y luego la resultante de F con F_3, o bien calculando primero $F' = F_2 + F_3$ y luego la resultante de F_1 y F'.

Una representación geométrica nos muestra que (figura 1.2):

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Propiedades de los espacios vectoriales».

Escuche la biografía de *Hermann Grassmann* en su multimedia de *Álgebra lineal.*

ii. $(F_1+F_2)+F_3 = F_1+(F_2+F_3).$

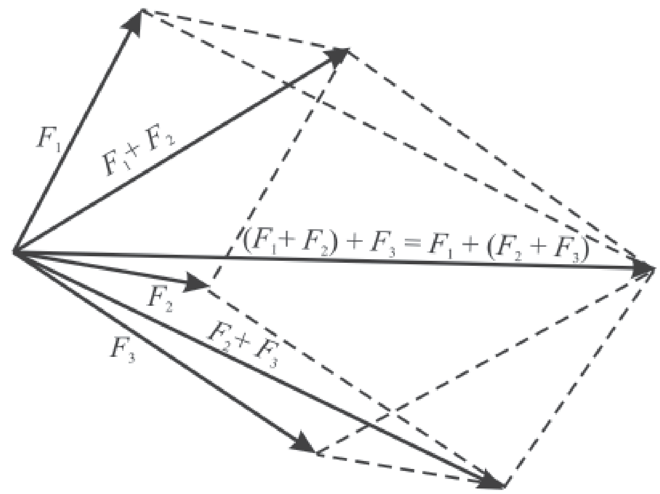


Figura 1.2

Por otra parte, existe una fuerza de magnitud cero, fuerza nula, que podemos designar 0 y tiene la propiedad de:

iii. $F+0=0+F=F.$

Ahora, si llamamos la opuesta de F a una fuerza F' que tiene la misma magnitud y dirección que F pero sentido contrario, encontramos que la resultante de F y F' es cero, es decir:

iv. Para cada F existe $F',$ tal que $F+F'=0.$ Esta F' se llamará $-F.$

Por último, es claro que el orden en el que se consideran las fuerzas no altera la resultante; esto es:

v. $F_1+F_2=F_2+F_1.$

Veamos ahora otra operación: si F es una fuerza y α un número real, llamamos αF la fuerza que tiene la misma dirección de F, su magnitud es $|\alpha|$ veces la de F y su sentido es el mismo de F si α es positivo, y opuesto a F si α es negativo.

Hemos entonces definido una operación cuyos «factores» son una fuerza y un número real y su resultado es otra fuerza. Esta operación se dice que es externa, ya que hace intervenir un elemento (el número α) que no pertenece al conjunto T.

Veamos algunas propiedades de esta operación:

vi. Si α es un número real y $F\in T,$ entonces $\alpha F\in T.$

Razonando geométricamente, podemos verificar que (figura 1.3):

vii. $\alpha(F_1 + F_2) = \alpha F_1 + \alpha F_2.$

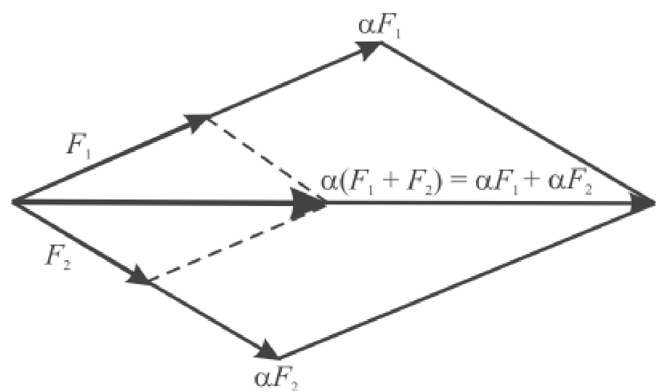


Figura 1.3

Aplicando la definición, se puede comprobar que:

viii. $(\alpha + \beta)F = \alpha F + \beta F.$

ix. $\alpha(\beta F) = (\alpha\beta)F.$

x. $1 \cdot F = F.$

1.1.2 Las parejas de números reales ($\mathbb{R}^2$)

Definimos en $\mathbb{R}^2$ una operación binaria interna (es decir, los elementos que intervienen en la operación pertenecen a $\mathbb{R}^2$) que llamaremos suma, así:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2,\ y_1 + y_2),$$

o sea que la suma de dos parejas de $\mathbb{R}^2$ es otra pareja de $\mathbb{R}^2$; se cumple entonces que:

i. Si (x_1, y_1) y $(x_2, y_2) \in \mathbb{R}^2$, entonces $(x_1, y_1) + (x_2, y_2) \in \mathbb{R}^2$.

Hay también una ley asociativa, es decir,

$$\begin{aligned}\left[(x_1, y_1) + (x_2, y_2)\right] + (x_3, y_3) &= (x_1 + x_2,\ y_1 + y_2) + (x_3, y_3) \\ &= (x_1 + x_2 + x_3,\ y_1 + y_2 + y_3).\end{aligned}$$

Por otra parte,

$$\begin{aligned}(x_1, y_1)+[(x_2, y_2)+(x_3, y_3)] &= (x_1, y_1)+(x_2+x_3,\ y_2+y_3)\\ &= (x_1+x_2+x_3,\ y_1+y_2+y_3),\end{aligned}$$

luego:

ii. $\left[(x_1, y_1)+(x_2, y_2)\right]+(x_3, y_3)=(x_1, y_1)+\left[(x_2, y_2)+(x_3, y_3)\right].$

Existe, además, un elemento de $\mathbb{R}^2$ que es el (0, 0) tal que:

iii. $(x, y)+(0,0)=(x, y)$ para todo (x, y) en $\mathbb{R}^2$.

Así mismo, para cada par (x, y) podemos tomar el par $(-x,-y)$, de modo que:

iv. $(x, y)+(-x,-y)=(0,0).$

Finalmente, hay una ley conmutativa, puesto que:

$$(x_1, y_1)+(x_2, y_2)=(x_1+x_2,\ y_1+y_2)=(x_2+x_1,\ y_2+y_1)=(x_2, y_2)+(x_1, y_1).$$

Entonces,

v. $(x_1, y_1)+(x_2, y_2)=(x_2, y_2)+(x_1, y_1).$

Definimos ahora una segunda operación, externa, entre un número real y un par, así:

$$\alpha\cdot(x, y)=(\alpha x,\ \alpha y),$$

cuyo resultado es un elemento de $\mathbb{R}^2$; con esta operación se verifican las siguientes propiedades:

vi. Si $\alpha\in\mathbb{R}$ y $(x, y)\in\mathbb{R}^2$, entonces $\alpha(x, y)\in\mathbb{R}^2$.

Veamos ahora que:

vii. $\alpha\left[(x_1, y_1)+(x_2, y_2)\right]=\alpha(x_1, y_1)+\alpha(x_2, y_2).$

Demostración

$$\begin{aligned}\alpha\left[(x_1, y_1)+(x_2, y_2)\right] &= \alpha(x_1+x_2,\ y_1+y_2)=\left(\alpha(x_1+x_2),\ \alpha(y_1+y_2)\right)\\ &= (\alpha x_1+\alpha x_2,\ \alpha y_1+\alpha y_2).\end{aligned}$$

Por otro lado,

$$\alpha(x_1, y_1)+\alpha(x_2, y_2)=(\alpha x_1,\ \alpha y_1)+(\alpha x_2,\ \alpha y_2)=(\alpha x_1+\alpha x_2,\ \alpha y_1+\alpha y_2).$$

viii. $(\alpha+\beta)(x,y)=\alpha(x,y)+\beta(x,y)$.

Demostración

$$(\alpha+\beta)(x,y)=\big((\alpha+\beta)x,\,(\alpha+\beta)y\big)=(\alpha x+\beta x,\,\alpha y+\beta y)\text{ y}$$
$$\alpha(x,y)+\beta(x,y)=(\alpha x,\,\alpha y)+(\beta x,\,\beta y)=(\alpha x+\beta x,\,\alpha y+\beta y).$$

ix. $\alpha\left[\beta(x,y)\right]=(\alpha\beta)(x,y)$.

Demostración

$$\alpha\left[\beta(x,y)\right]=\alpha(\beta x,\beta y)=(\alpha\beta x,\,\alpha\beta y)\text{ y}$$
$$(\alpha\beta)(x,y)=(\alpha\beta x,\,\alpha\beta y).$$

x. $1(x,y)=(x,y)$.

Demostración

$$1(x,y)=(1x,1y)=(x,y).$$

1.2 Definición

Los ejemplos anteriores nos sirven para obtener un conjunto de propiedades comunes a ellos. Todos los resultados que podamos deducir a partir de dichas propiedades serán válidos en los ejemplos realizados y en cualquier otro que goce de las mismas propiedades.

A estas propiedades comunes las llamamos axiomas en la siguiente definición.

Un *espacio vectorial* V es un conjunto de objetos llamados *vectores* con dos operaciones, *suma* y *producto por un escalar*, que satisface los siguientes axiomas:

i. Si **a** y **b** son elementos de V, entonces **a + b** pertenece a V; o de otra forma, la suma es una operación cerrada en V.

ii. $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$ para todo $\mathbf{a},\mathbf{b},\mathbf{c}\in V$.

iii. Existe un elemento **0** en V tal que $\mathbf{a}+\mathbf{0}=\mathbf{0}+\mathbf{a}=\mathbf{a}$ para toda $\mathbf{a}\in V$.

iv. Para cada $\mathbf{a}\in V$ existe un elemento $-\mathbf{a}\in V$ tal que $\mathbf{a}+(-\mathbf{a})=\mathbf{0}$.

v. $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$ para todo $\mathbf{a},\mathbf{b}$ en V.

vi. Si $\mathbf{a}\in V$ y α es un escalar ($\alpha\in\mathbb{R}$ o $\alpha\in\mathbb{C}$), entonces $\alpha\mathbf{a}\in V$; es decir, V es cerrado bajo la operación producto por un escalar.

vii. $\alpha(\mathbf{a}+\mathbf{b}) = \alpha\,\mathbf{a}+\alpha\,\mathbf{b}$ para todo escalar α y para todo $\mathbf{a}$ y $\mathbf{b} \in V$.

viii. $(\alpha+\beta)\mathbf{a} = \alpha\,\mathbf{a} + \beta\,\mathbf{a}$ para todo $\alpha,\ \beta$ escalares y $\mathbf{a} \in V$.

ix. $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$ para todo $\alpha,\ \beta$ escalares y $\mathbf{a} \in V$.

x. $1\mathbf{a} = \mathbf{a}$ para todo $\mathbf{a} \in V$.

Nota: cuando los escalares son números reales, decimos que se trata de un espacio vectorial real, y en caso de que sean números complejos hablamos de un espacio vectorial complejo.

En el axioma *viii* debe notarse que el signo + al lado izquierdo de la igualdad se refiere a la suma en el conjunto de escalares, y el signo + al lado derecho denota la suma en el conjunto de vectores *V*.

1.3 Ejemplos

1. $\mathbb{R}^n$, con las operaciones de suma y producto por un escalar definidas usualmente, tiene estructura de espacio vectorial.

2. Sea $V = \mathbb{P}_n = \left\{P \,/\, p(x) = a_n\,x^n + a_{n-1}\,x^{n-1} + \ldots + a_1\,x + a_0,\ a_i \in \mathbb{R}\right\}$ el conjunto de los polinomios de grado menor o igual que *n*.

   El elemento **0** (neutro) en $\mathbb{P}_n$ es $0_{(x)} = 0\,x^n + 0\,x^{n-1} + \ldots + 0\,x + 0.$

   Si $P(x) = a_n\,x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$ y
   $Q(x) = b_n\,x^n + b_{n-1}x^{n-1} + \ldots + b_1x + b_0,$

   entonces

   $$(P + Q)\,(x) = P(x) + Q(x) =$$
   $$(a_n + b_n)\,x^n + (a_{n-1} + b_{n-1})x^{n-1} + \ldots + (a_1 + b_1)\,x + (a_0 + b_0).$$

   Se ve claramente que la suma de dos polinomios de grado menor o igual que *n* es otro polinomio de grado menor o igual que *n*. Se pueden comprobar las propiedades *ii* y *v* a *x* con:

   $$(\alpha\,p)(x) = \alpha\left[\,p(x)\right]$$
   $$= \alpha\,a_n\,x^n + \alpha\,a_{n-1}\,x^{n-1} + \ldots + \alpha\,a_1\,x + \alpha a_0,\ \ \alpha\,p \in P_n.$$

   $$-p\,(x) = -a_n\,x^n - a_{n-1}x^{n-1} + \ldots - a_1x - a_0$$

   $$p(x) + \left[-p(x)\right] = 0\,(x).$$

3. Si $V = M_{m,n}$ (conjunto de matrices m por n) con la suma y la multiplicación escalar usuales, es fácil verificar que $M_{m,n}$ es un espacio vectorial con $\mathbf{0}$ como la matriz cero de $M_{m,n}$. Si $A = (a_{ij})_{(m,n)}$, entonces $-A = (-a_{ij})_{(m,n)}$, $-A \in M_{m,n}$.

4. Sea S_3 el conjunto de matrices invertibles de 3×3. Se define la suma $A \oplus B$ por $A \oplus B = AB$. Si A y B son invertibles, entonces AB es invertible y se cumple *i*.

   El axioma *ii* es la ley asociativa para la multiplicación de matrices. Los axiomas *iii* y *iv* se satisfacen con $\mathbf{0} = I_3$ y $-A = A^{-1}$. Sin embargo, $AB \neq BA$ y el axioma *v* no se cumple; por lo tanto, S_3 no es un espacio vectorial.

5. Sea $V = C[a,b]$ = conjunto de funciones continuas de valores reales definidas en el intervalo $[a,b]$. Se define:

$$(f+g)(x) = f(x) + g(x) \quad \text{y} \quad (\alpha f)(x) = \alpha[f(x)].$$

   Como la suma de funciones continuas es continua, el axioma *i* se cumple; de la misma forma el producto de un escalar por una función continua también es una función continua (propiedad *vi*). Los otros axiomas se verifican fácilmente con $\mathbf{0}$, la función 0 y $(-f)(x) = -f(x)$. Luego $C[a,b]$, con las operaciones definidas, tiene estructura de espacio vectorial.

6. Sea V el conjunto de los números reales ($\mathbb{R}$) con las operaciones $\mathbf{x} \oplus \mathbf{y} = \mathbf{x} - \mathbf{y}$ ($\oplus$ es la resta ordinaria) y la multiplicación por un escalar como la multiplicación ordinaria. Veamos si V es un espacio vectorial.

   $\mathbf{x} \oplus \mathbf{y}$ y $\alpha \mathbf{x}$ son números reales, lo cual verifica las condiciones *i* y *vi* de la definición.

   Veamos qué pasa con la propiedad conmutativa.

   Sea $\mathbf{x} = 2$ y $\mathbf{y} = 3$.

$$\mathbf{x} \oplus \mathbf{y} = 2 - 3 = -1,$$
$$\mathbf{y} \oplus \mathbf{x} = 3 - 2 = 1,$$

   luego:

$$\mathbf{x} \oplus \mathbf{y} \neq \mathbf{y} \oplus \mathbf{x}.$$

   Además, las propiedades *ii*, *iii* y *iv* tampoco se cumplen. Las propiedades *vii*, *ix* y *x* se cumplen, pero la propiedad *viii* no; veamos por qué:

   Sea $\alpha = 2$, $\beta = 3$, $\mathbf{x} = 4$.

$$(\alpha+\beta)\,\mathbf{x} = (2+3)4 = 5\times 4 = 20,$$

pero:

$$\alpha\mathbf{x}\oplus\beta\mathbf{x} = 2\times 4 - 3\times 4 = 8 - 12 = -4.$$

7. Sea $V = \{(x, y, z) \colon ax+by+cz = 0,\ a, b, c \in \mathbb{R}\}$.

Esto es, *V* es el conjunto de puntos en $\mathbb{R}^3$ que están en el plano con vector normal (*a*, *b*, *c*) y que pasa por el origen.

Veamos si se cumplen las dos propiedades de cerradura, es decir, *i* y *vi*.

Sean (x_1, y_1, z_1) y (x_2, y_2, z_2) elementos de *V*; veamos si

$$(x_1+x_2,\ y_1+y_2,\ z_1+z_2) \in V.$$

Como $(x_1, y_1, z_1) \in V$, (1)

$$ax_1+by_1+cz_1 = 0.$$

De igual forma,

$$ax_2+by_2+cz_2 = 0. \tag{2}$$

Al sumar y factorizar (1) y (2) obtenemos:

$$a(x_1+x_2)+b(y_1+y_2)+c(z_1+z_2) = 0.$$

Esta expresión significa que $(x_1+x_2,\ y_1+y_2,\ z_1+z_2) \in V$.

Ahora observemos qué sucede con $\alpha(x_1, y_1, z_1)$, $\alpha\in\mathbb{R}$.

$$a\,x_1+b\,y_1+c\,z_1 = 0,$$
$$\alpha(a\,x_1+b\,y_1+c\,z_1) = \alpha\,0 = 0,$$
$$a\,\alpha\,x_1+b\,\alpha\,y_1+c\,\alpha\,z_1 = 0.$$

Por lo tanto:

$$(\alpha\,x_1,\ \alpha\,y_1,\ \alpha z_1) \in V,$$

esto es:

$$\alpha(x_1, y_1\, z_1) \in V.$$

Además, si $\alpha = 0$, se tiene que (0, 0, 0) $\in V$, y con $\alpha = -1$ tenemos que:

$$-(x_1, y_1 z_1) \in V,$$

con lo cual verificamos las propiedades *iii* y *iv*. Las demás propiedades se comprueban fácilmente aplicando las propiedades de los números reales.

8. El siguiente ejemplo nos muestra que no necesariamente la suma vectorial ha de estar relacionada con la suma ordinaria, y que el vector cero, **0**, no tiene que relacionarse con el número real 0.

Sea $V = \{\mathbf{x} / \mathbf{x} \in \mathbb{R},\ \mathbf{x} > 0\}$. La suma y la multiplicación escalar se definen así:

$$\mathbf{x} \oplus \mathbf{y} = \mathbf{x}\mathbf{y} \quad \text{y} \quad \alpha \odot \mathbf{x} = \mathbf{x}^{\alpha},\quad \alpha \in \mathbb{R}.$$

Veamos que *V*, junto con estas operaciones, es un espacio vectorial. Comprobemos inicialmente las dos propiedades de cerradura *i* y *vi*.

Sean $\mathbf{x},\ \mathbf{y} \in V$; entonces, $\mathbf{x} \oplus \mathbf{y} = \mathbf{x}\mathbf{y} \in V$. El producto de dos reales positivos es un real positivo.

Ahora, sean $\alpha \in \mathbb{R}$ y $\mathbf{x} \in V$; entonces, $\alpha \odot \mathbf{x} = \mathbf{x}^{\alpha}$. Como **x** es un real positivo, al elevarlo a cualquier potencia real se obtiene un número real positivo.

Ejemplifiquemos algunas sumas y multiplicaciones escalares en *V*.

$$3 \oplus 4 = 3 \times 4 = 12$$
$$4 \oplus 8 = 32...$$

$$\underset{\text{Escalar}}{\underset{\downarrow}{2}} \odot \underset{\text{Vector}}{\underset{\downarrow}{3}} = 3^2 = 9$$

$$\underset{\text{Escalar}}{\underset{\downarrow}{-\frac{1}{2}}} \odot \underset{\text{Vector}}{\underset{\downarrow}{4}} = 4^{-1/2} = \frac{1}{4^{1/2}} = \frac{1}{2}$$

- Comprobemos la propiedad asociativa. Sean **x**, **y**, **z**, en *V*. Entonces,

$$(\mathbf{x} \oplus \mathbf{y}) \oplus \mathbf{z} \underset{\substack{\downarrow \\ \text{Definición de} \\ \text{la operación}}}{=} (\mathbf{x}\mathbf{y})\mathbf{z} \underset{\substack{\downarrow \\ \text{Propiedad asociativa} \\ \text{del producto entre reales}}}{=} \mathbf{x}(\mathbf{y}\mathbf{z}) = \mathbf{x} \oplus (\mathbf{y} \oplus \mathbf{z}).$$

- Idéntico aditivo:

Encuentre $\mathbf{0} \in V$ tal que

$$\mathbf{x} \oplus \mathbf{0} = \mathbf{x},$$
$$\mathbf{x}\,\mathbf{0} = \mathbf{x}.$$

Luego $\mathbf{0} = 1$; es decir, el número real positivo 1 es el idéntico aditivo o «cero» para este espacio.

- Inverso aditivo.

  Para todo $\mathbf{x} \in V$, $-\mathbf{x}$ debe satisfacer que

  $$\mathbf{x} \oplus (-\mathbf{x}) = 0,$$

  luego $\mathbf{x}(-\mathbf{x}) = 1$.

  Así que $-\mathbf{x} = \dfrac{1}{\mathbf{x}}$.

  Como $\mathbf{x} > 0$, $\dfrac{1}{\mathbf{x}} > 0$; por lo tanto,

  $$-\mathbf{x} \in V.$$

Las demás propiedades las dejamos como ejercicio.

1.4 Propiedades

Teorema 1

Sea V un espacio vectorial. Entonces:

i. El elemento neutro, $\mathbf{0}$, de V es único.

ii. Para todo $\mathbf{x} \in V$ existe un único $\mathbf{x}' \in V$ tal que $\mathbf{x} + \mathbf{x}' = \mathbf{0}$.

iii. $\alpha \mathbf{0} = \mathbf{0}$ para todo escalar α.

iv. $0\mathbf{x} = \mathbf{0}$ para todo $\mathbf{x} \in V$.

v. Si $\alpha \mathbf{x} = \mathbf{0}$, entonces $\alpha = 0$ o $\boldsymbol{x} = \mathbf{0}$ (o ambos).

vi. $(-1)\mathbf{x} = -\mathbf{x}$ para todo $\mathbf{x} \in V$.

Demostremos las partes *i*, *iii* y *vi* del teorema.

i. Supongamos que $\mathbf{0}$ y $\mathbf{0}'$ son elementos neutros de V; entonces,

   $\mathbf{0} + \mathbf{0}' = \mathbf{0}'$ y $\mathbf{0}' + \mathbf{0} = \mathbf{0}$, propiedad *iii*

   luego $\mathbf{0}' = \mathbf{0} + \mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0}$. propiedad *v*

iii. $\mathbf{0} + \mathbf{0} = \mathbf{0}$, propiedad *iii*

$\alpha(\mathbf{0}+\mathbf{0}) = \alpha\mathbf{0}$, multiplicando por α

$\alpha\mathbf{0} + \alpha\mathbf{0} = \alpha\mathbf{0}$, propiedad *vii*

$(\alpha\mathbf{0} + \alpha\mathbf{0}) + (-\alpha\mathbf{0}) = \alpha\mathbf{0} + (-\alpha\mathbf{0})$, (suma a ambos lados $-\alpha\mathbf{0}$)

$\alpha\mathbf{0} + (\alpha\mathbf{0} + (-\alpha\mathbf{0})) = \mathbf{0}$, propiedades *ii* y *iv*

$\alpha\mathbf{0} + \mathbf{0} = \mathbf{0}$, propiedad *iv*

$\alpha\mathbf{0} = \mathbf{0}$. propiedad *iii*

vi. $1 + (-1) = 0$,

$\mathbf{0} = 0\mathbf{x} = [1 + (-1)]\mathbf{x} = 1\mathbf{x} + (-1)\mathbf{x} = \mathbf{0}$,

$1\mathbf{x} + (-1)\mathbf{x} + (-\mathbf{x}) = -\mathbf{x}$,

$(\mathbf{x} + (-\mathbf{x})) + (-1)\mathbf{x} = -\mathbf{x}$,

$\mathbf{0} + (-1)\mathbf{x} = -\mathbf{x}$,

$(-1)\mathbf{x} = -\mathbf{x}$.

¿Qué propiedades se aplican en los pasos de la demostración?

Ejercicios del capítulo 1 (módulo 1)

1. Verifique con detalle que $\mathbb{R}^n$ es un espacio vectorial.

2. Verifique con detalle que $\mathbb{P}_n$ es un espacio vectorial.

3. En el ejemplo 7, verifique los axiomas *ii*, *v*, *vii*, *viii*, *ix* y *x* de la definición de espacio vectorial.

4. Verifique con detalle que M_{mn} es un espacio vectorial.

5. En el ejemplo 8 verifique el cumplimiento de los axiomas *v*, *vii*, *viii*, *ix* y *x* de la definición de espacio vectorial.

En los ejercicios 6 a 16 determine si el conjunto dado, junto con las operaciones dadas, es un espacio vectorial. Si no lo es, mencione al menos un axioma de la definición que no se cumple.

6. Sea $V = \{(x, y, z) \in \mathbb{R}^3 / z = 0, \quad x, y \in \mathbb{R}\}$ con las operaciones usuales de $\mathbb{R}^3$.

7. Sea $V = \mathbb{R}$, con las operaciones ordinarias de suma y multiplicación en $\mathbb{R}$.

8. Sea V el conjunto de matrices simétricas reales de $n \times n$ con las operaciones matriciales usuales.

9. Sea V el conjunto de matrices antisimétricas reales de $n \times n$ con las operaciones matriciales usuales.

10. Sea V el conjunto de matrices invertibles $n \times n$ con las operaciones matriciales ordinarias.

11. Sea V el conjunto de matrices no invertibles $n \times n$ con las operaciones matriciales ordinarias.

12. Sea $V = \{A_{n\times n} / AC = 0, \ C \text{ es una matriz constante de } n \times n\}$ con las operaciones matriciales ordinarias.

13. Sea V el conjunto de todos los pares ordenados de números reales (x, y) con $x \leq 0$, con las operaciones usuales en $\mathbb{R}^2$.

14. Sea $V = \{C\}$ (un conjunto con un solo elemento) con las operaciones de suma y producto por un escalar definidas por $C + C = C$ y $\alpha C = C$ para todo $\alpha \in \mathbb{R}$.

15. Sea V el conjunto $\mathbb{R}^3$ con la operación de suma usual y la operación producto por un escalar definida por:

$$\alpha \odot (x, y, z) = (x, 1, z), \quad \alpha \in \mathbb{R}.$$

16. Sea V el conjunto $\mathbb{R}^3$ con la operación de suma dada por $(x_1, y_1, z_1) \oplus (x_2, y_2, z_2) = (x_2, y_1 + y_2, z_2)$ y la operación producto por un escalar usual.

17. Demuestre las partes *ii*, *iv* y *v* del teorema 1.

18. Si **x** y **y** son vectores de un espacio vectorial V, demuestre que existe un vector único $\mathbf{z} \in V$ tal que $\mathbf{x} + \mathbf{z} = \mathbf{y}$.

Módulo 2

Subespacios

En $\mathbb{R}^3$ las rectas y los planos que pasan por el origen son subespacios.

Introducción

En muchas aplicaciones es necesario emplear subconjuntos de espacios vectoriales, que son a su vez espacios vectoriales. En el módulo 1 vimos que el conjunto de ternas ordenadas que forman un plano que pasa por el origen es un espacio vectorial, así que tanto $\mathbb{R}^3$ como este subconjunto son espacios vectoriales.

Objetivos

1. Reconocer en *V* aquellos subconjuntos que tienen estructura de espacio vectorial.
2. Aplicar el criterio de subespacio para determinar subespacios en un espacio vectorial *V*.
3. Aplicar las propiedades de los subespacios.

Preguntas básicas

1. ¿Qué es un subespacio?
2. Si *H* es un subespacio de *V*, ¿el elemento neutro de *V* pertenece a *H*?
3. ¿Es la intersección de subespacios de *V* un subespacio de *V*? ¿Lo es la unión?
4. ¿En todo espacio vectorial hay subespacios?

Contenidos

2.1 Definición y criterio de subespacio
2.2 Ejemplos
2.3 Propiedades

Vea el módulo 2 del programa de televisión *Álgebra lineal*.

2.1 Definición y criterio de subespacio

Sea H un subconjunto no vacío de un espacio vectorial V y supóngase que H en sí mismo es un espacio vectorial con las operaciones de suma y multiplicación por un escalar definidas en V. Entonces se dice que H es un *subespacio* de V.

El siguiente teorema nos mostrará un resultado que simplificará notablemente el trabajo para determinar si un subconjunto de V es o no un subespacio de V.

Teorema 1: Criterio de subespacio

Un subconjunto no vacío H del espacio vectorial V es un subespacio de V si se cumplen las dos reglas de cerradura. Esto es:

1. Si $\mathbf{x} \in H$ e $\mathbf{y} \in H$, entonces $\mathbf{x} + \mathbf{y} \in H$.
2. Si $\mathbf{x} \in H$, entonces $\alpha\mathbf{x} \in H$ para todo escalar α.

La demostración del teorema consiste en verificar que al darse las dos reglas de cerradura se verifican los demás axiomas de la definición de espacio vectorial.

Como los vectores de H también están en V, las leyes de asociatividad, conmutatividad, de distribución y de identidad multiplicativa (axiomas *ii*, *v*, *vii*, *viii*, *ix* y *x*) se cumplen.

Ahora, si $\mathbf{x} \in H$, entonces $0\mathbf{x} \in H$ por (2) y $0\mathbf{x} = \mathbf{0}$ (teorema 1, módulo 1). De igual manera, $(-1)\mathbf{x} \in H$ y $(-1)\mathbf{x} = -\mathbf{x}$ (teorema 1, módulo1), verificándose el axioma *iv*. Con esto queda completa la demostración.

El teorema anterior establece que para probar si un conjunto H, $H \subset V$ es o no un subespacio de V, basta con probar las dos reglas de cerradura.

Como consecuencia del teorema podemos expresar lo siguiente:

i. Todo subespacio de un espacio vectorial V contiene el $\mathbf{0}$.

ii. Todo espacio vectorial tiene dos *subespacios triviales*: el mismo espacio ($V \subset V$) y el conjunto que tiene como único elemento el módulo. ¿Por qué?

Los subespacios distintos a V y $\{\mathbf{0}\}$ se llaman *subespacios propios*.

2.2 Ejemplos

1. Sea $H = \{(x, y) / y = mx,\ m \text{ es un real fijo y } x \in \mathbb{R}\}$.

$$H \subset \mathbb{R}^2 \text{ y } H \neq \varnothing.$$

Sean (x_1, y_1) y $(x_2\ y_2)$ elementos de H; entonces,

$$(x_1, y_1) = (x_1,\ m\,x_1) \text{ y } (x_2\,, y_2) = (x_2, m\,x_2).$$

$$\begin{aligned}(x_1 + x_2\,,\ y_1 + y_2) &= (x_1 + x_2\,,\ mx_1 + mx_2) \\ &= (x_1 + x_2\,,\ m(x_1 + x_2)).\end{aligned}$$

Luego:

$$(x_1 + x_2\,,\ y_1 + y_2) \in H.$$

Ahora, sea $(x, y) \in H$ y veamos si $\alpha\,(x, y) \in H$.

$$(x, y) = (x, m\,x),$$
$$\alpha(x, y) = \alpha(x, mx) = (\alpha x,\ \alpha mx) = \big(\alpha x,\ m(\alpha\,x)\big),$$
$$(\alpha x,\ m(\alpha x)) \in H.$$

Por lo tanto:

$$\alpha(x, y) \in H.$$

En consecuencia, el conjunto de puntos del plano que están sobre una recta que pasa por el origen forma un subespacio de $\mathbb{R}^2$.

2. Sea $H = \left\{(x, y) \in \mathbb{R}^2 \,/\, y = mx + b\right\},\ H \subset \mathbb{R}^2,\ H \neq \varnothing$

Veamos si *H* es un subespacio de $\mathbb{R}^2$.

H está formado por los puntos sobre una recta de pendiente *m* e intercepto con el eje *y* igual a *b;* es decir, una recta que no pasa por el origen.

Sean $(x_1\,, y_1)$ y (x_2, y_2) elementos de *H*. Veamos si $(x_1, y_1) + (x_2, y_2) \in H$.

$$(x_1\,, y_1) = (x_1,\ m\,x_1 + b), \quad (x_2\,, y_2) = (x_2\,,\ m\,x_2 + b).$$

$$\begin{aligned}(x_1\,, y_1) + (x_2\,, y_2) = (x_1 + x_2\,,\ y_1 + y_2) &= (x_1 + x_2\,,\ m\,x_1 + b + mx_2 + b) \\ &= (x_1 + x_2\,,\ m\,(x_1 + x_2) + 2b).\end{aligned}$$

Luego $(x_1, y_1) + (x_2, y_2) \notin H$; por lo tanto, *H* no es un subespacio de $\mathbb{R}^2$.

3. Sea $H = \left\{(x, y, z) \in \mathbb{R}^3 \,/\, x = at,\ \ y = bt,\ \ z = ct;\ \ a,\, b,\, c,\, t \in \mathbb{R}\right\}$.

H es el conjunto de puntos de $\mathbb{R}^3$ que están sobre una recta del espacio que pasa por el origen. Comprobemos para *H* las dos reglas de cerradura.

Sean $(x_1, y_1, z_1)\ y\ (x_2, y_2, z_2) \in H;$ entonces,

$$(x_1\,, y_1, z_1) = (at_1, bt_1, ct_1),\ t_1 \in \mathbb{R}\,,$$
$$(x_2\,, y_2, z_2) = (at_2, bt_2, ct_2),\ t_2 \in \mathbb{R}\,,$$

$$(x_1, y_1, z_1) + (x_2, y_2, z_2) = (at_1 + at_2, bt_1 + bt_2, ct_1 + ct_2)$$
$$= (a(t_1 + t_2), b(t_1 + t_2), c(t_1 + t_2)), (t_1 + t_2) \in \mathbb{R}.$$

Luego $((x_1, y_1, z_1) + (x_2, y_2, z_2)) \in H$ y

$$\alpha(x_1, y_1, z_1) = (\alpha(at_1), \alpha(bt_1), \alpha(ct_1))$$
$$= (a(\alpha t_1), b(\alpha t_1), c(\alpha t_1)) \in H, (\alpha t_1 \in \mathbb{R}).$$

Así que H es un subespacio de $\mathbb{R}^3$.

En $\mathbb{R}^2$ tenemos los siguientes subespacios: los subespacios triviales $\mathbb{R}^2$ y $\{0\}$, y un subespacio propio, las rectas que pasan por el origen.

Los subespacios propios de $\mathbb{R}^3$ son las rectas y los planos que pasan por el origen.

4. Sabemos que todo polinomio es una función continua, así que el conjunto de polinomios de grado menor o igual que n en el intervalo cerrado $[a,b]$, $\mathbb{P}_n[a,b]$ es un subconjunto del conjunto de funciones continuas dentro del mismo intervalo.

$$\mathbb{P}_n[a,b] \subset C[a,b].$$

Como $\mathbb{P}_n$ es un espacio vectorial para todo entero n, entonces es un subespacio de $C[a,b]$.

5. Sea $C'[a,b]$ el conjunto de funciones con primera derivada continua, definidas en $[a,b]$.

Como toda función derivable es continua, entonces $C'[a,b] \subset C[a,b]$. Además, la suma de funciones diferenciables es una función diferenciable, y un múltiplo constante de una función diferenciable es diferenciable. Se ve entonces que $C'[a,b]$ es un subespacio propio de $C[a,b]$, ya que no toda función continua es diferenciable.

2.3 Propiedades

Teorema 2

Sean H_1 y H_2 subespacios de un espacio vectorial V. Entonces $H_1 \cap H_2$ es un subespacio de V.

Demostración

$H_1 \cap H_2 \neq \varnothing$ ya que $\mathbf{0} \in H_1 \cap H_2$.

Sean $\mathbf{x}_1$ y $\mathbf{x}_2$ elementos de $H_1 \cap H_2$, luego $\mathbf{x}_1 \in H_1$ y $\mathbf{x}_1 \in H_2$. Igualmente, $\mathbf{x}_2 \in H_1$ y $\mathbf{x}_2 \in H_2$. Como H_1 y H_2 son subespacios, se cumple que

$$(\mathbf{x}_1 + \mathbf{x}_2) \in H_1 \quad \text{y} \quad (\mathbf{x}_1 + \mathbf{x}_2) \in H_2.$$

Es decir:

$$(\mathbf{x}_1 + \mathbf{x}_2) \in H_1 \cap H_2.$$

Ahora, $\alpha\mathbf{x}_1 \in H_1$ y $\alpha\mathbf{x}_1 \in H_2$, ya que $\mathbf{x}_1 \in H_1$ y $\mathbf{x}_1 \in H_2$.

Luego:

$$\alpha\mathbf{x}_1 \in H_1 \cap H_2.$$

Por lo tanto, se cumplen las dos reglas de cerradura y $H_1 \cap H_2$ es un subespacio.

La intersección de subespacios es un subespacio, pero la unión $(H_1 \cup H_2)$ no necesariamente es un subespacio.

Ejemplo

$H_1 = \{(x, y) \in \mathbb{R}^2 \,/\, y = x\}$.

$H_2 = \{(x, y) \in \mathbb{R}^2 \,/\, y = 2x\}$.

$(1, 1) \in H_1$ y $(1, 2) \in H_2$.

Entonces:

$$(1, 1) \in (H_1 \cup H_2) \text{ y } (1, 2) \in H_1 \cup H_2.$$

Pero $(1, 1) + (1, 2) = (2, 3) \notin (H_1 \cup H_2)$ ya que $(2, 3) \notin H_1$ y $(2, 3) \notin H_2$. Así que $H_1 \cup H_2$ no es cerrado bajo la suma ; por lo tanto, no es un subespacio.

Ejercicios del capítulo 1 (módulo 2)

En los ejercicios 1 a 12 se dan un espacio V y un subconjunto W. Determine si W es un subespacio.

1. $V = \mathbb{R}^2,\ W = \{(x, y) : x = y\}$.
2. $V = M_{nn},\ W = \{A \in M_{nn} : A \text{ es simétrica}\}$.
3. $V = \mathbb{R}^3,\ W = \{(x, y, z) \in \mathbb{R}^3 : (x, y, z) \text{ es perpendicular a } (a, b, c)\}$.
4. $V = \mathbb{P}_4,\ W = \{p \in \mathbb{P}_4 : p(0) = 0\}$.
5. $V = \mathbb{P}_4,\ W = \{p \in \mathbb{P}_4 : p(0) = 1\}$.
6. $V = \mathbb{P}_4,\ W = \{p \in \mathbb{P}_4 : \text{ grado de } p = 4\}$.
7. La traza de una matriz $A_{n\times n}$ se define como:

$$\text{tr } A = a_{11} + a_{22} + \ldots + a_{nn}.$$

Sean $V = M_{nn}$ y $W = \{A \in M_{nn} : \text{ tr } A = 0\}$.

8. $V = M_{22},\ W = \left\{A \in M_{22} : A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}\right\}$.
9. $V = M_{23},\ W = \left\{A \in M_{23} : A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}, \text{ donde } a = 2c + 1\right\}$.
10. $V = M_{23},\ W = \left\{A \in M_{23} : A = \begin{bmatrix} a & b & c \\ d & 0 & 0 \end{bmatrix}, \text{ donde } b = a + c\right\}$.
11. $V = C[0, 1],\ W = \{f \in C[0, 1] : f(0) = f(1) = 0\}$.
12. $V = C[0, 1],\ W = \{f \in C[0, 1] : f(0) = 1\}$.
13. Sean $V = M_{23},\ W_1 = \left\{A \in M_{23} : A = \begin{bmatrix} 0 & b & c \\ d & e & f \end{bmatrix}\right\}$ y $W_2 = \left\{A \in M_{23} : A = \begin{bmatrix} a & b & c \\ d & e & 0 \end{bmatrix} \text{ con } b = a + c\right\}$.
    a. Demuestre que W_1 y W_2 son subespacios.
    b. Describa el subconjunto $W = W_1 \cap W_2$ y demuestre que es un subespacio.

14. Sea $W = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{X} = \mathbf{0},\ \text{con } A \in M_{m\times n}\}$.

Demuestre que W es un subespacio de $\mathbb{R}^n$. W se llama *espacio nulo* de la matriz A.

15. Sean $\mathbf{x} = (1, 2, 4)$ e $\mathbf{y} = (-3, 2, 0)$ dos vectores en $\mathbb{R}^3$ y sea

$$W = \left\{\mathbf{u} \in \mathbb{R}^3 : \mathbf{u} = \alpha\mathbf{x} + \beta\mathbf{y},\ \text{con } \alpha, \beta \in \mathbb{R}\right\}.$$

Demuestre que W es un subespacio de $\mathbb{R}^3$.

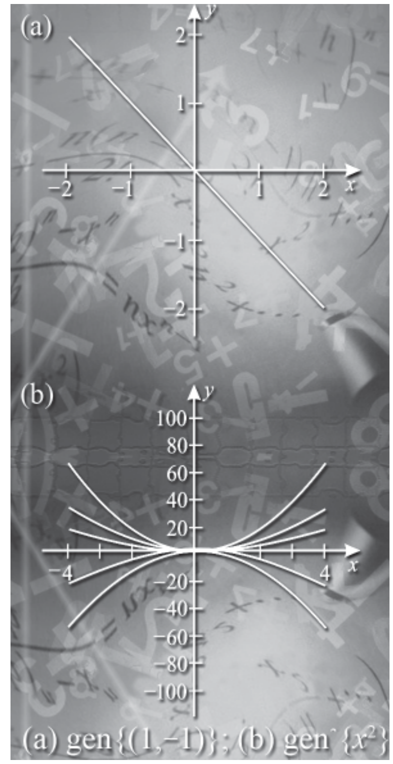
16. Sean W_1 y W_2 subespacios de un espacio vectorial V.

Sea $W_1 + W_2 = \{\mathbf{v} : \mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 \text{ con } \mathbf{v}_1 \in W_1 \text{ y } \mathbf{v}_2 \in W_2\}$.

Demuestre que $W_1 + W_2$ es un subespacio de V.

Módulo 3

Combinación lineal y subespacio generado



La gráfica representa los subespacios generados por $\left\{\begin{pmatrix}1\\-1\end{pmatrix}\right\}$ y $\left\{x^2\right\}$.

Introducción

Iniciamos el módulo definiendo qué es combinación lineal, concepto necesario para la construcción de subespacios a partir de subconjuntos finitos del espacio vectorial V. Al subconjunto dado lo llamamos conjunto generador, y al subespacio construido, subespacio generado.

Objetivos

1. Definir una combinación lineal.
2. Generar un subespacio de un espacio vectorial a partir de un conjunto.
3. Diferenciar un conjunto generador del subespacio generado por él.

Preguntas básicas

1. ¿Qué es una combinación lineal?
2. ¿Cuándo un conjunto genera un espacio vectorial?
3. ¿Cómo se genera un subespacio de V a partir de un subconjunto de V?

Contenidos

3.1 Definición y ejemplos de combinación lineal
3.2 Conjunto generador de un espacio vectorial
3.3 Espacio generado por un conjunto de vectores

Vea el módulo 3 del programa de televisión *Álgebra Lineal.*

3.1 Definición y ejemplos de combinación lineal

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Combinación lineal».

Definición 1

Sean $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ vectores en un espacio vectorial *V*. Decimos que **x** es *una combinación lineal* de los vectores $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ si existen escalares $a_1, a_2, ..., a_n$, tales que:

$$\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + ... + a_n\mathbf{v}_n$$
$$= \sum_{i=1}^{n} a_i\mathbf{v}_i .$$

Observaciones

1. Toda combinación lineal de vectores de un espacio dado es otro vector del mismo espacio.

2. El módulo de un espacio vectorial es combinación lineal de cualquier conjunto de vectores del mismo espacio.

$$\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + ... + 0\mathbf{v}_n .$$

3. Todo vector es combinación lineal de sí mismo.

$$\mathbf{x}_1 = 1\mathbf{x}_1 .$$

4. Todo vector es combinación lineal del conjunto que contiene dicho vector.

$$\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + ... + 0\mathbf{v}_n .$$

Ejemplos

1. La figura 3.1 muestra un vector **v** en $\mathbb{R}^2$ o $\mathbb{R}^3$ como combinación lineal de los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$.

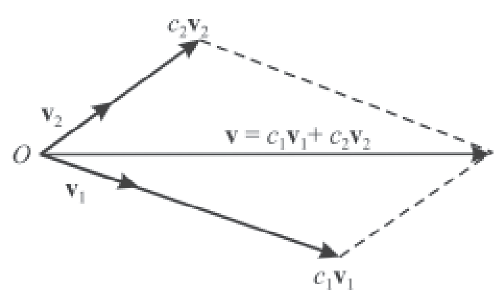


Figura 3.1

2. Si $\vec{a}$ y $\vec{b}$ son vectores libres, $\Pr(\vec{a}/\vec{b})$, la proyección de $\vec{a}$ sobre $\vec{b}$, es una combinación lineal de $\vec{b}$, y $\Pr(\vec{b}/\vec{a})$ es una combinación lineal de $\vec{a}$.

3. En $\mathbb{R}^3$, $\begin{pmatrix}-5\\-2\\-7\end{pmatrix}$ es una combinación lineal de los vectores $\begin{pmatrix}4\\1\\2\end{pmatrix}$ y $\begin{pmatrix}1\\0\\-1\end{pmatrix}$ ya que:

$$\begin{pmatrix}-5\\-2\\-7\end{pmatrix} = -2\begin{pmatrix}4\\1\\2\end{pmatrix} + 3\begin{pmatrix}1\\0\\-1\end{pmatrix}.$$

4. En $\mathbb{R}^2$, sean $\mathbf{v}_1 = (2, 1)$ y $\mathbf{v}_2 = (-4, -2)$. Veamos si $\mathbf{v} = (6, 3)$ es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$.

Debemos entonces determinar si existen escalares a_1 y a_2 tales que

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2.$$
$$(6, 3) = a_1(2,1) + a_2(-4, -2)$$

Luego

$$6 = 2a_1 - 4a_2,$$
$$3 = a_1 - 2a_2,$$
$$\begin{bmatrix}2 & -4\\1 & -2\end{bmatrix}\begin{bmatrix}a_1\\a_2\end{bmatrix} = \begin{bmatrix}6\\3\end{bmatrix}.$$

La matriz aumentada es:

$$\left[\begin{array}{cc|c}2 & -4 & 6\\1 & -2 & 3\end{array}\right] \xrightarrow[\text{se reduce a}]{\text{operaciones elementales}} \left[\begin{array}{cc|c}1 & -2 & 3\\0 & 0 & 0\end{array}\right].$$

El sistema tiene infinitas soluciones dadas por $a_1 = 3 + 2a_2$ y $a_2 \in \mathbb{R}$. Por ejemplo, si $a_2 = 1$, $a_1 = 5$ y $(6, 3) = 5(2, 1) + (-4, -2)$; pero si hacemos $a_2 = 0$, $a_1 = 3$ y $(6, 3) = 3(2, 1) + 0(-4, -2)$, etc.

De modo que **v** es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ de muchas maneras.

Podemos entonces observar que si un vector **v** es una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$, entonces esa combinación lineal puede no ser única.

5. En $\mathbb{P}_3$, sea $A = \{-1 + x,\ 1 + x^2,\ 2x + x^3\}$. Veamos si $Q(x) = 3 + x^2 + x^3$ es una combinación lineal de *A*.

Expresemos a $Q(x)$ como una combinación lineal de los polinomios de *A*.

$$3+x^2+x^3 = a_1(-1+x)+a_2(1+x^2)+a_3(2x+x^3).$$

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Mezcla de concreto».

Al igualar los coeficientes de los términos del mismo grado en ambos lados de la igualdad, tenemos:

(1) $3 = -a_1 + a_2.$

(2) $0 = a_1 + 2a_3.$

(3) $1 = a_2.$

(4) $1 = a_3.$

Con el valor $a_2 = 1$ en (1), $3 = -a_1 + 1$, entonces $a_1 = -2.$

Con $a_1 = -2$ y $a_3 = 1$, la ecuación (2) se expresa como $0 = -2 + 2(1)$.

Por lo tanto, hemos determinado escalares a_1, a_2, a_3 tales que:

$$3 + x^2 + x^3 = -2(-1 + x) + (1 + x^2) + (2x + x^3).$$

3.2 Conjunto generador de un espacio vectorial

Definición 2

Sean $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ vectores en un espacio vectorial V, y suponga que todo vector $\mathbf{v} \in V$ se puede expresar como una combinación lineal de estos vectores. Entonces se dice que $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es un *conjunto generador de V*, o que los vectores $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ generan a V.

Es decir, para todo $\mathbf{v} \in V$, existen $a_1, a_2, ..., a_n$ reales, tales que:

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + ... + a_n\mathbf{v}_n.$$

Ejemplos

6. Los vectores $\begin{pmatrix}1\\0\end{pmatrix}$ y $\begin{pmatrix}0\\1\end{pmatrix}$ generan $\mathbb{R}^2$.

Por ejemplo, $\begin{pmatrix}2\\4\end{pmatrix} = 2\begin{pmatrix}1\\0\end{pmatrix} + 4\begin{pmatrix}0\\1\end{pmatrix}$,

y los vectores $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$ generan $\mathbb{R}^3$.

7. En $\mathbb{P}_n$ todo polinomio se puede expresar como combinación lineal de los monomios $1, x, x^2, \ldots, x^n$; luego generan a $\mathbb{P}_n$.

8. En M_{22}: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

Entonces, $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ generan M_{22}.

3.3 Espacio generado por un conjunto de vectores

Definición 3

Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ n vectores en un espacio vectorial V. El *espacio generado* por $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es el conjunto de las combinaciones lineales de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Esto es: $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} = \{\mathbf{v} : \mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n\}$, donde $a_1, a_2, \ldots, a_n$ son escalares.

Del conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ se dice que es *generador*.

Teorema 1

Si $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son n vectores en un espacio vectorial V, el $\text{gen}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es un subespacio de V.

La demostración del teorema se deja como ejercicio.

Ejemplos

9. En $\mathbb{R}$, sea $A = \{1\}$.

$\text{gen } A = \{x / x = a \cdot 1,\ a \in \mathbb{R}\} = \mathbb{R}$.

$\{1\}$ es generador para $\mathbb{R}$.

10. En $\mathbb{R}^3$, sea $A = \{(1, -2, 0), (3, 1, -1)\}$.

$\text{gen}(A) = \{(x, y, z) / (x, y, z) = a_1(1, -2, 0) + a_2(3, 1, -1);\ a_1, a_2 \in \mathbb{R}\}$.

Determinemos una relación entre x, y, z.

$$\begin{aligned} (x, y, z) &= (a_1 + 3a_2, -2a_1 + a_2, -a_2), \\ x &= a_1 + 3a_2, \\ y &= -2a_1 + a_2, \\ z &= -a_2. \end{aligned}$$

$$\left[\begin{array}{rr|r} 1 & 3 & x \\ -2 & 1 & y \\ 0 & -1 & z \end{array}\right] \begin{array}{l} R_2 \to R_2 + 2R_1 \\ R_3 \to -R_3 \end{array} \left[\begin{array}{rr|c} 1 & 3 & x \\ 0 & 7 & y+2x \\ 0 & 1 & -z \end{array}\right] \begin{array}{l} R_2 \to \frac{1}{7}R_2 \\ R_3 \to R_3 - R_2 \end{array} \left[\begin{array}{rr|c} 1 & 3 & x \\ 0 & 1 & \frac{y+2x}{7} \\ 0 & 0 & -z-\frac{y+2x}{7} \end{array}\right]$$

Luego, los elementos de gen (A) deben satisfacer

$$-z - \frac{(y+2x)}{7} = 0,$$

o sea:

$$2x + y + 7z = 0.$$

$$\text{gen}\,(A) = \{(x, y, z) / 2x + y + 7z = 0\} \text{ (plano que pasa por el origen).}$$

Nota: el espacio generado por dos vectores no nulos en $\mathbb{R}^3$ y que no sean paralelos es un plano que pasa por el origen.

Teorema 2

Sean $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$, $\mathbf{v}_{n+1}$, $n+1$ vectores que están en un espacio vectorial V. Si $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ genera a V, entonces $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n, \mathbf{v}_{n+1}$ también genera a V. Esto es, la adición de uno o más vectores a un conjunto generador de V da por resultado otro conjunto generador de V.

Demostración

Sea $\mathbf{v} \in V$; como $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ genera a V, entonces existen escalares $a_1, a_2, ..., a_n$ tales que $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + ... + a_n\mathbf{v}_n$.

Así que $\mathbf{v}$ también puede expresarse como

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + ... + a_n\mathbf{v}_n + 0\mathbf{v}_{n+1},$$

lo cual muestra que $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n, \mathbf{v}_{n+1}\}$ también genera V.

Ejemplo

El espacio vectorial $\mathbb{R}^2$ puede ser generado por el conjunto $A = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$.

También puede ser generado por un conjunto mayor $A' = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$.

Ejercicios del capítulo 1 (módulo 3)

En los ejercicios 1 a 3 determine si el vector dado $\mathbf{w}$ es combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$. Si lo es, encuentre a_1 y a_2 tales que $\mathbf{w} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2$.

1. $\mathbf{v}_1 = (2, -1)$ y $\mathbf{v}_2 = (-4, 2)$.

   a. $\mathbf{w} = (-6, 3)$.
   b. $\mathbf{w} = (1, 1)$.
   c. $\mathbf{w} = (0, 0)$.

2. $\mathbf{v}_1 = (1, -3)$ y $\mathbf{v}_2 = (-2, 6)$.

   a. $\mathbf{w} = (-3, 9)$.
   b. $\mathbf{w} = (3, 6)$.
   c. $\mathbf{w} = (0, 3)$.

3. $\mathbf{v}_1 = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix}$.

   a. $\mathbf{w} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.
   b. $\mathbf{w} = \begin{bmatrix} -3 & 2 \\ -4 & 1 \end{bmatrix}$.
   c. $\mathbf{w} = \begin{bmatrix} -13 & -6 \\ 1 & -3 \end{bmatrix}$.

4. Sean $p_1 = x + x^2$, $p_2 = x + x^3$ y $p_3 = x + x^2 + x^3$.

   Determine cuál o cuáles de los siguientes polinomios son combinación lineal de p_1, p_2 y p_3.

   a. $2x + x^2$
   b. $2 - 3x + 4x^2 + x^3$
   c. 0
   d. x

En los ejercicios 5 a 10 describa el espacio generado por los vectores dados.

5. $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$

6. $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}.$

7. $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 6 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 7 \\ 0 \\ 2 \end{bmatrix}.$

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

9. $p_1 = 2x + 3, \;\; p_2 = -3x - 5.$

10. $p_1 = 3x + 4x^2, \;\; p_2 = 2x - 5x^2, \;\; p_3 = x + x^2.$

En los ejercicios 11 a 14 determine si el conjunto dado de vectores genera el espacio vectorial dado.

11. En $\mathbb{R}^2$; $\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \end{pmatrix}.$

12. En $\mathbb{R}^2$; $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \begin{pmatrix} -3 \\ -6 \end{pmatrix}.$

13. En $\mathbb{R}^3$; $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}.$

14. En $\mathbb{R}^3$; $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ 9 \\ 3 \end{pmatrix}.$

15. Muestre que un conjunto de dos vectores de $\mathbb{R}^3$ no puede generar a $\mathbb{R}^3$.

16. Demuestre que si $\mathbf{u}$ y $\mathbf{v}$ están en $\text{gen}\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$, entonces $\mathbf{u} + \mathbf{v}$ y $\alpha\mathbf{u}$ están en $\text{gen}\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$.

Módulo 4

Independencia lineal

Si $\mathbf{v}_1$ y $\mathbf{v}_2$ son dos vectores no paralelos se dice que son LI, o sea que uno no es múltiplo escalar del otro y generan un plano.

Si los vectores son paralelos ya no pueden generar un plano, sino sólo una recta, en este caso son LD.

Introducción

En el módulo anterior vimos que pueden tenerse infinitud de conjuntos generadores de un espacio vectorial. Es claro, por razones de economía, que lo que se desea es hallar los conjuntos generadores más pequeños posibles para los espacios vectoriales. Para lograr esto, se requiere el concepto de *independencia lineal*.

Mostraremos el significado de independencia lineal y su relación con la teoría de sistemas homogéneos de ecuaciones y determinantes.

Objetivos

1. Establecer los conceptos de dependencia e independencia lineal.
2. Interpretar geométricamente el significado de la dependencia lineal en $\mathbb{R}^3$.
3. Vincular la teoría de espacios vectoriales con los sistemas de ecuaciones lineales y los determinantes.

Preguntas básicas

1. ¿Cuándo un conjunto es LD?
2. ¿Cuándo un conjunto es LI?
3. ¿Qué significa que un subconjunto de tres vectores de $\mathbb{R}^3$ sea LD?
4. ¿Tiene sentido hablar de un vector LI?
5. ¿Por qué los conceptos de independencia y dependencia lineal se vinculan con la solución de sistemas homogéneos de ecuaciones lineales?

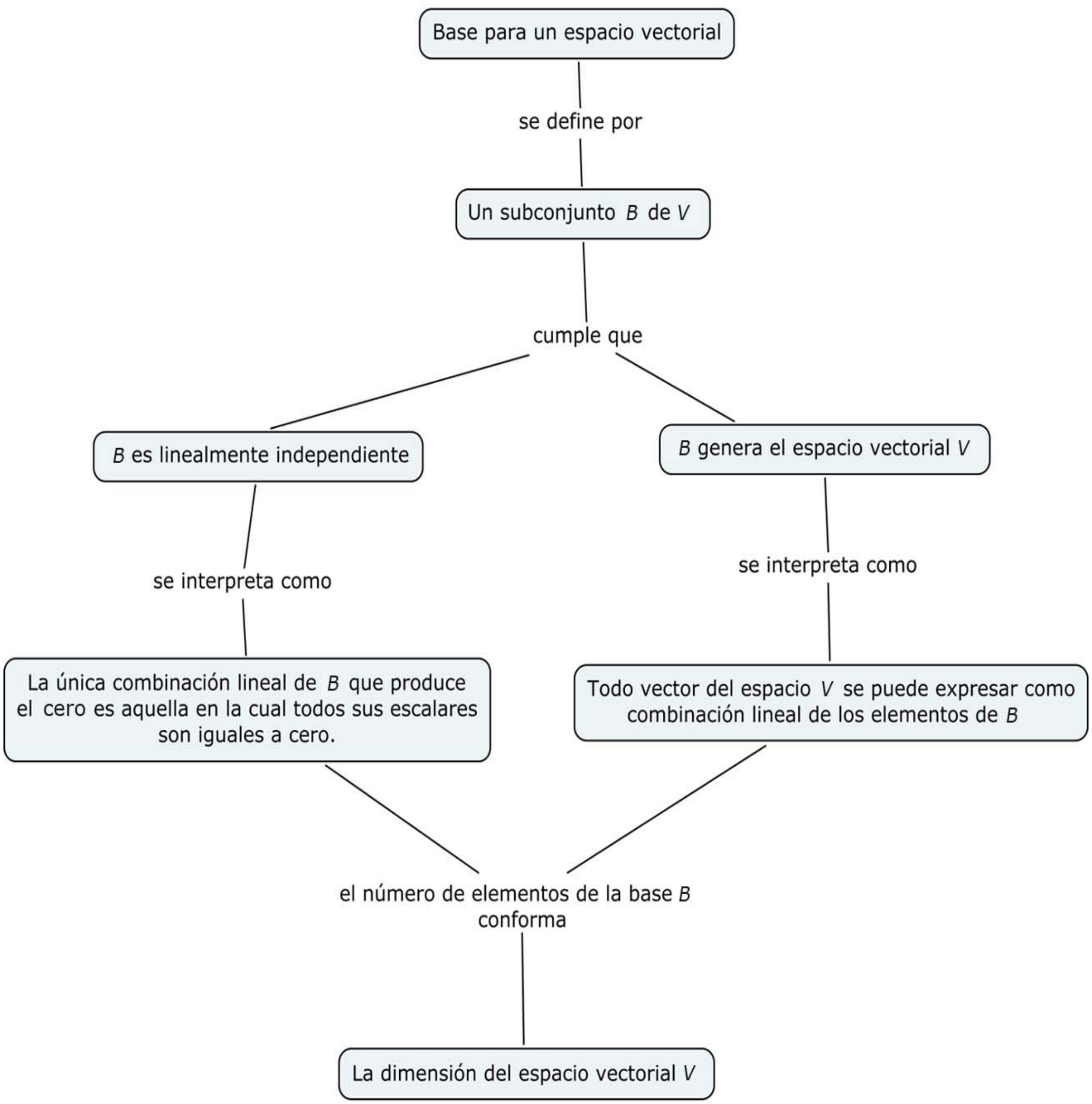
Contenidos

4.1 Definición y ejemplos de conjuntos linealmente independientes (LI) y linealmente dependientes (LD)

4.2 Interpretación geométrica de la dependencia lineal en $\mathbb{R}^3$

4.3 Propiedades

Vea el módulo 4 del programa de televisión *Álgebra lineal*.



Mapa 2: módulos 4 y 5

4.1 Definición y ejemplos de conjuntos linealmente independientes (LI) y linealmente dependientes (LD)

Definición 1

Sean $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$, n vectores diferentes de un espacio vectorial V. Se dice que los vectores son *linealmente dependientes*, LD, si existen n escalares $c_1, c_2, ..., c_n$, no todos cero, tales que:

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n = \mathbf{0}. \tag{1}$$

Si los vectores no son linealmente dependientes, se dice que son *linealmente independientes*, LI, es decir, la ecuación:

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n = \mathbf{0}$$

es válida sólo si $c_1 = c_2 = ... = c_n = 0$.

Si los vectores $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ son distintos y formamos el conjunto $A = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$, entonces también decimos que el conjunto A es LD o LI según el caso.

Es claro que la ecuación (1) siempre se satisface si elegimos todos los escalares $c_1, c_2, ..., c_n$ iguales a cero. Lo importante es ver si es posible satisfacer la ecuación con al menos uno de los escalares diferente de cero.

Ejemplos

1. En $\mathbb{R}^3$, sea $A = \{(1, 7, -2), (-1, 1, 0), (3, 5, -2)\}$. Veamos si A es LD o LI.

Solución

Comenzamos planteando la ecuación

$$c_1(1, 7, -2) + c_2(-1, 1, 0) + c_3(3, 5, -2) = (0, 0, 0),$$

la cual, al igualar los componentes en ambos lados de la ecuación, nos lleva al siguiente sistema homogéneo de ecuaciones:

$$\begin{aligned} c_1 - c_2 + 3c_3 &= 0 \\ 7c_1 + c_2 + 5c_3 &= 0 \\ -2c_1 \qquad - 2c_3 &= 0 \end{aligned} \qquad A\mathbf{x} = \mathbf{0}, \qquad \mathbf{x} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}.$$

La matriz de coeficientes del sistema $\begin{bmatrix} 1 & -1 & 3 \\ 7 & 1 & 5 \\ -2 & 0 & -2 \end{bmatrix}$ se reduce por medio de operaciones elementales a $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$. El sistema equivalente es:

$$\begin{aligned} c_1 &= -c_3. \\ c_2 &= 2c_3. \\ c_3 &= \ \ c_3. \end{aligned}$$

La fila nula de esta matriz nos indica que el sistema tiene infinitas soluciones, es decir, hay soluciones con $\mathbf{x} \neq \mathbf{0}$ $(c_3 \neq 0)$; por lo tanto, A es LD.

2. En $\mathbb{P}_2$, sea $B = \{-1 + x^2,\ 2 + x,\ x^2\}$. ¿Es B LD o LI?

Solución

Dada la ecuación

$$c_1(-1 + x^2) + c_2(2 + x) + c_3x^2 = 0 + 0x + 0x^2,$$

veamos cómo son c_1, c_2, c_3.

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

(2) $c_2 = 0.$

(3) $c_1 + c_3 = 0.$

(2) en (1), $-c_1 = 0,$ o sea $c_1 = 0$ (4).

(4) en (3), $c_3 = 0.$

Luego $c_1 = c_2 = c_3 = 0$; por lo tanto, B es LI.

Proposiciones

1. Dos vectores en un espacio vectorial V son LD si y sólo si uno de ellos es un múltiplo escalar del otro.
2. Todo conjunto que tenga como elemento el módulo de un espacio vectorial es LD.
3. Todo conjunto unitario cuyo elemento no sea el módulo es LI.

4.2 Interpretación geométrica de la dependencia lineal en $\mathbb{R}^3$

Suponga que $\mathbf{u}, \mathbf{v}, \mathbf{w}$ son tres vectores LD en $\mathbb{R}^3$; entonces existen escalares c_1, c_2, c_3, no todos cero, tales que:

$$c_1\mathbf{u} + c_2\mathbf{v} + c_3\mathbf{w} = \mathbf{0}.$$

Suponga que $c_3 \neq 0$; entonces $\mathbf{w} = -\dfrac{c_1}{c_3}\mathbf{u} - \dfrac{c_2}{c_3}\mathbf{v} = \alpha\mathbf{u} + \beta\mathbf{v}$,

con $\alpha = -\dfrac{c_1}{c_3}$ y $\beta = -\dfrac{c_2}{c_3}$.

Veamos entonces que $\mathbf{u}, \mathbf{v}$ y $\mathbf{w}$ son coplanares:

$$\begin{aligned}\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) &= (\alpha\mathbf{u} + \beta\mathbf{v}) \cdot (\mathbf{u} \times \mathbf{v}) \\ &= \alpha(\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v})) + \beta(\mathbf{v} \cdot (\mathbf{u} \times \mathbf{v})) \\ &= \alpha 0 + \beta 0 = 0.\end{aligned}$$

Recuerde que tres vectores del espacio son coplanares si y sólo si su producto triple es igual a cero.

¿Por qué $\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v})$ y $\mathbf{v} \cdot (\mathbf{u} \times \mathbf{v})$ son iguales a cero?

En conclusión: si $\mathbf{u}, \mathbf{v}, \mathbf{w}$ son tres vectores LD en $\mathbb{R}^3$, ellos son coplanares.

¿Será cierta la recíproca de esta implicación? Es decir, si $\mathbf{u}, \mathbf{v}, \mathbf{w}$ son tres vectores coplanares de $\mathbb{R}^3$, entonces $\mathbf{u}, \mathbf{v}, \mathbf{w}$ son LD.

4.3 Propiedades

Teorema 1

Un conjunto de n vectores de $\mathbb{R}^m$ siempre es LD si $n > m$.

Demostración

Sean $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ n vectores en $\mathbb{R}^m$.

Examinemos la ecuación

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n = \mathbf{0},$$

donde $\mathbf{v}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \; \mathbf{v}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, ..., \; \mathbf{v}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$

Entonces la ecuación se convierte en:

$$\begin{aligned} a_{11}c_1 + a_{12}c_2 + ... + a_{1n}c_n &= 0 \\ a_{21}c_1 + a_{22}c_2 + ... + a_{2n}c_n &= 0 \\ \vdots \quad\quad \vdots \quad\quad\quad \vdots & \\ a_{m1}c_1 + a_{m2}c_2 + ... + a_{mn}c_n &= 0. \end{aligned}$$

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Conjunto de vectores coplanares».

Como $n > m$ (nº de incógnitas > nº de ecuaciones), el sistema tiene infinitas soluciones. Por tanto, existen escalares $c_1, c_2, ..., c_n$, no todos cero, que satisfacen la ecuación y $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es LD.

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Verificación de independencia lineal».

Corolario

Un conjunto de vectores LI en $\mathbb{R}^n$ contiene a lo más *n* vectores.

Teorema 2

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

Entonces las columnas de *A* consideradas como vectores son LD si y sólo si el sistema $A\mathbf{x} = \mathbf{0}$ tiene soluciones no triviales (diferentes a $\mathbf{x} = \mathbf{0}$).

Teorema 3

Sean $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$, *n* vectores en $\mathbb{R}^n$, y sea *A* una matriz $n \times n$ cuyas columnas son $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$. Entonces $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ son LI si y sólo si la única solución al sistema homogéneo $A\mathbf{x} = \mathbf{0}$ es la trivial $\mathbf{x} = \mathbf{0}$.

Teorema 4

Sea *A* una matriz $n \times n$. Entonces $\det A \neq 0$ si y sólo si las columnas de *A* son LI.

Teorema 5

Todo conjunto de *n* vectores LI en $\mathbb{R}^n$ genera $\mathbb{R}^n$.

Demostración

Suponga que $\mathbf{v}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}$, ..., $\mathbf{v}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix}$ son LI y sea $\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ un vector de $\mathbb{R}^n$.

Veamos que existen escalares $c_1, c_2, ..., c_n$ tales que $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n$.

$$\begin{array}{rlllll} x_1 = & c_1a_{11} + & c_2a_{12} + & \cdots + & c_na_{1n} \\ x_2 = & c_1a_{21} + & c_2a_{22} + & \cdots + & c_na_{2n} \\ & \vdots & \vdots & & \vdots \\ x_n = & c_1a_{n1} + & c_2a_{n2} + & \cdots + & c_na_{nn} \end{array} \qquad \mathbf{C} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

Luego, el sistema $A\mathbf{C} = \mathbf{v}$ tiene matriz $\underset{n\times n}{A}$, tal que $\det A \neq 0,$ ya que las columnas de A son LI; por lo tanto, el sistema tiene solución única $\mathbf{C}$ y el teorema queda establecido.

Teorema 6

Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\} \subset V.$

S es LD si y sólo si existe al menos un vector de *S* que pueda expresarse como combinación lineal de los vectores restantes de *S*.

Demostración

i. Supongamos que *S* es LD.

Existen $a_1, a_2, ..., a_n$ escalares, no todos nulos, tales que:

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + ... + a_i\mathbf{v}_i + ... \; a_n\mathbf{v}_n = \mathbf{0}.$$

Sea $a_i \neq 0,$ entonces,

$$\mathbf{v}_i = -\frac{a_1}{a_i}\mathbf{v}_1 - \frac{a_2}{a_i}\mathbf{v}_2 - ... - \frac{a_n}{a_i}\mathbf{v}_n \,.$$

Luego $\mathbf{v}_i$ es una combinación lineal de los restantes n – 1 vectores de *S*.

ii. Si $\mathbf{v}_i$ es una combinación lineal de los restantes vectores de *S*, entonces:

$$\mathbf{v}_i = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + ... + \lambda_{i-1}\mathbf{v}_{i-1} + \lambda_{i+1}\mathbf{v}_{i+1} + ... + \lambda_n\mathbf{v}_n.$$

Por lo tanto,

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + ... + \lambda_{i-1}\mathbf{v}_{i-1} + (-1)\mathbf{v}_i + \lambda_{i+1}\mathbf{v}_{i+1} + ... + \lambda_n\mathbf{v}_n = \mathbf{0}.$$

En esta combinación lineal de los vectores de *S* igualada a cero, al menos el coeficiente de $\mathbf{v}_i\,(-1)$ es diferente de cero; por lo tanto, *S* es LD.

Observación

Si un conjunto *S* LD genera un espacio vectorial *V*, el conjunto que resulta de eliminar en *S* un vector que sea combinación lineal del resto de vectores de *S*, también genera el espacio vectorial *V*.

Ejemplo 3

Sea el conjunto de vectores $S=\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4\}\subset\mathbb{R}^4$

con $\mathbf{v}_1=\begin{bmatrix}1\\1\\0\\0\end{bmatrix}$, $\mathbf{v}_2=\begin{bmatrix}1\\0\\1\\0\end{bmatrix}$, $\mathbf{v}_3=\begin{bmatrix}0\\1\\1\\0\end{bmatrix}$ y $\mathbf{v}_4=\begin{bmatrix}2\\1\\1\\0\end{bmatrix}$,

y sea $W=\text{gen}\,S$. Como $\mathbf{v}_4=\mathbf{v}_1+\mathbf{v}_2$, entonces $W=\text{gen}\,S_1$, donde $S_1=\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$.

Ejercicios del capítulo 1 (módulo 4)

En los ejercicios 1 a 4 muestre por inspección que los vectores son linealmente dependientes (LD).

1. $\mathbf{u}_1 = (2,\ 3),\quad \mathbf{u}_2 = (-4,\ -6),$ en $\mathbb{R}^2$.

2. $\mathbf{u}_1 = (5,\ 2,\ 3),\ \mathbf{u}_2 = (0,\ 3,\ -1),\ \mathbf{u}_3 = (-1,\ 4,\ 0),\ \mathbf{u}_4 = (5,\ 7,\ -2),$ en $\mathbb{R}^3$.

3. $P_1 = -1 + 4x,\quad P_2 = \dfrac{1}{2} - 2x,$ en $\mathbb{P}_1$.

4. $A_1 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix},\quad A_2 = \begin{bmatrix} 0 & 1 \\ 4 & 2 \end{bmatrix},\quad A_3 = \begin{bmatrix} 1 & 3 \\ 4 & 1 \end{bmatrix},$ en M_{22}.

En los ejercicios 5 a 11 determine si el conjunto de vectores dado es LI o LD.

5. $(2,\ -1,\ 3),\ (3,\ 4,\ 1),\ (2,\ -3,\ 4),$ en $\mathbb{R}^3$.

6. $(3,\ 1,\ 4,\ 1),\ (1,\ 1,\ 0,\ 0),\ (0,\ 0,\ 2,\ 2),$ en $\mathbb{R}^4$.

7. $(3,\ 0,\ 2,\ -2),\ (5,\ 0,\ 3,\ -1),\ (1,\ -2,\ 1,\ 1),\ (0,\ 4,\ -1,\ 1),$ en $\mathbb{R}^4$.

8. $P_1(x) = 1 + x + x^2,\quad P_2(x) = 2 - x + 3x^2,\quad P_3(x) = -1 + 5x - 3x^2,$ en $\mathbb{P}_2$.

9. $P_1(x) = 1 + x,\quad P_2(x) = x^2 + x^3,\quad P_3(x) = -2 - 2x + 3x^2 + 3x^3,$ en $\mathbb{P}_3$.

10. $A_1 = \begin{bmatrix} -1 & 3 & 2 \\ 0 & 0 & 0 \end{bmatrix},\quad A_2 = \begin{bmatrix} 0 & 6 & 4 \\ 4 & 6 & 2 \end{bmatrix},\quad A_3 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 1 \end{bmatrix},$ en M_{23}.

11. $\begin{bmatrix} -1 & 0 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 0 & 6 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix},$ en M_{22}.

12. Si $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto LI, demuestre que $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, $\{\mathbf{u}_1\}$, $\{\mathbf{u}_2\}$ son LI.

13. Justifique las proposiciones 1, 2 y 3 de este módulo.

14. Construya argumentaciones para justificar los teoremas 2, 3 y 4.

15. Si $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ es LI, pruebe que cualquier subconjunto no vacío de este conjunto es LI.

16. Si $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ es LD, pruebe que $\{\mathbf{u}_1, \ldots, \mathbf{u}_n, \mathbf{u}_{n+1}, \ldots, \mathbf{u}_k\}$ también es LD.

17. Sea A una matriz cuadrada $n \times n$, cuyas columnas son los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Demuestre que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son LI si y sólo si la forma escalonada reducida por renglones de A no contiene un renglón de ceros.

En los ejercicios 18 y 19 escriba la solución del sistema homogéneo dado en términos de uno o más vectores LI.

18. $$\begin{aligned} x_1 - x_2 + 7x_3 - x_4 &= 0 \\ 2x_1 + 3x_2 - 8x_3 + x_4 &= 0 \end{aligned}$$

19. $$\begin{aligned} x_1 + 2x_2 - x_3 &= 0 \\ 2x_1 + 5x_2 + 4x_3 &= 0 \end{aligned}$$

20. Demuestre que cualesquiera cuatro polinomios en $\mathbb{P}_2$ son LD.

21. Demuestre que cualesquiera $n+2$ polinomios en $\mathbb{P}_n$ son LD.

22. Demuestre que cualesquiera siete matrices en M_{32} son LD.

23. Sea $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ un conjunto LI.

    Demuestre que los vectores $\mathbf{v}_1,\ \mathbf{v}_1 + \mathbf{v}_2,\ \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3,\ ...,\ \mathbf{v}_1 + \mathbf{v}_2 + ... + \mathbf{v}_n$ son LI.

24. Suponga que $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ es un conjunto LI y que $\mathbf{v}_{k+1}$ no está en $\text{gen}\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$. Demuestre que $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k, \mathbf{v}_{k+1}\}$ es un conjunto LI.

25. ¿Para cuáles valores de α son LD los vectores $(-1, 0, -1)$, $(2, 1, 2)$ y $(1, 1, \alpha)$ en $\mathbb{R}^3$?

26. ¿Para cuáles valores de λ son LD los vectores $3 + x$ y $2 + \lambda^2 + 2x$ en $\mathbb{P}_1$?

27. Encuentre un conjunto de tres vectores LI en $\mathbb{R}^3$ que contenga los vectores $(1, 2, 4)$, $(-1, 0, 2)$.

28. Si $\mathbf{u}, \mathbf{v}$ y $\mathbf{w}$ son tres vectores coplanares en $\mathbb{R}^3$, demuestre que $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es un conjunto LD.

Módulo 5

Bases y dimensión

Base estándar $\{1, x, x^2, x^3\}$ de $\mathbb{P}_3$.

Introducción

En el proceso de construcción de espacios vectoriales es importante escoger conjuntos generadores lo más pequeños posible, esto es, que generen el subespacio donde se va a trabajar sin tomar más vectores de los necesarios; para ello se deben eliminar del conjunto generador aquellos vectores que son combinaciones lineales del resto y que, por lo tanto, no agregan nuevos vectores al subespacio.Con este procedimiento se obtiene un conjunto generador *linealmente independiente* (LI). Partiendo de lo anterior, el presente tema cubre la ejecución del procedimiento que permite identificar las *bases* de los espacios vectoriales, en donde el número de elementos de la base es la *dimensión* del espacio.

Objetivos

1. Construir una base para un espacio vectorial.
2. Diferenciar un conjunto generador de una base.
3. Identificar la dimensión del espacio vectorial, con el número de vectores en la base del espacio.
4. Relacionar los conceptos de independencia lineal, generador y dimensión para la construcción de bases en un espacio vectorial.
5. Reconocer la importancia de trabajar con bases en la generación de subespacios, ya que esto simplifica el procedimiento algebraico.

Preguntas básicas

1. ¿Qué es una base para un espacio vectorial?
2. ¿Cómo se construye una base?
3. ¿Cómo se relacionan los conceptos de base y dimensión?
4. ¿Cuándo se dice que un espacio tiene dimensión infinita?
5. ¿Un conjunto generador de $\mathbb{R}^n$ con n vectores es una base para $\mathbb{R}^n$?
6. ¿Un conjunto LI de n vectores de $\mathbb{R}^n$ es una base para $\mathbb{R}^n$?

Contenidos

5.1 Definición y ejemplos de bases
5.2 Propiedades de las bases
5.3 Dimensión. Definición, ejemplos y propiedades
5.4 El problema de la base

Vea el módulo 5 del programa de televisión *Álgebra lineal*.

5.1 Definición y ejemplos de bases

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Bases».

Definición 1

Un conjunto finito de vectores $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es una *base* para un espacio vectorial V si:

i. $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es linealmente independiente.

ii. $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ genera a V.

Ejemplos

1. Los vectores $\mathbf{e}_1 = (1,0)$ y $\mathbf{e}_2 = (0,1)$ forman una base de $\mathbb{R}^2$. Los vectores $\mathbf{e}_1 = (1,0,0)$, $\mathbf{e}_2 = (0,1,0)$, $\mathbf{e}_3 = (0,0,1)$ forman una base de $\mathbb{R}^3$ y, en general, los vectores

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, ..., \mathbf{e}_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}, ..., \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

forman una base para $\mathbb{R}^n$. Cada uno de estos conjuntos recibe el nombre de base natural, estándar o canónica para $\mathbb{R}^2$, $\mathbb{R}^3$ y $\mathbb{R}^n$, respectivamente.

Ya sabemos que todo conjunto de n vectores LI en $\mathbb{R}^n$ genera a $\mathbb{R}^n$; por lo tanto, todo conjunto de n vectores LI en $\mathbb{R}^n$ es una base de $\mathbb{R}^n$.

2. En $\mathbb{P}_n$ la base estándar está formada por el conjunto $\{1, x, x^2, ..., x^n\}$. ¿Por qué?

3. En M_{22} el conjunto de matrices $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ es una base para el espacio vectorial M_{22}. Ésta es la base estándar o canónica.

4. Sea $\Pi = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : 3x - 2y + z = 0 \right\}$ un subespacio de $\mathbb{R}^3$. Vamos a determinar una base para Π.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \Pi \Leftrightarrow z = -3x + 2y.$$

Así que los vectores de Π tienen la forma $\begin{pmatrix} x \\ y \\ -3x+2y \end{pmatrix}$,

siendo *x* y *y* valores arbitrarios.

$$\begin{pmatrix} x \\ y \\ -3x+2y \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix},$$

lo que muestra que estos dos vectores generan Π; además, son LI, ya que uno no es un múltiplo escalar del otro.

Es decir, $\left\{\begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}\right\}$ es una base para Π.

5.2 Propiedades de las bases

Teorema 1

Si $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es una base de *V* y si $\mathbf{v} \in V$, entonces existe un único conjunto de escalares $c_1, c_2, ..., c_n$ tales que $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n$.

Demostración

Suponga que **v** pudiera expresarse de dos formas como combinación lineal de los vectores de la base.

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + ... + d_n\mathbf{v}_n,$$
$$(c_1 - d_1)\mathbf{v}_1 + (c_2 - d_2)\mathbf{v}_2 + ... + (c_n - d_n)\mathbf{v}_n = 0.$$

Como $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es LI, entonces:

$$c_1 - d_1 = c_2 - d_2 = ... = c_n - d_n = 0 \quad \Rightarrow \quad c_i = d_i, \ \forall i.$$

¿Contienen todas las bases de un espacio vectorial el mismo número de vectores? El siguiente teorema nos da la respuesta a esta pregunta.

Teorema 2

Si $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m\}$ y $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ son bases en un espacio vectorial V, entonces $m = n$.

Demostración

Sean $S_1 = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m\}$ y $S_2 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ dos bases para el espacio vectorial V. Suponga que $m \neq n$, entonces $m > n$ o $n > m$.

Si $m > n$, vemos que S_1 es LD.

Cada vector de S_1 se puede expresar como combinación lineal de los vectores de la base S_2.

$$\begin{array}{lllll}
\mathbf{u}_1 = & a_{11}\mathbf{v}_1 + & a_{12}\mathbf{v}_2 + & \cdots + & a_{1n}\mathbf{v}_n \\
\mathbf{u}_2 = & a_{21}\mathbf{v}_1 + & a_{22}\mathbf{v}_2 + & \cdots + & a_{2n}\mathbf{v}_n \\
 & \vdots & \vdots & & \vdots \\
\mathbf{u}_m = & a_{m1}\mathbf{v}_1 & a_{m2}\mathbf{v}_2 & \cdots + & a_{mn}\mathbf{v}_n .
\end{array}$$

Para demostrar que S_1 es LD veamos que en la combinación lineal

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_m\mathbf{u}_m = \mathbf{0} \tag{1}$$

existe $c_i \neq 0$.

$$c_1(a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + ... + a_{1n}\mathbf{v}_n) + c_2(a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + ... + a_{2n}\mathbf{v}_n) + ...$$
$$+ c_m(a_{m1}\mathbf{v}_1 + a_{m2}\mathbf{v}_2 + ... + a_{mn}\mathbf{v}_n) = \mathbf{0}.$$

Reagrupamos los términos de la ecuación así:

$$(c_1a_{11} + c_2a_{21} + ... + c_ma_{m1})\mathbf{v}_1 + (c_1a_{12} + c_2a_{22} + ... + c_ma_{m2})\mathbf{v}_2 + ...$$
$$+ (c_1a_{1n} + c_2a_{2n} + ... + c_ma_{mn})\mathbf{v}_n = \mathbf{0}.$$

Como $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es LI, entonces:

$$\begin{array}{lllll}
a_{11}c_1 + & a_{21}c_2 + & \cdots + & a_{m1}c_m & = 0 \\
a_{12}c_1 + & a_{22}c_2 + & \cdots + & a_{m2}c_m & = 0 \\
\vdots & \vdots & & \vdots & \\
a_{1n}c_1 + & a_{2n}c_2 + & \cdots + & a_{m\,n}c_m & = 0,
\end{array}$$

el cual es un sistema homogéneo de n ecuaciones con m incógnitas, con $m > n$; luego, tiene infinitas soluciones ; por lo tanto, existen escalares $c_1, c_2, ..., c_m$, no todos cero, que verifican (1), y en consecuencia S_1 es LD.

De forma análoga se razona para el caso en que $n > m$.

5.3 Dimensión. Definición, ejemplos y propiedades

Definición 2

La *dimensión* de un espacio vectorial no nulo es el número de vectores en una base para *V*. Como el conjunto $\{0\}$ es linealmente dependiente, se dice que el espacio vectorial $\{0\}$ tiene dimensión cero.

Cuando la base del espacio vectorial *V* es un conjunto finito de vectores, se dice que *V* es de *dimensión finita*. Sin embargo, existen muchos espacios vectoriales para los cuales no hay una base con un número finito de vectores; estos espacios son de *dimensión infinita*, por ejemplo el espacio vectorial $\mathbb{P}$ de todos los polinomios y el espacio $C[a,b]$ de funciones continuas en $[a,b]$.

La dimensión de *V* se denota por dim *V*.

Ejemplos

5. $\dim \mathbb{R}^n = n.$
6. $\dim \mathbb{P}_n = n+1.$
7. $\dim\ M_{mn} = mn.$

Teorema 3

Si dim $V = n$ y $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m\}$ es un conjunto LI de *m* vectores de *V*, entonces $m \leq n.$

La dimensión de un espacio vectorial es el máximo número de vectores LI que hay en el espacio vectorial.

Teorema 4

Sea *H* un subespacio del espacio vectorial *V* de dimensión finita; entonces, $\dim\ H \leq \dim V$.

Demostración

Sea $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ una base para *V*, es decir, dim $V = n$. Cualquier conjunto LI de vectores de *V* tiene a lo más *n* vectores. Como $H \subset V$, todo conjunto LI de *H* también es LI en *V*. Luego *H* es de dimensión finita y $\dim H \leq n.$

Ejemplos

8. Sea $\mathbb{P}[0,1]$ el conjunto de polinomios definidos en el intervalo $[0,1]$. Como todo polinomio es una función continua, entonces $\mathbb{P}[0, 1] \subset C[0, 1]$ (funciones continuas en el intervalo $[0, 1]$). Si $C[0, 1]$ fuese de dimensión finita, entonces $\mathbb{P}[0, 1]$ tendría también dimensión finita, pero no existe un conjunto finito de polinomios que genere $\mathbb{P}[0, 1]$, luego $C[0, 1]$ es de dimensión infinita.

   Todo espacio vectorial que contenga un subespacio de dimensión infinita es de dimensión infinita.

9. Sea A una matriz $m \times n$ y sea $S = \{\mathbf{x} \in R^n : A\mathbf{x} = \mathbf{0}\}$.

   Suponga que $\mathbf{x}_1 \in S$ y $\mathbf{x}_2 \in S$, entonces $A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$ y $A(\alpha\mathbf{x}_1) = \alpha(A\mathbf{x}_1) = \alpha\mathbf{0} = \mathbf{0}$. Luego, S es un subespacio de $\mathbb{R}^n$ y $\dim S \leq n$. A S se le llama espacio solución del sistema homogéneo $A\mathbf{x} = \mathbf{0}$ o *núcleo* de la matriz A.

10. Determinemos una base para el espacio solución del sistema homogéneo

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

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

Luego:

$$x = -2z,$$
$$y = -3z.$$
$$z = 1z.$$

Todo vector del espacio solución es múltiplo escalar del vector $\begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix}$.

Por lo tanto:

$$B = \left\{ \begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix} \right\}.$$

Un conjunto de vectores en un espacio vectorial V es una base para V si genera a V y es LI. Ahora, si se sabe que la dimensión del espacio es n, basta verificar una de las dos condiciones.

Teorema 5

Sea V un espacio vectorial de dimensión n y sea $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ un conjunto de n vectores en V.

a. Si S es LI, entonces es una base para V.

b. Si S genera a V, entonces es una base para V.

Demostración

a. Como S es LI, debemos demostrar que S genera a V.

Si S no genera a V, existe un $\mathbf{v}_{n+1} \in V$ tal que $\mathbf{v}_{n+1} \notin \text{gen}\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$, luego $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n, \mathbf{v}_{n+1}\}$ es subconjunto de V, LI con $n + 1$ vectores; por lo tanto, dim $V > n$, lo cual es absurdo.

Para determinar si un subconjunto S de $\mathbb{R}^n$ es una base para $\mathbb{R}^n$, primero contamos el número de elementos de S. Si S tiene n elementos, cualquiera de las partes del teorema 5 nos permite determinar si S es base. Si S no tiene n elementos, no es base para $\mathbb{R}^n$. De la misma forma se procede en cualquier espacio o subespacio vectorial cuya dimensión sea conocida.

Teorema 6

Si S es un conjunto LI de vectores en un espacio vectorial V de dimensión finita, entonces existe una base B para V, que contiene a S.

El teorema nos dice que un conjunto LI de vectores en un espacio vectorial V se puede extender a una base para V.

5.4 El problema de la base

Este problema puede presentarse en una de estas dos formas:

1. Construir una base para V, tomando los vectores de V. Si puede escoger un conjunto de vectores que genere a V, puede eliminar los vectores LD, si los hay, y obtener una base para V.

2. Dado un conjunto S de vectores de V, construir una base para V añadiendo, o bien eliminando, algunos vectores de S, o ambas cosas.

   Si S genera a V, se procede como en el caso anterior. Si no, se aumenta S hasta obtener un generador y luego se procede como en el caso anterior.

Ejemplo

11. Sea $S = \{(1, -1, 1),\ (0, 1, -1)\}$. Obtenga una base de $\mathbb{R}^3$ que contenga a S.

Solución

Como $\mathbb{R}^3$ tiene dimensión tres, la base para $\mathbb{R}^3$ debe contener tres vectores.

El conjunto *S* es LI, ya que los vectores de *S* no son uno un múltiplo escalar del otro; luego, lo que hace falta es añadir otro vector **v** tal que $\mathbf{v} \notin \text{gen}\, S$.

$$\begin{aligned}\text{gen}\, S &= \{\mathbf{x} / \mathbf{x} = a(1, -1, 1) + b(0, 1, -1)\} \\ &= \{\mathbf{x} / \mathbf{x} = (a, -a+b, a-b)\}.\end{aligned}$$

Si $\mathbf{v} = (x_1, x_2, x_3)$ perteneciera a $\text{gen}\, S$, entonces:

$$\begin{aligned}x_1 &= a, \\ x_2 &= -a+b, \\ x_3 &= a-b;\end{aligned}$$

por lo tanto:

$$\left[\begin{array}{rr|c} 1 & 0 & x_1 \\ -1 & 1 & x_2 \\ 1 & -1 & x_3 \end{array}\right] \xrightarrow[\text{elementales}]{\text{operaciones}} \left[\begin{array}{rr|c} 1 & 0 & x_1 \\ 0 & 1 & x_1 + x_2 \\ 0 & -1 & x_3 - x_1 \end{array}\right] \rightarrow \left[\begin{array}{rr|c} 1 & 0 & x_1 \\ 0 & 1 & x_1 + x_2 \\ 0 & 0 & \underbrace{x_3 - x_1 + x_1 + x_2}_{x_3 + x_2} \end{array}\right].$$

Así que si $x_3 + x_2 = 0$, $\mathbf{v} \in \text{gen}\, S$, entonces tomamos **v** de tal manera que $x_3 + x_2 \neq 0$; por ejemplo, $\mathbf{v} = (0, 1, 0)$.

Una base para $\mathbb{R}^3$ será $B = \{(1, -1, 1), (0, 1, -1), (0, 1, 0)\}$.

Otra forma de solución

Si $\mathbf{v} = (x_1, x_2, x_3)$ es tal que $\{(1, -1, 1), (0, 1, -1), (x_1, x_2, x_3)\}$ fuera LD, entonces:

$$\begin{vmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ x_1 & x_2 & x_3 \end{vmatrix} = 0$$

$$1(x_3 + x_2) + x_1(1-1) = 0 \;\Rightarrow\; x_3 + x_2 = 0.$$

Luego, para que el conjunto sea LI, lo que se requiere es que $x_3 + x_2 \neq 0$.

Otra manera podría ser probar los vectores de la base canónica hasta que uno de ellos funcione.

Ejercicios del capítulo 1 (módulo 5)

En los ejercicios 1 a 4 explique mediante inspección por qué los vectores dados no forman una base del espacio vectorial dado.

1. $\mathbf{u}_1 = (3, 2), \quad \mathbf{u}_2 = (-1, 2), \quad \mathbf{u}_3 = (2, 4),$ para $\mathbb{R}^2$.

2. $\mathbf{w}_1 = (8, 3, 5, 4), \quad \mathbf{w}_2 = (0, 5, 7, 1), \quad \mathbf{w}_3 = (-1, 3, 2, 0),$ para $\mathbb{R}^4$.

3. $P_1 = 1 + 2x + 3x^2, \quad P_2 = 2x + x^2, \quad P_3 = 5 - 3x, \quad P_4 = 3 - 5x + x^2$, para $\mathbb{P}_2$.

4. $M_1 = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}, \ M_2 = \begin{bmatrix} 5 & 7 \\ -1 & 0 \end{bmatrix}, \ M_3 = \begin{bmatrix} 9 & 7 \\ 2 & 5 \end{bmatrix},$ en M_{22}.

En los ejercicios 5 a 11 determine cuándo los vectores dados forman una base del espacio vectorial dado.

5. $\mathbf{u}_1 = (3, 5), \ \mathbf{u}_2 = (4, 8),$ en $\mathbb{R}^2$.

6. $\mathbf{u}_1 = (1, 1), \ \mathbf{u}_2 = (-2, -2),$ en $\mathbb{R}^2$.

7. $\mathbf{u}_1 = (1, 1, 1), \ \mathbf{u}_2 = (1, 1, 0), \ \mathbf{u}_3 = (1, 0, 0),$ en $\mathbb{R}^3$.

8. $\mathbf{u}_1 = (2, 1, 2), \ \mathbf{u}_2 = (1, -2, -3), \ \mathbf{u}_3 = (5, 0, 1),$ en $\mathbb{R}^3$.

9. $P_1 = 1 + x, \quad P_2 = 3 - x,$ para $\mathbb{P}_1$.

10. $q_1 = 1 + x + x^2, \quad q_2 = x + 2x^2, \quad q_3 = 3x^2,$ para $\mathbb{P}_2$.

11. $M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad M_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$ en M_{22}.

En los ejercicios 12 a 17 encuentre una base del espacio vectorial dado y determine su dimensión.

12. Todos los vectores en $\mathbb{R}^2$ cuyas componentes suman cero.

13. Todas las matrices simétricas de 3×3.

14. Todas las matrices antisimétricas de 3×3.

15. Todos los vectores de $\mathbb{R}^3$ que están en el plano $2x - y - z = 0$.

16. Todos los vectores de $\mathbb{R}^3$ que están en la recta $x = 3t, \quad y = -2t, \quad z = t$.

17. Todos los polinomios de $\mathbb{P}_2$ de la forma $a_0 + a_1 x + a_2 x^2$, con $a_0 = a_2 - a_1$.

18. Determine una base para $\mathbb{R}^3$ que incluya los vectores

a. $(1,\ 1,\ 2)$

b. $(1, 1, 2)$, $(3,\ 0,\ 1)$.

19. Determine una base para $\mathbb{R}^4$ que incluya a los vectores $(1,\ 0, 1,\ 0)$ y $(0, 1, -1,\ 0)$.

20. Determine todos los valores de a para los cuales $\{(a^2,\ 0, 1), (0,\ a,\ 2), (1,\ 0,\ 1)\}$ es una base para $\mathbb{R}^3$.

21. Proporcione un ejemplo de un subespacio de dimensión dos de $\mathbb{R}^4$.

En los ejercicios 22 a 25 encuentre una base para el espacio solución del sistema homogéneo dado.

22.
$$\begin{aligned} x + y - z &= 0 \\ 2x - y + 2z &= 0 \end{aligned}$$

23.
$$\begin{aligned} x - y &= 0 \\ -3x + 3y &= 0 \end{aligned}$$

24.
$$\begin{aligned} x - 3y + 2z &= 0 \\ -2x + y + 3z &= 0 \\ 3x - 4y + 5z &= 0 \end{aligned}$$

25.
$$\begin{aligned} -x + 3y - 2z &= 0 \\ -3x + 9y - 6z &= 0 \\ 2x - 6y + 4z &= 0 \end{aligned}$$

26. Elabore una argumentación que justifique el teorema 3.

27. Elabore una argumentación que justifique el teorema 5*b*.

28. Elabore una argumentación que justifique el teorema 6.

Módulo 6

Subespacios asociados con una matriz

Subespacios asociados con la matriz

$$A=\begin{bmatrix} 3 & 0 & -1 \\ 3 & 0 & -1 \\ 4 & 0 & 5 \end{bmatrix}, \quad \begin{array}{l} N_A: \text{eje } x_2 \\ R_A: \text{plano } x_1x_3 \\ C_A: \text{plano de ecuación} \\ \qquad x_1 - x_2 = 0 \end{array}$$

Introducción

En una matriz se pueden definir algunos subespacios, como núcleo, espacio renglón y espacio columna. En este tema presentaremos estos subespacios y estableceremos relaciones entre ellos; además introduciremos los conceptos de rango y nulidad asociados con la solución de sistemas de ecuaciones lineales.

Partiendo de lo anterior, avanzaremos en el procedimiento para construir las bases. Cuando enfrentamos este problema requerimos hacer una descripción clara del subespacio correspondiente, lo cual podemos hacer de dos maneras: la primera consiste en dar un conjunto de vectores que generen el espacio; es así como describimos los espacios fila y columna. En la segunda se dan restricciones para determinar el subespacio, esto es, nos dicen las propiedades que deben satisfacer; de esta forma se define el espacio nulo. En ninguno de los dos casos es posible dar una base por simple inspección; es necesario desarrollar un procedimiento sistemático.

Objetivos

1. Determinar los subespacios asociados con una matriz.
2. Establecer las diferentes relaciones que existen entre estos subespacios.
3. Utilizar estos subespacios para determinar bases de subespacios generados por un conjunto de vectores.
4. Utilizar el rango de una matriz para determinar cuándo un sistema de ecuaciones lineales no homogéneo tiene solución.

Preguntas básicas

1. ¿Cuáles son los subespacios asociados con A?

En las siguientes preguntas, la matriz B es la matriz escalonada equivalente por renglones a A.

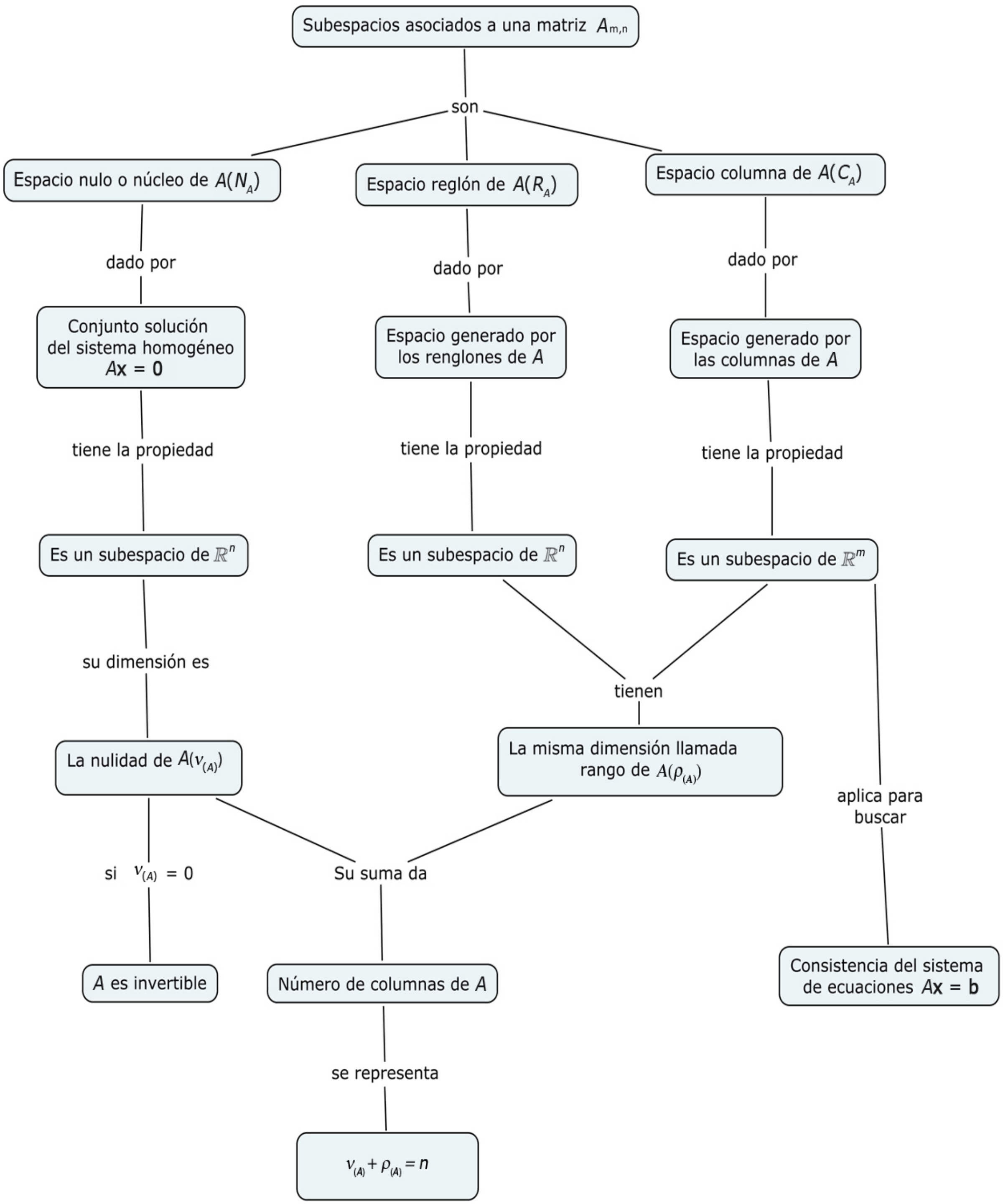
2. ¿Cómo son los espacios nulo y renglón de las matrices A y B?
3. ¿Qué relación existe entre los espacios renglón y columna de A?
4. ¿Cómo se relacionan los espacios nulo y renglón de A?
5. ¿Qué es el rango de una matriz?
6. ¿Cómo se relacionan los rangos de A y B?

Vea el módulo 6 del programa de televisión *Álgebra lineal*.

7. ¿Qué condición debe cumplir el término b en la ecuación matricial $Ax = b$ para que el sistema tenga solución?
8. ¿Cómo se relacionan el rango y la nulidad de A con el número de columnas de A?

Contenidos

6.1 Espacios nulo, renglón y columna de una matriz A
6.2 Propiedades
6.3 Relaciones con el espacio nulo o núcleo de una matriz A
6.4 Relaciones con el rango de una matriz A



Mapa 3: módulo 6

6.1 Espacios nulo, renglón y columna de una matriz *A*

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Base para el espacio nulo de *A*».

Definición 1

Sea *A* una matriz $m \times n$. El conjunto

$$N_A = \left\{ \mathbf{x} \in \mathbb{R}^n / A\mathbf{x} = \mathbf{0} \right\}$$

se llama *espacio nulo de A* o *núcleo de A* y se denota por N_A. En el ejemplo 9 del módulo 5 mostramos que este conjunto es un subespacio de $\mathbb{R}^n$.

A la dimensión de N_A se le llama *nulidad de A* y se denota $\nu_{(A)}$.

Ejemplo 1

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

Determine el espacio nulo de *A*.

Solución

La matriz *A* es la matriz de coeficientes del sistema homogéneo $A\mathbf{x} = \mathbf{0}$. Hacemos operaciones elementales de renglón para hallar las soluciones de este sistema.

$$\begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 0 \end{bmatrix} R_2 \to R_2 - 3R_1 \begin{bmatrix} 1 & -1 & 2 \\ 0 & 4 & -6 \end{bmatrix} R_2 \to \frac{1}{4} R_2 \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -3/2 \end{bmatrix}$$

$$R_1 \to R_1 + R_2 \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -3/2 \end{bmatrix} \qquad \begin{aligned} x_1 &= -\frac{1}{2} x_3 \\ x_2 &= \frac{3}{2} x_3 \\ x_3 &= x_3 \end{aligned}$$

$$N_A = \text{gen}\left\{ \begin{pmatrix} -1/2 \\ 3/2 \\ 1 \end{pmatrix} \right\}, \ \upsilon_{(A)} = 1.$$

Definición 2

$$\text{Sea } \underset{m \times n}{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

i. El espacio generado por los renglones de A,

$$\mathbf{r}_1 = (a_{11}, a_{12}, ..., a_{1n}),\ \ \mathbf{r}_2 = (a_{21}, a_{22}, ..., a_{2n})\ , ...,\ \mathbf{r}_m = (a_{m1}, a_{m2}, ..., a_{mn}),$$

es un subespacio de $\mathbb{R}^n$ llamado *espacio renglón de A* y se denota R_A.

ii. El espacio generado por las columnas de A

$$\mathbf{c}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix},\ \mathbf{c}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, ...,\ \mathbf{c}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

es un subespacio de $\mathbb{R}^m$ llamado *espacio columna de A* y se denota C_A.

6.2 Propiedades

Teorema 1

Si A y B son matrices $m \times n$ equivalentes por renglones, entonces los espacios generados por los renglones de A y de B son iguales.

Demostración

Si A y B son equivalentes por renglones, entonces B se obtiene de A por operaciones elementales de renglón. Es claro que la operación de intercambio de renglones no afecta el espacio renglón. Las otras operaciones CR_i y $R_j + CR_i$ son combinaciones lineales entre los renglones de A, las cuales pertenecen al espacio renglón A; por lo tanto, el espacio generado por los renglones de B está contenido en el espacio generado por los renglones de A, $R_B \subset R_A$.

De forma análoga se ve que $R_A \subset R_B$; en consecuencia, $R_A = R_B$.

Este resultado nos permite determinar una base para un espacio vectorial generado por un conjunto de vectores dado de $\mathbb{R}^n$.

Ejemplo 2

$$\text{Sea } S = \left\{ \begin{pmatrix} 2 \\ -3 \\ 4 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 4 \\ -6 \\ 5 \\ 3 \\ -4 \end{pmatrix}, \begin{pmatrix} -10 \\ 15 \\ -14 \\ 3 \\ 7 \end{pmatrix}, \begin{pmatrix} -8 \\ 12 \\ -10 \\ 4 \\ 9 \end{pmatrix} \right\}.$$

Determine una base para $\text{gen } S$.

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Espacio de renglones y columnas».

Solución

Tomamos los vectores de S como renglones de una matriz A y por medio de operaciones elementales obtenemos una matriz equivalente B en forma escalonada. Los renglones no nulos de B serán LI y formarán una base para R_B, que es el mismo R_A. Estos vectores tomados como columnas serán una base para $\text{gen } S$.

$$A=\begin{bmatrix} 2 & -3 & 4 & 1 & 2 \\ 4 & -6 & 5 & 3 & -4 \\ -10 & 15 & -14 & 3 & 7 \\ -8 & 12 & 10 & 4 & 9 \end{bmatrix} \xrightarrow{\text{operaciones elementales}} B=\begin{bmatrix} 1 & -3/2 & 2 & 1/2 & 1 \\ 0 & 0 & 1 & -1/3 & 8/3 \\ 0 & 0 & 0 & 1 & 1/10 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

$$\text{Base para gen } S=\left\{\begin{pmatrix} 1 \\ -3/2 \\ 2 \\ 1/2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1/3 \\ 8/3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1/10 \end{pmatrix}\right\}.$$

Observación

Si B es una matriz escalonada, entonces $\dim R_B$ es igual al número de pivotes de B.

Teorema 2

Sea B una matriz $m \times n$ en la forma escalonada por renglones. Las columnas de B que contienen los pivotes forman una base para C_B ; por lo tanto, $\dim C_B$ es igual al número de pivotes de B.

Demostración

Las columnas de B que contienen pivotes son LI porque sus entradas distintas de cero están en forma escalonada (en el ejemplo 2 las columnas 1, 3 y 4 son LI) .

Veamos que estas columnas generan C_B. Sea ρ el número de pivotes de B. Los primeros ρ renglones de B son diferentes de cero y los últimos $m-\rho$ son todos cero.

Sea $\mathbf{b} \in C_B$, $\boldsymbol{b}$ es combinación lineal de las columnas de B, luego los últimos $m-\rho$ componentes de $\boldsymbol{b}$ son todos cero.

Formemos una matriz $T_{m\times\rho}$ (ρ : # de renglones de B diferente de cero) tomando las columnas de B que contienen los pivotes y las colocamos ordenadamente:

$$T_{m\times\rho} = \begin{bmatrix} * & \cdots & \cdots & \\ 0 & * & \cdots & \\ & \vdots & & \\ 0 & 0 & \cdots & * \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} \rho \\ \\ \\ \end{array}$$

Las entradas señaladas con * son los pivotes de B; por lo tanto, diferentes de cero. Por la forma especial de T y **b**, la ecuación $\underset{\rho\times 1}{T\,\mathbf{x}} = \underset{m\times 1}{\mathbf{b}}$ puede resolverse por sustitución regresiva. Luego, **b** es combinación lineal de las columnas de T; por lo tanto, de las columnas de B que contienen pivotes.

Teorema 3

Sea $\underset{m\times n}{A}$ y $\underset{m\times n}{B}$ la matriz que resulta al reducir A a su forma escalonada. Los vectores columna de B que contienen los pivotes forman una base para C_B, entonces los vectores columna correspondientes de A formarán una base para C_A y, en consecuencia,

$$\dim C_A = \dim C_B .$$

Ahora podemos reunir los siguientes resultados:

$R_A = R_B$.
ρ : número de pivotes de B.
$\dim R_B = \dim R_A = \rho$.
$\dim C_B = \rho$.
$\dim C_B = \dim C_A$.

Luego $\dim R_A = \dim C_A = \rho$.

Definición 3

El *rango de A*, denotado por $\rho_{(A)}$, es la dimensión de los espacios fila y columna de A.

Consulte el apéndice «Una aplicación interesante de los espacios vectoriales» al final de este texto.

6.3 Relaciones con el espacio nulo o núcleo de una matriz *A*

Sea $\underset{m\times n}{A}$ y $\underset{m\times n}{B}$ la matriz que resulta al reducir *A* a su forma escalonada.
Sabemos que:

$N_A = N_B$.

$R_A = R_B$.

La dimensión de N_A es igual al número de parámetros en la solución del sistema equivalente reducido $B\mathbf{x} = \mathbf{0}$.

La dimensión del espacio renglón de *B* es igual al número de variables principales del sistema $B\mathbf{x} = \mathbf{0}$.

El número de variables principales más el número de parámetros es *n*. De todos estos hechos se deduce el siguiente teorema.

Teorema 4

Si *A* es una matriz $m\times n$, entonces:

$$\left.\begin{aligned}\dim R_A + \dim N_A = n\\ \dim C_A + \dim N_A = n\end{aligned}\right\rangle \text{ o } \rho_{(A)} + \nu_{(A)} = n.$$

Teorema 5

Sea $\underset{m\times n}{A}$. Cada vector del espacio renglón de *A*, R_A es ortogonal a todo vector del espacio nulo de *A*; o sea $N_A.(R_A \perp N_A)$.

Demostración

Consideremos la ecuación $A\mathbf{x} = \mathbf{0}$.

$$\begin{array}{c} \mathbf{r}_1 \rightarrow \\ \mathbf{r}_2 \rightarrow \\ \\ \mathbf{r}_m \rightarrow \end{array}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Si $\mathbf{x} \in N_A$, $A\mathbf{x} = \mathbf{0}$; por lo tanto, si $\mathbf{r}_i$ es el *i*-ésimo renglón de *A*, $\mathbf{r}_i \cdot \mathbf{x} = 0$, para $i = 1,\ldots,m$.

Ahora, si $\mathbf{y} \in R_A$, entonces

$\mathbf{y} = C_1\mathbf{r}_1 + C_2\mathbf{r}_2 + \ldots + C_m\mathbf{r}_m$ y

$$\mathbf{y} \cdot \mathbf{x} = C_1 \mathbf{r}_1 \cdot \mathbf{x} + C_2 \mathbf{r}_2 \cdot \mathbf{x} + \ldots + C_m \mathbf{r}_m \cdot \mathbf{x} = 0,$$

luego, todo vector del espacio renglón es ortogonal a todo vector del espacio nulo.

Escuche la biografía de *Richard Hamming* en su multimedia de *Álgebra lineal.*

Ejemplo 3

Sea $A = \begin{bmatrix} 0 & 2 & -3 & 1 & 2 \\ 0 & -2 & 3 & 3 & 1 \\ 0 & 4 & -6 & 6 & 7 \end{bmatrix}.$

Determinemos $N_A,\ C_A,\ \rho_{(A)},\ \nu_{(A)}$.

Solución

$$\begin{bmatrix} 0 & 2 & -3 & 1 & 2 \\ 0 & -2 & 3 & 3 & 1 \\ 0 & 4 & -6 & 6 & 7 \end{bmatrix} \xrightarrow{\text{operaciones elementales}} \begin{bmatrix} 0 & 1 & -3/2 & 1/2 & 1 \\ 0 & 0 & 0 & 1 & 3/4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Solución del sistema homogéneo:

Variables principales x_2 y x_4.

$$x_4 = -\frac{3}{4} x_5\,,$$

$$x_2 = \frac{3}{2} x_3 - \frac{1}{2} x_4 - x_5\,.$$

Aplicando sustitución regresiva:

$$x_2 = \frac{3}{2} x_3 - \frac{1}{2}\left(-\frac{3}{4} x_5\right) - x_5 = \frac{3}{2} x_3 - \frac{5}{8} x_5\,.$$

Parámetros: x_1, x_3 y x_5.

Entonces:

$$\begin{aligned} x_1 &= x_1, \\ x_2 &= \frac{3}{2} x_3 - \frac{5}{8} x_5, \\ x_3 &= x_3, \\ x_4 &= -\frac{3}{4} x_5, \\ x_5 &= x_5. \end{aligned}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \in N_A \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 3/2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} 0 \\ -5/8 \\ 0 \\ -3/4 \\ 1 \end{pmatrix}.$$

$$N_A = \text{gen}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3/2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -5/8 \\ 0 \\ -3/4 \\ 1 \end{pmatrix} \right\}, \ \nu_{(A)} = 3.$$

Observamos que la nulidad de A es igual al número de parámetros en la solución del sistema homogéneo $A\mathbf{x} = \mathbf{0}$.

$$R_A = \text{gen}\left\{ \left(0,\ 1,\ -\frac{3}{2},\ \frac{1}{2},\ 1\right),\ \left(0,\ 0,\ 0,\ 1,\ \frac{3}{4}\right) \right\}.$$

Los pivotes se encuentran en la 2.ª y 4.ª columnas de la matriz escalonada equivalente a A, luego en A las columnas 2.ª y 4.ª son LI y constituyen una base para C_A.

$$C_A = \text{gen}\left\{ \begin{pmatrix} 2 \\ -2 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 6 \end{pmatrix} \right\}, \quad \rho_{(A)} = 2.$$

El rango de A es el número de variables principales en la solución del sistema homogéneo $A\mathbf{x} = \mathbf{0}$.

Podemos comprobar que cada vector de R_A es ortogonal a cualquier vector de N_A.

6.4 Relaciones con el rango de una matriz *A*

Teorema 6

Sea $\underset{n\times n}{A}$. A es invertible si y sólo si

a. $\nu_{(A)} = 0.$

b. $\rho_{(A)} = n.$

Teorema 7

Sea $\underset{m\times n}{A}$. El sistema $A\mathbf{x}=\mathbf{b}$ tiene solución si y sólo si $\mathbf{b}\in C_A$.

Demostración

$A\mathbf{x}=\mathbf{b}$ tiene solución si y sólo si

$$[\mathbf{c}_1,\mathbf{c}_2,...,\mathbf{c}_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=\mathbf{b} \text{ tiene solución.}$$

Esto es: $\mathbf{c}_1x_1+\mathbf{c}_2x_2+...+\mathbf{c}_nx_n=\mathbf{b}$ ($\mathbf{c}_i$: columna i de A) tiene solución.

De modo que **b** es combinación lineal de las columnas de A; es decir, que $\mathbf{b}\in C_A$.

Nota: el teorema anterior establece que para que $A\mathbf{x}=\mathbf{b}$ tenga solución es necesario y suficiente que:

$$\rho_{(A)}=\rho_{[A\vdots b]}$$
$$\rho_{(A)}=\rho_{[A\vdots b]} \Leftrightarrow \mathbf{b}\in C_A$$
$$\rho_{(A)}\neq\rho_{[A\vdots b]} \Leftrightarrow \mathbf{b}\notin C_A \text{ y el sistema será inconsistente.}$$

Ejemplo 4

Veamos si el sistema

$$\begin{aligned} x+y-z&=7\\ 6x+y+3z&=20\\ 4x-y+5z&=4 \end{aligned}$$

tiene solución.

Solución

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 7\\ 6 & 1 & 3 & 20\\ 4 & -1 & 5 & 4 \end{array}\right]\to\left[\begin{array}{ccc|c} 1 & 1 & -1 & 7\\ 0 & -5 & 9 & -22\\ 0 & -5 & 9 & -24 \end{array}\right]\to\left[\begin{array}{ccc|c} 1 & 1 & -1 & 7\\ 0 & -5 & 9 & -22\\ 0 & 0 & 0 & -2 \end{array}\right]$$

$$\to\left[\begin{array}{ccc|c} 1 & 1 & -1 & 7\\ 0 & 1 & -9/5 & 22/5\\ 0 & 0 & 0 & 1 \end{array}\right] \begin{array}{l}\rho_{(A)}=2\\ \rho_{(A\vdots b)}=3\end{array},\ \mathbf{b}\notin C_A$$

El sistema no tiene solución.

Ejercicios del capítulo 1 (módulo 6)

En los ejercicios 1 a 3 se da una matriz en forma escalonada. Encuentre una base para su espacio renglón, una base para su espacio columna y determine su rango.

1. $\begin{bmatrix} 1 & 5 & 3 & 0 & 1 \\ 0 & 0 & 1 & 7 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$

2. $\begin{bmatrix} 1 & 4 & 5 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$

3. $\begin{bmatrix} 1 & 4 & 0 & 5 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$

En los ejercicios 4 a 9 encuentre bases para los espacios nulo, renglón y columna de la matriz dada. Determine, además, el rango y la nulidad, y verifique en cada caso que $\rho_{(A)} + \nu_{(A)} = n.$

4. $\begin{bmatrix} 3 & 2 & 5 \\ 4 & 3 & 2 \\ -1 & 7 & 8 \\ 0 & 6 & 1 \end{bmatrix}.$

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

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

7. $\begin{bmatrix} 0 & 2 & -3 & 1 & 2 \\ 0 & -2 & 3 & 3 & 1 \\ 0 & 4 & -6 & 6 & 7 \end{bmatrix}.$

8. $\begin{bmatrix} 1 & 4 & -1 & 0 \\ 2 & -2 & 3 & 1 \\ 1 & 6 & 0 & 2 \\ 2 & 8 & 9 & 0 \end{bmatrix}.$

9. $\begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & 0 & 2 & 3 \\ 4 & 0 & -2 & 1 \\ 3 & -1 & 0 & 4 \end{bmatrix}.$

En los ejercicios 10 a 12 encuentre una base para el espacio generado por los conjuntos de vectores dados.

10. $(2, 3, 5),\ (-1, 0, -2),\ (5, 6, 12),\ (2, 1, 0).$

11. $\begin{bmatrix} 2 \\ 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 3 \\ 5 \\ 3 \end{bmatrix}.$

12. $(-2, 6, 4, 1),\ (1, 0, -3, 5),\ (0, 4, -2, 1).$

En los ejercicios 13 a 15 determine si el sistema dado tiene solución.

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

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

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

16. Sea A una matriz diagonal. Demuestre que $\rho_{(A)}$ es el número de componentes diferentes de cero en la diagonal.

17. Demuestre que para cualquier matriz A, $\rho_{(A)} = \rho_{(A')}$.

18. Sea A una matriz triangular $n\times n$ con ceros en la diagonal. Demuestre que $\rho_{(A)} < n$.

19. Sea A una matriz $n\times n$. Demuestre que $\rho_{(A)} < n$ si y sólo si existe un vector $\mathbf{x}\in\mathbb{R}^n$ tal que $\mathbf{x}\neq\mathbf{0}$ y $A\mathbf{x}=\mathbf{0}$.

20. ¿Los siguientes enunciados son verdaderos o falsos? Justifique su respuesta.

    a. Si A es una matriz de $m\times n$, entonces $R = C$
    b. Si A es una matriz 5×3, entonces las columnas de A deben ser LI.
    c. Si A es una matriz 3×5, las columnas de A no pueden ser LI.
    d. Si A es una matriz de $m\times n$ y las columnas de A son LI, entonces $A\mathbf{x}=\mathbf{b}$ puede o no tener solución. Pero si tiene solución, ésta es única.

21. Sea A una matriz $m\times n$.

    a. Si las columnas de A son LI, ¿cuál es el rango de A y cuál es la relación entre m y n?
    b. Si las columnas de A generan $\mathbb{R}^m$, ¿cuál es el rango de A y cuál es la relación entre m y n?
    c. Si las columnas de A forman una base de $\mathbb{R}^m$, ¿cuál es la relación entre m y n?

Módulo 7

Coordenadas y cambio de base

$$B_1 = \{\mathbf{v}_1, \mathbf{v}_2\},\ [\mathbf{v}]_{B_1} = \begin{bmatrix} 5.5 \\ 4.5 \end{bmatrix}$$

$$B_2 = \{\mathbf{v}_1', \mathbf{v}_2'\},\ [\mathbf{v}]_{B_2} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Coordenadas respecto a diferentes bases.

Introducción

En muchas aplicaciones, un problema descrito en un sistema coordenado puede ser más fácilmente resuelto si se cambia a otro sistema, lo cual se realiza por medio de un cambio de variables. En álgebra lineal, una base proporciona un sistema coordenado para el espacio vectorial, introduciendo el concepto de vector coordenado. En cada problema debe seleccionarse la base correcta, que será aquella que simplifique la expresión de los vectores. Así, por ejemplo, en cristalografía la medición de las longitudes en los enlaces entre los átomos, los ángulos entre ellos, etc., se facilitará si se introducen unas coordenadas que se ajusten a las direcciones de los enlaces y se emplean las herramientas del álgebra lineal. La base estándar y los ejes x, y, z asociados con ella no siempre son la mejor elección.

En este tema se muestra la forma de convertir la expresión de un vector dado en una base a otra, partiendo del estudio de los vectores coordenados hasta llegar a la construcción del algoritmo que permite efectuar el cambio de base.

Objetivos

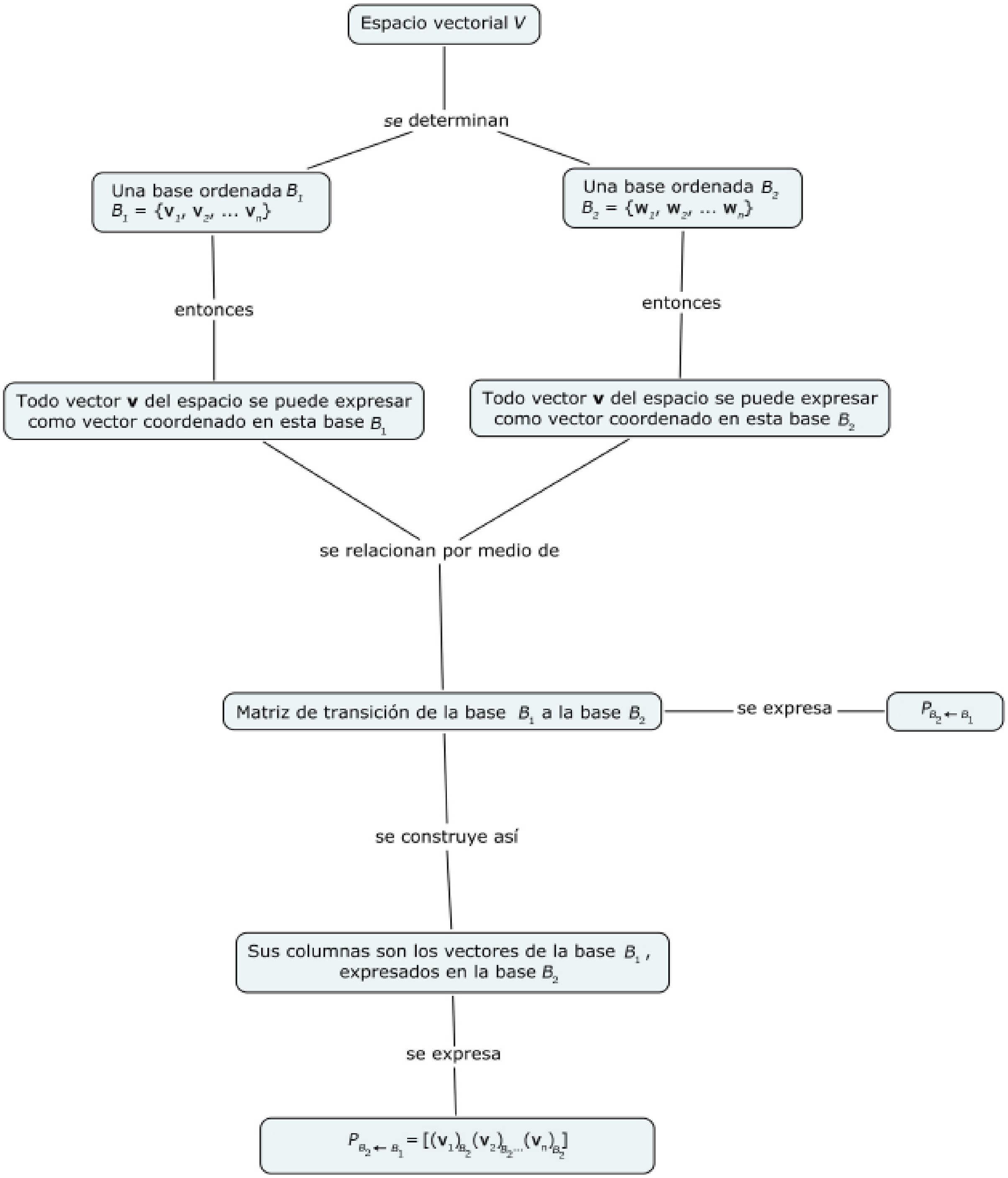
1. Introducir bases ordenadas para los espacios vectoriales.
2. Expresar los vectores como vectores de coordenadas.
3. Establecer el manejo algebraico de los vectores de coordenadas.
4. Determinar la matriz de transición de una base a otra en un espacio vectorial.
5. Deducir un algoritmo que permita pasar la expresión de un vector de una base a otra.
6. Utilizar la expresión de un vector como vector coordenado para determinar cuándo un conjunto de vectores es LI.

Preguntas básicas

1. ¿Cómo se expresa un vector como vector coordenado en una base B?
2. ¿Cómo se realizan las operaciones de suma y producto por un escalar entre los vectores de coordenadas de un espacio V?
3. ¿Cómo están formadas las columnas de la matriz de transición de una base B_1 a una base B_2?
4. ¿Cómo se relacionan las matrices de transición de B_1 a B_2 y de B_2 a B_1?
5. ¿Cómo se realiza el cambio de la expresión de un vector de una base B_1 a una base B_2?
6. ¿Cómo se emplean los vectores coordenados para determinar si un conjunto es LI?

Vea el módulo 7 del programa de televisión *Álgebra lineal*.

Contenidos



Mapa 4: módulo 7

7.1 Vectores coordenados

Definición 1

Sean V un espacio vectorial de dimensión n y $B=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$ una base para V. Entonces, para cada $\mathbf{v}\in V$ existen escalares únicos $a_1,a_2,\ldots,a_n$ tales que:

$$\mathbf{v}=a_1\mathbf{v}_1+a_2\mathbf{v}_2+\ldots+a_n\mathbf{v}_n\,.$$

El vector cuyos componentes son los escalares $a_1,a_2,\ldots,a_n$ se llama *vector de coordenadas* o *vector coordenado de* **v** *respecto a B*, y se escribe:

$$[\mathbf{v}]_B=\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix}.$$

Es importante observar que $[\mathbf{v}]_B$ depende del orden de los elementos de B. Luego, al dar la base B ésta siempre debe asumirse como una base ordenada.

Un vector de coordenadas es una n-tupla ordenada.

Ejemplo 1

En $\mathbb{R}^2$, la base canónica está compuesta por los vectores $\mathbf{e}_1=\begin{pmatrix}1\\0\end{pmatrix}$ y $\mathbf{e}_2=\begin{pmatrix}0\\1\end{pmatrix}$, pero dos vectores de $\mathbb{R}^2$ LI también forman una base para $\mathbb{R}^2$, así: los vectores $\mathbf{v}_1=\begin{pmatrix}1\\2\end{pmatrix}$ y $\mathbf{v}_2=\begin{pmatrix}3\\5\end{pmatrix}$ forman otra base de $\mathbb{R}^2$.

Sea $\mathbf{x}=\begin{pmatrix}-1\\-1\end{pmatrix}$; entonces, $\mathbf{x}=-1\mathbf{e}_1-1\mathbf{e}_2$.

Si $B_1=\{\mathbf{e}_1,\mathbf{e}_2\}$; entonces, $(\mathbf{x})_{B_1}=\begin{pmatrix}-1\\-1\end{pmatrix}$.

Cuando no se dice explícitamente en qué base está expresado **x**, se asume que el vector está en la base estándar.

Sea $B_2=\left\{\begin{pmatrix}1\\2\end{pmatrix},\begin{pmatrix}3\\5\end{pmatrix}\right\}$,
$$\mathbf{x}=2\begin{pmatrix}1\\2\end{pmatrix}-1\begin{pmatrix}3\\5\end{pmatrix}$$
$$\mathbf{x}=2\mathbf{v}_1-1\mathbf{v}_2$$

luego $(\mathbf{x})_{B_2}=\begin{pmatrix}2\\-1\end{pmatrix}$ (figura 7.1).

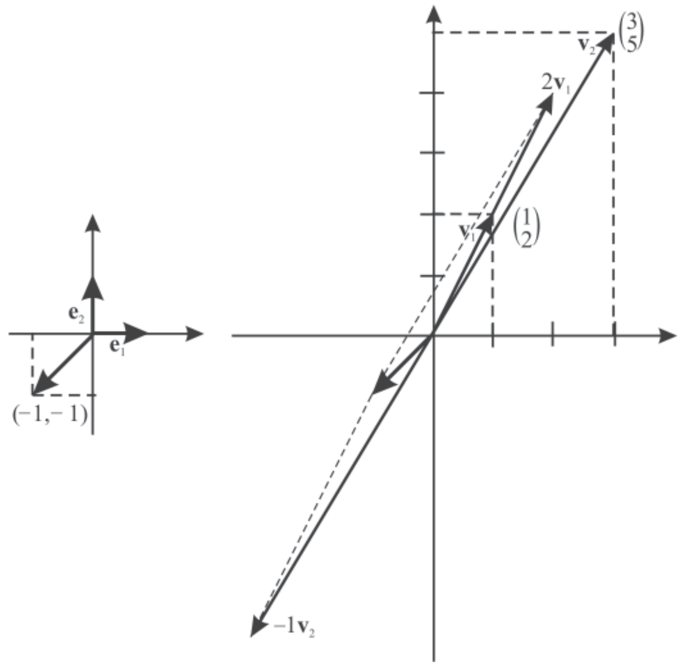


Figura 7.1

Ejemplo 2

En $\mathbb{P}_2$, consideremos el polinomio $P(x) = 3 + 2x + x^2$, el vector de coordenadas de $P(x)$:

i. En la base estándar, $B_1 = \{1, x, x^2\}$, es: $[P(x)]_{B_1} = \begin{pmatrix}3\\2\\1\end{pmatrix}$.

ii. En la base $B_2 = \{1 + x,\ 1 - x^2,\ 1 + x + x^2\}$ es (?)

Debemos calcular las constantes c_1, c_2, c_3 tales que:

$$P(x) = c_1(1 + x) + c_2(1 - x^2) + c_3(1 + x + x^2),$$

lo cual no es más que un problema de combinación lineal. Entonces,

$$3 + 2x + x^2 = (c_1 + c_2 + c_3) + (c_1 + c_3)x + (-c_2 + c_3)x^2,$$
$$c_1 + c_2 + c_3 = 3,$$
$$c_1 + c_3 = 2,$$
$$-c_2 + c_3 = 1.$$

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Cambio de base por rotación en $\mathbb{R}^2$».

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

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right],$$

luego $c_1 = 0,\ c_2 = 1$ y $c_3 = 2$,
así que:

$$3 + 2x + x^2 = 0(1+x) + (1-x^2) + 2(1+x+x^2),$$

$$\left[P(x)\right]_{B_2} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

Teorema 1

Sea B una base de un espacio vectorial V, $\mathbf{u}$ y $\mathbf{v}$ dos vectores de V y α un escalar, entonces:

$$[\mathbf{u}]_B + [\mathbf{v}]_B = [\mathbf{u}+\mathbf{v}]_B\,,$$
$$\alpha[\mathbf{u}]_B = [\alpha\mathbf{u}]_B\,.$$

Observación

Con los vectores de coordenadas se opera de la misma forma que en $\mathbb{R}^n$.

7.2 Matrices de transición y cambio de base

Dado un espacio vectorial V con base B_1 podemos expresar cualquier $\mathbf{v} \in V$ en B_1. Si se tiene una segunda base B_2 para V, ¿puede calcularse $(\mathbf{v})_{B_2}$ empleando $(\mathbf{v})_{B_1}$?

Veamos el problema en $\mathbb{R}^3$. Sean $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ y $B_2 = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ dos bases para $\mathbb{R}^3$. Si $\mathbf{v} \in \mathbb{R}^3$, entonces $(\mathbf{v})_{B_1} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$, es decir,

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3. \tag{1}$$

Ahora, cada uno de los vectores $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ podemos expresarlo en la base B_2, así:

$$\begin{aligned}
\mathbf{v}_1 &= a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + a_{31}\mathbf{w}_3 \,. \\
\mathbf{v}_2 &= a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + a_{32}\mathbf{w}_3 \,. \qquad (2) \\
\mathbf{v}_3 &= a_{13}\mathbf{w}_1 + a_{23}\mathbf{w}_2 + a_{33}\mathbf{w}_3 \,.
\end{aligned}$$

Al sustituir (2) en (1) tenemos:

$$\begin{aligned}
\mathbf{v} &= c_1(a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + a_{31}\mathbf{w}_3) + c_2(a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + a_{32}\mathbf{w}_3) \\
&\quad + c_3(a_{13}\mathbf{w}_1 + a_{23}\mathbf{w}_2 + a_{33}\mathbf{w}_3) \\
&= (c_1a_{11} + c_2a_{12} + c_3a_{13})\mathbf{w}_1 + (c_1a_{21} + c_2a_{22} + c_3a_{23})\mathbf{w}_2 \\
&\quad + (c_1a_{31} + c_2a_{32} + c_3a_{33})\mathbf{w}_3 \,.
\end{aligned}$$

Luego:

$$(\mathbf{v})_{B_2} = \begin{pmatrix} c_1a_{11} + c_2a_{12} + c_3a_{13} \\ c_1a_{21} + c_2a_{22} + c_3a_{23} \\ c_1a_{31} + c_2a_{32} + c_3a_{33} \end{pmatrix} = \begin{array}{c} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ \begin{array}{ccc} \downarrow & \downarrow & \downarrow \\ (\mathbf{v}_1)_{B_2} & (\mathbf{v}_2)_{B_2} & (\mathbf{v}_3)_{B_2} \end{array} \end{array} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

$$\begin{aligned}
&= \left[(\mathbf{v}_1)_{B_2}\ (\mathbf{v}_2)_{B_2}\ (\mathbf{v}_3)_{B_2} \right] (\mathbf{v})_{B_1} \\
&= P_{B_2 \leftarrow B_1}\ (\mathbf{v})_{B_1} \,.
\end{aligned}$$

Definición 2

La matriz $P_{B_2 \leftarrow B_1}$ cuyas columnas son los vectores de la base B_1 expresadas en la base B_2 recibe el nombre de *matriz de transición* de la base B_1 a la base B_2.

Si $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ y $B_2 = \{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n\}$,

$$P_{B_2 \leftarrow B_1} = \begin{array}{c} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \\ \begin{array}{cccc} \downarrow & \downarrow & & \downarrow \\ (\mathbf{v}_1)_{B_2} & (\mathbf{v}_2)_{B_2} & & (\mathbf{v}_n)_{B_2} \end{array} \end{array}$$

Teorema 2

Sean B_1 y B_2 bases de un espacio vectorial V. Sea $P_{B_2 \leftarrow B_1}$ la matriz de transición de B_1 a B_2; entonces, para todo $\mathbf{x} \in V$,

$$(\mathbf{x})_{B_2} = P_{B_2 \leftarrow B_1} (\mathbf{x})_{B_1} \,.$$

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Cambio de base por rotación en $\mathbb{R}^3$».

Es importante resaltar que la matriz de transición $P_{B_2 \leftarrow B_1}$ es invertible. ¿Por qué?

Teorema 3

Si P es la matriz de transición de B_1 a B_2, entonces P^{-1} es la matriz de transición de B_2 a B_1.

Demostración

Sabemos que $(\mathbf{x})_{B_2} = P(\mathbf{x})_{B_1}$.

Como P es invertible, multipliquemos por P^{-1} a la izquierda, en cada lado de la ecuación.

$$\begin{aligned} P^{-1}(\mathbf{x})_{B_2} &= P^{-1}P(\mathbf{x})_{B_1} \\ &= I(\mathbf{x})_{B_1} = (\mathbf{x})_{B_1}, \end{aligned}$$

luego $(\mathbf{x})_{B_1} = P^{-1}(\mathbf{x})_{B_2}$.

Este teorema proporciona una forma simple de encontrar la matriz de transición, $P_{B_2 \leftarrow B_1}$, cuando B_1 es la base estándar de $\mathbb{R}^n$. Si $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es otra base cualquiera, la matriz de transición $P_{B_1 \leftarrow B_2}$ es la matriz cuyas columnas son los vectores de B_2 expresados en B_1; como los vectores de B_2 están dados en la base estándar, entonces:

$$P_{B_1 \leftarrow B_2} = [\mathbf{v}_1 \ \mathbf{v}_2 ... \mathbf{v}_n],$$

o sea que $P_{B_1 \leftarrow B_2}$ se forma escribiendo la base B_2 como una matriz. Luego, aplicando el teorema:

$$P_{B_2 \leftarrow B_1} = (P_{B_1 \leftarrow B_2})^{-1}.$$

Ejemplo 3

En $\mathbb{R}^3$, sean $B_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ y $B_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$.

Expresemos el vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ en términos de la base B_2.

Solución

La matriz de transición de B_2 a B_1, $P_{B_1 \leftarrow B_2}$, está dada por $P_{B_1 \leftarrow B_2} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

Encontremos $P_{B_2 \leftarrow B_1}$ aplicando el algoritmo para la obtención de la inversa, así:

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

$$P_{B_2 \leftarrow B_1}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_{B_2} = P_{B_2 \leftarrow B_1} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{B_1}$$

$$= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y \\ y-z \\ z \end{bmatrix}.$$

Ejemplo 4

Escribamos el polinomio $a_0 + a_1 x + a_2 x^2$ de $\mathbb{P}_2$ en términos de la base:

$$B_2 = \{1,\ x-1,\ x^2-1\}.$$

Empleando la base estándar, $B_1 = \{1, x, x^2\}$, se tiene:

$$(1)_{B_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad (x-1)_{B_1} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad (x^2-1)_{B_1} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

La matriz de transición $P_{B_1 \leftarrow B_2}$ es:

$$P_{B_1 \leftarrow B_2} = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Determinemos $P_{B_2 \leftarrow B_1}$:

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

$$P_{B_2 \leftarrow B_1}$$

$a_0 + a_1\mathbf{x} + a_2\mathbf{x}^2$ expresado como vector coordenado en la base B_1 es: $\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}_{B_2} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_0 + a_1 + a_2 \\ a_1 \\ a_2 \end{bmatrix}.$$

Podemos verificar el resultado así:

$$a_0 + a_1 x + a_2 x^2 = (a_0 + a_1 + a_2) + a_1(x-1) + a_2(x^2 - 1).$$

Ejemplo 5

En $\mathbb{R}^2$ supongamos que $(\mathbf{x})_{B_1} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, donde $B_1 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\}$; expresemos $\mathbf{x}$ en términos de la base $B_2 = \left\{ \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 5 \\ -1 \end{pmatrix} \right\}$.

$$(\mathbf{x})_{B_2} = P_{B_2 \leftarrow B_1} (\mathbf{x})_{B_1}.$$

Las columnas de $P_{B_2 \leftarrow B_1}$ son los vectores de B_1 expresados en B_2.

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = a_{11} \begin{pmatrix} 0 \\ 3 \end{pmatrix} + a_{21} \begin{pmatrix} 5 \\ -1 \end{pmatrix},$$

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} = a_{12} \begin{pmatrix} 0 \\ 3 \end{pmatrix} + a_{22} \begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

A cada una de estas ecuaciones vectoriales la llevamos a un sistema de dos ecuaciones lineales, así:

$$\begin{aligned} 5a_{21} &= 1 \\ 3a_{11} - 1a_{21} &= 1 \end{aligned} \qquad \text{y} \qquad \begin{aligned} 5a_{22} &= 2 \\ 3a_{12} - 1a_{22} &= 3 \end{aligned}$$

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

Estos dos sistemas poseen la misma matriz de coeficientes; por lo tanto, podemos formar una sola estructura con la matriz de coeficientes al lado izquierdo y los términos independientes al lado derecho.

$$\left[\begin{array}{cc|cc} 0 & 5 & 1 & 2 \\ 3 & -1 & 1 & 3 \end{array}\right] \xrightarrow{\text{operaciones elementales}} \left[\begin{array}{cc|cc} 1 & 0 & 2/5 & 17/15 \\ 0 & 1 & 1/5 & 2/5 \end{array}\right]$$

$$\downarrow \qquad \downarrow$$

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}_{B_2} \begin{pmatrix} 2 \\ 3 \end{pmatrix}_{B_2}$$

Luego:

$$P_{B_2 \leftarrow B_1} = \begin{bmatrix} 2/5 & 17/15 \\ 1/5 & 2/5 \end{bmatrix}.$$

Entonces:

$$\begin{aligned} (\mathbf{x})_{B_2} &= \begin{bmatrix} 2/5 & 17/15 \\ 1/5 & 2/5 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 1/3 \\ 0 \end{bmatrix}. \end{aligned}$$

Podemos observar que para obtener la matriz de transición de B_1 a B_2 hacemos lo siguiente:

$$\left[B_2 \mid B_1\right] \xrightarrow{\text{operaciones elementales}} \left[I \mid P_{B_2 \leftarrow B_1}\right].$$

7.3 Vectores coordenados e independencia lineal

Teorema 4

Sea $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ una base del espacio vectorial V. Supongamos que:

$$(\mathbf{x}_1)_{B_1} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \ (\mathbf{x}_2)_{B_1} = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, ..., \ (\mathbf{x}_n)_{B_1} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix}.$$

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

Entonces $\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n\}$ es LI si y sólo si $\det A \neq 0.$

Este teorema se desprende fácilmente del hecho de que los vectores en coordenadas son n-tuplas ordenadas, o sea elementos de $\mathbb{R}^n$, y del teorema 4 del módulo 4, el cual dice que $\det A \neq 0$ si y sólo si las columnas de A son LI.

Ejemplo 6

En $\mathbb{P}_3$ determine si los polinomios $1+x^2$, $-1-3x+4x^2+5x^3$, $2+5x-6x^3$, $4+6x+3x^2+7x^3$ son LD o LI.

Usando la base estándar de $\mathbb{P}_3$, $B_1=\{1,x,x^2,x^3\}$ se tiene:

$$(1+x^2)_{B_1}=\begin{pmatrix}1\\0\\1\\0\end{pmatrix},$$

$$(-1-3x+4x^2+5x^3)_{B_1}=\begin{pmatrix}-1\\-3\\4\\5\end{pmatrix},$$

$$(2+5x-6x^3)_{B_1}=\begin{pmatrix}2\\5\\0\\-6\end{pmatrix} \quad \text{y}$$

$$(4+6x+3x^2+7x^3)_{B_1}=\begin{pmatrix}4\\6\\3\\7\end{pmatrix}.$$

Entonces:

$$\det A=\begin{vmatrix}1&-1&2&4\\0&-3&5&6\\1&4&0&3\\0&5&-6&7\end{vmatrix}\neq 0.$$

Luego, los polinomios son LI.

Ejercicios del capítulo 1 (módulo 7)

En los ejercicios siguientes, todas las bases son bases ordenadas.

En los ejercicios 1 a 7 calcule el vector coordenado de **v** respecto a la base B dada para el espacio vectorial V.

1. V es $\mathbb{R}^2$, $B=\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\}$, $\mathbf{v}=\begin{pmatrix}3\\1\end{pmatrix}$.

2. V es $\mathbb{R}^3$, $B=\left\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$, $\mathbf{v}=\begin{pmatrix}4\\-1\\0\end{pmatrix}$.

3. V es $\mathbb{R}^3$, $B=\left\{\begin{pmatrix}1\\-1\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}1\\0\\2\end{pmatrix}\right\}$, $\mathbf{v}=\begin{pmatrix}2\\3\\-1\end{pmatrix}$.

4. V es $\mathbb{P}_1$, $B=\{1+x,\ -1+2x\}$, $\mathbf{v}=3+4x$.

5. V es $\mathbb{P}_2$, $B=\{1,\ x-1,\ x^2-1\}$, $\mathbf{v}=1+2x-x^2$.

6. V es $\mathbb{P}_2$, $B=\{1-x+x^2, 1+x, 1+x^2\}$, $\mathbf{v}=3-2x+4x^2$.

7. V es M_{22}, $B=\left\{\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}1&1\\0&0\end{bmatrix},\begin{bmatrix}1&1\\1&0\end{bmatrix},\begin{bmatrix}1&1\\1&1\end{bmatrix}\right\}$ $\mathbf{v}=\begin{bmatrix}2&0\\-1&3\end{bmatrix}$.

En los ejercicios 8 a 10, sean:

$$B_1=\left\{\begin{pmatrix}1\\3\end{pmatrix},\begin{pmatrix}-2\\1\end{pmatrix}\right\},\ B_2=\left\{\begin{pmatrix}1/\sqrt{2}\\-1/\sqrt{2}\end{pmatrix},\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}\right\},\ B_3=\left\{\begin{pmatrix}-2\\1\end{pmatrix},\begin{pmatrix}1\\3\end{pmatrix}\right\},\ B_4=\left\{\begin{pmatrix}\sqrt{3}/2\\1/2\end{pmatrix},\begin{pmatrix}-1/2\\\sqrt{3}/2\end{pmatrix}\right\}$$

bases para $\mathbb{R}^2$.

8. Halle las matrices de transición:

   a. de B_1 y B_2.
   b. de B_2 y B_3.
   c. de B_1 y B_3.

   Multiplique las matrices de a y b en ambas formas. ¿Está alguno de estos productos relacionado con la matriz de c?

9. Sea $\mathbf{x}=\begin{pmatrix}1\\1\end{pmatrix}$. Exprese $(\mathbf{x})_{B_1}$. Usando las matrices del ejercicio 8, calcule $(\mathbf{x})_{B_2}, (\mathbf{x})_{B_3}$.

10. Calcule las matrices de transición de B_2 a B_4 y de B_4 a B_2.

En los ejercicios 11 a 13 calcule el vector **v** si el vector de coordenadas $[\mathbf{v}]_B$ está dado respecto a la base B de V.

11. V es $\mathbb{R}^2$, $B=\left\{\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}\right\}$, $(\mathbf{v})_B=\begin{pmatrix}1\\2\end{pmatrix}$.

12. V es $\mathbb{P}_2$, $B=\{1-x,\ 1,\ 1+x+x^2\}$, $(\mathbf{v})_B=\begin{bmatrix}-1\\1\\2\end{bmatrix}$.

13. V es $\mathbb{P}_3$, $B=\{-1+x^2,\ 2+2x+x^3,\ 1+2x-x^2+3x^3,\ 2x^2+3x^3\}$, $(\mathbf{v})_B=\begin{bmatrix}2\\1\\-1\\3\end{bmatrix}$.

14. Sean $B_1=\{(1,3),(-1,2)\}$, $B_2=\{(0,1),(-2,3)\}$ bases para $\mathbb{R}^2$, y sean $\mathbf{v}=(3,4)$, $\mathbf{w}=(-4,5)$.

    a. Determine los vectores de coordenadas de **v** y **w** respecto a la base B_2.
    b. Determine la matriz de transición $P_{B_1\leftarrow B_2}$ de la base B_2 en la base B_1.
    c. Determine los vectores de coordenadas de **v** y **w** respecto a B_1 utilizando $P_{B_1\leftarrow B_2}$.
    d. Determine los vectores de coordenadas de **v** y **w** respecto a B_1 de manera directa.
    e. Determine la matriz de transición $P_{B_2\leftarrow B_1}$ de la base B_1 en la base B_2.
    f. Determine los vectores de coordenadas de **v** y **w** respecto a B_2 utilizando $P_{B_2\leftarrow B_1}$. Compare las respuestas con las del literal *a*.

15. Sean $B_1=\{1+x^2,\ -2+x,\ 3+x\}$ y $B_2=\{x+2x^2,\ 3+x^2,\ x\}$ bases para $\mathbb{P}_2$. Sean $\mathbf{v}=6-4x+8x^2$ y $\mathbf{w}=9-x+7x^2$. Responda los literales del ejercicio 14.

16. Sean $B_1=\{\mathbf{v}_1,\mathbf{v}_2\}$ y $B_2=\{\mathbf{w}_1,\mathbf{w}_2\}$ bases para $\mathbb{P}_1$, donde $\mathbf{w}_1=x$, $\mathbf{w}_2=1+x$. Si la matriz de transición de B_1 a B_2 es $\begin{bmatrix}2&3\\-1&2\end{bmatrix}$, determine B_1.

17. Sean $B_1=\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ y $B_2=\{\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3\}$ bases para $\mathbb{R}^3$, donde $\mathbf{v}_1=(1,0,1)$, $\mathbf{v}_2=(1,1,0)$, $\mathbf{v}_3=(0,0,1)$. Si la matriz de transición de B_2 en B_1 es $\begin{bmatrix}1&1&2\\2&1&1\\-1&-1&1\end{bmatrix}$, determine B_2.

18. Sean $B_1 = \{\mathbf{v}_1, \mathbf{v}_2\}$ y $B_2 = \{\mathbf{w}_1, \mathbf{w}_2\}$ bases para $\mathbb{R}^2$, donde $\mathbf{v}_1 = (1, 2)$, $\mathbf{v}_2 = (0, 1)$. Si la matriz de transición de B_2 en B_1 es $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, determine B_2.

19. Suponga que los ejes x y y en el plano se rotan un ángulo θ en sentido antihorario generando nuevos ejes x', y'.

    a. Determine las coordenadas x, y de los vectores **i** y **j** rotados.

    b. Demuestre que la matriz de cambio de coordenadas está dada por $\begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}$.

En los ejercicios 20 a 24 determine si el conjunto de vectores dado es LI o LD.

20. En $\mathbb{P}_2 : 3 + 2x,\ 1 - x + x^2,\ 2x - x^2$.

21. En $\mathbb{P}_2 : -2 + 2x,\ 2x + x + 12x^2,\ x + 4x^2$.

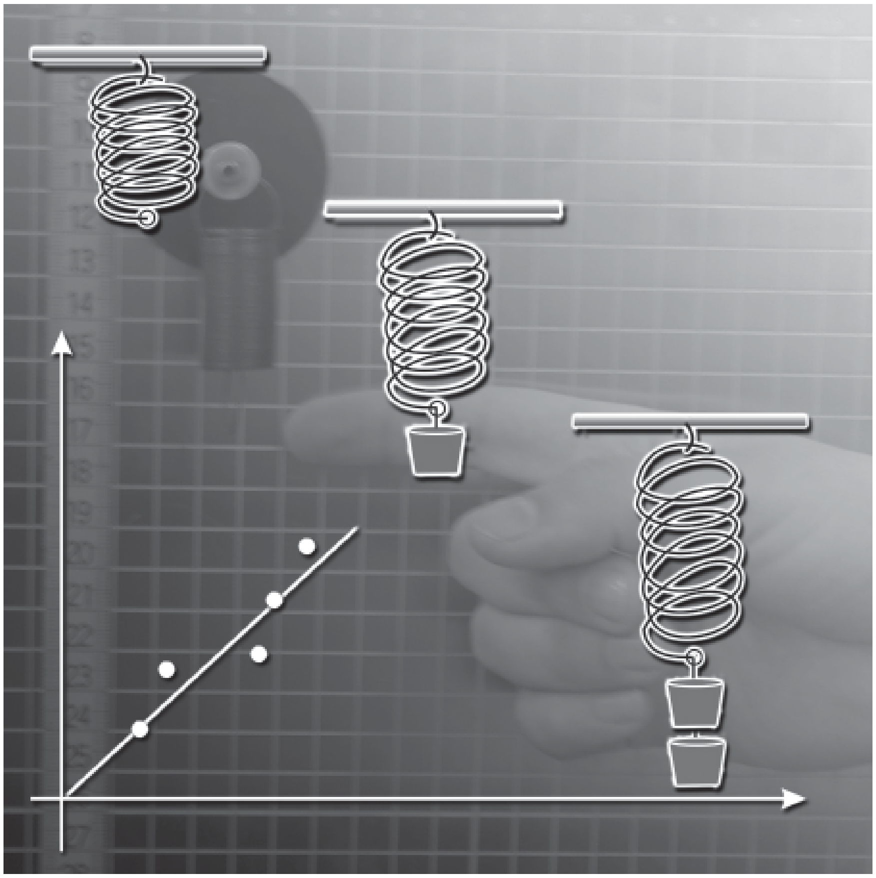
22. En $M_{22} : \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 7 \\ 9 & 6 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 4 & -2 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix}$.

23. En $\mathbb{P}_n : \{ p_1, p_2, ..., p_{n+1} : p_i(0) = 0,\ i = 1, ..., n+1 \}$.

24. En $M_{mn} : \{A_1, A_2, ..., A_{mn}$: la primera componente de cada matriz es cero$\}$.

Capítulo 2

Ortogonalidad



El experimentador desea comprobar la relación existente entre la fuerza aplicada al resorte y su deformación. Para ello elabora una tabla donde registra el peso aplicado y la deformación sufrida por el resorte, y obtiene una nube de puntos que ajusta a una recta mediante el método de mínimos cuadrados.

Contenido breve

Presentación

La ortogonalidad tiene su importancia práctica en el método de aproximación por mínimos cuadrados, con el cual se busca la mejor solución a un sistema de ecuaciones lineales inconsistente, evaluando el error cometido. Sin embargo, la ortogonalidad tiene una importancia más allá de esto y es que nos proporciona, cuando trabajamos en el espacio tridimensional, unos ejes coordenados ortogonales (perpendiculares).

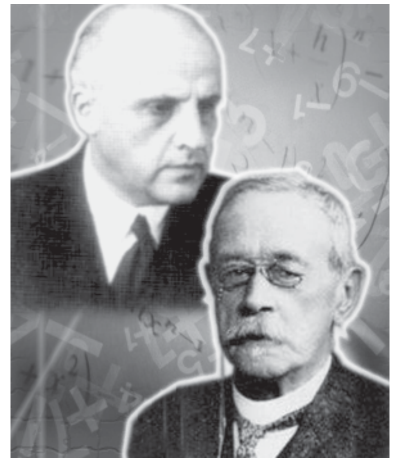
Si la idea de *base* es un paso clave para conectar la geometría de un espacio vectorial con el álgebra, ésta se hace de una manera mucho más intuitiva cuando la base es ortogonal. Necesitamos una *base* para convertir cada construcción geométrica en un cálculo algebraico, y necesitamos una base ortogonal para que dicho cálculo sea sencillo. Una manera de optimizar este procedimiento es volver unitarios los vectores de la base ortogonal; esto convierte la base ortogonal en base ortonormal.

Las ideas introducidas en los párrafos anteriores serán las que desarrollaremos en este tema. Éstas serán el punto de partida para la resolución de los problemas algebraicos que involucren bases ortonormales y proyecciones en $\mathbb{R}^n$.

Finalizado el capítulo, haremos una generalización de este tema a otros espacios vectoriales introduciendo el concepto de producto interno.

Módulo 8

Bases ortonormales y proyecciones en $\mathbb{R}^n$



El matemático danés Jörgen Pedersen Gram (1850-1916) es conocido por el método de ortogonalización, aunque se presume que no fue él quien primero lo utilizó. Aparentemente fue ideado por Pierre Simon de Laplace y utilizado también por Augustin Louis Cauchy en 1836. Gram murió arrollado por una bicicleta a la edad de 61 años.

El matemático alemán Erhard Schmidt (1876-1959) fundó el primer instituto de matemáticas aplicadas de Berlín. Alumno de David Hilbert, Schmidt hizo sus mayores contribuciones en ecuaciones integrales y teoría de funciones en el espacio de Hilbert.

Introducción

En $\mathbb{R}^n$ la base estándar $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n\}$ tiene la propiedad de que $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ para $i \neq j$ y además $|\mathbf{e}_i| = 1,\ i = 1, \ldots, n$.

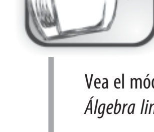
Hemos podido apreciar lo cómoda que resulta esta base en el manejo algebraico de los vectores de $\mathbb{R}^n$. En este módulo mostraremos la forma de dotar los subespacios de bases con las características de la base estándar de $\mathbb{R}^n$. A estas bases las llamaremos *bases ortonormales*. Además, veremos el concepto de proyección ortogonal de un vector de $\mathbb{R}^n$ sobre un subespacio de éste, concepto que nos será de gran utilidad cuando se trata de aproximar la solución de un sistema inconsistente $y = Ax$ de manera que se minimice el error cometido.

Objetivos

1. Construir bases ortonormales en los subespacios de $\mathbb{R}^n$, con el fin de facilitar el cálculo algebraico en un problema determinado.
2. Encontrar la proyección ortogonal de un vector sobre un subespacio de $\mathbb{R}^n$.
3. Descomponer un vector de $\mathbb{R}^n$ como suma de dos vectores, uno en un subespacio de $\mathbb{R}^n$ y otro en su complemento ortogonal.
4. Establecer propiedades relativas a las magnitudes de los vectores en $\mathbb{R}^n$.

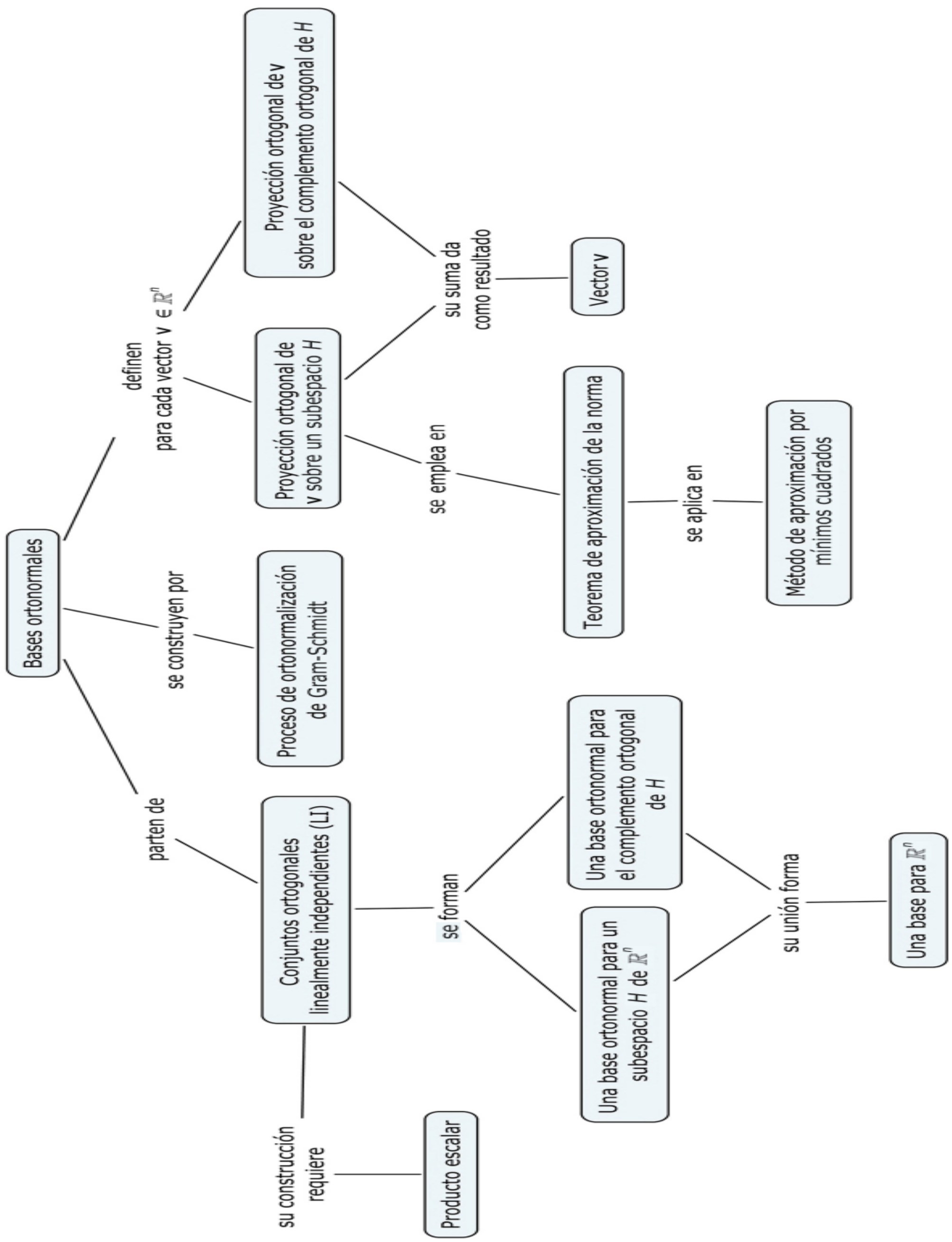
Preguntas básicas

1. ¿Qué es un conjunto ortogonal?
2. ¿Qué propiedades tienen los conjuntos ortogonales?
3. ¿Cómo se construye una base ortonormal para un subespacio de $\mathbb{R}^n$?
4. ¿Cómo se calcula la proyección ortogonal de un vector sobre un subespacio?
5. ¿Cómo se encuentra el complemento ortogonal de un subespacio H de $\mathbb{R}^n$?
6. ¿Cómo se relacionan un subespacio H y su complemento ortogonal $H^{\perp}$ en $\mathbb{R}^n$?
7. En la descomposición de un vector $\mathbf{a} \in \mathbb{R}^n$, como $\mathbf{a} = \mathbf{p} + \mathbf{q}$ con $\mathbf{p} \in H$ y $\mathbf{q} \in H^{\perp}$, ¿a qué son iguales $\mathbf{p}$ y $\mathbf{q}$?
8. ¿Cuál es la mejor aproximación de un vector $\mathbf{a} \in \mathbb{R}^n$ a los vectores de un subespacio H de $\mathbb{R}^n$?



Vea el módulo 8 del programa de televisión *Álgebra lineal*.

Contenidos



Mapa 5: módulos 8 y 9

8.1 Conjuntos ortonormales

Definición 1

Sea $S=\{\mathbf{u}_1,\mathbf{u}_2,...,\mathbf{u}_k\}$ un subconjunto de $\mathbb{R}^n$.

Se dice que *S* es un *conjunto ortonormal* si:

1. $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ para $i \neq j$.
2. $|\mathbf{u}_i| = 1$.

Cuando se cumple la condición 1, se dice que el conjunto es *ortogonal*.

Como los vectores de *S* son de magnitud unitaria, se dice que son vectores normalizados. De ahí que un conjunto que cumpla las dos condiciones se llame ortonormal. Recordemos que si $\mathbf{x}\in\mathbb{R}^n$, $\mathbf{x}=(x_1,x_2,...,x_n)$ y la magnitud de **x**, $|\mathbf{x}|$, está dada por:

$$|\mathbf{x}| = \sqrt{\mathbf{x}\cdot\mathbf{x}} = \sqrt{x_1^2+x_2^2+...+x_n^2}.$$

Si $|\mathbf{u}_i|=1$, entonces $\mathbf{u}_i \cdot \mathbf{u}_i = 1$.

Nota: los conceptos de ortogonalidad y magnitud de los vectores de $\mathbb{R}^n$ están dados en términos del producto escalar. Es importante, antes de seguir adelante, recordar las propiedades del producto escalar.

Ejemplo 1

Sean $\mathbf{a}=(1,\ 3,\ 0)$, $\mathbf{b}=(3,\ -1,\ 1)$ y $\mathbf{c}=(3,\ -1,\ 10)$; entonces $\{\mathbf{a},\mathbf{b},\mathbf{c}\}$ es un conjunto ortogonal en $\mathbb{R}^3$ ya que $\mathbf{a}\cdot\mathbf{b}=\mathbf{a}\cdot\mathbf{c}=\mathbf{b}\cdot\mathbf{c}=0$.

$$|\mathbf{a}|=\sqrt{10}\ ,\ |\mathbf{b}|=\sqrt{11}\ \text{ y } |\mathbf{c}|=\sqrt{110}\ .$$

Los vectores:

$$\mathbf{u}_1=\frac{\mathbf{a}}{|\mathbf{a}|}=\frac{1}{\sqrt{10}}(1,\ 3,\ 0),\ \ \mathbf{u}_2=\frac{\mathbf{b}}{|\mathbf{b}|}=\frac{1}{\sqrt{11}}(3,\ -1,\ 1)\ \text{ y }\ \mathbf{u}_3=\frac{\boldsymbol{c}}{|\boldsymbol{c}|}=\frac{1}{\sqrt{110}}(3,\ -1,\ 10)$$

son vectores unitarios en las direcciones de **a**, **b** y **c**, luego $\{\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3\}$ es un conjunto ortonormal.

Además, $\text{gen}\{\mathbf{a},\mathbf{b},\mathbf{c}\}=\text{gen}\{\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3\}$.

Teorema 1

Sea $S=\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_k\}$ un conjunto ortogonal de vectores no nulos de $\mathbb{R}^n$. Entonces S es linealmente independiente.

Demostración

Planteamos la ecuación:

$$c_1\mathbf{u}_1+c_2\mathbf{u}_2+\ldots+c_i\mathbf{u}_i+\ldots+c_k\mathbf{u}_k=\mathbf{0}.$$

A ambos lados de la igualdad multiplicamos escalarmente por $\mathbf{u}_i$, $i=1,\ldots,k$

$$(c_1\mathbf{u}_1+c_2\mathbf{u}_2+\ldots+c_i\mathbf{u}_i+\ldots+c_k\mathbf{u}_k)\cdot\mathbf{u}_i=\mathbf{0}\cdot\mathbf{u}_i=0.$$

Aplicando la propiedad de distribución del producto escalar, tenemos:

$$c_1(\mathbf{u}_1\cdot\mathbf{u}_i)+c_2(\mathbf{u}_2\cdot\mathbf{u}_i)+\ldots+c_i(\mathbf{u}_i\cdot\mathbf{u}_i)+\ldots+c_k(\mathbf{u}_k\cdot\mathbf{u}_i)=0.$$

Ahora, $\mathbf{u}_i\cdot\mathbf{u}_j=0$ para $i\neq j$, luego el lado izquierdo de la ecuación se reduce a $c_i(\mathbf{u}_i\cdot\mathbf{u}_i)$, de donde $c_i\left|\mathbf{u}_i\right|^2=0.$

Como $\mathbf{u}_i\neq\mathbf{0}$, entonces $c_i=0$, $i=1,\ldots,k.$

Corolario 1

Un conjunto ortonormal de vectores en $\mathbb{R}^n$ es linealmente independiente.

8.2 Otra forma del problema de la base. Proceso de ortonormalización de Gram-Schmidt

Dada una base $S=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k\}$ para un subespacio de $\mathbb{R}^n$, hallar una base ortonormal $B=\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_k\}$. La solución a este problema nos la brinda el siguiente teorema, conocido como proceso de ortonormalización de Gram-Schmidt.

Teorema 2

Sea H un subespacio de dimensión k de $\mathbb{R}^n$; entonces H tiene una base ortonormal.

Demostración

Supongamos que $\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k\}$ es una base para H.

Paso 1

Construcción de un vector unitario en la dirección de $\mathbf{v}_1$, $\mathbf{v}_1 \neq \mathbf{0}$, ya que $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ es linealmente independiente.

Sea $\mathbf{u}_1 = \dfrac{\mathbf{v}_1}{|\mathbf{v}_1|}$.

Paso 2

Construcción de un vector ortogonal a $\mathbf{u}_1$.

Sea:

$$\mathbf{v}_2' = \mathbf{v}_2 - \text{proy}_{u_1} \mathbf{v}_2.$$

$$\text{proy}_{u_1} \mathbf{v}_2 = \frac{(\mathbf{v}_2 \cdot \mathbf{u}_1)}{|\mathbf{u}_1|^2} \mathbf{u}_1;$$

como $|\mathbf{u}_1| = 1$,

$$\mathbf{v}_2' = \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1.$$

Veamos que $\mathbf{v}_2'$ es ortogonal a $\mathbf{u}_1$:

$$\mathbf{v}_2' \cdot \mathbf{u}_1 = \mathbf{v}_2 \cdot \mathbf{u}_1 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\underbrace{\mathbf{u}_1 \cdot \mathbf{u}_1}_{1} = 0.$$

Lo anterior se representa en la figura 8.1.

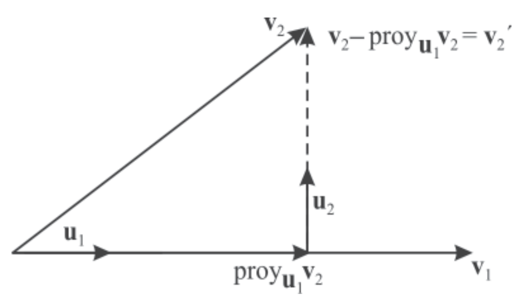


Figura 8.1

Paso 3

Construcción de un vector unitario en la dirección de $\mathbf{v}_2'$.

Sea $\mathbf{u}_2 = \dfrac{\mathbf{v}_2'}{|\mathbf{v}_2'|}$.

Entonces, $\{\mathbf{u}_1, \mathbf{u}_2\}$ es un conjunto ortonormal.

Supongamos que se ha construido $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p\}$; $p < k$ un conjunto ortonormal. Veamos cómo agregar un nuevo vector al conjunto, de modo que siga siendo ortonormal.

Paso 4

Construcción de un vector ortogonal a $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p\}$.

Sea $\mathbf{v}'_{p+1} = \mathbf{v}_{p+1} - (\mathbf{v}_{p+1} \cdot \mathbf{u}_1)\mathbf{u}_1 - ... - (\mathbf{v}_{p+1} \cdot \mathbf{u}_i)\mathbf{u}_i - ... - (\mathbf{v}_{p+1} \cdot \mathbf{u}_p)\mathbf{u}_p$.

$\mathbf{v}'_{p+1}$ se construye quitándole al vector $\mathbf{v}_{p+1}$ sus proyecciones ortogonales sobre los vectores del conjunto ortonormal construido. Este vector es ortogonal a todos los vectores del conjunto. Veámoslo:

$$\mathbf{v}'_{p+1} \cdot \mathbf{u}_i = \cancel{\mathbf{v}_{p+1} \cdot \mathbf{u}_i} - (\mathbf{v}_{p+1} \cdot \mathbf{u}_1)\underbrace{(\mathbf{u}_1 \cdot \mathbf{u}_i)}_{0} ... - (\cancel{\mathbf{v}_{p+1} \cdot \mathbf{u}_i})\underbrace{(\mathbf{u}_i \cdot \mathbf{u}_i)}_{1} - ... - (\mathbf{v}_{p+1} \cdot \mathbf{u}_p)\underbrace{(\mathbf{u}_p \cdot \mathbf{u}_i)}_{0}$$

$$= 0.$$

Paso 5

Construcción de un vector unitario en la dirección de $\mathbf{v}'_{p+1}$.

Sea $\mathbf{u}_{p+1} = \dfrac{\mathbf{v}'_{p+1}}{|\mathbf{v}'_{p+1}|}$.

El proceso continúa hasta completar los k vectores de la base ortonormal.

8.3 Proyecciones ortogonales

Definición 2

Sea H un subespacio de dimensión k de $\mathbb{R}^n$ con base ortonormal $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ y $\mathbf{v}$ un vector de $\mathbb{R}^n$. Entonces la *proyección ortogonal de* $\mathbf{v}$ *sobre H*, denotada $\text{proy}_H\, \mathbf{v}$, es un vector de H dado por:

$$\text{proy}_{\boldsymbol{H}}\, \mathbf{v} = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\mathbf{u}_2 + ... + (\mathbf{v} \cdot \mathbf{u}_k)\mathbf{u}_k.$$

Ejemplo 2

Sea $W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x - y + 2z = 0 \right\}$ un subespacio en $\mathbb{R}^3$.

i. Determine una base y la dimensión de W.
ii. Determine una base ortonormal para W.

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Proyección sobre el plano en R^3».

iii. Si $\mathbf{a}=\begin{pmatrix}2\\1\\-1\end{pmatrix}\in\mathbb{R}^3$, determine $\operatorname{proy}_W \mathbf{a}$.

Solución

i. $\begin{pmatrix}x\\y\\z\end{pmatrix}\in W \Leftrightarrow \begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}y-2z\\y\\z\end{pmatrix}=y\begin{pmatrix}1\\1\\0\end{pmatrix}+z\begin{pmatrix}-2\\0\\1\end{pmatrix}.$

$W=\text{gen}\left\{\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}-2\\0\\1\end{pmatrix}\right\}$. Como los vectores son LI, entonces:

$\left\{\mathbf{v}_1=\begin{pmatrix}1\\1\\0\end{pmatrix},\ \mathbf{v}_2=\begin{pmatrix}-2\\0\\1\end{pmatrix}\right\}$ es una base para W. $\dim W=2$.

A partir de esta base, podemos construir una base ortonormal para W.

ii. Designemos $\{\mathbf{u}_1,\mathbf{u}_2\}$ la base ortonormal.

1. $\mathbf{u}_1=\dfrac{\mathbf{v}_1}{|\mathbf{v}_1|}=\dfrac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}.$

2. $\mathbf{v}_2'=\mathbf{v}_2-(\mathbf{v}_2\cdot\mathbf{u}_1)\mathbf{u}_1$

$$=\begin{pmatrix}-2\\0\\1\end{pmatrix}-\left[\begin{pmatrix}-2\\0\\1\end{pmatrix}\cdot\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}\right]\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}=\begin{pmatrix}-1\\1\\1\end{pmatrix}.$$

3. $\mathbf{u}_2=\dfrac{\mathbf{v}_2'}{|\mathbf{v}_2'|}=\dfrac{1}{\sqrt{3}}\begin{pmatrix}-1\\1\\1\end{pmatrix}.$

Base ortonormal de $W=\left\{\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix},\begin{pmatrix}-1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{pmatrix}\right\}.$

iii. $\text{proy}_W\ \mathbf{a} = (\mathbf{a}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{a}\cdot\mathbf{u}_2)\mathbf{u}_2.$

$$= \left[\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}\right] \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} + \left[\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}\right] \begin{pmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} = \begin{pmatrix} 13/6 \\ 5/6 \\ -2/3 \end{pmatrix}.$$

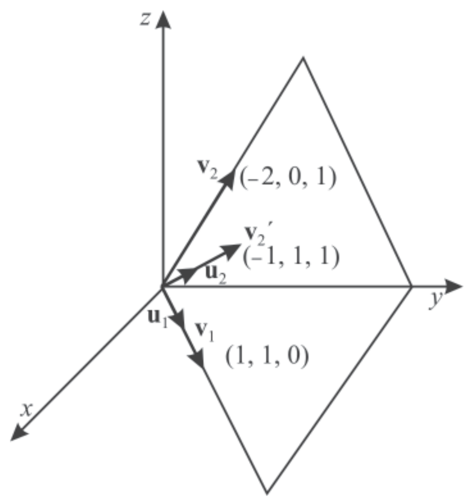


Figura 8.2

Los vectores $\mathbf{u}_1$ y $\mathbf{u}_2$ forman una base ortonormal para el plano generado por los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ (figura 8.2).

Calcular las coordenadas de un vector relativas a una base puede ser un problema relativamente difícil; ahora, si la base es ortonormal, esto es bastante sencillo, como se ve en el siguiente teorema.

Teorema 3

Sea $B = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ una base ortonormal de $\mathbb{R}^n$ y sea $\mathbf{v} \in \mathbb{R}^n$. Entonces:

$$\mathbf{v} = (\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2 + ... + (\mathbf{v}\cdot\mathbf{u}_n)\mathbf{u}_n.$$

Es decir, el vector **v** es la suma de las proyecciones ortogonales sobre los vectores de la base ortonormal de $\mathbb{R}^n$:

$$\mathbf{v} = \text{proy}_{R^n}\,\mathbf{v}.$$

Demostración

Ya que B es una base de $\mathbb{R}^n$, **v** se puede expresar de manera única como:

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_i\mathbf{u}_i + ... + c_n\mathbf{u}_n.$$

Multiplicando escalarmente por $\mathbf{u}_i$, $i = 1, ..., n$, a ambos lados de la igualdad, tenemos:

$$\begin{aligned}\mathbf{v}\cdot\mathbf{u}_i &= c_1\underbrace{(\mathbf{u}_1\cdot\mathbf{u}_i)}_{0} + c_2\underbrace{(\mathbf{u}_2\cdot\mathbf{u}_i)}_{0} + ... + c_i\underbrace{(\mathbf{u}_i\cdot\mathbf{u}_i)}_{1} + ... + c_n\underbrace{(\mathbf{u}_n\cdot\mathbf{u}_i)}_{0}\\ &= c_i.\end{aligned}$$

Como esto es válido para todo i, la demostración queda completa.

Ejemplo 3

Sea $B = \left\{\mathbf{u}_1 = \left(\frac{1}{\sqrt{5}},\ 0,\ \frac{2}{\sqrt{5}}\right),\ \mathbf{u}_2 = \left(\frac{-2}{\sqrt{5}},\ 0,\ \frac{1}{\sqrt{5}}\right),\ \mathbf{u}_3 = (0, 1, 0)\right\}$ una base ortonormal para $\mathbb{R}^3$. Expresemos el vector $\mathbf{v} = (2, -3, 1)$ como una combinación lineal de los vectores de B.

$$\begin{aligned}\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3 \text{ tal que: } & c_1 = \mathbf{v}\cdot\mathbf{u}_1 = 4/\sqrt{5},\\ & c_2 = \mathbf{v}\cdot\mathbf{u}_2 = -3/\sqrt{5},\\ & c_3 = \mathbf{v}\cdot\mathbf{u}_3 = -3.\end{aligned}$$

$$(2, -3, 1) = \frac{4}{\sqrt{5}}\left(\frac{1}{\sqrt{5}},\ 0,\ \frac{2}{\sqrt{5}}\right) - \frac{3}{\sqrt{5}}\left(\frac{-2}{\sqrt{5}},\ 0,\ \frac{1}{\sqrt{5}}\right) - 3(0, 1, 0).$$

El siguiente teorema establece la unicidad de la proyección ortogonal de **v** sobre el subespacio H.

Teorema 4

Sean H un subespacio de $\mathbb{R}^n$ y $\mathbf{v}$ en $\mathbb{R}^n$. Suponga que se tienen en H dos bases ortonormales $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m\}$ y $B_2 = \{\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_m\}$; entonces:

$$\begin{aligned}\text{proy}_H\,\mathbf{v} &= (\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2 + ... + (\mathbf{v}\cdot\mathbf{u}_m)\mathbf{u}_m\\ &= (\mathbf{v}\cdot\mathbf{t}_1)\mathbf{t}_1 + (\mathbf{v}\cdot\mathbf{t}_2)\mathbf{t}_2 + ... + (\mathbf{v}\cdot\mathbf{t}_m)\mathbf{t}_m.\end{aligned}$$

8.4 Complemento ortogonal

Definición 3

Sea H un subespacio de $\mathbb{R}^n$. Un vector **u** en $\mathbb{R}^n$ es ortogonal a H si es ortogonal a todo vector de H. El conjunto de todos los vectores de $\mathbb{R}^n$ que son ortogonales a H se llama *complemento ortogonal de H* en $\mathbb{R}^n$ y se denota $H^{\perp}$.

$$H^{\perp} = \{\mathbf{u}\in\mathbb{R}^n : \mathbf{u}\cdot\mathbf{h} = 0,\ \text{para todo } \mathbf{h}\in H\}.$$

Ejemplo 4

Sea $W=\{(x,y,z)/(x,y,z)=t(1,2,3),\ t\in\mathbb{R}\}$.

$W^{\perp}$ está formado por todos los vectores ortogonales a (1, 2, 3); se puede mostrar que $W^{\perp}$ es el plano que pasa por el origen con vector normal (1, 2, 3).

Teorema 5

Sea H un subespacio de $\mathbb{R}^n$; entonces:

a. $H^{\perp}$ es un subespacio de $\mathbb{R}^n$.

b. $H\cap H^{\perp}=\{\mathbf{0}\}$.

c. $\dim H^{\perp}=n-\dim H$.

Demostración

a. $H^{\perp}\neq\Phi$, $(0,\ 0,\dots,\ 0)\in H^{\perp}$ ya que $\forall\mathbf{h}\in H$

$$(0,\ 0,\dots,\ 0)\cdot\mathbf{h}=0.$$

Sean $\mathbf{x}_1,\mathbf{x}_2\in H^{\perp}$, $\mathbf{x}_1\cdot\mathbf{h}=0$ y $\mathbf{x}_2\cdot\mathbf{h}=0$, $\forall\mathbf{h}\in H$.
Luego:

$$(\mathbf{x}_1+\mathbf{x}_2)\cdot\mathbf{h}=\mathbf{x}_1\cdot\mathbf{h}+\mathbf{x}_2\cdot\mathbf{h}=0+0=0.$$

Ahora:

$$\alpha\mathbf{x}_1\cdot\mathbf{h}=\alpha(\mathbf{x}_1\cdot\mathbf{h})=\alpha 0=0,$$

lo que prueba que $H^{\perp}$ es un subespacio de $\mathbb{R}^n$.

b. $\{(0,\ 0,\dots,\ 0)\}\subset H\cap H^{\perp}$.

Probemos que $H\cap H^{\perp}\subset\{(0,\ 0,\dots,0)\}$.

Supongamos que $\mathbf{x}\in H\cap H^{\perp}$, $\mathbf{x}\in H\wedge\mathbf{x}\in H^{\perp}$; entonces $\mathbf{x}\cdot\mathbf{x}=0$.
Luego:

$$\mathbf{x}=(0,\ 0,\dots,\ 0)\ ;$$

por lo tanto:

$$H\cap H^{\perp}=\{(0,\ 0,\dots,\ 0)\}\ .$$

c. Sea $\{\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_k\}$ una base ortonormal de H. Esta base puede ampliarse hasta formar B, una base para $\mathbb{R}^n$, $B=\{\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_k,\mathbf{v}_{k+1},\dots,\mathbf{v}_n\}$. Mediante el proceso de Gram-Schmidt, B se puede transformar en una base ortonormal de $\mathbb{R}^n$.

$$B_1 = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k, \mathbf{u}_{k+1}, ..., \mathbf{u}_n\}.$$

Debemos demostrar que $\{\mathbf{u}_{k+1}, ..., \mathbf{u}_n\}$ es una base para $H^\perp$.

Como los vectores son LI, debe mostrarse que generan $H^\perp$.

Sea $\mathbf{x} \in H^\perp$; entonces $\mathbf{x} \in \mathbb{R}^n$ y

$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{u}_1)\mathbf{u}_1 + ... + (\mathbf{x} \cdot \mathbf{u}_k)\mathbf{u}_k + (\mathbf{x} \cdot \mathbf{u}_{k+1})\mathbf{u}_{k+1} + ... + (\mathbf{x} \cdot \mathbf{u}_n)\mathbf{u}_n.$$

Como los vectores $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k$ están en H, $\mathbf{x} \cdot \mathbf{u}_i = 0$, $i = 1, ..., k$, entonces:

$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{u}_{k+1})\mathbf{u}_{k+1} + ... + (\mathbf{x} \cdot \mathbf{u}_n)\mathbf{u}_n.$$

Luego $\{\mathbf{u}_{k+1}, ..., \mathbf{u}_n\}$ es una base para $H^\perp$ y $\dim H^\perp = n - k$.

Cualquier vector de $\mathbb{R}^n$ se podrá expresar de manera única como suma de un vector de H con un vector de $H^\perp$.

Teorema 6: Teorema de la proyección

Sean H un subespacio de $\mathbb{R}^n$ y $\mathbf{v} \in \mathbb{R}^n$.

Entonces existe un par único de vectores **p** y **q** tales que:

$$\mathbf{p} \in H \text{ y } \mathbf{q} \in H^\perp \text{ y } \mathbf{v} = \mathbf{p} + \mathbf{q},$$

donde $\mathbf{p} = \text{proy}_H \mathbf{v}$ y $\mathbf{q} = \text{proy}_{H^\perp} \mathbf{v}$.

Ejemplo 5

En el plano del ejemplo 2 determinamos la base ortonormal:

$$B = \left\{ \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \right\}$$

y para $\mathbf{a} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$ encontramos $\text{proy}_W \mathbf{a} = \begin{pmatrix} 13/6 \\ 5/6 \\ -2/3 \end{pmatrix}$.

Podemos expresar **a** como la suma de dos vectores **p** y **q** tales que:

$$\mathbf{p} = \text{proy}_W \mathbf{a} \text{ y } \mathbf{q} = \text{proy}_{W^\perp} \mathbf{a},$$

$$\mathbf{a} = \mathbf{p} + \mathbf{q} \Rightarrow \mathbf{q} = \mathbf{a} - \mathbf{p} = \mathbf{a} - \text{proy}_W\ \mathbf{a},$$

$$\mathbf{q} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} - \begin{pmatrix} 13/6 \\ 5/6 \\ -2/3 \end{pmatrix} = \begin{pmatrix} -1/6 \\ 1/6 \\ -3/3 \end{pmatrix} = \text{proy}_{W^\perp}\ \mathbf{a},$$

$$W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x - y + 2z = 0 \right\}.$$

$\dim W = 2$; como $\dim W + \dim W^\perp = 3$, entonces $\dim W^\perp = 1$. Luego el vector $\mathbf{q}$ puede ser una base para $W^\perp$ y $\dfrac{\mathbf{q}}{|\mathbf{q}|}$ una base ortonormal para $W^\perp$.

$$|\mathbf{q}| = \frac{1}{\sqrt{6}},$$

entonces:

$$\frac{\mathbf{q}}{|\mathbf{q}|} = \sqrt{6} \begin{pmatrix} -1/6 \\ 1/6 \\ -1/3 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix}.$$

$$\text{Base ortonormal de } W^\perp = \left\{ \begin{pmatrix} -1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix} \right\}.$$

No es necesario averiguar el valor de **q** para determinar la base de $W^\perp$. De la ecuación de W sabemos que (1, –1, 2) es un vector normal a W; por lo tanto, pertenece a $W^\perp$.

Luego, una base de $W^\perp$ puede ser $\left\{ \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \right\}$ y una base ortonormal de $W = \left\{ \begin{pmatrix} 1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix} \right\}$.

Teorema 7: Teorema de aproximación de la norma

Sea H un subespacio de $\mathbb{R}^n$. Entonces para cualquier vector $\mathbf{v} \in \mathbb{R}^n$, el vector de H más cercano a **v** es $\text{proy}_H\ \mathbf{v}$. Es decir, si **h** es cualquier vector de H, $|\mathbf{v} - \mathbf{h}|$ es mínima cuando $\mathbf{h} = \text{proy}_H\ \mathbf{v}$.

Demostración

Sea **h** un vector cualquiera de H. Entonces:

$$\mathbf{v} - \mathbf{h} = (\mathbf{v} - \text{proy}_H\ \mathbf{v}) + (\text{proy}_H\ \mathbf{v} - \mathbf{h}).$$

Ahora:

$$(\mathbf{v}-\text{proy}_H\ \mathbf{v})\in H^{\perp}\quad \text{y}\quad (\text{proy}_H\ \mathbf{v}-\mathbf{h})\in H,$$

y entonces:

$$\begin{aligned}
&(\mathbf{v}-\text{proy}_H\ \mathbf{v})\cdot(\text{proy}_H\ \mathbf{v}-\mathbf{h})=0.\\
&|\mathbf{v}-\mathbf{h}|^2=\big((\mathbf{v}-\text{proy}_H\ \mathbf{v})+(\text{proy}_H\ \mathbf{v}-\mathbf{h})\big)\cdot\big((\mathbf{v}-\text{proy}_H\ \mathbf{v})+(\text{proy}_H\ \mathbf{v}-\mathbf{h})\big)\\
&\qquad\quad=|\mathbf{v}-\text{proy}_H\ \mathbf{v}|^2+|\text{proy}_H\ \mathbf{v}-\mathbf{h}|^2.
\end{aligned}$$

Si $\mathbf{h}\neq\text{proy}_H\ \mathbf{v}$,

$$|\text{proy}_H\ \mathbf{v}-\mathbf{h}|^2>0\quad \text{y}\quad |\mathbf{v}-\mathbf{h}|^2>|\mathbf{v}-\text{proy}_H\ \mathbf{v}|^2\ ;$$

por lo tanto:

$$|\mathbf{v}-\mathbf{h}|>|\mathbf{v}-\text{proy}_H\ \mathbf{v}|.$$

Así que $\text{proy}_H\ \mathbf{v}$ es el vector que minimiza $|\mathbf{v}-\mathbf{h}|$.

En el ejemplo 5, $\mathbf{p}=\text{proy}_W\ \mathbf{a}=\begin{pmatrix}13/6\\5/6\\-2/3\end{pmatrix}$ es el vector en W más cercano a $\mathbf{a}=\begin{pmatrix}2\\1\\-1\end{pmatrix}$.

8.5 Desigualdad de Cauchy-Schwarz en $\mathbb{R}^n$

Finalizamos el capítulo con un resultado importante relativo a las magnitudes de los vectores, la desigualdad de Cauchy-Schwarz en $\mathbb{R}^n$, a partir de la cual puede deducirse la desigualdad triangular en $\mathbb{R}^n$.

Teorema 8: Desigualdad de Cauchy-Schwarz en $\mathbb{R}^n$

Sean $\mathbf{a}$ y $\mathbf{b}$ vectores en $\mathbb{R}^n$. Entonces:

i. $|\mathbf{a}\cdot\mathbf{b}|\leq|\mathbf{a}||\mathbf{b}|$.

ii. $|\mathbf{a}\cdot\mathbf{b}|=|\mathbf{a}||\mathbf{b}|$ si y sólo si $\mathbf{a}=\mathbf{0}$ o $\mathbf{b}=\lambda\mathbf{a}$ para algún real λ.

Demostración

i. Tomemos dos vectores $\mathbf{a}$ y $\mathbf{b}$ de $\mathbb{R}^n$ tales que uno no es múltiplo escalar del otro, o sea que están sobre rectas en distintas direcciones (figura 8.3).

La proyección ortogonal de **b** sobre **a** está dada por $\text{proy } \mathbf{b}/\mathbf{a} = \left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|^2}\right)\mathbf{a}.$

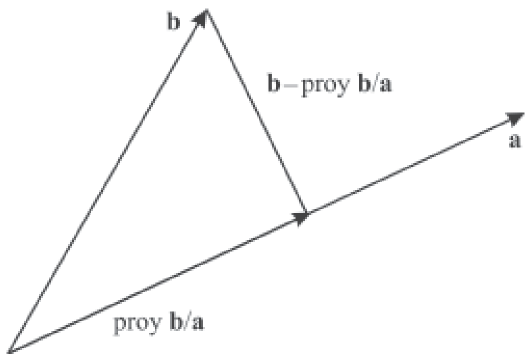


Figura 8.3

La distancia (al cuadrado) del punto **b** a la recta que lleva la dirección de **a** es:

$$\left|\mathbf{b}-\left(\frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|^2}\right)\mathbf{a}\right|^2 = |\mathbf{b}|^2-\frac{2(\mathbf{a}\cdot\mathbf{b})^2}{|\mathbf{a}|^2}+\left(\frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|^2}\right)^2|\mathbf{a}|^2$$

$$=\frac{|\mathbf{b}|^2|\mathbf{a}|^2-(\mathbf{a}\cdot\mathbf{b})^2}{|\mathbf{a}|^2}.$$

Como esta distancia es mayor que cero, entonces el numerador en la expresión anterior debe ser positivo, luego:

$$|\mathbf{b}|^2|\mathbf{a}|^2-(\mathbf{a}\cdot\mathbf{b})^2>0,$$

$$|\mathbf{b}|^2|\mathbf{a}|^2>(\mathbf{a}\cdot\mathbf{b})^2.$$

Extrayendo raíz cuadrada a ambos lados, tenemos que:

$$|\mathbf{b}||\mathbf{a}|>|\mathbf{a}\cdot\mathbf{b}|.$$

Cuando $\mathbf{a}$ y $\mathbf{b}$ tienen la misma dirección, o sea que $\mathbf{b}=\lambda\mathbf{a}$ para algún escalar λ, la distancia calculada será igual a cero; por lo tanto, se tendrá que:

$$|\mathbf{a}\cdot\mathbf{b}|=|\mathbf{a}||\mathbf{b}|.$$

Luego, en general, tenemos que $|\mathbf{a}\cdot\mathbf{b}|\le|\mathbf{a}||\mathbf{b}|$.

Ejercicios del capítulo 2 (módulo 8)

1. Verifique que las siguientes son bases ortogonales para $\mathbb{R}^3$. Obtenga bases ortonormales a partir de ellas.

   a. $\{(0, 0, 1), (-1, 1, 0), (1, 1, 0)\}$.

   b. $\{(0, 1, -1), (1, -1/2, -1/2), (1, 1, 1)\}$.

2. Utilice el proceso de Gram-Schmidt para transformar la base de un subespacio de $\mathbb{R}^3$, dada por $\{(1, 1, -1), (0, 1, -1)\}$, en una base ortonormal.

3. Igual que en el ejercicio 2, con la base $\{(1, -1, 0), (2, 0, 1)\}$.

4. Utilice el proceso de Gram-Schmidt para determinar una base ortonormal para el subespacio de $\mathbb{R}^4$ con base $\{(1, -1, 0, 1), (0, 0, 1, 0), (2, 0, 0, -1)\}$.

5. Sea $S = \left\{\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right), \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)\right\}$ una base ortonormal de $\mathbb{R}^3$. Escriba el vector $(1, 2, -1)$ en esta base.

6. Sea W el subespacio de $\mathbb{R}^3$ con base ortonormal $\left\{(0, 1, 0), \left(\frac{1}{\sqrt{5}}, 0, \frac{2}{\sqrt{5}}\right)\right\}$. Escriba el vector $\mathbf{v} = (1, 2, -1)$ como $\mathbf{p} + \mathbf{q}$ con $\mathbf{p} \in W$ y $\mathbf{q} \in W^{\perp}$.

7. Determine la distancia del punto $(2, 3, -1)$ al plano $3x - 2y + z = 0$. (Sugerencia: encuentre la proyección ortogonal de $(2, 3, -1)$ sobre el plano).

8. Construya una base ortonormal para el espacio solución del sistema homogéneo:

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

9. Encuentre una base ortonormal en $\mathbb{R}^3$ que incluya los vectores $\mathbf{u_1} = \left(\frac{1}{\sqrt{5}}, 0, \frac{2}{\sqrt{5}}\right)$ y $\mathbf{u_2} = (0, 1, 0)$.

En los ejercicios 10 a 12 se dan un subespacio H y un vector $\mathbf{v}$.

a. Calcule $\text{proy}_H\ \mathbf{v}$.

b. Encuentre una base ortonormal para $H^{\perp}$.

c. Escriba $\mathbf{v} = \mathbf{p} + \mathbf{q}$ con $\mathbf{p} \in H$ y $\mathbf{q} \in H^{\perp}$.

10. $H = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : 2x - y = 0 \right\}$; $\mathbf{v} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

11. $H = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : 3x - 2y + 6z = 0 \right\}$; $\mathbf{v} = \begin{bmatrix} -3 \\ 1 \\ 4 \end{bmatrix}$.

12. $H = \left\{ \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \in \mathbb{R}^4 : x = 2y,\ w = -y \right\}$; $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix}$.

13. Si $\mathbf{a}$ y $\mathbf{b}$ son vectores de $\mathbb{R}^n$ tales que $\mathbf{a} = \mathbf{0}$ o $\mathbf{b} = \lambda\mathbf{a}$ para $\lambda \in \mathbb{R}$, demuestre que $|\mathbf{a} \cdot \mathbf{b}| = |\mathbf{a}||\mathbf{b}|$.

14. Usando la desigualdad de Cauchy-Schwarz, pruebe la desigualdad triangular $|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|$.

Módulo 9

Método de aproximación por mínimos cuadrados

Un avión que toma fotografías infrarrojas utiliza un detector que registra la imagen electrónicamente; esta imagen sufre distorsiones que deben corregirse y así obtener las coordenadas reales; para ello se utiliza el método de mínimos cuadrados.

Introducción

Muchos experimentos relacionados con ciencias físicas, biológicas o sociales se proponen encontrar la relación existente entre las variables presentes en un determinado fenómeno, por medio de una ley matemática.

En estos procesos se trata de ajustar una curva a los diversos puntos obtenidos experimentalmente. En general, el resultado obtenido es un sistema de ecuaciones $A\mathbf{x} = \mathbf{y}$ inconsistente. El problema, entonces, es encontrar un $\overline{\mathbf{x}}$ en $\mathbb{R}^n$ tal que $A\overline{\mathbf{x}}$ sea tan cercano a $\mathbf{y}$ como sea posible.

Objetivos

1. Establecer la mejor aproximación de un vector de $\mathbb{R}^n$ a un subespacio del mismo, aplicando el método de aproximación por mínimos cuadrados.
2. Encontrar el error cometido al hacer una aproximación con el método de mínimos cuadrados para poder valorar la confiabilidad de los resultados obtenidos.
3. Valorar la utilidad de la temática abordada, transfiriéndola y aplicándola en la resolución de situaciones problemáticas que se presentan en la vida diaria.

Preguntas básicas

1. ¿En qué consiste el método de aproximación por mínimos cuadrados?
2. ¿Cuándo se aplica este método?
3. ¿Por qué cuando se da $\mathbf{y} = A\mathbf{x},$ decimos que $\mathbf{y} \in C_A$?
4. ¿Por qué el vector **x** que minimiza el error cometido es tal que $A\mathbf{x}$ es igual a la proyección ortogonal de **y** sobre C_A?
5. ¿Cómo se determina el error cometido en el cálculo?

Vea el módulo 9 del programa de televisión *Álgebra lineal*.

Contenidos

9.1 Aproximación por una recta
9.2 Aproximación cuadrática

9.1 Aproximación por una recta

Supongamos que se busca la recta $y = b + mx$ que mejor se ajusta a los n datos $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ (figura 9.1).

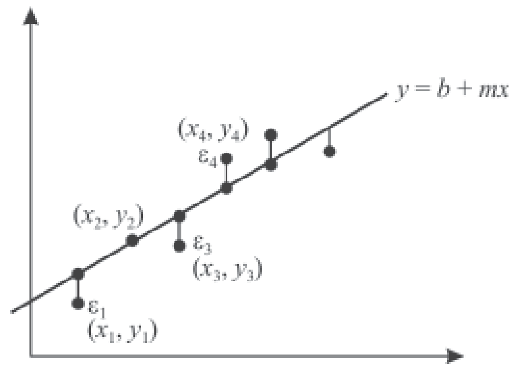


Figura 9.1

La gráfica ilustra la ubicación de algunos puntos obtenidos del experimento y sus distancias (errores ε_i) a los respectivos puntos sobre la recta $y = b + mx$. El problema puede plantearse así: encontrar m y b de la recta, tal que el error cometido sea mínimo.

Una forma de minimizar los errores es hacer mínima la suma de los cuadrados de los errores $(\varepsilon_1^2 + \varepsilon_2^2 + ... + \varepsilon_n^2)$.

$$\begin{aligned}
\varepsilon_1 &= y_1 - (b + mx_1) \\
\varepsilon_2 &= y_2 - (b + mx_2) \\
\vdots \quad & \quad \vdots \\
\varepsilon_n &= y_n - (b + mx_n).
\end{aligned}$$

Un enunciado más preciso del problema es: encuentre m y b tales que:

$$(y_1 - (b + mx_1))^2 + (y_2 - (b + mx_2))^2 + ... + (y_n - (b + mx_n))^2 \text{ sea mínima.}$$

Desarrollaremos un método matricial para encontrar esta aproximación de mínimos cuadrados.

Si los puntos $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ estuvieran sobre la recta $y = b + mx$ se cumpliría que:

$$\begin{aligned}
y_1 &= (b + mx_1) \\
y_2 &= (b + mx_2) \qquad \text{o} \quad \mathbf{Y} = A\mathbf{X} \\
\vdots \quad & \quad \vdots \\
y_n &= (b + mx_n)
\end{aligned}$$

con $\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$ y $\mathbf{X} = \begin{pmatrix} b \\ m \end{pmatrix}.$

Si los puntos no están todos sobre la recta, el sistema $\mathbf{Y} = A\mathbf{X}$ es inconsistente, es decir, $\mathbf{Y} - A\mathbf{X} \neq \mathbf{0}$, y el problema será:

encontrar un vector $\mathbf{X}$ tal que $|\mathbf{Y} - A\mathbf{X}|$ sea mínima.

$(\mathbf{Y} - A\mathbf{X})$ es un vector de $\mathbb{R}^n$ cuyas componentes son los errores cometidos en la obtención de los *n* datos experimentales.

$$\mathbf{Y} - A\mathbf{X} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}, \quad |\mathbf{Y} - A\mathbf{X}| = \sqrt{\varepsilon_1^2 + \varepsilon_2^2 + \ldots + \varepsilon_n^2}.$$

Así que hacer mínima $|\mathbf{Y} - A\mathbf{X}|$ equivale a minimizar $\varepsilon_1^2 + \varepsilon_2^2 + \ldots + \varepsilon_n^2$.

Si $\mathbf{Y} = A\mathbf{X}$, entonces **Y** es una combinación lineal de las columnas de *A*, o sea que $\mathbf{Y} \in C_A$; como $\mathbf{Y} = A\mathbf{X}$ es inconsistente, $\mathbf{Y} \notin C_A$. Por el teorema de aproximación de la norma en $\mathbb{R}^n$, $\mathbf{Y} - A\mathbf{X}$ es mínimo cuando:

$$A\mathbf{X} = \operatorname{proy}_{C_A} \mathbf{Y}.$$

Sea $\overline{\mathbf{X}}$ el vector, tal que $A\overline{\mathbf{X}} = \operatorname{proy}_{C_A} \mathbf{Y}$.

Podemos hacer una representación geométrica en $\mathbb{R}^3$.

El espacio columna de *A*, C_A, puede ser un plano o una recta que pasa por el origen. Representemos C_A como un plano que pasa por el origen (figura 9.2).

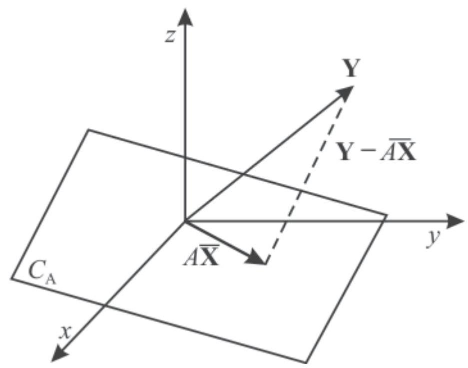


Figura 9.2

$\mathbf{Y} - A\overline{\mathbf{X}}$ es un vector ortogonal a C_A, o sea que para todo $A\mathbf{X} \in C_A$ se tiene:

$$A\mathbf{X} \cdot (\mathbf{Y} - A\overline{\mathbf{X}}) = \mathbf{0},$$
$$(A\mathbf{X})^T (\mathbf{Y} - A\overline{\mathbf{X}}) = \mathbf{0},$$
$$\mathbf{X}^T A^T (\mathbf{Y} - A\overline{\mathbf{X}}) = \mathbf{0},$$
$$\mathbf{X}^T (A^T \mathbf{Y} - A^T A\overline{\mathbf{X}}) = \mathbf{0} \qquad \forall \mathbf{X} \in \mathbb{R}^2.$$

Luego $A^T \mathbf{Y} - A^T A\overline{\mathbf{X}} = \mathbf{0}$,
$A^T A\overline{\mathbf{X}} = A^T \mathbf{Y}$.

Esta ecuación matricial representa dos ecuaciones con dos incógnitas, que frecuentemente se denominan *ecuaciones normales.*

Si $A^T A$ es invertible (se puede demostrar que esto sucede cuando los n datos no son colineales), $\overline{\mathbf{X}}$ tiene solución única dada por:

$$\overline{\mathbf{X}} = (A^T A)^{-1} A^T \mathbf{Y}.$$

Ejemplo 1

Encuentre la recta que da el mejor ajuste para los datos (2, 1), (3, 2), (4, 3), (5, 2).

En este caso:

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 2 \end{bmatrix}, \quad \mathbf{X} = \begin{pmatrix} b \\ m \end{pmatrix}$$

$$A^T A\overline{\mathbf{X}} = A^T \mathbf{Y}$$

$$=\begin{bmatrix}1&1&1&1\\2&3&4&5\end{bmatrix}\begin{bmatrix}1&2\\1&3\\1&4\\1&5\end{bmatrix}\overline{\mathbf{X}}=\begin{bmatrix}1&1&1&1\\2&3&4&5\end{bmatrix}\begin{bmatrix}1\\2\\3\\2\end{bmatrix}.$$

$$\begin{bmatrix}4&14\\14&54\end{bmatrix}\begin{bmatrix}b\\m\end{bmatrix}=\begin{bmatrix}8\\30\end{bmatrix}$$

$$\left[\begin{array}{cc|c}4&14&8\\14&54&30\end{array}\right]\xrightarrow{\text{operaciones elementales}}\left[\begin{array}{cc|c}1&0&3/5\\0&1&2/5\end{array}\right].$$

Ecuación de la recta: $y=\frac{3}{5}+\frac{2}{5}x.$

El error cometido está dado por (figura 9.3):

$$\left|\mathbf{Y}-A\overline{\mathbf{X}}\right|=\left|\begin{pmatrix}1\\2\\3\\2\end{pmatrix}-\begin{bmatrix}1&2\\1&3\\1&4\\1&5\end{bmatrix}\begin{bmatrix}3/5\\2/5\end{bmatrix}\right|=\left|\begin{pmatrix}1\\2\\3\\2\end{pmatrix}-\begin{pmatrix}7/5\\9/5\\11/5\\13/5\end{pmatrix}\right|$$

$$\left|\mathbf{Y}-A\overline{\mathbf{X}}\right|=\left|\begin{pmatrix}-2/5\\1/5\\4/5\\-3/5\end{pmatrix}\right|=\sqrt{\frac{4+1+16+9}{25}}=\sqrt{\frac{6}{5}}\approx 1.1.$$

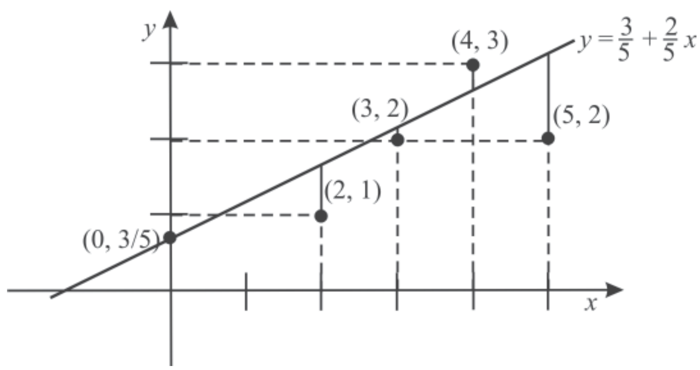


Figura 9.3

9.2 Aproximación cuadrática

Si lo que se busca es una curva cuadrática $y=a+bx+cx^2$ que sea el mejor ajuste en el sentido de mínimos cuadrados a n datos $(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)$, el procedimiento que se debe seguir es exactamente análogo al método desarrollado

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Crecimiento de población».

anteriormente para el ajuste lineal (figura 9.4).

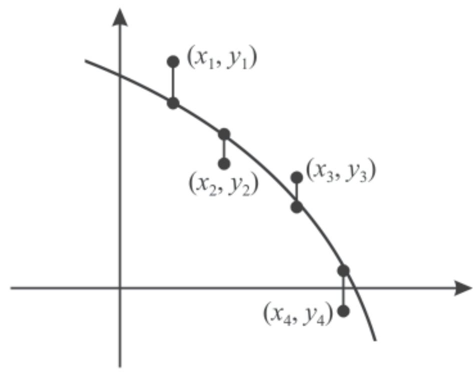


Figura 9.4

Si los n puntos estuvieran sobre la parábola $y = a + bx + cx^2$, la ecuación se verificaría así:

$$\begin{aligned} y_1 &= a + bx_1 + cx_1^2 \\ y_2 &= a + bx_2 + cx_2^2 \\ \vdots & \qquad \vdots \\ y_n &= a + bx_n + cx_n^2 \end{aligned} \qquad \text{o} \qquad \mathbf{Y} = A\mathbf{X}$$

con:

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \qquad A = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} \quad \text{y} \quad \mathbf{X} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Como los puntos no satisfacen las ecuaciones, el sistema $\mathbf{Y} - A\mathbf{X}$ es inconsistente; por lo tanto, el problema será encontrar un vector $\overline{\mathbf{X}}$ en $\mathbb{R}^3$ tal que $\left|\mathbf{Y} - A\overline{\mathbf{X}}\right|$ tenga un valor mínimo; luego $A\overline{\mathbf{X}} = \text{proy}_{C_A}\,\mathbf{Y}$. Si llamamos $\overline{\mathbf{X}}$ el vector minimizador, entonces:

$$A^T\mathbf{Y} = A^T A\overline{\mathbf{X}}.$$

Ejemplo 2

El método de ajuste cuadrático se puede usar para hacer estimaciones de las constantes físicas. Veámoslo:

Un cubito se desliza sin fricción sobre el plano inclinado de la figura 9.5 y parte del punto 0.

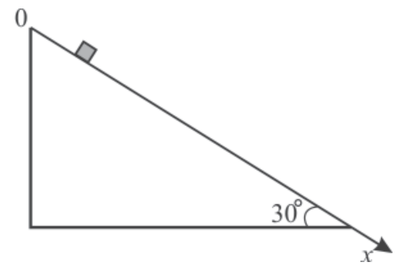


Figura 9.5

Su ley de movimiento está dada por: $x(t) = v_0 t + \frac{g}{4} t^2$.

Se tienen las siguientes mediciones:

$$\begin{array}{ll} t = 1\,\text{s} & x = 12.4\,\text{m} \\ t = 2\,\text{s} & x = 29.8\,\text{m} \\ t = 3\,\text{s} & x = 52.0\,\text{m} \\ t = 4\,\text{s} & x = 79.21\,\text{m} \end{array}$$

A partir de estos datos podemos hacer estimaciones para los valores de v_0 y g; además, calcular el error cometido.

Sean:

$$\mathbf{Y} = \begin{pmatrix} 12.4 \\ 29.8 \\ 52.0 \\ 79.21 \end{pmatrix}, \qquad A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \\ 3 & 9 \\ 4 & 16 \end{bmatrix}, \qquad \mathbf{X} = \begin{pmatrix} v_0 \\ g/4 \end{pmatrix}.$$

$$A^T \mathbf{Y} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \end{bmatrix} \begin{bmatrix} 12.4 \\ 29.8 \\ 52.0 \\ 79.21 \end{bmatrix} = \begin{bmatrix} 544.84 \\ 1866.96 \end{bmatrix},$$

$$A^T A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 4 \\ 3 & 9 \\ 4 & 16 \end{bmatrix} = \begin{bmatrix} 30 & 100 \\ 100 & 354 \end{bmatrix}.$$

Resolvemos ahora el sistema $\left(A^T A\right)\overline{\mathbf{X}} = A^T \mathbf{Y}$.

$$\left[\begin{array}{cc|c} 30 & 100 & 544.84 \\ 100 & 354 & 1866.96 \end{array}\right] \xrightarrow{\text{operaciones elementales}} \left[\begin{array}{cc|c} 1 & 0 & 9.96 \\ 0 & 1 & 2.46 \end{array}\right].$$

En los ejercicios 1 a 4 determine la recta de mínimos cuadrados para los datos dados.

$$\overline{\mathbf{X}} = \begin{pmatrix} 9.96 \\ 2.46 \end{pmatrix} = \begin{pmatrix} v_0 \\ g/4 \end{pmatrix}.$$

$$g \approx 9.84\,\text{m/s}^2.$$

El error cometido está dado por $\left|\mathbf{Y} - A\overline{\mathbf{X}}\right|$.

$$A\overline{\mathbf{X}} = \begin{bmatrix} 1 & 1 \\ 2 & 4 \\ 3 & 9 \\ 4 & 16 \end{bmatrix} \begin{bmatrix} 9.96 \\ 2.46 \end{bmatrix} = \begin{bmatrix} 12.42 \\ 29.76 \\ 52.02 \\ 79.2 \end{bmatrix}.$$

$$\varepsilon = \begin{vmatrix} 12.4 - 12.42 \\ 29.8 - 29.76 \\ 52.0 - 52.02 \\ 79.21 - 79.20 \end{vmatrix} = \begin{vmatrix} -0.02 \\ 0.04 \\ 0.02 \\ 0.01 \end{vmatrix} = \sqrt{(0.02)^2 + (0.04)^2 + (0.02) + (0.01)^2}$$

$$= 0.05.$$

Ejercicios del capítulo 2 (módulo 9)

En los ejercicios 1a 4 determine la recta de mínimos cuadrados para los datos dados.

1. $(0,\ 0),\ (1,1),\ (3,\ 2),\ (4,\ 5).$
2. $(1,\ 4),\ (-2,\ 5),\ (3,\ -1),\ (4,\ 1).$
3. $(-2,\ -2),\ (-1,\ 0),\ (0,\ -2),\ (1,\ 0).$
4. $(0,\ 2),\ (1,\ 2),\ (2,\ 0).$

En los ejercicios 5 y 6 determine el polinomio cuadrático de mínimos cuadrados para los puntos dados.

5. $(-7,\ 3),\ (2,\ 8),\ (1,\ 5).$
6. $(0,\ 3.2),\ (0.5,\ 1.6),\ (1,\ 2),\ (2,\ -0.4),\ (2.5,\ -0.8),\ (3,\ -1.6),\ (4,\ 0.3),\ (5,\ 2.2).$
7. Sean $A = \begin{bmatrix} 1 & -2 \\ -1 & 2 \\ 0 & 3 \\ 2 & 5 \end{bmatrix}$ e $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \\ -4 \\ 2 \end{bmatrix}.$

   a. Encuentre una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{y}$.
   b. Calcule el error de mínimos cuadrados asociado con la solución encontrada en *a*.

8. Encuentre el mejor ajuste cuadrático para los datos del ejemplo 2. ¿Qué tipo de ajuste ocasiona el menor error?
9. Se tienen dos magnitudes relacionadas cuadráticamente *x* y *y*, es decir, existen constantes *a*, *b* y *c* tales que $y = a + bx + cx^2$. En las mediciones experimentales de estas magnitudes se obtuvieron los siguientes datos:

x	0	0.5	1	2	2.5	3	4	5
y	3.2	1.6	2	–0.4	–0.8	–1.6	0.3	2.2

   Encuentre la parábola que mejor ajuste estos datos.

10. Una cinta elástica se estira a longitudes de $L = 5$, 6 y 7 pies por la aplicación de fuerzas $F = 1$, 2 y 4 kg respectivamente. Asumiendo la ley de Hooke, $L = a + bF$, encuentre la longitud normal a de la cinta, por el método de mínimos cuadrados.

Módulo 10

Espacios con producto interno

Una aplicación interesante de los espacios con producto interno es la aproximación mediante series de Fourier a funciones continuas. La gráfica ilustra aproximaciones de Fourier de órdenes 3 y 4 a la función $f(t) = t$.

Introducción

En el mundo cotidiano, euclidiano tridimensional, los conceptos de la geometría euclidiana rigen nuestra manera de ver el mundo. Para todos es obvio que la distancia más corta entre dos puntos es el segmento de recta que une esos dos puntos; sin embargo, si nos encontráramos en un taxi, en un sitio de una ciudad, y quisiéramos ir a otro lugar de la misma, nuestro taxista no mediría la distancia 'a vuelo de pájaro', como corresponde a la versión euclidiana de distancia. Si nos permitiéramos pensar en distancia de una manera más flexible, abriríamos las puertas para concebir una 'distancia' entre polinomios, funciones, matrices y muchos otros objetos que existen en álgebra lineal. Es precisamente en estas situaciones donde se requiere la operación denominada *producto interno*.

Objetivos

1. Determinar magnitudes de vectores en espacios vectoriales donde hay definido un producto interno.
2. Construir bases ortonormales en espacios vectoriales con producto interno.

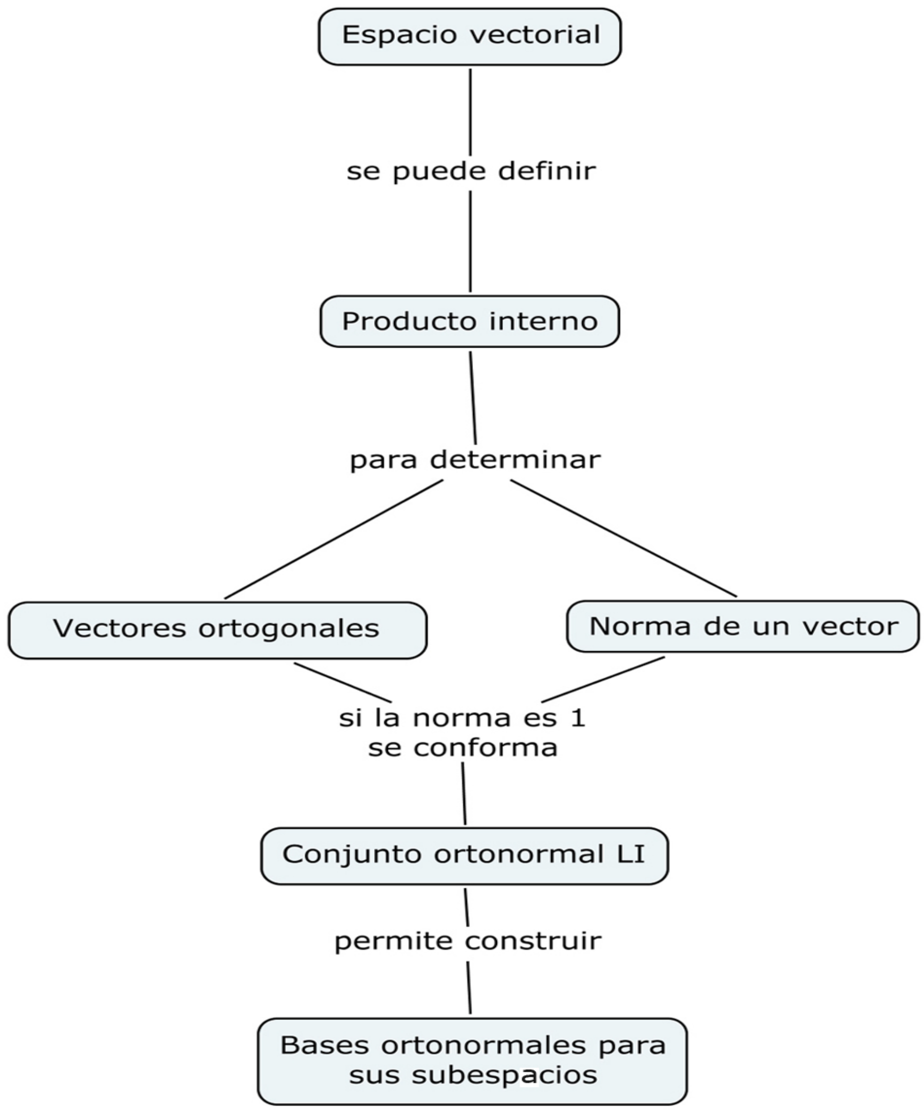
Preguntas básicas

1. ¿Qué es un producto interno?
2. ¿Por qué es importante definir un producto interno en un espacio vectorial?
3. ¿Cómo se construye una base ortonormal en un subespacio vectorial con producto interno?

Contenidos

10.1 Producto interno
10.2 Norma y ortogonalidad en un espacio vectorial con producto interno
10.3 Última forma del problema de la base
10.4 Proyección ortogonal en espacios vectoriales con producto interno

Vea el módulo 10 del programa de televisión *Álgebra lineal*.



Mapa 6: módulo 10

10.1 Producto interno

Definición 1

Sea V un espacio vectorial real o complejo.

Un producto interno sobre V es una función tal que a todo par de vectores $\mathbf{x}$ y $\mathbf{y}$ de V asocia un único número real o complejo $(\mathbf{x}, \mathbf{y})$ que satisface las siguientes propiedades:

Si $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ están en V y α es un escalar, entonces:

i. $(\mathbf{x}, \mathbf{x}) \geq 0$.

ii. $(\mathbf{x}, \mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$.

iii. $(\mathbf{x}, \mathbf{y} + \mathbf{z}) = (\mathbf{x}, \mathbf{y}) + (\mathbf{x}, \mathbf{z})$.

iv. $(\mathbf{x} + \mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z})$.

v. $(\mathbf{x}, \mathbf{y}) = (\overline{\mathbf{y}, \mathbf{x}})$.

vi. $(\alpha\mathbf{x}, \mathbf{y}) = \alpha(\mathbf{x}, \mathbf{y})$.

vii. $(\mathbf{x}, \alpha\mathbf{y}) = \overline{\alpha}(\mathbf{x}, \mathbf{y})$.

Del espacio vectorial V se dice que es un *espacio con producto interno*.

Nota: en las condiciones *v* y *vii* la barra indica conjugación compleja.

Si $(\mathbf{x}, \mathbf{y})$ es un número real, $(\overline{\mathbf{x}, \mathbf{y}}) = (\mathbf{x}, \mathbf{y})$.

Ejemplo 1

En $\mathbb{R}^n$ el producto escalar entre *n*-tuplas es un producto interno. Veamos.

Sean: $\mathbf{x} = (x_1, x_2, ..., x_n)$ e $\mathbf{y} = (y_1, y_2, ..., y_n)$.

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + ... + x_n y_n .$$

El producto escalar satisface la siguientes condiciones:

i. $\mathbf{x} \cdot \mathbf{x} \geq 0$.

ii. $\mathbf{x} \cdot \mathbf{x} = 0$ si y sólo si $\mathbf{x} = \mathbf{0}$.

iii. $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$.

iv. $(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$.

v. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$.

vi. $\alpha\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot \alpha\mathbf{y} = \alpha(\mathbf{x} \cdot \mathbf{y})$.

Las propiedades *i* a *iv* del producto escalar corresponden a las propiedades del producto interno.

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Espacios con producto interno».

Escuche la biografía de *Jean-Baptiste Joseph Fourier* en su multimedia de *Álgebra lineal*.

La propiedad *v* del producto escalar se convierte en propiedad conmutativa, ya que $\overline{(\mathbf{y}\cdot\mathbf{x})} = (\mathbf{y}\cdot\mathbf{x})$ debido a que el producto escalar es un número real.

Del mismo modo, la propiedad *vii* del producto interno queda integrada con la propiedad *vi*, ya que los escalares son números reales y $\overline{\alpha}(\mathbf{x}\cdot\mathbf{y}) = \alpha(\mathbf{x}\cdot\mathbf{y})$.

Ejemplo 2

En $\mathbb{C}^n$, espacio vectorial de *n*-tuplas ordenadas de números complejos, puede definirse un producto interno como $(\mathbf{x},\mathbf{y}) = x_1\overline{y}_1 + x_2\overline{y}_2 + ... + x_n\overline{y}_n$, con:

$$\mathbf{x} = (x_1, x_2, ..., x_n) \text{ e } \mathbf{y} = (y_1, y_2, ..., y_n)\,.$$

i. $(\mathbf{x},\mathbf{x}) \geq 0$, ya que:

$$(\mathbf{x},\mathbf{x}) = x_1\overline{x}_1 + x_2\overline{x}_2 + ... + x_j\overline{x}_j + ... + x_n\overline{x}_n$$

$$\Downarrow$$

$$\begin{aligned} &(a_j + b_j i)\,(a_j - b_j i) \\ &a_j^2 - b_j^2 i^2 \quad (i^2 = -1) \\ &a_j^2 + b_j^2 = \left|x_j\right|^2 \end{aligned}$$

$$= \left|x_1\right|^2 + \left|x_2\right|^2 + ... + \left|x_j\right|^2 + ... + \left|x_n\right|^2 \geq 0.$$

ii. $(\mathbf{x},\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$.

Las condiciones *iii* y *iv* se deducen del hecho de que $z_1(z_2 + z_3) = z_1z_2 + z_1z_3$ para cualesquier números complejos z_1, z_2, z_3 .

Veamos la condición *v*.

$$\begin{aligned} \overline{(\mathbf{y},\mathbf{x})} &= \overline{y_1\overline{x}_1 + y_2\overline{x}_2 + ... + y_n\overline{x}_n} \\ &= \overline{y_1\overline{x}_1} + \overline{y_2\overline{x}_2} + ... + \overline{y_n\overline{x}_n} \\ &= \overline{y}_1\overline{\overline{x}}_1 + \overline{y}_2\overline{\overline{x}}_2 + ... + \overline{y}_n\overline{\overline{x}}_n \\ &= \overline{y}_1x_1 + \overline{y}_2x_2 + ... + \overline{y}_nx_n \\ &= x_1\overline{y}_1 + x_2\overline{y}_2 + ... + x_n\overline{y}_n = (\mathbf{x},\mathbf{y}). \end{aligned}$$

La condición *vi* se deduce inmediatamente.

Para *vii*, tenemos:

$$(\mathbf{x}, \alpha\mathbf{y}) \underset{v)}{=} \overline{(\alpha\mathbf{y}, \mathbf{x})} = (\overline{\alpha\mathbf{y}}, \overline{\mathbf{x}}) \underset{vi)}{=} \bar{\alpha}(\overline{\mathbf{y}}, \overline{\mathbf{x}}) \underset{v)}{=} \bar{\alpha}(\mathbf{x}, \mathbf{y}).$$

Ejemplo 3

En $M_{3\times 2}$ si:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \text{ y } B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix},$$

entonces puede definirse un producto interno como:

$$(A, B) = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22} + a_{31}b_{31} + a_{32}b_{32}\,.$$

Ejemplo 4

En $C_{[a,b]}$, espacio de funciones continuas de variable real en el intervalo $[a,b]$, se define:

$$(f\,,\,g) = \int_a^b f(t)\; g(t)\; dt.$$

El anterior es un producto interno, ya que:

i. $(f\,,f) = \int_a^b f^2(t)\; dt \geq 0.$ Del cálculo sabemos que si $h \in C_{[a,b]}$ y $h \geq 0$ sobre $[a,b]$, entonces $\int_a^b h(t)\,dt \geq 0.$ Ahora, si $\int_a^b h(t)\; dt = 0,$ entonces $h = 0$ sobre $[a,b]$. Con esto se prueban i y ii.

Las propiedades de las integrales definidas permiten deducir iii a vii.

En este espacio trabajamos con escalares reales, luego las propiedades se dan en la misma forma que cuando se trabaja con el producto escalar.

Ejemplo 5

En $\mathbb{C}^2$ sean $\mathbf{x} = (3 + 2i,\ 7 - 4i)$ e $\mathbf{y} = (1 - i,\ 2 + i)$; entonces:

$$\begin{aligned}
(\mathbf{x}, \mathbf{y}) &= (3+2i)(1+i) + (7-4i)(2-i) \\
&= 3 + 5i + 2i^2 + 14 - 15i + 4i^2 \\
&= 17 - 10i + 6i^2 \\
&= 17 - 10i - 6 \qquad (i^2 = -1) \\
&= 11 - 10i.
\end{aligned}$$

Ejemplo 6

En $C_{[0,1]}$, sean:

$$\begin{aligned} f(t) &= t^2 + 3, \\ g(t) &= 2t^2 - t. \end{aligned}$$

Entonces:

$$\begin{aligned} (f, g) &= \int_0^1 (t^2 + 3)(2t^2 - t)\, dt \\ &= \int_0^1 (2t^4 - t^3 + 6t^2 - 3t)\, dt \\ &= 2\frac{t^5}{5} - \frac{t^4}{4} + 6\frac{t^3}{3} - 3\frac{t^2}{2}\Big]_0^1 = \frac{13}{20}. \end{aligned}$$

10.2 Norma y ortogonalidad en un espacio vectorial con producto interno

Definición 2

Sea V un espacio vectorial con producto interno y suponga que $\mathbf{x}$ y $\mathbf{y}$ están en V; entonces:

i. La *norma* (longitud) de $\mathbf{x}$ se denota por $\|\mathbf{x}\|$ y se define como:

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}.$$

ii. Si $\mathbf{x}$ y $\mathbf{y}$ son diferentes de cero, se dice que son *ortogonales* cuando:

$$(\mathbf{x}, \mathbf{y}) = 0.$$

Ejemplo 7

En $\mathbb{C}^2$, sea $\mathbf{x} = (2+i,\ -1+3i)$. Calculemos $\|\mathbf{x}\|$.

$$\begin{aligned} \|\mathbf{x}\| &= \sqrt{(\mathbf{x}, \mathbf{x})}, \\ (\mathbf{x}, \mathbf{x}) &= (2+i)(2-i) + (-1+3i)(-1-3i) = 15, \\ \|\mathbf{x}\| &= \sqrt{15}. \end{aligned}$$

Ejemplo 8

En $\mathbb{C}[0, 2\pi]$ las funciones sen t y cos t son ortogonales. Veámoslo:

$$(\operatorname{sen} t,\ \cos t) = \int_0^{2\pi} \operatorname{sen} t \cos t\ dt = \frac{1}{2}\int_0^{2\pi} \operatorname{sen} 2t\, dt = -\frac{\cos 2t}{4}\Big|_0^{2\pi} = 0.$$

Ejemplo 9

En $\mathbb{C}\,[0, 2\pi]$ encontremos $\|\cos t\|$.

$$\|\cos t\| = \sqrt{(\cos t, \cos t)} = \left[\int_0^{2\pi} \cos^2 t\, dt\right]^{1/2} = \left[\int_0^{2\pi} \frac{(1+\cos 2t)}{2}\, dt\right]^{1/2}$$

$$= \left(\frac{1}{2}\left(t + \frac{\operatorname{sen} 2t}{2}\right)_0^{2\pi}\right)^{1/2} = \sqrt{\pi}.$$

Los teoremas dados en el módulo 8 para $\mathbb{R}^n$ se extienden a cualquier espacio vectorial con producto interno.

Definición 3

Sea V un espacio vectorial con producto interno y sea $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\} \subset V$.

$\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ es un *conjunto ortonormal* si:

i. $(\mathbf{u}_i, \mathbf{u}_j) = 0$ para $i \neq j$.

ii. $\|\mathbf{u}_i\| = 1$.

Si sólo se cumple la condición *i* se dice que el conjunto es *ortogonal.*

10.3 Última forma del problema de la base

Dada una base $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ para un subespacio de un espacio vectorial con producto interno, hallar una base ortonormal $B = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ para el subespacio.

El problema se resuelve aplicando el proceso de Gram-Schmidt, teniendo en cuenta la definición particular del producto interno en ese espacio vectorial.

Ejemplo 10

Construyamos una base ortonormal en $\mathbb{P}_2\,[-1, 1]$. Podemos partir de la base estándar $\{1, x, x^2\}$; como todo polinomio es una función continua, $\mathbb{P}_2\,[-1, 1]$ es un subespacio de $C\,[-1, 1]$; por lo tanto, emplearemos el producto interno definido en funciones continuas.

Sea $\mathbf{u}_1 = \dfrac{1}{\|1\|}$.

$$\|1\| = \sqrt{(1,1)} = \left(\int_{-1}^{1} 1^2\,dx\right)^{1/2} = (x|_{-1}^{1})^{1/2} = 2^{1/2} = \sqrt{2}.$$

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}}.$$

$$\mathbf{v}_2' = \mathbf{v}_2 - (\mathbf{v}_2, \mathbf{u}_1)\mathbf{u}_1 = x - \left(x, \frac{1}{\sqrt{2}}\right)\frac{1}{\sqrt{2}}.$$

$$\left(x, \frac{1}{\sqrt{2}}\right) = \int_{-1}^{1} x\frac{1}{\sqrt{2}}dx = \frac{1}{\sqrt{2}}\frac{x^2}{2}\Big]_{-1}^{1} = 0.$$

$$\mathbf{v}_2' = x.$$

$$\mathbf{u}_2 = \frac{x}{\|x\|}, \qquad \|x\| = \sqrt{(x,x)} = \left(\int_{-1}^{1} x^2\,dx\right)^{1/2} = \left(\frac{x^3}{3}\Big]_{-1}^{1}\right)^{1/2} = \sqrt{\frac{2}{3}}.$$

$$= \sqrt{\frac{3}{2}}x.$$

$$\mathbf{v}_3' = \mathbf{v}_3 - (\mathbf{v}_3, \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3, \mathbf{u}_2)\mathbf{u}_2$$
$$= x^2 - \left(x^2, \frac{1}{\sqrt{2}}\right)\frac{1}{\sqrt{2}} - \left(x^2, \sqrt{\frac{3}{2}}x\right)\sqrt{\frac{3}{2}}x.$$

$$\left(x^2, \frac{1}{\sqrt{2}}\right) = \int_{-1}^{1} x^2 . \frac{1}{\sqrt{2}}dx = \frac{1}{\sqrt{2}}\frac{x^3}{3}\Big]_{-1}^{1} = \frac{\sqrt{2}}{3},$$
$$\left(x^2, \sqrt{\frac{3}{2}}x\right) = \int_{-1}^{1} x^2\sqrt{\frac{3}{2}}x\,dx = \sqrt{\frac{3}{2}}\frac{x^4}{4}\Big]_{-1}^{1} = 0.$$

Luego:

$$\mathbf{v}_3' = x^2 - \frac{\sqrt{2}}{3}\frac{1}{\sqrt{2}} = x^2 - \frac{1}{3}.$$

$$\|\mathbf{v}_3'\| = \left\|x^2 - \frac{1}{3}\right\| = \left[\int_{-1}^{1}\left(x^2 - \frac{1}{3}\right)^2 dx\right]^{1/2} = \frac{\sqrt{8}}{3\sqrt{5}},$$
$$\mathbf{u}_3 = \frac{3\sqrt{5}}{\sqrt{8}}\left(x^2 - \frac{1}{3}\right) = \frac{3\sqrt{5}}{\sqrt{8}}\frac{(3x^2 - 1)}{3} = \frac{\sqrt{5}}{\sqrt{8}}(3x^2 - 1).$$

Por lo tanto, una base ortonormal en $\mathbb{P}_2\ [-1, 1]$,es:

$$B = \left\{ \frac{1}{\sqrt{2}},\ \sqrt{\frac{3}{2}}x,\ \sqrt{\frac{5}{8}}(3x^2 - 1) \right\}.$$

10.4 Proyección ortogonal en espacios vectoriales con producto interno

Definición 4

Sea H un subespacio de un espacio V con producto interno y $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ una base ortonormal de H. Si $\mathbf{v} \in V$, entonces la proyección ortogonal de **v** sobre H, denotada $\text{proy}_H\,\mathbf{v}$, está dada por:

$$\text{proy}_H\,\mathbf{v} = (\mathbf{v}, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}, \mathbf{u}_2)\mathbf{u}_2 + ... + (\mathbf{v}, \mathbf{u}_k)\mathbf{u}_k.$$

Definición 5

Sea H un subespacio de un espacio V con producto interno. *El complemento ortogonal de H*, denotado por $H^{\perp}$, está dado por:

$$H^{\perp} = \{\mathbf{x} \in V : (\mathbf{x}, \mathbf{h}) = 0 \text{ para todo } \mathbf{h} \in H\}.$$

Nota: los teoremas referentes a $\dim H^{\perp}$, teorema de la proyección y teorema de aproximación de la norma tienen restricciones cuando V es un espacio de dimensión infinita. Veamos qué sucede en cada caso, cuando V es de dimensión infinita.

Si H es un subespacio de V de dimensión finita k, entonces:

i. $H^{\perp}$ es un subespacio de V de dimensión infinita.

ii. Todo vector $\mathbf{v} \in V$ se puede expresar de manera única como suma de dos vectores **p** y **q** tales que $\mathbf{p} \in H$ y $\mathbf{q} \in H^{\perp}$; **p** será $\text{proy}_H\,\mathbf{v}$ pero **q** no es $\text{proy}_{H^{\perp}}\,\mathbf{v}$, ya que ésta no está definida debido a que $H^{\perp}$ tiene dimensión infinita.

iii. El teorema de aproximación de la norma debe tener la restricción de que H sea de dimensión finita para poder construir $\text{proy}_H\,\mathbf{v}$ como la mejor aproximación del vector **v** de V al subespacio H.

Ejemplo 11

Sea $H \subset \mathbb{P}_3[0, 1]$ tal que $H = \text{gen}\{1, x^2\}$.

Encontremos una base para $H^{\perp}$.

Sea $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$ un elemento de $H^{\perp}$.

Como $p(x)$ debe ser ortogonal a los vectores de la base de H, entonces:

$$(p(x), 1) = 0 \text{ y } (p(x), x^2) = 0.$$

$$\begin{aligned}(p(x), 1) &= \int_0^1 (a_0 + a_1x + a_2x^2 + a_3x^3)\,dx \\ &= a_0x + a_1\frac{x^2}{2} + a_2\frac{x^3}{3} + a_3\frac{x^4}{4}\Big]_0^1 \\ &= a_0 + \frac{a_1}{2} + \frac{a_2}{3} + \frac{a_3}{4} = 0. \qquad (1)\end{aligned}$$

$$\begin{aligned}(p(x), x^2) &= \int_0^1 (a_0x^2 + a_1x^3 + a_2x^4 + a_3x^5)\,dx \\ &= a_0\frac{x^3}{3} + a_1\frac{x^4}{4} + a_2\frac{x^5}{5} + a_3\frac{x^6}{6}\Big]_0^1 \\ &= \frac{a_0}{3} + \frac{a_1}{4} + \frac{a_2}{5} + \frac{a_3}{6} = 0. \qquad (2)\end{aligned}$$

Debemos resolver simultáneamente las ecuaciones (1) y (2) para determinar la base de $H^{\perp}$.

$$a_0 + \frac{a_1}{2} + \frac{a_2}{3} + \frac{a_3}{4} = 0,$$
$$\frac{a_0}{3} + \frac{a_1}{4} + \frac{a_2}{5} + \frac{a_3}{6} = 0.$$

$$\begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 & 1/6 \end{bmatrix} \xrightarrow{\text{operaciones elementales}} \begin{bmatrix} 1 & 0 & -1/5 & -1/4 \\ 0 & 1 & 16/15 & 1 \end{bmatrix}.$$

Luego, $a_0 = \frac{1}{5}a_2 + \frac{1}{4}a_3,$

$$a_1 = -\frac{16}{15}a_2 - 1a_3,$$
$$a_2 = 1a_2 + 0a_3,$$
$$a_3 = 0a_2 + 1a_3.$$

Por lo tanto, los vectores $\begin{pmatrix} 1/5 \\ -16/15 \\ 1 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} 1/4 \\ -1 \\ 0 \\ 1 \end{pmatrix}$ forman una base para $H^{\perp}$.

Escribiendo los vectores como polinomios de $\mathbb{P}_3$, tenemos:

$$\text{base para } H^{\perp} = \left\{ \frac{1}{5} - \frac{16}{15}x + x^2, \frac{1}{4} - x + x^3 \right\}.$$

Ejercicios del capítulo 2 (módulo 10)

1. En $\mathbb{R}^2$, si $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, sea $(\mathbf{x}, \mathbf{y}) = 2x_1y_1 + 2x_2y_2$.

   Demuestre que $(\mathbf{x}, \mathbf{y})$ es un producto interno.

2. Con el producto interno definido en el ejercicio 1, sean $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ y $\mathbf{y} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. Calcule $(\mathbf{x}, \mathbf{y})$ y $\|\mathbf{x}\|$.

3. Considere una forma de «normar» el ascenso a una montaña. Los índices de dificultad son como se muestra:

Categoría	**Ángulo de ascenso**	**Índice de dificultad**
1	$0º \leq \theta < 10º$	1
2	$10º \leq \theta < 25º$	2
3	$25º \leq \theta < 45º$	4

   Una norma de dificultad para un viaje de x_1 millas en la categoría 1, x_2 millas en la categoría 2 y x_3 millas en la categoría 3 es $\|(x_1, x_2, x_3)\| = \sqrt{x_1^2 + 2x_2^2 + 4x_3^2}$, la cual se genera mediante el producto interno:

$$((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 + 2x_2y_2 + 4x_3y_3.$$

   De los ascensos siguientes, ¿cuál es el más difícil?

Viaje	**Categoría 1**	**Categoría 2**	**Categoría 3**	**(en millas)**
T_1	3	2	1	
T_2	0	6	0	
T_3	5	0	1	

4. Sea D_n el conjunto de matrices diagonales de $n \times n$ con componentes reales, si A y $B \in D_n$; entonces se define (A, B) así: $(A, B) = a_{11}b_{11} + a_{22}b_{22} + ... + a_{nn}b_{nn}$.

   Pruebe que D_n es un espacio con producto interno.

5. Encuentre una base ortonormal para D_n.

6. Encuentre una base ortonormal para $\mathbb{P}_2[0, 1]$.

7. En $\mathbb{C}^2$ encuentre una base ortonormal comenzando con la base $(1, i), (2 - i, 3 + 2i)$.

8. Si $A = (a_{ij})$ es una matriz real de $n \times n$, la traza de A, que se escribe $\text{tr } A$, es la suma de los componentes de la diagonal de A: $\text{tr } A = a_{11} + a_{22} + ... + a_{nn}$. En M_{nn} se define $(A, B) = \text{tr } (AB^t)$.

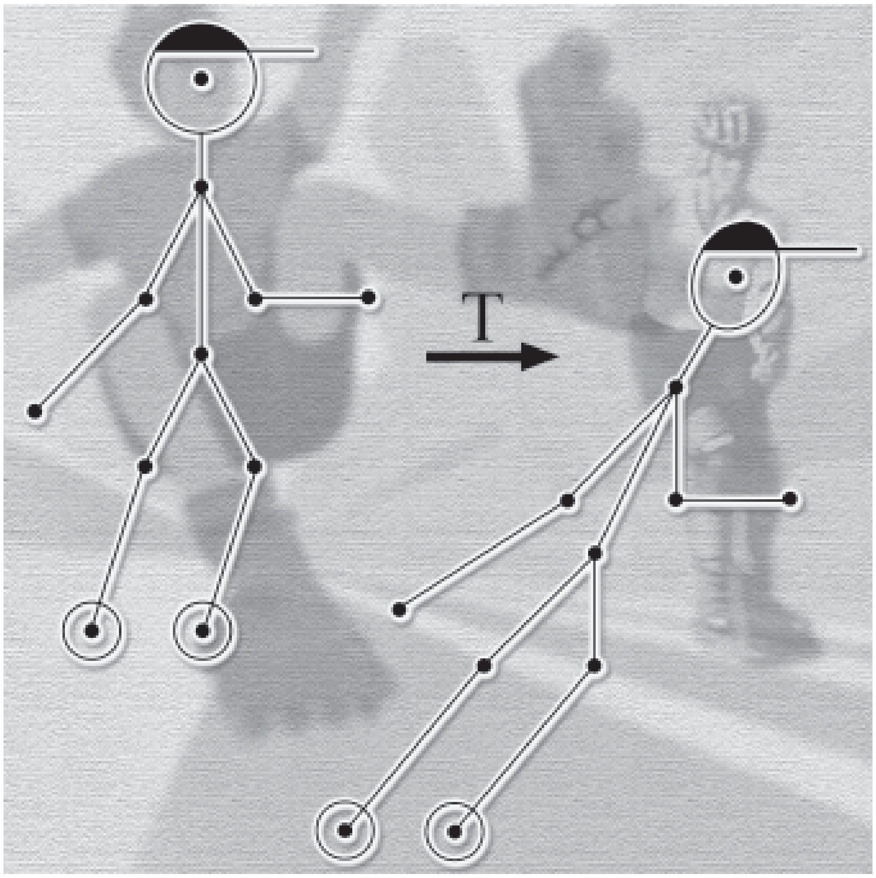
   Demuestre que con esta operación M_{nn} es un espacio con producto interno.

9. Encuentre una base ortonormal para M_{22}.

10. En $\mathbb{P}_3[0, 1]$, sea W el subespacio de $\mathbb{P}_3$ con base $B = \left\{x, x^2\right\}$. Determine una base ortonormal para W.

Capítulo 3

Transformaciones lineales



Un caricaturista emplea computadores y álgebra lineal para transformar la idea y dar la sensación de movimiento a la imagen que presenta. En el dibujo, la figura de la izquierda se transforma en la de la derecha; con esto, permite a su vez una transformación que consiste en multiplicar cada punto (vector) de la figura inicial por una matriz A: $(T(\mathbf{x}) = A\mathbf{x})$.

Contenido breve

Presentación

El concepto de función es uno de los más importantes en matemáticas. En particular, uno de los objetos principales del álgebra lineal es el análisis de las funciones lineales definidas entre espacios vectoriales de dimensión finita. Estas funciones reciben el nombre de *transformaciones lineales* y son de gran aplicación en múltiples problemas de las ciencias físicas, naturales, sociales y económicas.

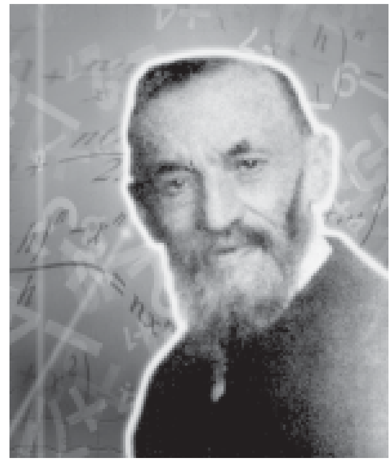
Nuestros ejemplos principales van a ser funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$ que están asociadas con matrices.

Como resultado fundamental del capítulo, vamos a asociar una matriz de transformación a cada transformación lineal definida entre espacios vectoriales de dimensión finita; este hecho lo hemos destacado como el tercer problema básico del álgebra lineal.

La asociación de las transformaciones lineales con sus respectivas matrices de transformación, nos permite estudiar estas funciones por medio del álgebra matricial.

Módulo 11

Definiciones, ejemplos y álgebra de las transformaciones lineales



Para una matriz fija $A_{m\times n}$, a cualquier vector **x** de $\mathbb{R}^n$ le corresponde un vector A**x** de $\mathbb{R}^m$. Esta correspondencia definida por el producto matricial A**x** es el principal ejemplo de una transformación lineal, cuya definición actual se debe al matemático italiano Giuseppe Peano (1858-1932).

Introducción

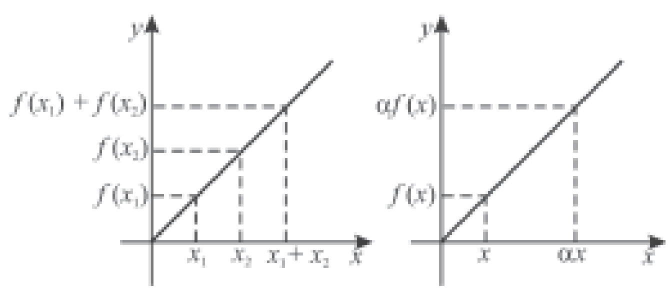
De los cursos anteriores de matemáticas sabemos que una *función* f consta de dos conjuntos A y B y una regla $f(x)$ que asigna a cada elemento de A un único elemento de B. El conjunto A se llama *dominio* de f y B *codominio* de f. Escribimos $f: A \to B$ para indicar que f es una función del conjunto A en el conjunto B.

Por lo general, se ha trabajado con funciones donde el dominio es el conjunto de números reales $\mathbb{R}$ y el codominio es algún subconjunto de $\mathbb{R}$. Un tipo especial de estas funciones son las llamadas *funciones lineales* , ya que la gráfica de la función es una línea recta.

Queremos destacar dos propiedades que poseen algunas funciones lineales, las rectas que pasan por el origen, esto es, $f:\mathbb{R}\to\mathbb{R}$, $f(x)=mx$, donde m es la pendiente de la recta. Veamos que:

1. $f(x_1+x_2)=m(x_1+x_2)=mx_1+mx_2 = f(x_1)+f(x_2)$.
2. $f(\alpha x)=m(\alpha x)=\alpha(mx)=\alpha f(x)$.

Es decir, la función f transforma «sumas en sumas» y «productos por escalares en productos por escalares». Geométricamente esto se ve en la siguiente figura:



Nos proponemos hacer una generalización algebraica de este tipo de funciones, donde tanto el dominio como el codominio son espacios vectoriales. A estas funciones las llamamos transformaciones lineales. Mostraremos, además, que el conjunto de las transformaciones lineales de un espacio vectorial V en un espacio vectorial W es cerrado a la suma y al producto por un escalar; es decir, el conjunto de transformaciones lineales de V en W tiene estructura de espacio vectorial.

Vea el módulo 11 del programa de televisión *Álgebra lineal*.

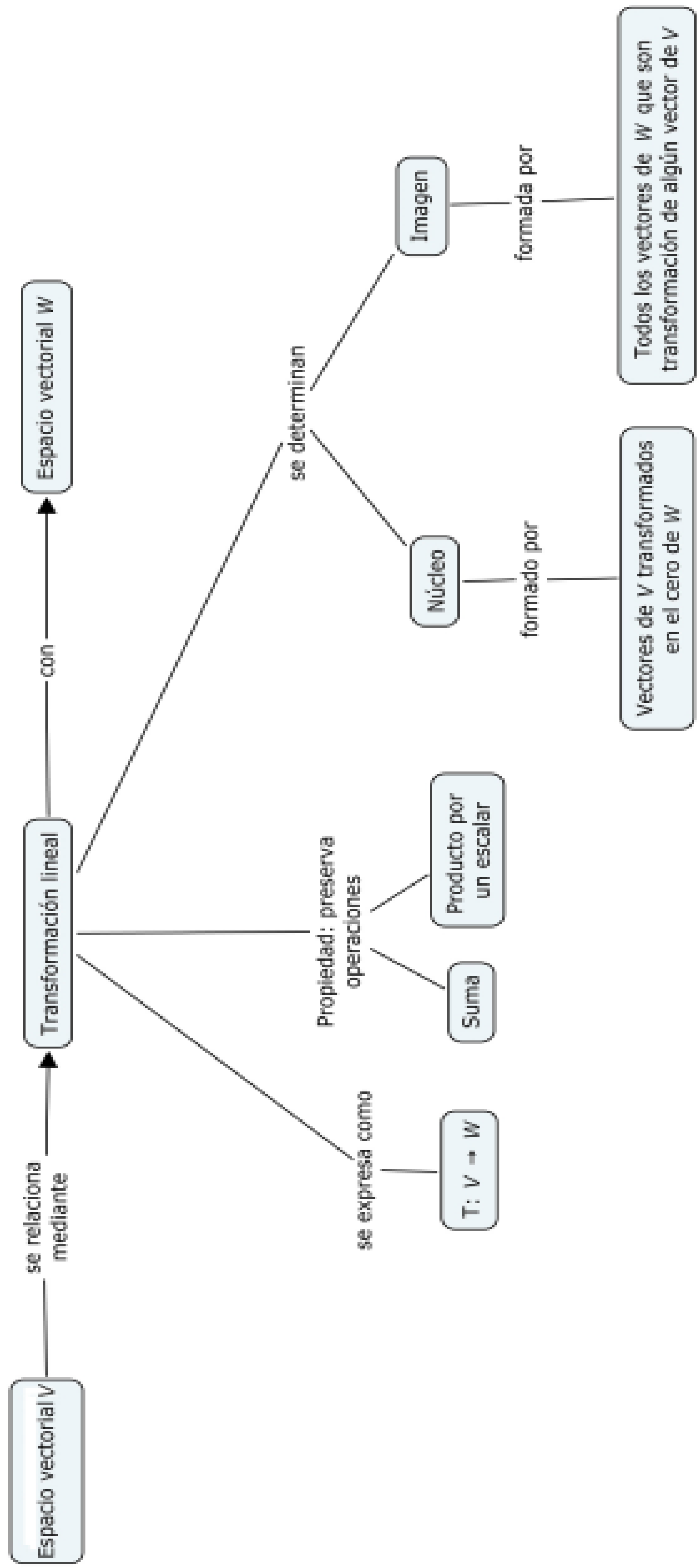
Objetivos

1. Establecer funciones entre espacios vectoriales que preservan las operaciones de suma y producto por un escalar.
2. Ilustrar las transformaciones lineales en los espacios vectoriales más usuales: $\mathbb{R}^n$, $\mathbb{P}_n$, M_{mn} y $C_{[a,b]}$.
3. Aprender el manejo algebraico de las transformaciones lineales.

Preguntas básicas

1. ¿Qué es una transformación lineal?
2. ¿Toda función lineal es una transformación lineal?
3. ¿Cualquier transformación $T:\mathbb{R}^n \to \mathbb{R}^m$ tal que $T(\mathbf{x}) = A\mathbf{x}$ es lineal?
4. ¿Cómo se suman transformaciones lineales?
5. ¿Cómo se realiza la operación producto por escalar con transformaciones lineales?
6. ¿Son las operaciones de suma y producto por un escalar cerradas en el conjunto de transformaciones de V en W?
7. ¿Cómo se realiza la composición de transformaciones lineales?

Contenidos



Mapa 7: módulos 11 y 12

11.1 Definición y ejemplos de transformaciones lineales

Definición 1

Sean V y W espacios vectoriales y sea T una función cuyo dominio es V y cuyo codominio es $W\,(T : V \to W)$. Se dice que T es una *transformación lineal* si:

a. Para todo **x, y** elementos de V.

$$T(\mathbf{x}+\mathbf{y}) \;=\; T(\mathbf{x})+T(\mathbf{y}).$$

b. Para todo $\mathbf{x} \in V$ y $\alpha \in \mathbb{R}$ (o $\alpha \in \mathbb{C}$).

$$T(\alpha\mathbf{x}) \;=\; \alpha\, T(\mathbf{x}).$$

A las transformaciones lineales también se les llama operadores lineales.

Una transformación lineal es, entonces, una función entre dos espacios vectoriales «que preserva las operaciones de espacio vectorial». Es decir, la imagen de la suma de dos vectores del dominio es la suma de las imágenes de cada uno de los vectores y la imagen del producto de un vector del dominio por un escalar es el producto de la imagen del vector por el escalar. Gráficamente es como se representa en la figura 11.1.

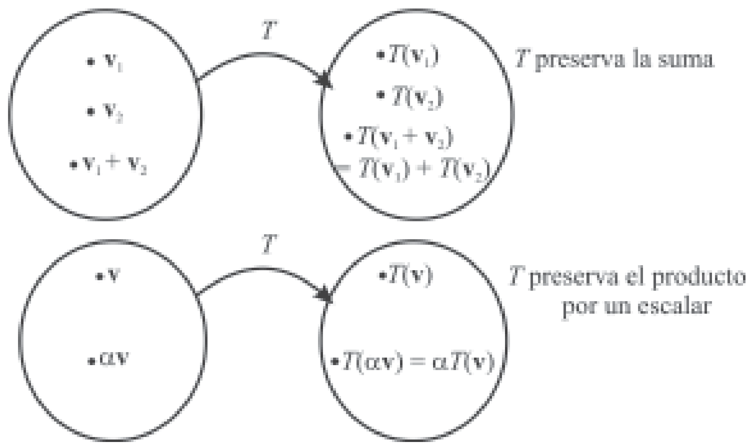


Figura 11.1

Escuche la biografía de *Giuseppe Peano* en su multimedia de *Álgebra lineal.*

Ejemplo 1

La función $f : \mathbb{R} \to \mathbb{R}$, tal que $f(\mathbf{x}) = m\mathbf{x}$ analizada en la introducción es una transformación lineal.

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Gráfica de una transformación lineal».

Ejemplo 2

Sea $T:\mathbb{R}\to\mathbb{R}$ tal que $T(\mathbf{x}) = m\mathbf{x}+b,$ donde m y b son números reales y $b\neq 0.$ Demostremos que T no es una transformación lineal.

Solución

Veamos si se cumple el primer requisito:

$$T(\mathbf{x}+\mathbf{y}) = m(\mathbf{x}+\mathbf{y})+b = m\mathbf{x}+m\mathbf{y}+b. \qquad (1)$$

$$T(\mathbf{x}) + T(\mathbf{y}) = (m\mathbf{x}+b)+(m\mathbf{y}+b) = m\mathbf{x}+m\mathbf{y}+2b. \qquad (2)$$

Podemos observar que (1) $\neq$ (2); por lo tanto, no se verifica la condición *a* de la definición; en consecuencia, T no es lineal.

Las funciones correspondientes a los ejemplos 1 y 2 son líneas rectas. Sin embargo, la recta dada en el ejemplo 2 no es una transformación lineal; la diferencia entre ellas es el término constante; en el ejemplo 1, donde b es cero, hay linealidad; si $b\neq 0$ como en el ejemplo 2, no hay linealidad. En conclusión, las únicas rectas que representan transformaciones lineales son las rectas que pasan por el origen.

Ejemplo 3

Sea A una matriz $m\times n$ y $T:\mathbb{R}^n\to\mathbb{R}^m$ una función tal que $T(\mathbf{x}) = A\mathbf{x}.$ Veamos que T es lineal.

a. $T(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}+\mathbf{y}) = A\mathbf{x}+A\mathbf{y} = T(\mathbf{x})+T(\mathbf{y}).$

b. $T(\alpha\mathbf{x}) = A(\alpha\mathbf{x}) = \alpha(A\mathbf{x}) = \alpha T(\mathbf{x}).$

De *a* y *b* concluimos que T es una transformación lineal.

A la función $f_A:\mathbb{R}^n\to\mathbb{R}^m$ tal que $f_A(\mathbf{x}) = A\mathbf{x}$ con A una matriz $m\times n$ se le llama *función inducida* por A.

Más adelante veremos que no sólo $T(\mathbf{x}) = A\mathbf{x}$ es una transformación lineal, sino que toda transformación lineal entre espacios vectoriales de dimensión finita se puede representar mediante una matriz.

Ejemplo 4

Sea $A=\begin{bmatrix} 3 & 2 & -1 \\ -1 & 0 & -2 \end{bmatrix},$ $T:\mathbb{R}^3\to\mathbb{R}^2$ tal que

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 & 2 & -1 \\ -1 & 0 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{es la función inducida por } A.$$

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x + 2y - z \\ -x - 2z \end{bmatrix} \quad \text{es una transformación lineal.}$$

Ejemplo 5

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

Entonces f_A gira el plano $\mathbb{R}^2$ alrededor del origen un ángulo θ (figura 11.2). Veamos por qué:

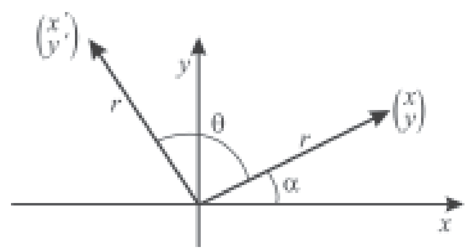


Figura 11.2

Sea $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ un vector en el plano xy. Supongamos que $\mathbf{v}$ se gira un ángulo θ en sentido antihorario. Sea $\mathbf{v}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ el vector rotado, el cual no cambia su magnitud al ser transformado, $|\mathbf{v}| = |\mathbf{v}'| = r.$ Entonces:

$$x = r\cos\alpha, \qquad x' = r\cos(\alpha+\theta).$$
$$y = r\operatorname{sen}\alpha, \qquad y' = r\operatorname{sen}(\alpha+\theta).$$

Calculando $f_A(\mathbf{v}) = A\mathbf{v} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$,

$$A\mathbf{v} = \begin{bmatrix} x\cos\theta & -y\operatorname{sen}\theta \\ x\operatorname{sen}\theta & y\cos\theta \end{bmatrix} = \begin{bmatrix} r\cos\alpha\cos\theta & - & r\operatorname{sen}\alpha\operatorname{sen}\theta \\ r\cos\alpha\operatorname{sen}\theta & + & r\operatorname{sen}\alpha\cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} r\cos(\alpha+\theta) \\ r\operatorname{sen}(\alpha+\theta) \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix} = \mathbf{v}'.$$

Ejemplo 6

Un fabricante produce tres artículos diferentes, para lo cual requiere dos materias primas. La tabla 11.1 nos muestra el número de unidades de cada materia prima que se requiere para la elaboración de cada artículo.

Tabla 11.1

Artículo / Mat. primas	A_1	A_2	A_3
M_1	2	1	2
M_2	1	2	3

Si se necesita producir cierta cantidad de los artículos A_1, A_2, A_3 debemos preguntarnos: ¿cuántas unidades de las materias primas M_1 y M_2 se requieren?

Se define un vector de producción $\mathbf{p}=\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$, donde a_1, a_2, a_3 denotan las cantidades que se deben producir de los artículos A_1, A_2, A_3, respectivamente.

Similarmente, definimos $\mathbf{m}=\begin{bmatrix} m_1 \\ m_2 \end{bmatrix}$, donde m_1 y m_2 denotan las cantidades de las materias primas M_1 y M_2 requeridas para la producción. Luego:

$$m_1 = 2a_1 + a_2 + 2a_3 \,,$$
$$m_2 = a_1 + 2a_2 + 3a_3 \,,$$

lo cual en forma matricial lo escribimos como:

$$\begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}=\begin{bmatrix} m_1 \\ m_2 \end{bmatrix},$$

o también:

$$A\mathbf{p} = \mathbf{m}.$$

La función inducida por A es tal que dado un vector de producción $\mathbf{p}$, lo transforma en un vector de materia prima $\mathbf{m}$. Así que $\mathbf{m} = T(\mathbf{p})$.

La ecuación matricial planteada tiene la forma $A\mathbf{x} = \mathbf{b}$, que es la expresión matricial de un sistema de ecuaciones lineales, donde A es una matriz $m \times n$, $\mathbf{x} \in \mathbb{R}^n$ y $\mathbf{b} \in \mathbb{R}^m$. Cuando se trata de resolver el sistema de ecuaciones lineales, debe hallarse $\mathbf{x}$ conocidos A y $\mathbf{b}$. En el ejemplo, el enfoque es diferente; acá la ecuación $A\mathbf{x} = \mathbf{b}$ «dice»: si se tiene un $\mathbf{x}$ de $\mathbb{R}^n$ podemos encontar $\mathbf{b} \in \mathbb{R}^m$ tal que $\mathbf{b} = T(\mathbf{x})$.

Ejemplo 7

Sea $D:\mathbb{P}_n \to \mathbb{P}_{n-1}$ tal que para $f \in \mathbb{P}_n$, $D(f) = f'$. Como $(f+g)' = f'+g'$ y $(\alpha f)' = \alpha f'$, ya que *f* y *g* son funciones polinómicas, las cuales son diferenciables, entonces *D* es una transformación lineal y se llama *operador diferencial*.

Ejemplo 8

Sea $L: C[a,b] \to \mathbb{R}$ definida por $L(f) = \int_a^b f(x)\,dx.$ Entonces, si f y g están en $C[a,b]$:

a. $$\int_a^b (f+g)(x)\ dx = \int_a^b (f(x)+g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\ dx$$

$$L(f+g) = L(f)+L(g).$$

b. $$\int_a^b (\alpha f)(x)\,dx = \int_a^b \alpha(f(x))\ dx = \alpha\int_a^b f(x)\,dx$$

$$L\ (\alpha f) = \alpha\, L(f).$$

Por lo tanto, L es lineal y se llama *operador integral*.

Ejemplo 9

Sean *V* y *W* espacios vectoriales y $T:\ V \to W$ una transformación tal que $T(\mathbf{v}) = \mathbf{0}$ para todo $\mathbf{v}$ que pertenece a *V*.

Entonces:

$$T(\mathbf{v}_1+\mathbf{v}_2) = \mathbf{0} = \mathbf{0}+\mathbf{0} = T(\mathbf{v}_1)+T(\mathbf{v}_2)\ \text{ y }\ T(\alpha\mathbf{v}) = \mathbf{0} = \alpha\mathbf{0} = \alpha T(\mathbf{v}).$$

Luego, la transformación es lineal y se llama *transformación cero.*

Ejemplo 10

Sea *V* un espacio vectorial. Definamos $I: V \to V$ una transformación tal que $I(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ en *V*. Es evidente que esta transformación es lineal y se denomina *transformación identidad* u *operador identidad.*

Ejemplo 11

Sea $T: M_{mn} \to M_{nm}$ tal que para $\underset{m\times n}{A}$, $T(A) = A^T$. Como $(A+B)^T = A^T + B^T$ y $(\alpha A)^T = \alpha A^T$, T es lineal y se llama *operador de transposición.*

Ejemplo 12

Sea $T: M_{22} \to M_{22}$ definida por:

$$T\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Veamos que T es una transformación lineal.

i. $$T\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix}\right) = T\left(\begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix}\right)$$

$$= \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix}\right)$$

$$= \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

$$= T\begin{bmatrix} a & b \\ c & d \end{bmatrix} + T\begin{bmatrix} e & f \\ g & h \end{bmatrix}.$$

ii. $$T\left(\alpha\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = T\left[\begin{pmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{pmatrix}\right] = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\alpha\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \alpha\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

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

11.2 Álgebra de las transformaciones lineales

11.2.1 Suma y producto por un escalar

Sean T_1 y T_2 transformaciones lineales del espacio vectorial V en el espacio vectorial *W;* entonces:

i. La *suma* de T_1 y T_2 está dada por $(T_1 + T_2)\mathbf{v} = T_1(\mathbf{v}) + T_2(\mathbf{v})$ para todo $\mathbf{v} \in V$; es claro que $T_1(\mathbf{v}) + T_2(\mathbf{v})$ es un elemento de W. Luego la suma es una transformación de V en W.

ii. Sea α un escalar; el *múltiplo escalar* αT_1 de T_1 por α es la transformación $\alpha T_1 : V \to W$ definida por $(\alpha T_1)(\mathbf{v}) = \alpha(T_1(\mathbf{v}))$.

Teorema 1

Sean T_1 y T_2 transformaciones lineales entre los espacios vectoriales V y W; entonces, $T_1 + T_2$ y αT_1 son transformaciones lineales de V en W.

Demostración

Sean $\mathbf{v_1}, \mathbf{v_2} \in V$ y sean α_1 y $\alpha_2 \in \mathbb{R}$. Entonces:

$$
\begin{aligned}
(T_1 + T_2)(\mathbf{v_1} + \mathbf{v_2}) &= T_1(\mathbf{v_1} + \mathbf{v_2}) + T_2(\mathbf{v_1} + \mathbf{v_2}) \\
&= T_1(\mathbf{v_1}) + T_1(\mathbf{v_2}) + T_2(\mathbf{v_1}) + T_2(\mathbf{v_2}) \\
&= (T_1(\mathbf{v_1}) + T_2(\mathbf{v_1})) + (T_1(\mathbf{v_2}) + T_2(\mathbf{v_2})) \\
&= (T_1 + T_2)(\mathbf{v_1}) + (T_1 + T_2)(\mathbf{v_2}).
\end{aligned}
$$

Ahora,

$$
\begin{aligned}
(T_1 + T_2)(\alpha \mathbf{v_1}) &= T_1(\alpha \mathbf{v_1}) + T_2(\alpha \mathbf{v_1}) \\
&= \alpha T_1(\mathbf{v_1}) + \alpha T_2(\mathbf{v_1}) \\
&= \alpha(T_1(\mathbf{v_1}) + T_2(\mathbf{v_1})) \\
&= \alpha((T_1 + T_2)\mathbf{v_1}).
\end{aligned}
$$

Luego, $T_1 + T_2$ es una transformación lineal.

La demostración de que αT_1 es una transformación lineal, se deja como ejercicio.

11.2.2 Composición de transformaciones lineales

Sean U, V y W espacios vectoriales, $T_2 : U \to V$ y $T_1 : V \to W$ transformaciones lineales. La *composición de* T_1 *con* T_2 es la transformación $T_1 \circ T_2 : U \to W$ definida por $T_1 \circ T_2\,(\mathbf{v}) = T_1(T_2(\mathbf{v}))$ para todo $\mathbf{v} \in U$ (figura 11.3).

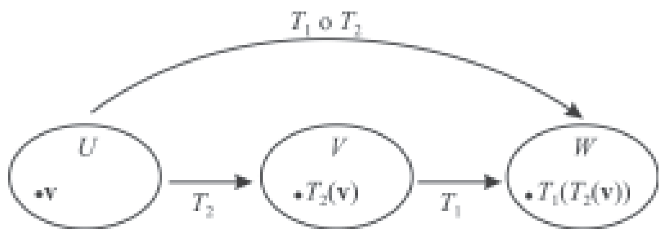


Figura 11.3

Teorema 2

Sean U, V, W espacios vectoriales, $T_2 : U \to V$ y $T_1 : V \to W$ transformaciones lineales; entonces, la transformación $T_1 \circ T_2 : U \to W$ es una transformación lineal.

Demostración

Sean $\mathbf{v_1}$ y $\mathbf{v_2}$ elementos de U y $\alpha \in \mathbb{R}$. Entonces:

i. $$\begin{aligned} T_1 \circ T_2\ (\mathbf{v_1}+\mathbf{v_2}) &= T_1(T_2(\mathbf{v_1}+\mathbf{v}_2)) \\ &= T_1\ (T_2(\mathbf{v_1})+T_2(\mathbf{v_2})) \\ &= T_1\ (T_2(\mathbf{v_1}))+\ T_1\ (T_2(\mathbf{v_2})) \\ &= T_1 \circ T_2\,(\mathbf{v_1}) +\ T_1 \circ \mathrm{T}_2\ (\mathbf{v_2}). \end{aligned}$$

ii. $$\begin{aligned} T_1 \circ T_2\ (\alpha \mathbf{v_1}) &= T_1(T_2(\alpha \mathbf{v_1})) \\ &= T_1\ (\alpha T_2(\mathbf{v_1})) \\ &= \alpha T_1\ (T_2(\mathbf{v_1})) \\ &= \alpha(T_1 \circ T_2\,(\mathbf{v_1})). \end{aligned}$$

De *i* y *ii* concluimos que $T_1 \circ T_2$ es una transformación lineal.

La composición $T \circ T$ suele escribirse $T^{\circ 2}$. En forma semejante, se escribe $T^{\circ 3}$ en vez de $T^2 \circ T$. También se define T^1 como T y T^0 como I, la transformación identidad.

Ejemplo 13

Sean $T_1\!:\ \mathbb{P}_2 \to \mathbb{P}_1$ definida por $T_1\,(a+bx+cx^2) = b+cx$, $T_2:\ \mathbb{P}_2 \to \mathbb{P}_1$ definida por $T_2\,(a+bx+cx^2)\ =\ c-ax$. Evalúe $T_1 + T_2$ y $5T_1$.

$$\begin{aligned}(T_1 + T_2)(a+bx+cx^2) &= T_1(a+bx+cx^2) + T_2(a+bx+cx^2)\\ &= (b+cx) + (c-ax)\\ &= (b+c) + (c-a)\,x.\end{aligned}$$

$$\begin{aligned}(5T_1)(a+bx+cx^2) &= 5\,(T_1(a+bx+cx^2))\\ &= 5\,(b+cx) = 5b+5cx.\end{aligned}$$

Ejemplo 14

Sean $T_2:\mathbb{R}^2 \to \mathbb{R}^3$ y $T_1:\mathbb{R}^3 \to \mathbb{R}^4$ transformaciones lineales definidas por:

$$T_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \\ 2x \end{pmatrix} \quad \text{y} \quad T_1\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x-y \\ x+y \\ x+z \\ 2z \end{pmatrix}.$$

Determine $T_1 \text{ o } T_2\begin{pmatrix} 2 \\ 4 \end{pmatrix}$, $T_1 \text{ o } T_2\begin{pmatrix} x \\ y \end{pmatrix}$.

$$T_1 \text{ o } T_2\begin{pmatrix} 2 \\ 4 \end{pmatrix} = T_1\left(T_2\begin{pmatrix} 2 \\ 4 \end{pmatrix}\right) = T_1\begin{pmatrix} 6 \\ -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \\ 10 \\ 8 \end{pmatrix}$$

$$\begin{aligned}T_1 \text{ o } T_2\begin{pmatrix} x \\ y \end{pmatrix} &= T_1\left(T_2\begin{pmatrix} x \\ y \end{pmatrix}\right) = T_1\begin{pmatrix} x+y \\ x-y \\ 2x \end{pmatrix} = \begin{pmatrix} x+y-x+y \\ x+y+x-y \\ x+y+2x \\ 4x \end{pmatrix}\\ &= \begin{pmatrix} 2y \\ 2x \\ 3x+y \\ 4x \end{pmatrix}.\end{aligned}$$

Ejercicios del capítulo 3 (módulo 11)

En los ejercicios 1 a 4 determine cuáles de las siguientes transformaciones $T:\mathbb{R}^2 \to \mathbb{R}$ son lineales.

1. $T(x,y) = 3x + y.$
2. $T(x,y) = x^2 - y.$
3. $T(x,y) = y.$
4. $T(x,y) = 2.$

En los ejercicios 5 a 8 determine cuáles de las siguientes transformaciones $T:\mathbb{R}^2 \to \mathbb{R}^2$ son lineales.

5. $T(x,y) = (x,0).$
6. $T(x,y) = (1,y).$
7. $T(x,y) = (x^2, y^2).$
8. $T(x,y) = (xy,\ x+y).$

En los ejercicios 9 a 11 determine cuáles de las siguientes transformaciones $T:\mathbb{P}_2 \to \mathbb{P}_3$ son lineales.

9. $T(a_0 + a_1x + a_2x^2) = a_0 + a_1x^2 + a_2x^3.$
10. $T(a_0 + a_1x + a_2x^2) = (a_0 + a_1)x - (2a_1 + a_2)x^3.$
11. $T(a_0 + a_1x + a_2x^2) = a_1a_2 + 3a_0a_2x^2 - x^3.$

Determine si las transformaciones dadas en los ejercicios 12 a 20 son lineales.

12. $T: M_{22} \to \mathbb{R}$ dada por $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$
13. $T: M_{22} \to \mathbb{R}$ dada por $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a + b + c + d.$
14. $T: M_{nn} \to \mathbb{R}$ dada por $T(A) = a_{11} + a_{22} + ... + a_{nn}.$
15. $T: M_{nn} \to \mathbb{R}$ dada por $T(A) = a_{11}\ a_{22} \dots a_{nn}.$
16. $T: M_{nn} \to \mathbb{R}$ dada por $T(A) = \det A.$
17. $T:\mathbb{R}^n \to \mathbb{R}$ dada por $T(x_1, x_2, ..., x_n) = a_1x_1 + a_2x_2 + ... + a_nx_n$, con $a_1, a_2, ..., a_n$ constantes dadas.
18. $T: C[0,1] \to \mathbb{R}$ dada por $T(f(x)) = f(x-1).$
19. $T: C[0,1] \to \mathbb{R}$ dada por $T(f(x)) = f(x) + 1.$
20. $T: C[0,1] \to \mathbb{R}$ dada por $T(f) = f(1).$

En los ejercicios 21 y 22 se describe geométricamente una transformación $T:\mathbb{R}^2 \to \mathbb{R}^2$. Defina analíticamente la transformación y muestre que es lineal.

21.

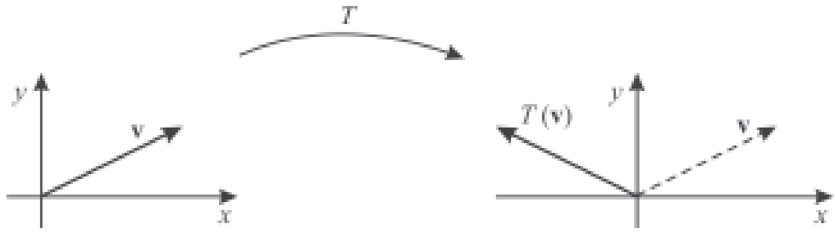
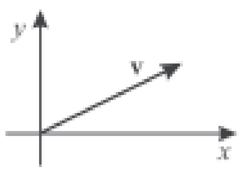


Figura 1 Figura 2

22.



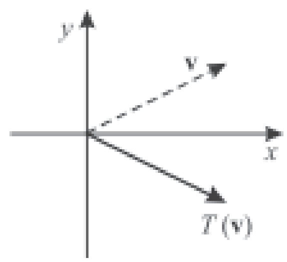


Figura 3 Figura 4

23. Sea B una matriz cuadrada de orden n. Considere la transformación $T: M_{n\times n} \to M_{n\times n}$ dada por $T(A) = AB - BA$ (B es una matriz fija de orden n). Demuestre que T es una transformación lineal.

24. Si $T:V \to W$ es una transformación lineal. Demuestre que αT también es una transformación lineal de V en W.

25. Sean T_1 y T_2 transformaciones lineales de $\mathbb{R}^n \to \mathbb{R}^m$ dadas por $T_1(\mathbf{x}) = \mathbf{A}\mathbf{x}$ y $T_2(\mathbf{x}) = B\mathbf{x}.$ Determine:

a. $(T_1 + T_2)(\mathbf{x})$.

b. $(T_1 \circ T_2)(\mathbf{x})$.

26. Sean T_1 y T_2 transformaciones lineales de $\mathbb{R}^3$ en $\mathbb{R}^2$ dadas por:

$$T_1\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y + z \\ x + y \end{bmatrix} \quad \text{y} \quad T_2\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x + y \\ 2x - y - z \end{bmatrix}.$$

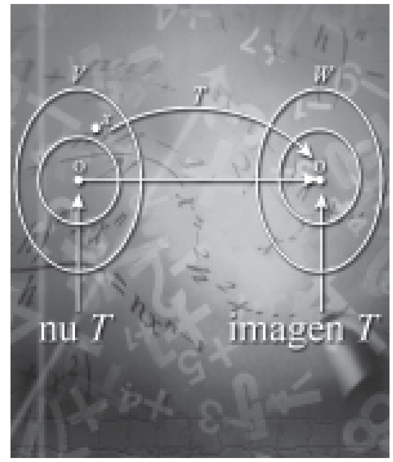
Calcule $T_1 + T_2,\ 4T_1$.

27. Sean $T_1:\mathbb{R}^3 \to \mathbb{R}^2$ tal que $T_1\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y+z \\ x+2z \end{bmatrix}$ y $T_2:\mathbb{R}^2 \to \mathbb{R}^3$ tal que $T_2\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x-4y \\ x-y \end{bmatrix}$.

Determine $(T_1 \circ T_2)$ y $(T_2 \circ T_1)$.

Módulo 12

Propiedades de las transformaciones lineales. Núcleo e imagen



El núcleo y la imagen de una transformación lineal nos muestran los ceros de la transformación y el conjunto de valores que ella toma.

Introducción

Cuando establecemos una función entre dos conjuntos nos interesa saber qué características particulares tiene, cuáles son los ceros de la función y cuál es el conjunto de valores que ella toma. En esta sección estudiaremos estas propiedades para las transformaciones lineales.

Objetivos

1. Estudiar las propiedades de las transformaciones lineales.
2. Determinar para cada transformación lineal dos subespacios: uno en el dominio, llamado núcleo, y otro en el codominio, llamado imagen.

Preguntas básicas

Sea $T: V \to W$ una transformación lineal:

1. ¿Cómo se transforma el «cero» de V?
2. ¿Cómo se calcula la transformación de una combinación lineal de vectores de V?
3. Si $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es una base de V, ¿cómo se calcula $T(\mathbf{v})$ para cualquier $\mathbf{v} \in V$?
4. ¿Cómo se determina el núcleo de T?
5. ¿Cómo se determina la imagen de T?
6. ¿Qué son la nulidad y el rango en una transformación lineal?

Contenidos

Vea el módulo 12 del programa de televisión *Álgebra lineal*.

12.1 Propiedades de las transformaciones lineales

Teorema 1

Sea $T:V\to W$ una transformación lineal; entonces:

a. $T(\mathbf{0})=\mathbf{0}.$

b. $T(\mathbf{x}-\mathbf{y})=T(\mathbf{x})-T(\mathbf{y})$ para todo $\mathbf{x}$, $\mathbf{y}$ en V.

c. $T(\alpha_1\,\mathbf{v}_1+\alpha_2\,\mathbf{v}_2+...+\alpha_n\,\mathbf{v}_n) = \alpha_1\,T(\mathbf{v}_1)+\alpha_2\,T(\mathbf{v}_2)+...+\alpha_n\,T(\mathbf{v}_n)$, siendo $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ y $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ conjuntos cualesquiera de escalares y elementos de V, respectivamente.

Demostración

a. $T(\mathbf{0})=T(0\mathbf{x}) = 0\,T(\mathbf{x})=\mathbf{0}.$

b.
$$\begin{aligned}T(\mathbf{x}-\mathbf{y})&=T(\mathbf{x}+(-1)\mathbf{y})=T(\mathbf{x})+T((-1)\mathbf{y})\\&=T(\mathbf{x})+(-1)T(\mathbf{y})=T(\mathbf{x})-T(\mathbf{y}).\end{aligned}$$

c. Razonemos por inducción sobre n:

Si $n=2$,
$$\begin{aligned}T(\alpha_1\,\mathbf{v}_1+\alpha_2\,\mathbf{v}_2)&=T(\alpha_1\,\mathbf{v}_1)+T(\alpha_2\,\mathbf{v}_2)\\&=\alpha_1\,T(\mathbf{v}_1)+\alpha_2\,T(\mathbf{v}_2).\end{aligned}$$

Supongamos que la proposición se cumple para $n = k$ y veamos que se verifica para $n = k + 1$.

$$\begin{aligned}&T(\alpha_1\,\mathbf{v}_1+\alpha_2\,\mathbf{v}_2+...+\alpha_k\,\mathbf{v}_k+\alpha_{k+1}\mathbf{v}_{k+1})=\\&T((\alpha_1\,\mathbf{v}_1+\alpha_2\,\mathbf{v}_2+...+\alpha_k\,\mathbf{v}_k)+\alpha_{k+1}\mathbf{v}_{k+1})=\\&T(\alpha_1\,\mathbf{v}_1+\alpha_2\,\mathbf{v}_2+...+\alpha_k\,\mathbf{v}_k)+T(\alpha_{k+1}\mathbf{v}_{k+1})=\\&\alpha_1\,T(\mathbf{v}_1)+\alpha_2\,T(\mathbf{v}_2)+...+\alpha_k\,T(\mathbf{v}_k)+\alpha_{k+1}\,T(\mathbf{v}_{k+1}),\end{aligned}$$

aplicando la suposición establecida y el hecho de que la transformación es lineal.

Observación

Es importante notar que en la parte *a* del teorema anterior, el cero del lado izquierdo de la igualdad es el cero de V y el del lado derecho es el de W; además, la parte *c* del teorema es una generalización de las partes *a* y *b*.

En general, cuando se define una función de V en W, ésta se especifica mediante una regla que asigna a cada elemento de V un único elemento de W, ya que sería imposible decir para cada uno de los elementos de V cuál es el asignado en W debido a que se trata de la asignación de una infinidad de elementos. Sin embargo, cuando se trata de una transformación lineal, es posible saber cómo está definida T en todo el espacio vectorial V, conociendo lo que hace T a una base de V. Así que en un espacio de dimensión finita, es posible describir T proporcionando sólo las imágenes de un conjunto finito de vectores.

Teorema 2

Sea $T: V \to W$ una transformación lineal de un espacio vectorial V de dimensión n en un espacio vectorial W. Sea $B = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ una base para V; entonces, para $\mathbf{v} \in V$, $T(\mathbf{v})$ está completamente determinada por $\{T(\mathbf{v}_1), T(\mathbf{v}_2), ..., T(\mathbf{v}_n)\}$.

Demostración

Sea $\mathbf{v} \in V$, $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + ... + \alpha_n \mathbf{v}_n$ con $\alpha_i \in \mathbb{R}$, $i = 1, 2, ..., n$.

$$T(\mathbf{v}) = T(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + ... + \alpha_n \mathbf{v}_n), \text{ y por el teorema 1}c,$$
$$T(\mathbf{v}) = \alpha_1 T(\mathbf{v}_1) + \alpha_2 T(\mathbf{v}_2) + ... + \alpha_n T(\mathbf{v}_n).$$

Ejemplo 1

Sea $T: \mathbb{R}^2 \to \mathbb{R}^3$ una transformación lineal tal que:

$$T\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}-1\\0\\2\end{pmatrix} \text{ y } T\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}0\\2\\-1\end{pmatrix}.$$

Los vectores $\begin{pmatrix}1\\2\end{pmatrix}$ y $\begin{pmatrix}2\\1\end{pmatrix}$ forman una base de $\mathbb{R}^2$, luego la transformación está completamente determinada en todo el espacio vectorial.

Veamos cómo está dada $T\begin{pmatrix}x\\y\end{pmatrix}$.

$$\begin{pmatrix}x\\y\end{pmatrix} = C_1\begin{pmatrix}1\\2\end{pmatrix} + C_2\begin{pmatrix}2\\1\end{pmatrix}; \quad \left[\begin{array}{cc:c}1 & 2 & x\\ 2 & 1 & y\end{array}\right] \xrightarrow{\text{operaciones elementales}} \left[\begin{array}{cc:c}1 & 0 & \dfrac{2y-x}{3}\\ 0 & 1 & \dfrac{2x-y}{3}\end{array}\right]$$
$$= \frac{2y-x}{3}\begin{pmatrix}1\\2\end{pmatrix} + \frac{2x-y}{3}\begin{pmatrix}2\\1\end{pmatrix}.$$

$$T\begin{pmatrix}x\\y\end{pmatrix} = \left(\frac{2y-x}{3}\right) T\begin{pmatrix}1\\2\end{pmatrix} + \left(\frac{2x-y}{3}\right) T\begin{pmatrix}2\\1\end{pmatrix}$$
$$= \left(\frac{2y-x}{3}\right)\begin{pmatrix}-1\\0\\2\end{pmatrix} + \left(\frac{2x-y}{3}\right)\begin{pmatrix}0\\2\\-1\end{pmatrix} = \frac{1}{3}\begin{pmatrix}x-2y\\4x-2y\\-4x+5y\end{pmatrix}.$$

En particular,

$$T\begin{pmatrix}0\\-1\end{pmatrix}=\frac{1}{3}\begin{pmatrix}2\\2\\-5\end{pmatrix}.$$

Teorema 3

Sea V un espacio vectorial de dimensión finita con base $B=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$ y sean $\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_n$ n vectores en un espacio vectorial W. Entonces existe una única transformación lineal $T:V\to W$ tal que $T(\mathbf{v}_i)=\mathbf{w}_i$ para $i=1,2,\ldots,n$.

La demostración de este teorema se deja como ejercicio.

12.2 Núcleo e imagen de una transformación lineal

Definición 1

Sean V y W espacios vectoriales y $T:V\to W$ una transformación lineal; entonces:

a. El *núcleo* de T, denotado nu T, está dado por $\text{nu}\,T=\{\mathbf{v}\in V\,/\,T(\mathbf{v})=\mathbf{0}\}$.

b. La *imagen* de T, denotada imagen T, está dada por:

$$\text{imagen}\,T=\{\mathbf{w}\in W\,/\,\mathbf{w}=T(\mathbf{v}),\ \text{para algún }\mathbf{v}\in V\}.$$

Teorema 4

Si $T:V\to W$ es una transformación lineal, entonces:

a. El núcleo de T es un subespacio de V.

b. La imagen de T es un subespacio de W.

Demostración (figura 12.1)

a. $\text{nu}\,T\neq\Phi$ ya que $T(\mathbf{0})=\mathbf{0}$.

Sean **x, y** elementos de nu T; entonces, $T(\mathbf{x})=\mathbf{0}$ y $T(\mathbf{y})=\mathbf{0}$, $T(\mathbf{x}+\mathbf{y})=T(\mathbf{x})+T(\mathbf{y})=\mathbf{0}+\mathbf{0}=\mathbf{0}$, luego $(\mathbf{x}+\mathbf{y})\in\text{nu}\,T$.

Ahora, $T(\alpha\mathbf{x})=\alpha T(\mathbf{x})=\alpha\,\mathbf{0}=\mathbf{0}$. Por tanto, $\alpha\mathbf{x}\in\text{nu}\,T$.

En consecuencia, nu T es un subespacio de V.

b. Sean $\mathbf{w}_1$ y $\mathbf{w}_2$ elementos de imagen T, luego existen $\mathbf{v}_1$ y $\mathbf{v}_2$ elementos de V tales que $T(\mathbf{v}_1)=\mathbf{w}_1$ y $T(\mathbf{v}_2)=\mathbf{w}_2$; entonces, $T(\mathbf{v}_1+\mathbf{v}_2)=T(\mathbf{v}_1)+T(\mathbf{v}_2)=\mathbf{w}_1+\mathbf{w}_2$.

Por lo tanto, $(\mathbf{w}_1 + \mathbf{w}_2) \in$ imagen T.

Además, $\alpha\,\mathbf{w}_1 = \alpha T(\mathbf{v}_1) = T(\alpha \mathbf{v}_1),$ es decir, $\alpha\,\mathbf{w}_1 \in$ imagen T y, en consecuencia, imagen T es un subespacio de W.

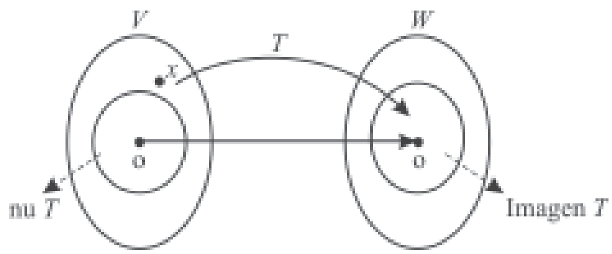


Figura 12.1

Definición 2

Sea T una transfomación lineal de V en W; entonces:

a. La dimensión del núcleo de T, denotada $\nu(T)$, se denomina *nulidad de T.*

b. La dimensión de la imagen T, denotada $\rho(T)$, se denomina *rango de T.*

Ejemplo 2

Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ dada por:

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+2y+z \\ -x+3y+z \end{pmatrix}.$$

Determine nu T, imagen T, $\nu(T)$ y $\rho(T)$.

Solución

Para hallar el conjunto de vectores $(x,\ y,\ z)$ de $\mathbb{R}^3$ tales que $T\,(x,\ y,\ z) = (0, 0)$, debemos resolver el sistema homogéneo

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

$$\begin{bmatrix} 1 & 2 & 1 \\ -1 & 3 & 1 \end{bmatrix} \xrightarrow{\text{operaciones elementales}} \begin{bmatrix} 1 & 0 & 1/5 \\ 0 & 1 & 2/5 \end{bmatrix} \qquad \begin{aligned} x &= -1/5z \\ y &= -2/5z \\ z &= z \end{aligned}$$

si $z = 5,\ \ x = -1,\ \ y = -2$.

Luego, una base para $\text{nu}\,T = \left\{\begin{pmatrix}-1\\-2\\5\end{pmatrix}\right\}$ y nu T = gen $\left\{\begin{pmatrix}-1\\-2\\5\end{pmatrix}\right\}$.

La dimensión de nu T es 1; $\nu(T)=1$.

Sea $\begin{pmatrix}a\\b\end{pmatrix}\in\mathbb{R}^2$ tal que $\begin{pmatrix}a\\b\end{pmatrix}\in$ imagen T, luego:

$$\begin{pmatrix}a\\b\end{pmatrix}=\begin{pmatrix}x+2y+z\\-x+3y+z\end{pmatrix}$$

$$\left[\begin{array}{ccc|c}1&2&1&a\\-1&3&1&b\end{array}\right]\xrightarrow{\text{operaciones elementales}}\left[\begin{array}{ccc|c}1&0&\frac{1}{5}&\frac{3a-2b}{5}\\0&1&\frac{2}{5}&\frac{a+b}{5}\end{array}\right]\qquad \begin{aligned}x&=\frac{3a-2b}{5}-\frac{1}{5}z\\ y&=\frac{a+b}{5}-\frac{2}{5}z\\ z&=z\end{aligned}$$

$$\begin{aligned}x&=\frac{1}{5}(3a-2b-z),\\ y&=\frac{1}{5}(a+b-2z),\\ z&=z.\end{aligned}$$

Así que para cualquier elemento $\begin{pmatrix}a\\b\end{pmatrix}$ de $\mathbb{R}^2$ se puede encontrar una terna $\begin{pmatrix}x\\y\\z\end{pmatrix}$ de $\mathbb{R}^3$ dando a z cualquier valor y determinando x y y como se indica en la solución del sistema. Luego, imagen $T=\mathbb{R}^2$ y $\rho(T)=2$.

Una forma de ver la acción de una transformación lineal consiste en hallar las imágenes determinadas por T sobre figuras geométricas, como polígonos o círculos.

Ejemplo 3

Sea T definida por $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{bmatrix}1&-1\\2&0\end{bmatrix}\begin{pmatrix}x\\y\end{pmatrix}$. Halle la imagen debida a T, del cuadrado de vértices $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$.

Solución

Se hallan las transformaciones de los vértices. Como los lados del cuadrado son segmentos de recta, la imagen del cuadrado resulta de unir las imágenes de los vértices mediante segmentos de recta.

$$T\begin{pmatrix}0\\0\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix},\quad T\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}-1\\0\end{pmatrix},\quad T\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\2\end{pmatrix},\quad T\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}0\\2\end{pmatrix}.$$

Gráficamente podemos ilustrar esto, así (figura 12.2):

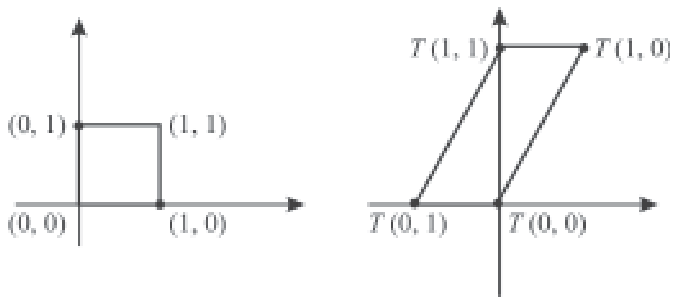


Figura 12.2

Ejemplo 4

Muestre que $T:\mathbb{R}^3 \to \mathbb{R}^3$ definida por:

$$T\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{bmatrix}a&0&0\\0&b&0\\0&0&c\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}\quad a, b, c>0,$$

transforma la esfera $x^2+y^2+z^2=R^2$ en un elipsoide.

Solución

$$T\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}ax\\by\\cz\end{pmatrix}=\begin{pmatrix}x'\\y'\\z'\end{pmatrix}\quad \text{y}$$

$$\frac{(ax)^2}{a^2}+\frac{(by)^2}{b^2}+\frac{(cz)^2}{c^2}=x^2+y^2+z^2=R^2.$$

Al dividir entre R^2 se obtiene:

$$\frac{x'^2}{(aR)^2}+\frac{y'^2}{(bR)^2}+\frac{z'^2}{(cR)^2}=1,$$

que es la ecuación del elipsoide que se ilustra en la figura 12.3:

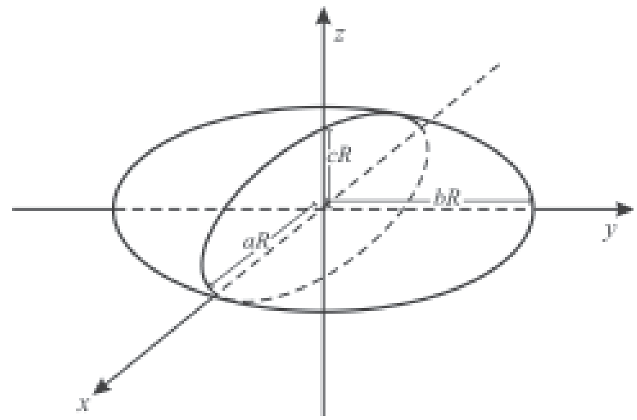


Figura 12.3

Ejemplo 5

Sea $T : M_{nn} \to M_{nn}$ tal que $T(A) = A^{T} + A$. Pruebe que T es una transformación lineal y determine núcleo e imagen.

Solución

T es lineal, ya que:

$$\begin{aligned} T(A+B) &= (A+B)^T + (A+B) \\ &= A^T + B^T + A + B = (A^T + A) + (B^T + B) = T(A) + T(B). \end{aligned}$$

$$T(\alpha A) = (\alpha A)^T + \alpha A = \alpha A^T + \alpha A = \alpha (A^T + A) = \alpha T(A).$$

Determinemos el núcleo de *T*.

$$\text{nu } T = \left\{ \underset{n\times n}{A} : \; A^T + A = 0_{(n\times n)} \right\}$$

$$A^T + A = 0_{(n\times n)} \Leftrightarrow A^T = -A.$$

Por lo tanto, $\text{nu } T = \left\{ \underset{n\times n}{A} : \; A \text{ es una matriz antisimétrica} \right\}$.

La imagen *T* está dada por:

$$\text{imagen } T = \left\{ \underset{n\times n}{C} : \; C = A^T + A \right\}.$$

Veamos qué conjunto constituye la imagen *T*.

$$C^T = (A^T + A)^T = (A^T)^T + A^T = A + A^T = A^T + A = C.$$

Esto es, $\text{imagen } T = \left\{ \underset{n\times n}{C} : \; C \text{ es una matriz simétrica} \right\}$.

Ejercicios del capítulo 3 (módulo 12)

En los ejercicios 1 a 7 encuentre núcleo, imagen, rango y nulidad de la transforrmación lineal dada:

1. $T:\mathbb{R}^2 \to \mathbb{R}^2$, $\quad T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x-y \end{bmatrix}.$

2. $T:\mathbb{R}^2 \to \mathbb{R}^3$, $\quad T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ 5x+5y \\ x-y \end{bmatrix}.$

3. $T:\mathbb{R}^3 \to \mathbb{R}^3$, $\quad T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ x+z \\ y+z \end{bmatrix}.$

4. $T:\mathbb{P}_3 \to \mathbb{P}_2$, $T(a_0 + a_1x + a_2x^2 + a_3x^3) = a_0 + a_1 + a_2x + a_3x^2.$

5. $T:\mathbb{P}_2 \to \mathbb{P}_3$, $T(p) = xp.$

6. $T:M_{3\times 3} \to \mathbb{R}$, $T(A) = t_rA.$

7. $T:C[0, 1] \to \mathbb{R}$, $T(f) = f(1).$

8. Sea $D:\mathbb{P} \to \mathbb{P}$ ($\mathbb{P}$ es el espacio vectorial de todos los polinomios) dada por $D(p) = p'$ (la derivada del polinomio p). Describa el núcleo de D.

9. Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal tal que $T(1, 1) = (0, 0)$ y $T(0, 1) = (1, 1)$. Demuestre que tanto el núcleo como la imagen de T son rectas en el plano xy que pasan por el origen. Encuentre las ecuaciones de estas rectas.

10. Sea $T:\mathbb{R}^3 \to \mathbb{R}^2$ una transformación lineal tal que $T(1, 1, 0) = T(0, 2, 1) = (0, 0)$ y $T(-1, 2, 4) = \mathbf{v} \neq \mathbf{0}.$ Demuestre que el núcleo de T es un plano en $\mathbb{R}^3$ que pasa por el origen. Encuentre su ecuación.

11. Sea $T:\mathbb{R}^n \to \mathbb{R}^m$ la transformación lineal $T(\mathbf{X}) = A\mathbf{X}$, donde A es una matriz $m\times n$. ¿Qué relación guardan el núcleo de T y el espacio solución del sistema homogéneo de ecuaciones lineales $A\mathbf{X} = \mathbf{0}$?

12. Sea $T:M_{2\times 2} \to M_{2\times 2}$ la transformación lineal $T(A) = AB - BA$ donde $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Describa el núcleo de T.

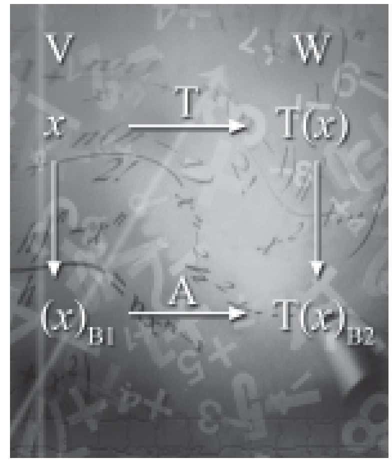
13. Encuentre una transformación lineal T de $\mathbb{R}^3 \to \mathbb{R}^3$ tal que $\text{nu}\,T = \{(x, y, z): 2x - y + z = 0\}$.

14. Demuestre el teorema 3.

15. Muestre que $T:\mathbb{R}^2 \to \mathbb{R}^2$, definida por $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, transforma el círculo unitario $x_1^{\,2} + x_2^{\,2} = 1$ en una elipse.

Módulo 13

El problema de la representación matricial de las transformaciones lineales



Podemos estudiar las transformaciones lineales por medio de sus matrices de transformación.

Introducción

En este tema abordaremos uno de los problemas centrales del álgebra lineal, el cual consiste en mostrar que toda transformación lineal entre espacios vectoriales de dimensión finita se puede expresar como el producto de una matriz por un vector y además deducir un algoritmo para obtener esta matriz. Teniendo este resultado, el problema de encontrar el núcleo y la imagen de una TL se reduce a determinar el núcleo y el espacio columna de la matriz respectiva. En otras palabras, podemos estudiar las transformaciones mediante el del álgebra de matrices.

Objetivos

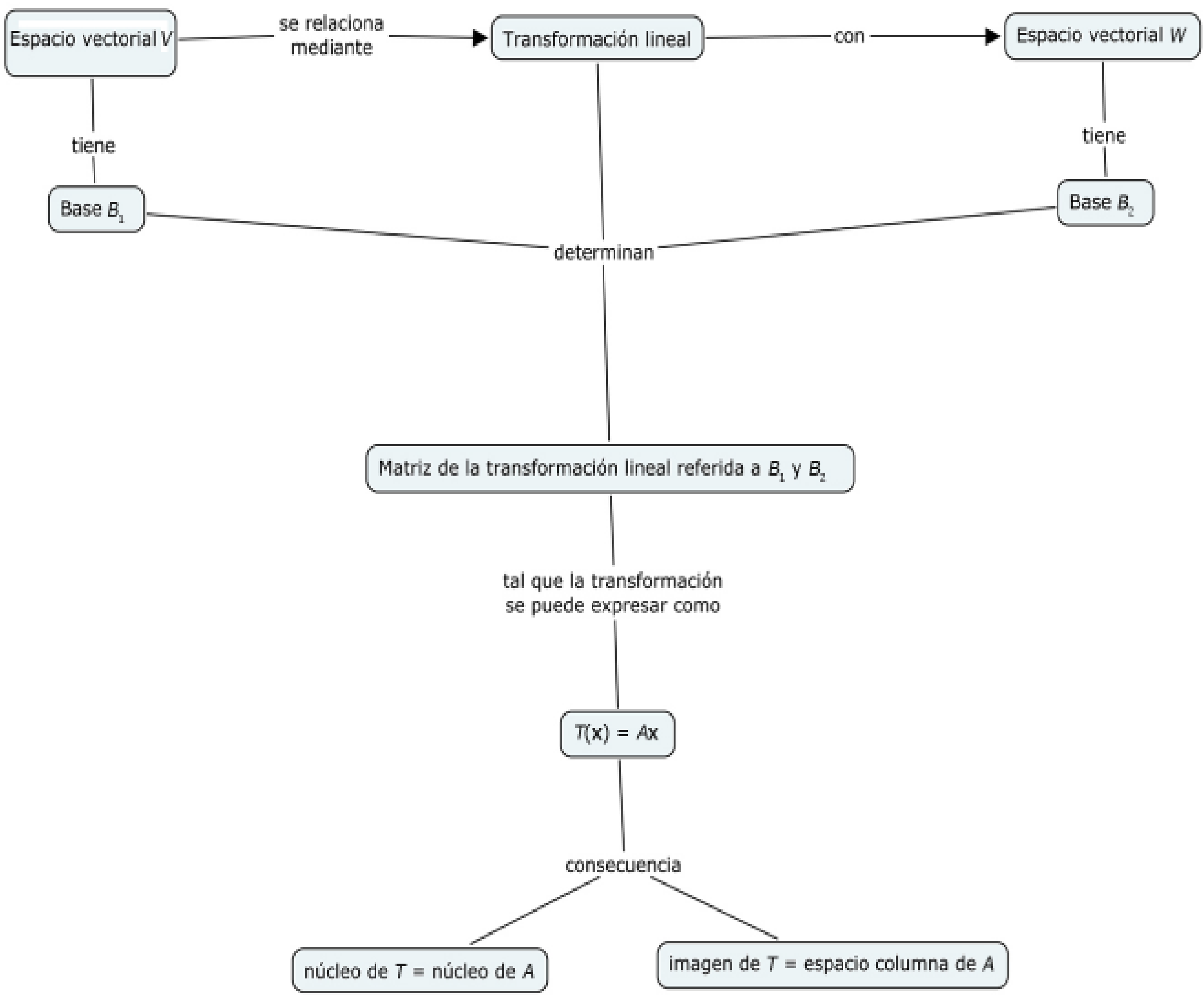
1. Caracterizar una transformación lineal entre dos espacios vectoriales *V* y *W* de dimensiones *n* y *m*, respectivamente, por medio de una matriz A_{mxn} .
2. Establecer la relación existente entre las distintas representaciones de una mis ma TL de modo que se pueda estudiar sus propiedades desde cualquiera de sus representaciones.
3. Trabajar las TL por medio de sus matrices de transformación y hacer el análisis de dichas funciones tomando como referente el álgebra matricial.
4. Reconocer la importancia práctica que tiene poder trabajar en forma matricial un problema de análisis de funciones.

Preguntas básicas

1. ¿Cómo está formada la matriz de una transformación lineal referida a un par de bases?
2. Si A_T es la matriz de una transformación referida a un par de bases B_1 y B_2, ¿cómo se obtiene $(T(\mathbf{x}))_{B_2}$?
3. ¿Qué procedimiento se sigue para hallar la matriz de una transformación lineal $T: V \to W$ referida a las bases B_1 y B_2 de *V* y *W*, respectivamente?
4. Si *A* y *B* son representaciones matriciales de la transformación lineal *T*, ¿qué relación hay entre *A* y *B*?
5. ¿Qué propiedades tienen las matrices similares?

Vea el módulo 13 del programa de televisión *Álgebra lineal*.

Contenidos



Mapa 8: módulo 13

13.1 Tercer problema básico del álgebra lineal

Dada una transformación $T: V \to W$, con dim $V = n$ y dim $W = m$, hallar una matriz A de $m \times n$ que represente a T.

Resolveremos primero el problema para $T: \mathbb{R}^n \to \mathbb{R}^m$ y luego lo extenderemos a espacios V y W de dimensiones n y m, respectivamente.

Teorema 1

Sea $T: \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal; entonces, existe una única matriz A de $m \times n$ tal que $T(\mathbf{x}) = A\mathbf{x}$, para $\mathbf{x}$ en $\mathbb{R}^n$.

Sea $\mathbf{x} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$ un vector de $\mathbb{R}^n$; entonces:

$$\mathbf{x} = c_1\mathbf{e_1} + c_2\mathbf{e_2} + \ldots + c_n\mathbf{e_n},$$

de modo que:

$$T(\mathbf{x}) = c_1 T(\mathbf{e_1}) + c_2 T(\mathbf{e_2}) + \ldots + c_n T(\mathbf{e}_n). \qquad (1)$$

Si A es la matriz cuya j-ésima columna es $T(\mathbf{e_j})$ con $j = 1, \ldots, n$, entonces:

$$A\mathbf{x} = [T(\mathbf{e_1})\; T(\mathbf{e_2}) \ldots T(\mathbf{e}_n)] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

$$= c_1 T(\mathbf{e_1}) + c_2 T(\mathbf{e_2}) + \ldots + c_n T(\mathbf{e_n}) = T(\mathbf{x}). \qquad (2)$$

De (1) y (2) tenemos que:

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

Ahora mostraremos que T es única. Suponga que $B_{m\times n}$ también cumple que $T(\mathbf{x}) = B\mathbf{x}$ para $\mathbf{x} \in \mathbb{R}^n$.

Si $\mathbf{x} = \mathbf{e_i} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$,

$$\begin{aligned} T(\mathbf{e_i}) &= A\mathbf{e_i} = \text{Col}_i(A) \\ &= B\mathbf{e_i} = \text{Col}_i(B) \end{aligned}$$

luego las columnas de A y B coinciden, por lo cual $A = B$.

La matriz $A = [T(\mathbf{e_1})T(\mathbf{e_2})\ldots T(\mathbf{e}_n)]$ es la *matriz de la transformación* referida a la base estándar o canónica.

Ejemplo 1

Sea $T: \mathbb{R}^3 \to \mathbb{R}^2$ una transformación lineal definida por:

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} x+y \\ y-2z \end{bmatrix}.$$

Encontremos la matriz A que representa la transformación referida a las bases estándar:

$$T(\mathbf{e_1}) = T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{bmatrix} 1+0 \\ 0-2.0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \text{Col}_1(A),$$

$$T(\mathbf{e_2}) = T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{bmatrix} 0+1 \\ 1-2.0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \text{Col}_2(A),$$

$$T(\mathbf{e_3}) = T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{bmatrix} 0+0 \\ 0-2.1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \end{bmatrix} = \text{Col}_3(A).$$

Por lo tanto:

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

Así que:

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ y-2z \end{bmatrix}.$$

Teorema 2

Sea $T: V \to W$ una transformación lineal de un espacio vectorial V de dimensión n en un espacio vectorial W de dimensión m y sean $B_1 = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v}_n\}$ y $B_2 = \{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w}_m\}$ bases de V y W respectivamente; entonces, existe una única matriz $A_{T\, m\times n}$ tal que:

$$(T(\mathbf{x}))_{B_2} = A_T(\mathbf{x})_{B_1}.$$

Demostración

Si $\mathbf{x} \in V$, $\mathbf{x} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n}$, o sea que $(\mathbf{x})_{B_1} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$.

Sean: $T(\mathbf{v_1}) = \mathbf{y}_1, T(\mathbf{v_2}) = \mathbf{y}_2, \ldots, T(\mathbf{v_n}) = \mathbf{y}_n$ y

$A_T = [(\mathbf{y}_1)_{B_2}\ (\mathbf{y}_2)_{B_2} \ldots (\mathbf{y}_n)_{B_2}]$.

Ahora:

$$(\mathbf{v_1})_{B_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, (\mathbf{v_n})_{B_1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Entonces:

$$A_T\ (\mathbf{v_i})_{B_1} = (\mathbf{y_i})_{B_2},$$

$$A_T(\mathbf{x})_{B_1} = [(\mathbf{y_1})_{B_2}\ (\mathbf{y_2})_{B_2} \ldots (\mathbf{y_n})_{B_2}] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

$$= c_1(\mathbf{y}_1)_{B_2} + c_2(\mathbf{y}_2)_{B_2} + \ldots + c_n(\mathbf{y}_n)_{B_2}. \qquad (1)$$

Por otro lado:

$$\begin{aligned} T(\mathbf{x}) &= T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n) \\ &= c_1 T(\mathbf{v}_1) + c_2 T(\mathbf{v}_2) + \ldots + c_n T(\mathbf{v}_n) \\ &= c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \ldots + c_n\mathbf{y}_n, \end{aligned}$$

de manera que:

$$[T(\mathbf{x})]_{B_2} = c_1(\mathbf{y_1})_{B_2} + c_2(\mathbf{y_2})_{B_2} + \ldots + c_n(\mathbf{y_n})_{B_2}. \qquad (2)$$

De (1) y (2) concluimos que:

$$[T(\mathbf{x})]_{B_2} = A_T(\mathbf{x})_{B_1}.$$

La unicidad de la matriz A_T se demuestra igual que en el teorema 1.

Definición 1

Sean V y W espacios vectoriales con bases $B_1 = \{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v}_n\}$ y $B_2 = \{\mathbf{w_1}, \mathbf{w_2}, ..., \mathbf{w}_m\}$, respectivamente, y $T: V \to W$ una transformación lineal tal que $[T(\mathbf{x})]_{B_2} = A_T(\mathbf{x})_{B_1}$. A la matriz A_T la denominamos *matriz de la transformación referida a* B_1 y B_2 o *representación matricial de T respecto a* B_1 y B_2.

La solución al tercer problema básico del álgebra lineal está dada en el procedimiento desarrollado en la prueba del teorema 2.

- **Procedimiento para hallar la matriz de una transformación lineal $T: V \to W$ referida a las bases B_1 de V y B_2 de W**

Sean: $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ y $B_2 = \{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m\}$.

Paso 1. Encontrar $T(\mathbf{v}_1), T(\mathbf{v}_2)...T(\mathbf{v}_n)$.

Paso 2. Expresar $T(\mathbf{v}_1), T(\mathbf{v}_2)...T(\mathbf{v}_n)$ en B_2.
Para ello, hacemos lo siguiente:

$$\left[\mathbf{w}_1, \mathbf{w}_2, ...\mathbf{w}_m \mid T(\mathbf{v}_1)\ T(\mathbf{v}_2), ..., T(\mathbf{v}_n)\right]$$

$$\xrightarrow{\text{operaciones elementales}} \left[I \mid (T(\mathbf{v}_1))_{B_2}\ (T(\mathbf{v}_2))_{B_2} \ldots (T(\mathbf{v}_n))_{B_2}\right].$$

Paso 3. La matriz $A_T = [(T(\mathbf{v_1}))_{B_2}\ (T(\mathbf{v_2}))_{B_2} \ldots (T(\mathbf{v}_n))_{B_2}]$.

$$\begin{tikzcd}
\mathbf{x} \arrow[r, "T"] \arrow[d] & T(\mathbf{x}) \arrow[d] \\
{[\mathbf{x}]_{B_1}} \arrow[r, "A_T"] & {[T(\mathbf{x})]_{B_2} = A_T(\mathbf{x})_{B_1}}
\end{tikzcd}$$

Figura 13.1

La figura 13.1 proporciona una interpretación gráfica de la ecuación $(T(\mathbf{x}))_{B_2} = A_T(\mathbf{x})_{B_1}$, donde se muestra que las transformaciones lineales podemos trabajarlas con matrices.

Ejemplo 2

Sea $T: \mathbb{R}^3 \to \mathbb{R}^2$ la transformación lineal definida en el ejemplo 1 y sean:

$B_1 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ y $B_2 = \{\mathbf{w}_1, \mathbf{w}_2\}$ bases de $\mathbb{R}^3$ y $\mathbb{R}^2$, respectivamente, con

$$\mathbf{v}_1 = \begin{bmatrix}1\\0\\1\end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix}0\\1\\1\end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix}1\\1\\1\end{bmatrix}, \quad \mathbf{w}_1 = \begin{bmatrix}1\\2\end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix}2\\0\end{bmatrix}.$$

Determinemos la matriz de T respecto a B_1 y B_2.

Solución

$$T(\mathbf{v}_1) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad T(\mathbf{v}_2) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad T(\mathbf{v}_3) = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Ahora expresemos estos vectores en B_2:

$$T(\mathbf{v}_1) = \begin{bmatrix} 1 \\ -2 \end{bmatrix} = a_1\mathbf{w}_1 + a_2\mathbf{w}_2 = a_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + a_2\begin{bmatrix} 2 \\ 0 \end{bmatrix},$$
$$T(\mathbf{v}_2) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = b_1\mathbf{w}_1 + b_2\mathbf{w}_2 = b_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + b_2\begin{bmatrix} 2 \\ 0 \end{bmatrix},$$
$$T(\mathbf{v}_3) = \begin{bmatrix} 2 \\ -1 \end{bmatrix} = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 = c_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2\begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Es decir, resolvemos tres sistemas lineales cuya matriz de coeficientes es la misma, los vectores de la base B_2. Entonces formamos la matriz:

$$\left[\begin{array}{cc|ccc} 1 & 2 & 1 & 1 & 2 \\ 2 & 0 & -2 & -1 & -1 \end{array}\right] \xrightarrow{\text{operaciones elementales}} \left[\begin{array}{cc|ccc} 1 & 0 & -1 & -1/2 & -1/2 \\ 0 & 1 & 1 & 3/4 & 5/4 \end{array}\right].$$

Luego la matriz A_T de T respecto a B_1 y B_2 es:

$$\begin{bmatrix} -1 & -1/2 & -1/2 \\ 1 & 3/4 & 5/4 \end{bmatrix}, \text{ y}$$

$$(T(\mathbf{x}))_{B_2} = \begin{bmatrix} -1 & -1/2 & -1/2 \\ 1 & 3/4 & 5/4 \end{bmatrix} (\mathbf{x})_{B_1}.$$

Ahora, si $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$,

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

empleando la matriz de la transformación referida a las bases estándar, encontrada en el ejemplo 1.

Expresemos **x** en la base B_1:

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{array}\right] \xrightarrow{\text{operaciones elementales}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

$$(T(\mathbf{x}))_{B_2} = \begin{bmatrix} -1 & -1/2 & -1/2 \\ 1 & 3/4 & 5/4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 5/2 \end{bmatrix}.$$

Entonces:

$$T(\mathbf{x}) = -2\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 5/2\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix},$$

lo cual coincide con el valor encontrado utilizando la matriz de T referida a las bases estándar.

Ejemplo 3

Sea $T : \mathbb{P}_2 \to \mathbb{P}_1$, tal que $T(1) = 6$, $T(x) = x$ y $T(x^2) = x+1$. Determine:

a. La matriz de T respecto a las bases estándar en $\mathbb{P}_2$ y $\mathbb{P}_1$.

b. La matriz de T respecto a las bases $B_1 = \{1+x^2,\ x+x^2,\ 1+x+x^2\}$ y $B_2 = \{1+x,\ 2-x\}$ de $\mathbb{P}_2$ y $\mathbb{P}_1$, respectivamente.

c. Si $P(x) = 2+3x+x^2$, calcule $T(P(x))$ utilizando las matrices halladas en *a* y *b*.

Solución

a. Expresemos las transformaciones de 1, x y x^2 como vectores coordenados de $\mathbb{P}_1$.

$$T(1) = \begin{bmatrix} 6 \\ 0 \end{bmatrix}, \quad T(x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad T(x^2) = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

luego:

$$C_T = \begin{bmatrix} 6 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

$$T(a_0 + a_1x + a_2x^2) = a_0T(1) + a_1T(x) + a_2T(x^2)$$

$$= a_0\begin{bmatrix} 6 \\ 0 \end{bmatrix} + a_1\begin{bmatrix} 0 \\ 1 \end{bmatrix} + a_2\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 6a_0 + a_2 \\ a_1 + a_2 \end{bmatrix}$$

$$= 6a_0 + a_2 + (a_1 + a_2)x.$$

b. Hallemos las transformaciones de los vectores de B_1.

$$T(1 + x^2) = 6(1) + 1 + (0 + 1)x = 7 + x,$$
$$T(x + x^2) = 6(0) + 1 + (1 + 1)x = 1 + 2x,$$
$$T(1 + x + x^2) = 6(1) + 1 + (1 + 1)x = 7 + 2x.$$

Ahora expresemos estas transformaciones en B_2.

$$\left[\begin{array}{cc|ccc} 1 & 2 & 7 & 1 & 7 \\ 1 & -1 & 1 & 2 & 2 \end{array}\right] \xrightarrow{\text{operaciones elementales}} \left[\begin{array}{cc|ccc} 1 & 0 & 3 & \frac{5}{3} & \frac{11}{3} \\ 0 & 1 & 2 & -\frac{1}{3} & \frac{5}{3} \end{array}\right].$$

Entonces la matriz de T respecto a B_1 y B_2 es:

$$A_T = \begin{bmatrix} 3 & \frac{5}{3} & \frac{11}{3} \\ 2 & -\frac{1}{3} & \frac{5}{3} \end{bmatrix}.$$

c. Si $P(x) = 2 + 3x + x^2$,

$$P = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix},$$

$$T(P(x)) = A(P(x))$$

$$= \begin{bmatrix} 6 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 13 \\ 4 \end{bmatrix}$$

$$= 13 + 4x.$$

Ahora expresemos $(P(x))_{B_1}$:

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 3 \\ 1 & 1 & 1 & 1 \end{array}\right] \xrightarrow{\text{operaciones elementales}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 4 \end{array}\right], \quad (P(x))_{B_1} = \begin{bmatrix} -2 \\ -1 \\ 4 \end{bmatrix},$$

$$(T(P(x)))_{B_2} = \begin{bmatrix} 3 & \frac{5}{3} & \frac{11}{3} \\ 2 & -\frac{1}{3} & \frac{5}{3} \end{bmatrix} \begin{bmatrix} -2 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 3 \end{bmatrix},$$

$$T(p(x)) = 7(1 + x) + 3(2 - x) = 13 + 4x.$$

Teorema 3

Sean V y W espacios vectoriales de dimensión finita con dim $V = n$ y $T : V \to W$ una transformación lineal con representación matricial A_T; entonces:

i. $\rho(T) = \rho(A_T)$.

ii. $\nu(T) = \nu(A_T)$.

iii. $\nu(T) + \rho(T) = n$.

Escuche la biografía de *Benoit Mandelbrot* en su multimedia de *Álgebra lineal*.

La demostración del teorema se deja como ejercicio.

El siguiente teorema muestra otra forma de calcular la matriz de una transformación lineal. Al aplicar este método, la matriz se obtiene como un producto de tres matrices.

Teorema 4

Sean V y W espacios vectoriales de dimensiones n y m, respectivamente, y $T : V \to W$ una transformación lineal.

Si C es la matriz de la transformación respecto a las bases canónicas S_n y S_m de V y W, respectivamente, P_1 la matriz de transición de B_1 a S_n en V, P_2 la matriz de transición de B_2 a S_m en W y A_T la matriz de la transformación respecto a B_1 y B_2, entonces:

$$A_T = P_2^{-1} C P_1.$$

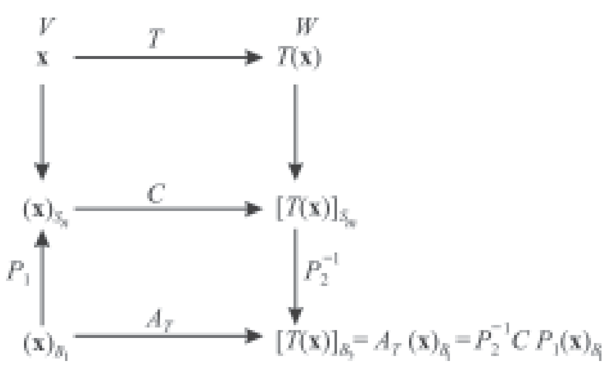


Figura 13.2

Demostración (figura 13.2)

Sea $(\mathbf{x})_{B_1}$; entonces, $P_1(\mathbf{x})_{B_1} = (\mathbf{x})_{S_n}$, $C(\mathbf{x})_{S_n} = (T(\mathbf{x}))_{S_m}$.

Ahora, como P_2 hace la transición de B_2 a S_m, P_2^{-1} hace la transición de S_m a B_2.

Luego $(T(\mathbf{x}))_{B_2} = P_2^{-1}(T(\mathbf{x}))_{S_m}$, y reemplazando:

$$= P_2^{-1}\, C\, P_1(\mathbf{x})_{B_1}.$$

Como la matriz de transformación de B_1 a B_2 es única,

$$A_T = P_2^{-1}\, C\, P_1.$$

Ejemplo 4

En la transformación $T : \mathbb{P}_2 \to \mathbb{P}_1$ tal que $T(1) = 6$, $T(x) = x$ y $T(x^2) = x+1$ planteada en el ejemplo 3, encontramos:

$$C_T = \begin{bmatrix} 6 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$ la matriz de la transformación referida a las bases estándar,

$$A_T = \begin{bmatrix} 3 & 5/3 & 11/3 \\ 2 & -1/3 & 5/3 \end{bmatrix}$$ la matriz de T respecto a B_1 y B_2,

con $B_1 = \{1+x^2,\ x+x^2,\ 1+x+x^2\}$ y $B_2 = \{1+x,\ 2-x\}$. Hallemos ahora A_T empleando el producto matricial establecido en el teorema 4 (figura 13.3).

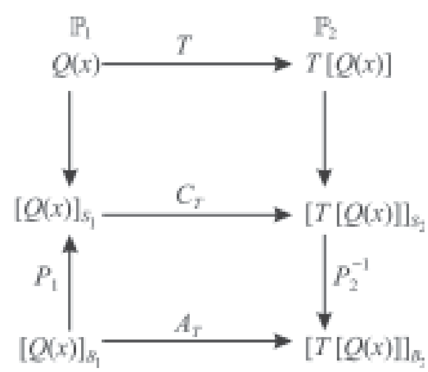


Figura 13.3

Vea la animación «Fractales» en su multimedia de *Álgebra lineal*.

$$P_2^{-1} = -\frac{1}{3}\begin{bmatrix} -1 & -2 \\ -1 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix},$$

$$A_T = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 6 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 5/3 & 11/3 \\ 2 & -1/3 & 5/3 \end{bmatrix}.$$

13.2 Matrices similares y cambio de base

Al construir una representación matricial A para una transformación T podemos analizar T trabajando con A. Tendremos una ventaja en este hecho si A es una matriz «sencilla», esto es, que simplifique la manipulación algebraica. Como bases diferentes dan por resultado matrices diferentes; la elección «correcta» de una base para obtener una matriz A sencilla es importante.

Ejemplo 5

Sea $T = \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2x-2y \\ -5x+y \end{bmatrix}$.

La matriz de T referida a la base estándar se obtiene así:

$$T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ -5 \end{bmatrix} \text{ y } T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix},$$

entonces:

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

Sea $B_1 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -5 \end{bmatrix} \right\}$ otra base para $\mathbb{R}^2$ y obtengamos la matriz de T referida a la base B_1.

$$A_T = \left[\left(T\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)_{B_1} \left(T\begin{bmatrix} 2 \\ -5 \end{bmatrix} \right)_{B_1} \right],$$

$$T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ -4 \end{bmatrix}, \quad T\begin{bmatrix} 2 \\ -5 \end{bmatrix} = \begin{bmatrix} 6 \\ -15 \end{bmatrix}.$$

Expresemos estas transformaciones en B_1:

$$\left[\begin{array}{cc:cc} 1 & 2 & -4 & 6 \\ 1 & -5 & -4 & -15 \end{array}\right] \xrightarrow{\text{operaciones elementales}} \left[\begin{array}{cc:cc} 1 & 0 & -4 & 0 \\ 0 & 1 & 0 & 3 \end{array}\right].$$

La matriz de T referida a B_1 es:

$$A_T = \begin{bmatrix} -4 & 0 \\ 0 & 3 \end{bmatrix}.$$

Por el álgebra matricial sabemos que, en ciertas operaciones, los cálculos son más fáciles si las matrices son diagonales, por ejemplo: hallar inversas, calcular determinantes, hacer potencias. Estas consideraciones son relevantes cuando se trabaja con matrices de gran tamaño.

No es evidente cuál es la base que debemos escoger para que A_T sea una matriz diagonal. Este problema lo resolveremos cuando abordemos el problema de la diagonalización en el siguiente capítulo.

La pregunta que podemos resolver en este momento es la siguiente: ¿qué relación existe entre las diferentes representaciones matriciales de una transformación lineal dada? Veámoslo para el caso en que la transformación lineal está definida de *V* en *V*.

Teorema 5

Sea $T : V \to V$ una transformación lineal con matriz A_T respecto a una base B_1 y matriz B_T respecto a una base B_2. Si *P* es la matriz de transición de la base B_2 a la base B_1, entonces:

$$B_T = P^{-1} A_T P .$$

Demostración (figura 13.4)

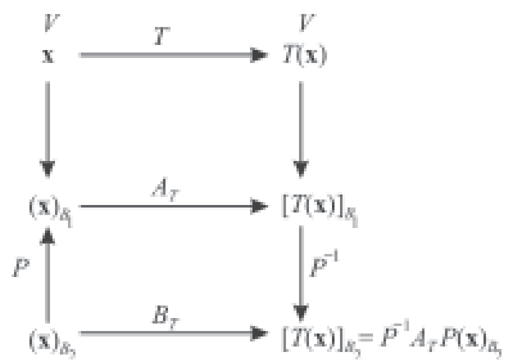


Figura 13.4

$$(\mathbf{x})_{B_1} = P(\mathbf{x})_{B_2},$$
$$(T(\mathbf{x}))_{B_1} = A_T (\mathbf{x})_{B_1} = A_T P(\mathbf{x})_{B_2},$$
$$(T(\mathbf{x}))_{B_2} = P^{-1} (T(\mathbf{x}))_{B_1} = P^{-1} A_T P(\mathbf{x})_{B_2}.$$

Como B_T es la matriz de la transformación respecto a la base B_2, entonces:

$$B_T = P^{-1} A_T P.$$

Definición 2

Dos matrices A y B de $n \times n$ son *similares* o *semejantes* si existe una matriz invertible P tal que $B = P^{-1} A P$.

Del teorema 5 y la definición 2 podemos establecer la siguiente proposición: sea $T : V \to V$ una transformación lineal. Dos matrices cualesquiera que representen a T son similares.

Teorema 6

Sean A, B y C matrices de $n \times n$. Entonces:

a. A es similar a A.

b. Si A es similar a B, entonces B es similar a A.

c. Si A es similar a B y B es similar a C, entonces A es similar a C.

d. Si A es similar a B, entonces $\det A = \det B$.

e. Si A es similar a B, entonces A^m es similar a B^m para cualquier entero positivo m.

f. Si A es similar a B, entonces A es invertible si y sólo si B es invertible. En ese caso A^{-1} es similar a B^{-1}.

Demostración

a. Como $A = I\,A\,I = I^{-1} A\,I$, A es similar a A.

b. Si A es similar a B, $B = P^{-1} A P$, luego:

$$P\,B\,P^{-1} = P(P^{-1} A P)P^{-1} = (PP^{-1})A(PP^{-1}) = A,$$

y entonces:

$$A = (P^{-1})^{-1}\; B(P^{-1})\,;$$

por lo tanto, B es similar a A.

d. Si A es similar a B, $B = P^{-1} A P$, entonces:

$$\begin{aligned}\det B &= \det(P^{-1} A P) = \det P^{-1} \det A \det P = \det(P^{-1}P)\det A \\ &= \det I \det A = \det A.\end{aligned}$$

e. Si A es similar a B, entonces $B = P^{-1} AP$.

$$B^2 = (P^{-1} AP)(P^{-1} AP) = P^{-1} A(PP^{-1})AP = P^{-1} AI\,AP$$
$$= P^{-1} A^2 P.$$

Luego, A^2 es similar a B^2.

Aplicando inducción, supóngase que A^k es similar a B^k, esto es, $B^k = P^{-1} A^k P$, y veamos que la propiedad se cumple para $k + 1$, o sea: $B^{k+1} = P^{-1} A^{k+1} P$.

$$B^{k+1} = B^k B = (P^{-1} A^k P)B = (P^{-1} A^k P)(P^{-1} AP)$$
$$= P^{-1} A^k (PP^{-1})AP = P^{-1} A^k I\,AP = P^{-1} A^{k+1} P.$$

Luego A^{k+1} es similar a B^{k+1} y la proposición queda demostrada para todo entero positivo m.

Las partes c y f se dejan como ejercicio.

Ejercicios del capítulo 3 (módulo 13)

1. Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ definida como $T(x, y) = (x-2y,\ x+2y)$. Sean B_1 la base estándar de $\mathbb{R}^2$ y $B_2 = \{(1,-1),\ (0, 1)\}$. Determine la matriz que representa a T respecto a:

   a. B_1.

   b. B_1 y B_2.

   c. B_2 y B_1.

   d. B_2.

   e. Calcule $T(2,-1)$ empleando la definición de T y las matrices obtenidas en *a*, *b*, *c* y *d*.

2. Sea $T:\mathbb{R}^3 \to \mathbb{R}^3$ definida como $T(x, y, z) = (x+2y+z,\ 2x-y,\ 2y+z)$. Sea B_1 la base natural para $\mathbb{R}^3$ y $B_2 = \{(1, 0, 1),\ (0, 1, 1),\ (0, 0, 1)\}$ otra base para $\mathbb{R}^3$. Determine la matriz de T respecto a:

   a. B_1.

   b. B_1 y B_2.

   c. B_2 y B_1.

   d. B_2.

   e. Calcule $T(1, 1, -2)$ empleando la definición de T y las matrices obtenidas en *a*, *b*, *c*, *d*.

3. Sea $T:\mathbb{R}^3 \to \mathbb{R}^2$ definida como $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ y-z \end{bmatrix}$. Sean B_1 y B_2 las bases naturales de $\mathbb{R}^3$ y $\mathbb{R}^2$, respectivamente; además, sean $B_1' \left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right\}$ y $B_2' \left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$ bases para $\mathbb{R}^3$ y $\mathbb{R}^2$, respectivamente. Determine la matriz de T respecto a:

   a. B_1 y B_2.

   b. B_1' y B_2'.

   c. Calcule $T\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ utilizando la definición de T y las matrices obtenidas en *a* y *b*.

4. Sea $T:\mathbb{P}_1 \rightarrow \mathbb{P}_2$ definida por $T(p(x)) = xp(x) + p(0)$.

Sean $B_1 = \{1, x\}$ y $B_1' = \{1+x, -1+x\}$ bases para $\mathbb{P}_1$.

Sean $B_2 = \{1, x, x^2\}$ y $B_2' = \{1+x^2, -1+x, 1+x\}$ bases para $\mathbb{P}_2$.

Determine la matriz de T respecto a:

a. B_1 y B_2.

b. B_1' y B_2'.

c. Determine $T(3-3x)$ utilizando la definición de T y las matrices obtenidas en *a* y *b*.

5. Sea $T:\mathbb{P}_1 \rightarrow \mathbb{P}_3$ definida por $T(p(x)) = x^2 p(x)$. Sean $B_1 = \{1, x\}$ y $B_1' = \{x, 1+x\}$ bases para $\mathbb{P}_1$. Sean $B_2 = \{1, x, x^2, x^3\}$ y $B_2' = \{1+x, x, -1+x^2, x^3\}$ bases para $\mathbb{P}_3$.

Determine la matriz de T con respecto a:

a. B_1 y B_2.

b. B_1' y B_2'.

6. Sea $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ y sea $T: M_{22} \rightarrow M_{22}$ la transformación lineal definida por $T(A) = AC - CA$ para A en M_{22}. Sean

$B_1 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ y $B_2 = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ bases para M_{22}. Determine la matriz de T respecto a:

a. B_1.

b. B_2.

c. B_1 y B_2.

d. B_2 y B_1.

7. Sea $T:\mathbb{R}^3 \rightarrow \mathbb{R}^3$ la transformación lineal cuya matriz respecto a las bases naturales para $\mathbb{R}^3$ es:

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

Determine:

a. $T\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

b. $T\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

8. Suponga que la matriz de $T:\mathbb{R}^3 \to \mathbb{R}^2$ respecto a las bases $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ y $B_2 = \{\mathbf{w}_1, \mathbf{w}_2\}$ es $A = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 0 \end{bmatrix}$,

donde $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{w}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

a. Calcule $(T(\mathbf{v}_1))_{B_2}$, $(T(\mathbf{v}_2))_{B_2}$, $(T(\mathbf{v}_3))_{B_2}$.

b. Calcule $T(\mathbf{v}_1)$, $T(\mathbf{v}_2)$ y $T(\mathbf{v}_3)$.

c. Calcule $T\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$.

d. Calcule $T\begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

9. Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal. Suponga que la matriz de T respecto a la base $B_1 = \{\mathbf{v}_1, \mathbf{v}_2\}$ es

$A = \begin{bmatrix} 2 & -3 \\ -1 & 4 \end{bmatrix}$, donde $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

a. Calcule $[T(\mathbf{v}_1)]_{B_1}$ y $[T(\mathbf{v}_2)]_{B_1}$.

b. Calcule $T(\mathbf{v}_1)$ y $T(\mathbf{v}_2)$.

c. Calcule $T = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$.

10. Sea $T:\mathbb{P}_1 \to \mathbb{P}_2$ una transformación lineal. Suponga que la matriz de T respecto a las bases $B_1 = \{\mathbf{v}_1, \mathbf{v}_2\}$ y

$B_2 = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ es $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ -1 & -2 \end{bmatrix}$, donde $\mathbf{v}_1 = 1 + x$, $\mathbf{v}_2 = -1 + x$, $\mathbf{w}_1 = 1 + x^2$, $\mathbf{w}_2 = x$, $\mathbf{w}_3 = -1 + x$.

a. Calcule $[T(\mathbf{v}_1)]_{B_2}$ y $[T(\mathbf{v}_2)]_{B_2}$.

b. Calcule $T(\mathbf{v}_1)$ y $T(\mathbf{v}_2)$.

c. Calcule $T(1 + 2x)$.

d. Calcule $T(b + ax)$.

11. En las transformaciones lineales definidas en los ejercicios 1, 3, 5 y 7 describa su núcleo y su imagen por medio de alguna de sus matrices asociadas.

12. Demuestre las partes *c* y *f* del teorema 6.

13. Se dan parejas de matrices A y B. En cada caso muéstrese que A y B no son similares.

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

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

c. $A=\begin{bmatrix}2 & 0\\ 1 & 2\end{bmatrix}$ $\quad B=\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix}$.

En los ejercicios del 14 al 18 considere $T:\mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal con representación matricial A_T.

14. La transformación realiza una *expansión o contracción a lo largo del eje x* multiplicando la coordenada x de cada pareja (x, y) por una constante $C > 0$.

    a. Defina esta transformación lineal. ¿En qué intervalo de valores debe estar la constante C para que el efecto sea respectivamente una expansión o una contracción?

    b. Determine la matriz de esta transformación lineal.

    c. Represente en el plano x, y el rectángulo de vértices (0, 0), (5, 0), (5, 2), (0, 2). Realice gráficamente una expansión (compresión) a lo largo del eje x para un valor de $C=2\ (C=\tfrac{1}{2})$.

15. La transformación realiza una *expansión o contracción a lo largo del eje y* multiplicando la coordenada y de cada pareja (x, y) por una constante $C > 0$.

    a. Defina esta transformación lineal. ¿En qué intervalo de valores debe estar la constante C para que el efecto sea respectivamente una expansión o una contracción?

    b. Determine la matriz de esta transformación lineal.

    c. Represente en el plano x, y el rectángulo de vértices (0, 0), (5, 0), (5, 2), (0, 2). Realice gráficamente una expansión (compresión) a lo largo del eje y para un valor de $C=2\ (C=\tfrac{1}{2})$.

16. La transformación realiza una *reflexión con respecto al eje x (al eje y)* cuando la coordenada y (x) de cada pareja (x, y) se cambia por $-y$ ($-x$).

   a. Defina, en cada caso, la transformación lineal.

   b. Determine, en cada caso, la matriz de la transformación lineal.

   c. Represente en el plano x, y el rectángulo de vértices $(0, 0)$, $(5, 0)$, $(5, 2)$, $(0, 2)$. Realice gráficamente una reflexión con respecto al eje x (al eje y).

17. La transformación realiza una *reflexión con respecto a la recta* $y = x$ cuando cada pareja (x, y) se cambia por la pareja (y, x).

   a. Defina esta transformación lineal.

   b. Determine la matriz de esta transformación lineal.

   c. Represente en el plano x, y el rectángulo de vértices $(0, 0)$, $(5, 0)$, $(5, 2)$, $(0, 2)$. Realice gráficamente una reflexión respecto a la recta $y = x$.

18. La transformación realiza un *corte a lo largo del eje x* [o del eje y] cuando cada pareja (x, y) se transforma en la pareja $(x + Cy, y)$ [o en la pareja $(x, y + Cx)$], respectivamente, donde C es un constante que puede ser positiva o negativa.

   a. Defina esta transformación lineal.

   b. Represente en el plano x, y el rectángulo de vértices $(0, 0)$, $(5, 0)$, $(5, 2)$, $(0, 2)$. Realice gráficamente un corte a lo largo del eje x (y) para un valor de $C = 2$.

19. Demuestre que toda matriz elemental E de 2×2 es uno de los siguientes casos:

   a. La representación matricial de un expansión a lo largo del eje x o y.

   b. La representación matricial de una compresión a lo largo del eje x o y.

   c. La representación matricial de una reflexión respecto del eje x o y.

   d. La representación matricial de una reflexión respecto a la recta $y = x$.

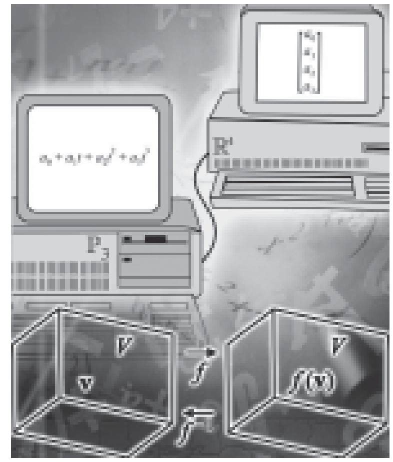
   e. La representación matricial de un corte a lo largo del eje x o y.

   f. El producto de la representación matricial de una reflexión respecto al eje x o y y la representación matricial de una expansión o compresión.

20. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que su representación matricial es invertible. Demuestre que T se puede obtener como una sucesión de expansiones, compresiones, cortes y reflexiones.

Módulo 14

Isomorfismos o transformaciones lineales invertibles



Una transformación lineal y su inversa. El espacio $\mathbb{P}_3$ es isomorfo a $\mathbb{R}^4$.

Introducción

En este módulo caracterizaremos un tipo especial de transformación lineal llamada *isomorfismo.* Esta transformación $T: V \to W$ «reproduce fielmente» la estructura de V en W, ya que establece una correspondencia biyectiva entre los elementos de V y W, y, por otra parte, T traslada sumas y productos por escalares —las operaciones de espacio vectorial— de los elementos de V a sumas y productos por escalares de los correspondientes elementos de W. Así pues, desde el punto de vista del álgebra lineal, cuando entre V y W hay un isomorfismo, esto es, cuando son *espacios vectoriales isomorfos,* V y W son indistinguibles.

Demostraremos que los espacios vectoriales de dimensión finita n son isomorfos, es decir, prácticamente todos ellos son «en esencia» $\mathbb{R}^n$.

Objetivos

1. Determinar cuándo una transformación lineal es uno a uno.
2. Determinar cuándo una transformación lineal es sobreyectiva.
3. Determinar cuándo una transformación lineal es un isomorfismo.
4. Estudiar a través de $\mathbb{R}^n$ las propiedades de otros espacios vectoriales de dimensión n.
5. Caracterizar los isomorfismos entre espacios vectoriales de dimensión finita por medio de su matriz de transformación.
6. Valorar la utilidad de establecer isomorfismos para trasladar una situación dada de un espacio a otro, con el fin de facilitar los procedimientos algebraicos y visualizar mejor las propiedades aplicadas.

Preguntas básicas

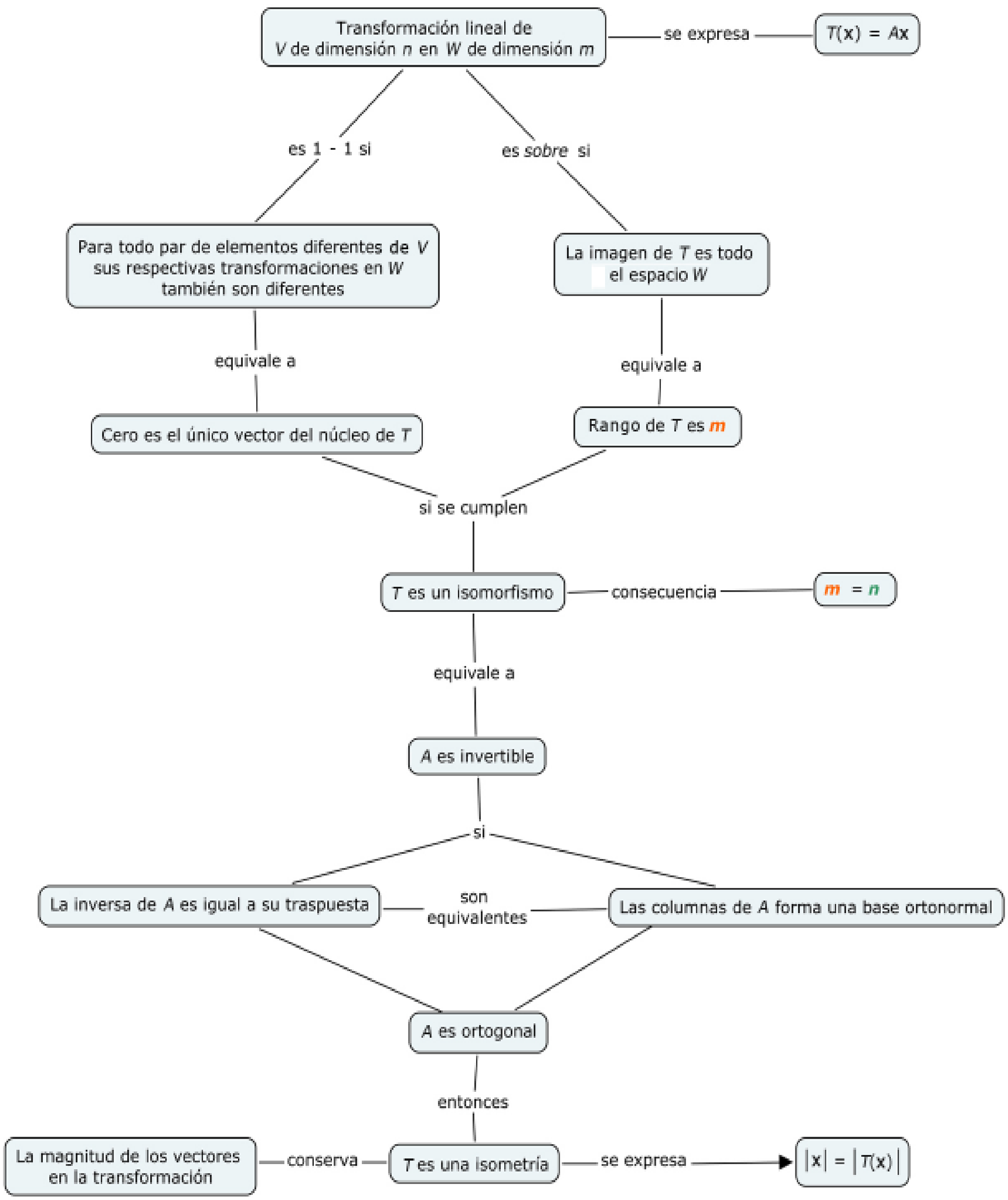
Sea $T: V \to W$ una tranformación lineal:

1. ¿Cuándo T es uno a uno?
2. ¿Cuándo T es sobreyectiva?
3. ¿Cuándo T es un isomorfismo?
4. Si $\dim V = n$ y $\dim W = m$, ¿cuándo T no puede ser uno a uno? ¿Cuándo T no puede ser sobre?
5. Si $\dim V = \dim W = n$, ¿qué condición será suficiente para que T sea un isomorfismo?

Vea el módulo 14 del programa de televisión *Álgebra lineal.*

6. Si T es un isomorfismo, ¿qué propiedades de V se trasladan a W?
7. Si T es un isomorfismo entre espacios vectoriales de dimensión finita, ¿cómo es la matriz de transformación de T?

Contenidos



Mapa 9: módulos 14 y 15

14.1 Transformaciones lineales uno a uno y sobreyectivas

Definición 1

Sea $T: V \to W$ una transformación lineal. T es *uno a uno* (1 - 1) o *inyectiva* si

$$T(\mathbf{v}_1) = T(\mathbf{v}_2) \to \mathbf{v}_1 = \mathbf{v}_2\,,$$

o bien, en forma equivalente,

$$\mathbf{v}_1 \neq \mathbf{v}_2 \to T(\mathbf{v}_1) \neq T(\mathbf{v}_2).$$

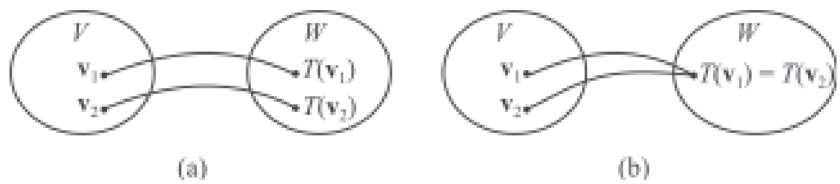


Figura 14.1

La figura 14.1*a* muestra una transformación lineal inyectiva y la 14.1*b*, una transformación lineal no inyectiva.

Definición 2

Sea $T: V \to W$ una transformación lineal. T se llama *sobre W* o *sobre*, si para todo $\mathbf{w} \in W$ existe al menos un $\mathbf{v} \in V$ tal que $T(\mathbf{v}) = \mathbf{w}$. En otras palabras, T es sobre si la imagen de T es todo el espacio W.

Imagen $T = W$.

El teorema siguiente dice que para demostrar si una transformación lineal es uno a uno, sólo se necesita examinar cuáles vectores se transforman en $\mathbf{0}$.

Teorema 1

Sea $T: V \to W$ una transformación lineal. T es uno a uno si y sólo si $\text{nu}\, T = \{\mathbf{0}\}$.

Demostración

a. Suponga que T es 1 - 1 y sea $\mathbf{x} \in \text{nu}\, T$; entonces $T(\mathbf{x}) = \mathbf{0}$. Como $T(\mathbf{0}) = \mathbf{0}$ y T es 1 - 1 se concluye que $\mathbf{x} = \mathbf{0}$, luego $\text{nu}\, T = \{\mathbf{0}\}$.

b. Suponga que $\text{nu}\, T = \{\mathbf{0}\}$ y sea $T(\mathbf{v}_1) = T(\mathbf{v}_2)$; entonces $T(\mathbf{v}_1) - T(\mathbf{v}_2) = \mathbf{0}$,

Vea la animación «Transformaciones uno a uno y sobre» en su multimedia de *Álgebra lineal*.

esto es, $T(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}$; es decir, $(\mathbf{v}_1 - \mathbf{v}_2) \in \text{nu } T$; por lo tanto, $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$, de donde $\mathbf{v}_1 = \mathbf{v}_2$ y la transformación es 1 - 1.

Ejemplo 1

Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \end{pmatrix}$. Para determinar si T es 1 - 1, encontramos nu T. Sea:

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} x+y \\ x-y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Resolvemos el sistema lineal $\begin{matrix} x+y=0 \\ x-y=0 \end{matrix}$

La única solución es $x = 0$ e $y = 0$, de modo que $\text{nu}\, T = \left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}$, y por el teorema 1, T es 1 - 1.

Ejemplo 2

Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-2y \\ y+z \end{bmatrix}$.

La matriz de la transformación referida a las bases estándar es:

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

Entonces, para determinar si T es 1 - 1 hallamos N_A:

$$\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow{\text{operaciones elementales}} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}.$$

La solución al sistema homogéneo es : $\begin{cases} x = -2z \\ y = -z \\ z = z \end{cases}$

Luego, $\text{nu}\, T = N_A = \text{gen}\left\{\begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}\right\}$.

Por lo tanto, T no es 1 - 1.

Veamos si T es sobre:

Examinando las columnas de A_T, vemos que cualesquiera dos columnas son LI, de modo que $C_A = \mathbb{R}^2$; por lo tanto, imagen $A_T = \mathbb{R}^2$. En consecuencia, T es sobre.

Ejemplo 3

Sea $T : \mathbb{R}^3 \to \mathbb{R}^4$ definida por:

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ y-z \\ z+x \\ y+2x \end{bmatrix}.$$

La matriz de T referida a las bases estándar es:

$$A_T = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \xrightarrow{\text{operaciones elementales}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} x &= 0 \\ y &= 0 \\ z &= 0 \end{aligned}$$

$$\text{nu}\ T\left\{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\right\};$$

luego, T es 1 - 1.

En la forma escalonada reducida de A_T podemos observar tres pivotes, luego las tres columnas de A_T son LI.

Entonces:

$$\text{imagen}\, T = C_A = \text{gen}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}\right\},$$

el cual es un subespacio de dimensión 3 de $\mathbb{R}^4$ y, por consiguiente, T no es sobre.

En los ejemplos podemos observar que una transformación lineal puede ser 1 - 1 y no ser sobre, o ser sobre y no ser 1 - 1. El siguiente teorema muestra que cada una de estas propiedades implica la otra, si los espacios V y W tienen la misma dimensión.

Teorema 2

Sea $T: V \to W$ una transformación lineal y sea $\dim V = \dim W$.

a. Si T es 1 - 1, entonces T es sobre.

b. Si T es sobre, entonces T es 1 - 1.

La demostración del teorema se deja como ejercicio.

Teorema 3

Sea $T: V \to W$ una transformación lineal. Suponga que $\dim V = n$ y $\dim W = m$; entonces:

a. Si $n > m$, T no es 1 - 1.

b. Si $m > n$, T no es sobre.

Demostración

a. Sea $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ una base para V. Sea $\mathbf{w}_i = T(\mathbf{v}_i)$ para $i = 1, 2, ..., n$. Puesto que $n > m$, entonces $\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n\}$ es LD, o sea que en la combinación lineal $C_1\mathbf{w}_1 + C_2\mathbf{w}_2 + ... C_n\mathbf{w}_n = \mathbf{0}$ existe por lo menos un $C_i \neq 0$. Por lo tanto, el vector $\mathbf{v} \in V$ tal que $\mathbf{v} = C_1\mathbf{v}_1 + C_2\mathbf{v}_2 + ... + C_n\mathbf{v}_n$ es diferente de $\mathbf{0}$ y:

$$\begin{aligned} T(\mathbf{v}) &= T(C_1\mathbf{v}_1 + C_2\mathbf{v}_2 + ... + C_n\mathbf{v}_n) \\ &= C_1T(\mathbf{v}_1) + C_2T(\mathbf{v}_2) + ... + C_nT(\mathbf{v}_n) \\ &= C_1\mathbf{w}_1 + C_2\mathbf{w}_2 + ... + C_n\mathbf{w}_n. \end{aligned}$$

Luego, $T(\mathbf{v}) = \mathbf{0}$, $\mathbf{v} \in \operatorname{nu} T$; por lo tanto, $\operatorname{nu} T \neq \{0\}$, es decir, T no es 1 - 1.

b. Suponga que $m > n$ y sea $\mathbf{v} \in V$, con:

$$\begin{aligned} \mathbf{v} &= a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + ... + a_n\mathbf{v}_n \\ T(\mathbf{v}) &= a_1T(\mathbf{v}_1) + a_2T(\mathbf{v}_2) + ... + a_nT(\mathbf{v}_n), \end{aligned}$$

luego $\{T(\mathbf{v}_1), T(\mathbf{v}_2), ..., T(\mathbf{v}_n)\}$ es un generador de imagen T; entonces $\rho(T) \leq n$, y como $m > n$, $\rho(T) < m$. Luego, imagen T es un subespacio de W diferente de W, esto es, T no es sobre.

Los ejemplos 2 y 3 ilustran el teorema anterior.

14.2 Isomorfismos

Definición 3

Sea $T: V \to W$ una transformación lineal. Entonces *T* es un *isomorfismo* si *T* es 1 - 1 y sobre. Se dice que los espacios vectoriales *V* y *W* son *isomorfos* si existe un isomorfismo *T* de *V* sobre *W*. Esto se denota como $V \cong W$.

Observación

Es importante notar en la definición anterior que la condición para que dos espacios vectoriales *V* y *W* sean isomorfos es que *exista* entre ellos una transformación lineal que sea isomorfismo. Así, entre espacios vectoriales isomorfos podemos definir transformaciones lineales que no sean isomorfismos.

El siguiente teorema nos muestra la similitud existente entre dos espacios vectoriales isomorfos. La palabra «isomorfismo» viene del griego «isomorphos», que significa «de igual forma».

Teorema 4

Sea $T: V \to W$ un isomorfismo.

a. Si $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ genera *V*, entonces $\{T(\mathbf{v}_1), T(\mathbf{v}_2), ..., T(\mathbf{v}_n)\}$ genera a *W*.

b. Si $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es LI en *V*, entonces $\{T(\mathbf{v}_1), T(\mathbf{v}_2), ..., T(\mathbf{v}_n)\}$ es linealmente independiente en *W*.

c. Si $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es base de *V*, entonces $\{T(\mathbf{v}_1), T(\mathbf{v}_2), ..., T(\mathbf{v}_n)\}$ es base de *W*.

d. Si *V* es de dimensión finita, entonces *W* es de dimensión finita y $\dim V = \dim W$.

Demostración

a. Supongamos que $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ genera *V*.

Sea $\mathbf{w} \in W$. Como *T* es sobre, $\mathbf{w} = T(\mathbf{v})$ para algún $\mathbf{v} \in V$:

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n \text{ y } T(\mathbf{v}) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + ... + c_nT(\mathbf{v}_n),$$

luego:

$$\mathbf{w} = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + ... + c_nT(\mathbf{v}_n);$$

por lo tanto:

$$\{T(\mathbf{v}_1), T(\mathbf{v}_2), ..., T(\mathbf{v}_n)\} \text{ genera a } W.$$

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Isomorfismos».

b. Supongamos que $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es LI.

Sea $c_1T(\mathbf{v}_1)+c_2T(\mathbf{v}_2)+...+c_nT(\mathbf{v}_n)=\mathbf{0}.$ Entonces,

$$T(c_1\mathbf{v}_1+c_2\mathbf{v}_2+...+c_n\mathbf{v}_n)=\mathbf{0},$$

es decir,

$$(c_1\mathbf{v}_1+c_2\mathbf{v}_2+...+c_n\mathbf{v}_n)\in \text{nu } T.$$

Como T es 1 - 1, $c_1\mathbf{v}_1+c_2\mathbf{v}_2+...+c_n\mathbf{v}_n=\mathbf{0},$ y como $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es LI, entonces $c_1=c_2=...=c_n=0$ y, en consecuencia,

$$\{T(\mathbf{v}_1), T(\mathbf{v}_2), ..., T(\mathbf{v}_n)\} \text{ es LI.}$$

c. Se deduce de *a* y *b*.

d. Se deduce de *c*.

Cuando se trata con espacios vectoriales reales de dimensión finita, el problema de saber si son o no isomorfos es bastante simple: sólo es necesario observar si tienen la misma dimensión, como se verá en el siguiente teorema.

Teorema 5

Sean V y W dos espacios vectoriales reales de dimensión finita con $\dim V=\dim W$. Entonces $V\cong W$.

Demostración

Dos espacios vectoriales son isomorfos si entre ellos existe una transformación lineal que sea un isomorfismo. Luego, la demostración consiste en construir dicha transformación lineal.

Sean $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ y $\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n\}$ bases para V y W, respectivamente.

Definimos $T:V\to W$ por $T(\mathbf{v}_i)=\mathbf{w}_i$ para $i=1,..., n$. Entonces, por el teorema 3, módulo 12, existe una única transformación lineal que satisface la condición dada.

Suponga que $\mathbf{v}\in \boldsymbol{V}$ y $T(\mathbf{v})=\mathbf{0}$; entonces:

$$\mathbf{v}=c_1\mathbf{v}_1+c_2\mathbf{v}_2+...+c_n\mathbf{v}_n,$$
$$T(\mathbf{v})=c_1T(\mathbf{v}_1)+c_2T(\mathbf{v}_2)+...+c_nT(\mathbf{v}_n),$$
$$c_1\mathbf{w}_1+c_2\mathbf{w}_2+...+c_n\mathbf{w}_n=\mathbf{0}.$$

Como $\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n\}$ es base para W, $c_1=c_2=...=c_n=0$; por lo tanto, $\mathbf{v}=\mathbf{0}$; en consecuencia, $\text{nu } T=\{\mathbf{0}\}$ y la transformación es 1 - 1. Como $\dim V=\dim W$, entonces T es sobre por el teorema 2. Luego T es un isomorfismo.

Teorema 6

Sean V y W espacios vectoriales reales de dimensión n y $T: V \rightarrow W$, tal que $T(\mathbf{x}) = A\mathbf{x}$ es una transformación lineal. Entonces A es invertible si y sólo si $T(\mathbf{x}) = A\mathbf{x}$ es un isomorfismo.

T es un isomorfismo si y sólo si T es 1 - 1, ya que como $\dim V = \dim W$, T también será sobre.

T es 1 - 1 si y sólo si $\operatorname{nu} T = \{\mathbf{0}\}$, es decir, $v(T) = 0$, y esto es equivalente a la proposición A es invertible.

Este teorema caracteriza los isomorfismos como transformaciones lineales cuya matriz de transformación es invertible, proporcionándonos una forma de determinar cuáles transformaciones lineales son isomorfismos al examinar su matriz de transformación.

Ejemplo 4

Sea $T: \mathbb{P}_n \rightarrow \mathbb{P}_n$ tal que $T(p) = p + p'$. Veamos que T es un isomorfismo.

Encontremos la matriz A_T de la transformación lineal respecto a la base estándar.

$$\begin{aligned}
T(1) &= 1 + (1)' = 1 + 0 = 1 \\
T(x) &= x + (x)' = x + 1 \\
T(x^2) &= x^2 + (x^2)' = x^2 + 2x \\
&\vdots \\
T(x^n) &= x^n + (x^n)' = x^n + nx^{n-1}.
\end{aligned}$$

$$\underset{(n+1)\times(n+1)}{A} = \begin{bmatrix} 1 & 1 & 0\cdots & 0 \\ 0 & 1 & 2\cdots & 0 \\ \vdots & 0 & 1\cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & n \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \det A = 1 \neq 0.$$

A es invertible, luego T es un isomorfismo.

Ejemplo 5

Los espacios vectoriales $\underset{2\times 2}{M}$, $\mathbb{P}_3$ y $\mathbb{R}^4$ son isomorfos, ya que todos tienen dimensión 4. Establezcamos entre ellos transformaciones lineales que sean isomorfismos.

$T_1: \underset{2\times 2}{M} \rightarrow P_3$ definida por:

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a + bx + cx^2 + dx^3.$$

$$T\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = 1 + 0x + 0x^2 + 0x^{3}.$$

$$T\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 0 + 1x + 0x^2 + 0x^3. \qquad A_T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = I_4$$

$$T\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = 0 + 0x + 1x^2 + 0x^3. \qquad \det A_T = 1$$

$$T\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = 0 + 0x + 0x^2 + 1x^3.$$

$$T_2 : \mathbb{P}_3 \to R^4 \text{ definida por } T(a + bx + cx^2 + dx^3) = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

Entonces $A_T = I_4$ (matriz identidad de orden 4); por lo tanto, T_2 es un isomorfismo.

Las transformaciones T_1 y T_2 definidas en el ejemplo nos muestran que estos espacios vectoriales son en «esencia» el mismo; en ambas, la matriz de transformación es la identidad.

Ejercicios del capítulo 3 (módulo 14)

1. Dé un ejemplo de una TL $T: V \to W$ que sea:

   a. Uno a uno pero no sobre.
   b. Sobreyectiva pero no uno a uno.
   c. Uno a uno y sobreyectiva.

2. Demuestre que $T: M_{mn} \to M_{nm}$ definida por $T(A) = A^T$ es un isomorfismo.

3. Encuentre un isomorfismo entre D_n, matrices diagonales de orden *n*, y $\mathbb{R}^n$.

4. Sea $T: \mathbb{R}^3 \to \mathbb{R}^3$ definida como $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$. Determine si *T* es un isomorfismo.

5. Para cada una de las siguientes transformaciones determine si es un isomorfismo a partir de la información dada.

   a. $T: \mathbb{R}^4 \to \mathbb{R}^4,\ \rho_{(T)} = 4.$
   b. $T: \mathbb{R}^4 \to \mathbb{R}^4,\ \nu_{(T)} = 2.$
   c. $T: \mathbb{P}_2 \to \mathbb{P}_2,\ \nu_{(T)} = 1.$
   d. $T: \mathbb{P}_3 \to \mathbb{P}_3,\ \rho_{(T)} = 4.$

6. Sea $V = \mathbb{P}_4$ y $W = \{p \in \mathbb{P}_5 : p(0) = 0\}$. Demuestre que $V \cong W$.

7. Sea $T: \mathbb{P}_n \to \mathbb{P}_n$ tal que $Tp(x) = xp'(x)$. Determine si *T* es un isomorfismo.

8. Demuestre que si $T: V \to W$ es un isomorfismo, entonces existe un isomorfismo $L: W \to V$ tal que $L(T(\mathbf{v})) = \mathbf{v}$. A *L* se le llama *transformación inversa de T* y se denota T^{-1}.

9. Demuestre que si $T: \mathbb{R}^n \to \mathbb{R}^n$ está definido por $T(\mathbf{x}) = A\mathbf{x}$ y T es un isomorfismo, entonces T^{-1} está dado por $T^{-1}(\mathbf{x}) = A^{-1}\mathbf{x}$.

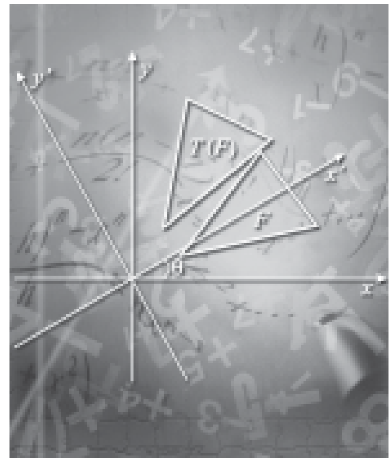
10. Sea $T: \mathbb{R}^3 \to \mathbb{R}^3$ una TL dada por $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + 4y - 6z \\ 5y - 2z \\ 9z \end{bmatrix}$.

    Demuestre que *T* es un isomorfismo y determine T^{-1}.

11. Sea $T: M_{n\times n} \to M_{n\times n}$ la transformación lineal dada por $T(A) = BA$, donde *B* es una matriz fija de orden *n*. Demuestre que *T* es un isomorfismo si y sólo si *B* es una matriz invertible. En tal caso, describa T^{-1}.

Módulo 15

Isometrías



Una rotación igual a un ángulo θ alrededor del origen es una transformación lineal que conserva la magnitud de los vectores.

Introducción

La palabra *isometría* significa igual medida. De esto se trata en esta sección: de transformaciones lineales que conservan la magnitud de los vectores.

Antes de iniciar el estudio de las isometrías, definiremos las matrices ortogonales y veremos algunas propiedades de estas matrices. Además, deduciremos una propiedad del producto escalar cuando en uno de los factores hay una matriz A. Estos resultados nos serán de gran utilidad en el tratamiento de las isometrías.

Objetivos

1. Mostrar un tipo de transformación lineal que conserva la magnitud de los vectores.
2. Caracterizar las isometrías por medio de las matrices ortogonales.
3. Caracterizar las isometrías en el plano.
4. Mostrar que conservar la magnitud de los vectores en una transformación en $\mathbb{R}^n$ es equivalente a conservar el producto escalar.

Preguntas básicas

1. ¿Cuándo una matriz es ortogonal?
2. ¿Cómo son las columnas en una matriz ortogonal?
3. ¿Qué es una isometría?
4. ¿Qué operación se conserva en una isometría?
5. ¿Cómo es la matriz de la transformación T, cuando T es isometría?
6. ¿Cómo son las transformaciones en $\mathbb{R}^2$ que son isometrías?
7. ¿Qué relación existe entre los isomorfismos y las isometrías?

Contenidos

15.1 Matrices ortogonales
15.2 Isometrías
15.3 Isometrías en $\mathbb{R}^2$

Vea el módulo 15 del programa de televisión *Álgebra lineal*.

15.1 Matrices ortogonales

Definición 1

Una matriz Q de $n \times n$ se llama *ortogonal* si Q es invertible y $Q^{-1} = Q^T$.

Si $Q^{-1} = Q^T$, entonces $Q^T Q = I_n$.

En el módulo 8 estudiamos la forma de constituir bases ortonormales en $\mathbb{R}^n$. El siguiente teorema nos da una relación entre las matrices ortogonales y las bases ortonormales de $\mathbb{R}^n$.

Teorema 1

La matriz $Q\, n \times n$ es ortogonal si y sólo si las columnas de Q forman una base ortonormal de $\mathbb{R}^n$.

Demostración

Sea:

$$Q = \begin{bmatrix} a_{11} & a_{12} \cdots & a_{1j} \cdots & a_{1n} \\ a_{21} & a_{22} \cdots & a_{2j} \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} \cdots & a_{nj} \cdots & a_{nn} \end{bmatrix} \qquad Q^T = \begin{bmatrix} a_{11} & a_{21} \cdots & a_{n1} \\ a_{12} & a_{22} \cdots & a_{n2} \\ \vdots & \vdots & \vdots \\ a_{1i} & a_{2i} \cdots & a_{ni} \\ \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} \cdots & a_{nn} \end{bmatrix}$$

Sea $B = (b_{ij}) = Q^T Q$.

$$b_{ij} = a_{1i}\, a_{1j} + a_{2i}\, a_{2j} + \ldots + a_{ni}\, a_{nj} = \mathbf{c}_i \cdot \mathbf{c}_j, \qquad (1)$$

donde $\mathbf{c}_i$ y $\mathbf{c}_j$ son las columnas i y j de Q.

Si las columnas de Q son ortonormales, entonces, de (1):

$$b_{ij} = 0 \quad \text{si} \quad i \neq j, \quad b_{ij} = 1 \quad \text{si} \quad i = j. \qquad (2)$$

Es decir, $B = I$ y entonces $Q^T = Q^{-1}$; por lo tanto, Q es ortogonal.

Recíprocamente, si Q es ortogonal $Q^T = Q^{-1}$, entonces $B = I$, de modo que:

$$b_{ij} = \mathbf{c}_i \cdot \mathbf{c}_j = 0 \ \text{ si } \ i \neq j \ \text{ y } \ \mathbf{c}_i \cdot \mathbf{c}_j = 1 \ \text{ si } \ i = j.$$

Luego, las columnas de Q son ortonormales.

Ejemplo 1

Los vectores

$$\begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}, \begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix}, \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$$ forman una base ortonormal de $\mathbb{R}^3$ (verifíquelo),

así que la matriz $Q = \begin{bmatrix} 2/3 & 1/3 & -2/3 \\ 1/3 & 2/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}$ es ortogonal.

Podemos verificarlo haciendo $Q^{T} Q$:

$$Q^T Q = \begin{bmatrix} 2/3 & 1/3 & 2/3 \\ 1/3 & 2/3 & -2/3 \\ -2/3 & 2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 2/3 & 1/3 & -2/3 \\ 1/3 & 2/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

15.2 Isometrías

Antes de abordar el tema, veamos una propiedad interesante del producto escalar.

Teorema 2

Sea A una matriz de $m \times n$ con componentes reales. Entonces, para cualesquier dos vectores $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{y} \in \mathbb{R}^m$:

$$(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot A^T\mathbf{y}.$$

Demostración

$$A\mathbf{x} \cdot \mathbf{y} = (A\mathbf{x})^T \mathbf{y} = (\mathbf{x}^T A^T)\mathbf{y} = \mathbf{x}^T (A^T \mathbf{y}) = \mathbf{x} \cdot A^T \mathbf{y}.$$

El teorema nos dice que cuando en uno de los factores de un producto escalar hay una matriz, ésta puede pasar al otro factor transponiéndola. Este resultado puede generalizarse para cuando se tiene un producto interno en un espacio vectorial diferente a $\mathbb{R}^n$; en este caso el teorema se expresa así:

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^T\mathbf{y}).$$

Definición 2

Una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ es una isometría si para todo $\mathbf{x}$ en $\mathbb{R}^n$

$$|T(\mathbf{x})| = |\mathbf{x}|,$$

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Isometrías».

esto es, una isometría en $\mathbb{R}^n$ es una transformación lineal que conserva la longitud en $\mathbb{R}^n$.

Si hay una transformación lineal que sea una isometría, entonces:

$$|T(\mathbf{x})-T(\mathbf{y})| = |T(\mathbf{x}-\mathbf{y})| = |\mathbf{x}-\mathbf{y}|.$$

En las isometrías se conserva la magnitud de los vectores; recordemos que la magnitud de un vector en $\mathbb{R}^n$ está definida en términos del producto escalar, el cual es el producto interno en $\mathbb{R}^n$. Es, entonces, esperable que en las isometrías se conserve el producto escalar.

Teorema 3

Sea T una isometría de $\mathbb{R}^n \to \mathbb{R}^n$ y sean $\mathbf{x}$ e $\mathbf{y}$ en $\mathbb{R}^n$; entonces:

$$T(\mathbf{x}) \cdot T(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}.$$

Demostración

$$\begin{aligned} |T(\mathbf{x})-T(\mathbf{y})|^2 &= (T(\mathbf{x})-T(\mathbf{y})) \cdot (T(\mathbf{x})-T(\mathbf{y})) \\ &= |T(\mathbf{x})|^2 - 2T(\mathbf{x}) \cdot T(\mathbf{y}) + |T(\mathbf{y})|^2. \end{aligned} \qquad (1)$$

$$\begin{aligned} |\mathbf{x}-\mathbf{y}|^2 &= (\mathbf{x}-\mathbf{y}) \cdot (\mathbf{x}-\mathbf{y}) \\ &= |\mathbf{x}|^2 - 2\mathbf{x} \cdot \mathbf{y} + |\mathbf{y}|^2. \end{aligned} \qquad (2)$$

En (1) y (2) sabemos que $|T(\mathbf{x})|^2 = |\mathbf{x}|^2$, $|T(\mathbf{y})|^2 = |\mathbf{y}|^2$ y

$$|T(\mathbf{x})-T(\mathbf{y})|^2 = |\mathbf{x}-\mathbf{y}|^2.$$

Entonces:

$$\begin{aligned} -2T(\mathbf{x}) \cdot T(\mathbf{y}) &= -2\mathbf{x} \cdot \mathbf{y}, \\ T(\mathbf{x}) \cdot T(\mathbf{y}) &= \mathbf{x} \cdot \mathbf{y}. \end{aligned}$$

El siguiente teorema relaciona las isometrías con las matrices ortogonales.

Teorema 4

Una transformación lineal $T:\mathbb{R}^n \to \mathbb{R}^n$ es una isometría si y sólo si la representación matricial de T es ortogonal.

Demostración

a. Supongamos que T es una isometría y que A es su representación matricial; entonces, para $\mathbf{x}$ e $\mathbf{y}$ en $\mathbb{R}^n$:

$$\mathbf{x}\cdot\mathbf{y} \underset{T.3}{\underset{\downarrow}{=}} T(\mathbf{x})\cdot T(\mathbf{y}) = A\mathbf{x}\cdot A\mathbf{y} \underset{T.2}{\underset{\downarrow}{=}} \mathbf{x}\cdot A^T A\mathbf{y},$$

$$\mathbf{x}\cdot\mathbf{y} = \mathbf{x}\cdot A^T A\mathbf{y},$$

$$\mathbf{x}\cdot\mathbf{y} - \mathbf{x}\cdot A^T A\mathbf{y} = 0,$$

$$\mathbf{x}\cdot(\mathbf{y} - A^T A\mathbf{y}) = 0, \quad \text{para todo } \mathbf{x}\in\mathbb{R}^n; \text{ entonces}$$

$$(\mathbf{y} - A^T A\mathbf{y}) \in \mathbb{R}^{n\perp} \quad \text{(ver definición de complemento ortogonal)},$$

$$\mathbf{y} - A^T A\mathbf{y} = \mathbf{0}, \quad \mathbf{y} = A^T A\mathbf{y};$$

por lo tanto, $A^T A = I$, de modo que $A^T = A^{-1}$ y A es ortogonal.

b. Supongamos que $T(\mathbf{x}) = A\mathbf{x}$ es una transformación lineal tal que A es ortogonal; entonces $T(\mathbf{x})\cdot T(\mathbf{y}) = A\mathbf{x}\cdot A\mathbf{y} = \mathbf{x}\cdot A^T A\mathbf{y} = \mathbf{x}\cdot I\,\mathbf{y} = \mathbf{x}\cdot\mathbf{y}$

En particular, si $\mathbf{x} = \mathbf{y}$ $T(\mathbf{x})\cdot T(\mathbf{x}) = \mathbf{x}\cdot\mathbf{x}$, o sea $|T(\mathbf{x})|^2 = |\mathbf{x}|^2$, o $|T(\mathbf{x})| = |\mathbf{x}|$, de donde T es una isometría.

Observaciones

En el teorema anterior hemos mostrado que las isometrías se caracterizan porque su representación matricial es ortogonal. Como toda matriz ortogonal es invertible, se desprende como consecuencia la siguiente proposición:

Toda isometría es un isomorfismo.

Ahora, las columnas de A_T son las transformaciones de los vectores de la base, y como éstas forman una base ortonormal de $\mathbb{R}^n$, para constituir isometrías los vectores asignados a una base de $\mathbb{R}^n$ deben ser vectores que formen una base ortonormal. El siguiente teorema nos muestra cómo hacer esto.

Teorema 5

Sean $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ y $\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n\}$ dos bases ortonormales de $\mathbb{R}^n$ y sea $T:\mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal tal que $T(\mathbf{u}_i) = \mathbf{w}_i$ para $i = 1,..., n$; entonces T es una isometría.

Demostración

Sea $\mathbf{x}\in\mathbb{R}^n$, $\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n$; entonces

$$|\mathbf{x}|^2 = \mathbf{x}\cdot\mathbf{x} = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n)\cdot(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n).$$

Como $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ es una base ortonormal,

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0 \quad \text{para } i \neq j \quad \text{y}$$

$$|\mathbf{u}_i|^2 = \mathbf{u}_i \cdot \mathbf{u}_i = 1,$$

entonces:

$$|\mathbf{x}|^2 = c_1^2 + c_2^2 + ... + c_u^2. \qquad (1)$$

$$\begin{aligned} T(\mathbf{x}) &= c_1T(\mathbf{u}_1) + c_2T(\mathbf{u}_2) + ... + c_nT(\mathbf{u}_n) \\ &= c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + ... + c_n\mathbf{w}_n. \end{aligned}$$

$$|T(\mathbf{x})|^2 = T(\mathbf{x}) \cdot T(\mathbf{x}) = (c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + ... + c_n\mathbf{w}_n) \cdot (c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + ... + c_n\mathbf{w}_n).$$

Como $\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n\}$ es una base ortonormal,

$$|T(\mathbf{x})|^2 = c_1^2 + c_2^2 + ... + c_n^2. \qquad (2)$$

De (1) y (2) concluimos que $|\mathbf{x}|^2 = |T(\mathbf{x})|^2$ o $|\mathbf{x}| = |T(\mathbf{x})|$; por lo tanto, T es una isometría.

El recíproco del teorema anterior también se verifica, como veremos en el siguiente teorema.

Teorema 6

Sea $T: \mathbb{R}^n \to \mathbb{R}^n$ una isometría; entonces, si $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ es una base ortonormal de $\mathbb{R}^n$, $\{T(\mathbf{u}_1), T(\mathbf{u}_2), ..., T(\mathbf{u}_n)\}$ es una base ortonormal de $\mathbb{R}^n$.

Demostración

Sea $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ una base ortonormal de $\mathbb{R}^n$; entonces, $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ para $i \neq j$ y $|\mathbf{u}_i| = 1$.

Como T es una isometría:

$$T(\mathbf{u}_i) \cdot T(\mathbf{u}_j) = \mathbf{u}_i \cdot \mathbf{u}_j = 0 \quad \text{para } i \neq j \quad \text{y} \quad |T(\mathbf{u}_i)| = |\mathbf{u}_i| = 1.$$

Luego $\{T(\mathbf{u}_1), T(\mathbf{u}_2), ..., T(\mathbf{u}_n)\}$ es una base ortonormal.

15.3 Isometrías en $\mathbb{R}^2$

Sea T una isometría de $\mathbb{R}^2 \to \mathbb{R}^2$. Sean $\mathbf{u}_1 = T\begin{pmatrix}1\\0\end{pmatrix}$ y $\mathbf{u}_2 = T\begin{pmatrix}0\\1\end{pmatrix}$.

Los vectores $\begin{pmatrix}1\\0\end{pmatrix}$ y $\begin{pmatrix}0\\1\end{pmatrix}$ forman una base ortonormal en $\mathbb{R}^2$, luego $\mathbf{u}_1$ y $\mathbf{u}_2$ también forman una base ortonormal; $|\mathbf{u}_1| = |\mathbf{u}_2| = 1$ y $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$ $(\mathbf{u}_1 \perp \mathbf{u}_2)$ (figura 15.1).

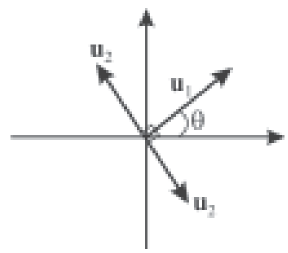


Figura 15.1

$$\mathbf{u}_1 = \begin{pmatrix}\cos\theta\\ \operatorname{sen}\theta\end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix}\cos(\theta \pm \pi/2)\\ \operatorname{sen}(\theta \pm \pi/2)\end{pmatrix}.$$

Entonces:

$$\mathbf{u}_2 = \begin{pmatrix}\cos(\theta + \pi/2)\\ \operatorname{sen}(\theta + \pi/2)\end{pmatrix} = \begin{pmatrix}-\operatorname{sen}\theta\\ \cos\theta\end{pmatrix} \text{ o } \mathbf{u}_2 = \begin{pmatrix}\cos(\theta - \pi/2)\\ \operatorname{sen}(\theta - \pi/2)\end{pmatrix} = \begin{pmatrix}\operatorname{sen}\theta\\ -\cos\theta\end{pmatrix}.$$

Luego la representación matricial de T es:

$$Q_1 = \begin{bmatrix}\cos\theta & -\operatorname{sen}\theta\\ \operatorname{sen}\theta & \cos\theta\end{bmatrix} \text{ o } Q_2 = \begin{bmatrix}\cos\theta & \operatorname{sen}\theta\\ \operatorname{sen}\theta & -\cos\theta\end{bmatrix}.$$

Si $T(\mathbf{x}) = Q_1\,\mathbf{x}$, entonces la transformación es una rotación de un ángulo θ en sentido antihorario.

Si $T(\mathbf{x}) = Q_2\,\mathbf{x}$,

$$Q_2 = \begin{bmatrix}\cos\theta & \operatorname{sen}\theta\\ \operatorname{sen}\theta & -\cos\theta\end{bmatrix} = \begin{bmatrix}\cos\theta & -\operatorname{sen}\theta\\ \operatorname{sen}\theta & \cos\theta\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}.$$

Entonces Q_2 es la representación matricial de una reflexión con respecto al eje x, seguida de una transformación de rotación.

Estos resultados los consignamos en el siguiente teorema.

Teorema 7

Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ una isometría. Entonces *T* es:

a. Una transformación de rotación.

b. Una reflexión con respecto al eje *x* seguida de una transformación de rota ción.

El concepto de isometría puede extenderse a espacios vectoriales de dimensión finita, donde haya definido un producto interno; acá las *isometrías* serán las transformaciones lineales que conservan la norma de los vectores; por lo tanto, el producto interno.

Ejemplo 2

Sea $T:\mathbb{R}^3 \to \mathbb{R}^3$ una transformación definida por $T(\mathbf{x}) = A\mathbf{x}$, donde:

$$A = \begin{bmatrix} \operatorname{sen}\theta & \cos\theta & 0 \\ \cos\theta & -\operatorname{sen}\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Veamos que *T* es una isometría. Para esto basta con demostrar que *A* es una matriz ortogonal y esto lo comprobamos realizando el producto AA^T.

$$\begin{bmatrix} \operatorname{sen}\theta & \cos\theta & 0 \\ \cos\theta & -\operatorname{sen}\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \operatorname{sen}\theta & \cos\theta & 0 \\ \cos\theta & -\operatorname{sen}\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Ejemplo 3

Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ una isometría. Veamos que *T* preserva los ángulos entre los vectores.

Sean $\mathbf{x}$ e $\mathbf{y}$ en $\mathbb{R}^2$

$$\text{ángulo entre } \mathbf{x} \text{ e } \mathbf{y} = \cos^{-1} \frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}|\ |\mathbf{y}|}, \tag{1}$$

$$\text{ángulo entre } T(\mathbf{x}) \text{ y } T(\mathbf{y}) = \cos^{-1} \frac{T(\mathbf{x})\cdot T(\mathbf{y})}{|T(\mathbf{x})|\ |T(\mathbf{y})|}. \tag{2}$$

Como *T* es isometría, $\mathbf{x}\cdot\mathbf{y} = T(\mathbf{x})\cdot T(\mathbf{y})$ y $|\mathbf{x}| = |T(\mathbf{x})|$, $|\mathbf{y}| = |T(\mathbf{y})|$;

por lo tanto, (1) = (2).

Ejercicios del capítulo 3 (módulo 15)

1. Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal tal que:

$$T\left(\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}}\right)=\left(\frac{1}{\sqrt{2}},\ -\frac{1}{\sqrt{2}}\right),$$

$$T\left(-\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}}\right)=\left(\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}}\right).$$

Demuestre que T es una isometría.

2. Considere la transformación lineal $T:M_{2\times 2} \to M_{2\times 2}$ para la cual:

$$T\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}=\begin{bmatrix}1/2 & 1/2\\ 1/2 & 1/2\end{bmatrix}, \qquad T\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}=\begin{bmatrix}1/2 & 1/2\\ -1/2 & -1/2\end{bmatrix},$$

$$T\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}=\begin{bmatrix}1/2 & -1/2\\ 1/2 & -1/2\end{bmatrix}, \qquad T\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1/2 & -1/2\\ -1/2 & 1/2\end{bmatrix}.$$

Compruebe que T es una isometría.

3. Determine el vector $(a,\ b)\in\mathbb{R}^2$ tal que la transformación $T:\mathbb{R}^2 \to \mathbb{R}^2$ para la cual $T(1,0)=\left[1/\sqrt{2},\ -1/\sqrt{2}\right]$, $T(0,1)=(a,b)$ sea una isometría.

4. Sea $B=\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ una base ortonormal de $\mathbb{R}^3$. Determine una isometría $T:\mathbb{R}^3 \to \mathbb{R}^3$, para la cual:

$$T(\mathbf{v}_1)=\frac{1}{3}(2\mathbf{v}_1+\mathbf{v}_2+2\mathbf{v}_3),\quad T(\mathbf{v}_2)=\frac{1}{\sqrt{2}}(\mathbf{v}_1-\mathbf{v}_3).$$

5. Dé un ejemplo de una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$ que preserve los ángulos y no sea una isometría.

6. Sea $T:\mathbb{R}^n \to \mathbb{R}^n$ una isometría y sea $T(\mathbf{x}) = A\mathbf{x}$. Demuestre que $S(\mathbf{x}) = A^{-1}\mathbf{x}$ es una isometría.

Un ingeniero trata de que las frecuencias naturales de su puente, valores característicos de la estructura, estén alejadas de las del viento para evitar que el puente comience a oscilar y se desplome. Los valores característicos son los rasgos más importantes de cualquier sistema dinámico.

Capítulo 4

Valores característicos, vectores característicos, diagonalización y formas canónicas

Contenido breve

Presentación

El movimiento de rotación de la Tierra alrededor del eje que determinan sus polos es una transformación lineal en el espacio tridimensional. En esta rotación, los puntos situados sobre este eje permanecen invariantes; así, una flecha que parta del centro de la Tierra hacia el Polo Sur conserva su magnitud y dirección; no ocurre lo mismo con una flecha que parta del centro hacia un punto sobre el Ecuador. Un *vector propio o característico* de esta transformación sería esa flecha a lo largo del eje polar y, como su longitud no se altera por la rotación, entonces el *valor propio* o *característico* correspondiente es 1.

En álgebra lineal, los *vectores característicos* o *propios* de una transformación lineal son los vectores no nulos que, cuando son transformados, dan lugar a un múltiplo escalar de sí mismos, es decir, no cambian su dirección. A este múltiplo escalar se le llama *valor propio*. Una transformación lineal, en un espacio vectorial, exhibe más claramente sus propiedades cuando se expresa con respecto a una base compuesta por vectores característicos, que son los que muestran los modos normales de funcionamiento del sistema estudiado. Cuando esto es posible, la matriz de la transformación es una matriz diagonal, y los elementos de la diagonal son los valores característicos de la transformación.

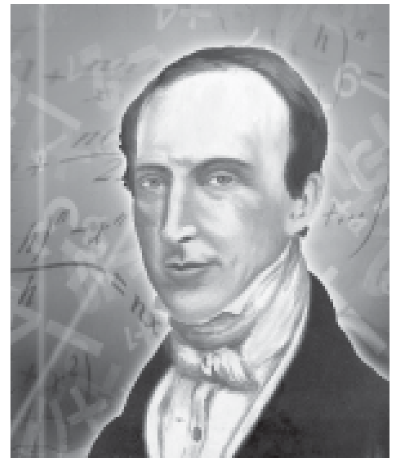
En este capítulo resolveremos dos problemas básicos del álgebra lineal. En primer lugar abordaremos el problema de la determinación de los valores y vectores característicos de una matriz $\underset{n\times n}{A}$, y en segundo término, el problema de la diagonalización de una matriz, esto es: encontrar una matriz diagonal que sea similar a una matriz dada A. Este problema está íntimamente relacionado con el primero.

Por último, como no toda matriz es diagonalizable, determinamos para cualquier matriz $\underset{n\times n}{A}$ su forma canónica de Jordan, la cual es una matriz mucho más «simple» que la matriz A. Estos procedimientos tendientes a determinar matrices similares a A «más simples» responden a la necesidad de simplificar el manejo algebraico cuando A es una matriz de tamaño grande.

En el desarrollo del capítulo se dará gran importancia a aplicaciones como los sistemas dinámicos descritos por un sistema de ecuaciones en diferencias y procesos de Markov.

Módulo 16

Valores y vectores característicos



Fue el matemático francés Augustin Louis Cauchy (1789-1857) quien, en 1840, usó por primera vez los términos «valores característicos» y «ecuación característica» para indicar los valores propios y la ecuación polinominal básica que satisfacen.

Introducción

Iniciamos nuestro estudio definiendo los valores y vectores característicos de una transformación lineal $T : V \to V$ o de la matriz A correspondiente a dicha transformación. Ilustramos con una matriz $A_{2\times 2}$ lo que sucede cuando referimos la transformación a la base formada con dos vectores característicos, y vemos que la matriz de la transformación en este caso es diagonal; esto hace que el efecto de la transformación se evidencie sobre las direcciones de los vectores característicos.

Se desarrolla un algoritmo para la obtención de los valores y vectores característicos y se establece una condición necesaria y suficiente para que una matriz tenga *n* vectores característicos linealmente independientes, es decir, para tener una base del espacio compuesta por vectores característicos.

Objetivos

1. Caracterizar en una transformación lineal $T : V \to V$ aquellos vectores que se transforman paralelamente a sí mismos y encontrar la magnitud de la transformación.
2. Desarrollar un algoritmo para encontrar los valores y vectores característicos de una matriz $A_{n\times n}$.
3. Establecer una condición suficiente y necesaria para que una matriz $A_{n\times n}$ tenga *n* vectores característicos linealmente independientes.
4. Relacionar los valores característicos de A con los valores característicos de A^T, αA, A^m y A^{-1} (cuando A^{-1} existe).

Preguntas básicas

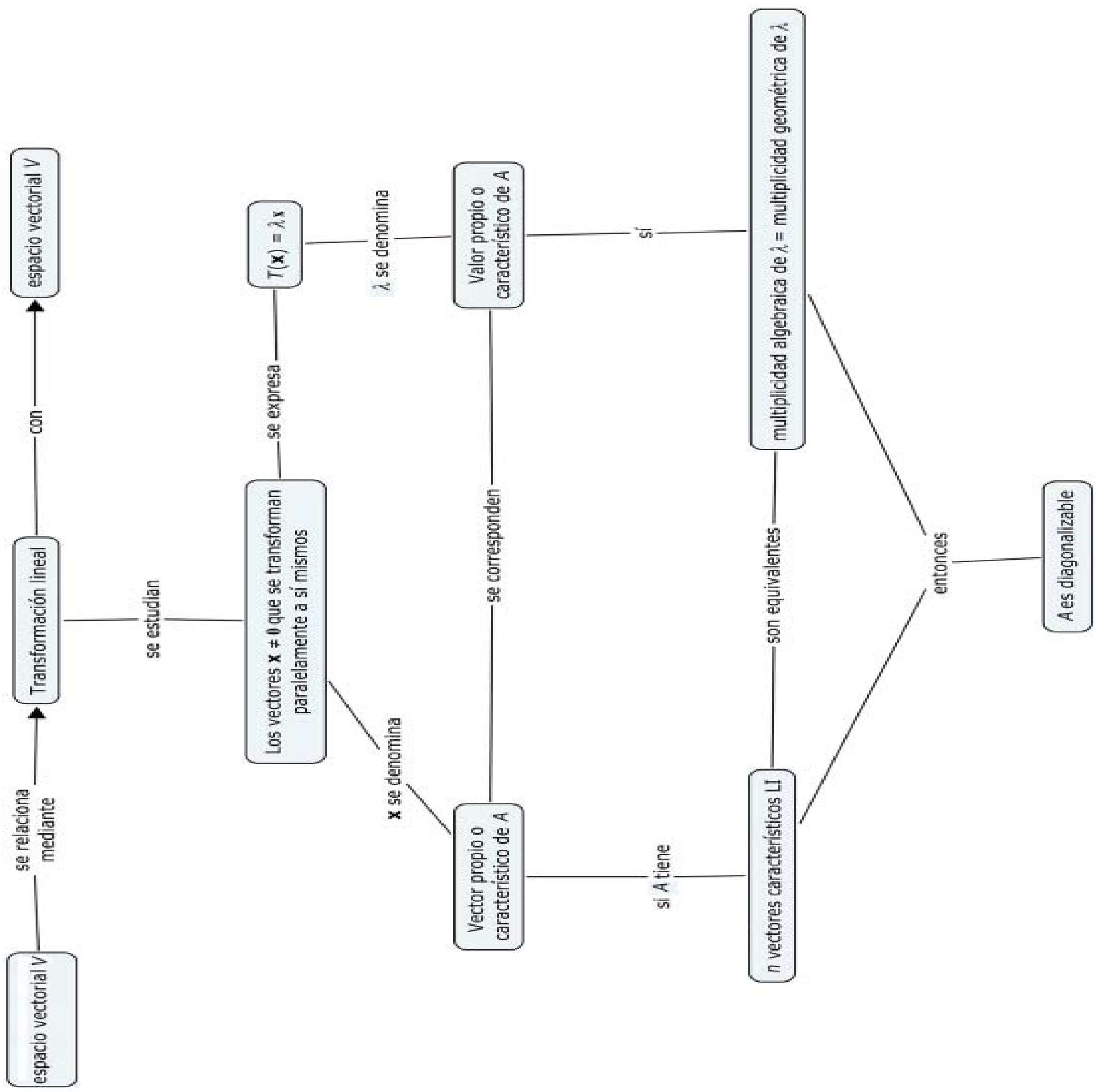
Sea $T : V \to V$ una transformación lineal y $A_{n\times n}$ una representación matricial de T.

1. ¿Qué es un vector característico de A?
2. ¿Qué es un valor característico de A?
3. ¿Cómo se encuentra el polinomio característico de A?
4. ¿Cómo se determinan los valores y vectores característicos de A?
5. ¿Qué es el espacio característico de un valor característico?
6. ¿Qué son las multiplicidades algebraica y geométrica de un valor característico?

Vea el módulo 16 del programa de televisión *Álgebra lineal*.

7. ¿Cómo se relacionan las multiplicidades algebraica y geométrica de cada valor característico?
8. ¿Qué condiciones deben cumplir los valores característicos de A para que A sea invertible?
9. ¿Cuándo la matriz $A_{n\times n}$ tiene n vectores característicos linealmente independientes?
10. Si $\lambda_1, \lambda_2, ..., \lambda_k$ son los valores característicos de A, ¿cuáles son los valores característicos de A^T, αA, A^m y A^{-1} (cuando A^{-1} existe)?

Contenidos



Mapa 10: módulos 16 y 17

16.1 Valores y vectores característicos

Definición 1

Sea *A* una matriz de $n \times n$ con componentes reales. El vector $\mathbf{x} \neq \mathbf{0}$, $\mathbf{x}$ en $\mathbb{C}^n$ tal que:

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

se llama *vector característico* de *A*, y al escalar λ (real o complejo) se le llama *valor característico* de *A* correspondiente al vector característico **x.**

Nota: los vectores característicos se toman en $\mathbb{C}^n$, *n*-tuplas ordenadas de números complejos, ya que estos pueden tener componentes imaginarios, cuando λ es un número complejo.

En la definición se dice que el valor característico λ corresponde al vector característico **x.** También se puede expresar al contrario: el vector característico **x** corresponde al valor característico λ.

A los valores y vectores característicos también se les llama valores y vectores *propios* o *autovalores* y *autovectores*.

16. 2 El problema de los valores y vectores característicos

Dada una matriz *A* de $n \times n$ con componentes reales, hallar todos los escalares λ y todos los vectores **x,** $\mathbf{x} \neq \mathbf{0}$, en $\mathbb{C}^n$, tales que $A\mathbf{x} = \lambda \mathbf{x}$.

Ejemplo 1

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

Entonces:

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

Así, $\lambda = 1$ es un valor característico de *A* correspondiente al vector característico:

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Similarmente:

$$A\begin{pmatrix}2\\3\end{pmatrix}=\begin{bmatrix}3&2\\3&4\end{bmatrix}\begin{bmatrix}2\\3\end{bmatrix}=\begin{bmatrix}12\\18\end{bmatrix}=6\begin{bmatrix}2\\3\end{bmatrix}.$$

Entonces $\lambda=6$ es un valor característico de A correspondiente al vector característico $\begin{bmatrix}2\\3\end{bmatrix}$.

La matriz $A=\begin{bmatrix}3&2\\3&4\end{bmatrix}$ es la representación matricial de la transformación lineal $T:\mathbb{R}^2\to\mathbb{R}^2$ definida por $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{bmatrix}3&2\\3&4\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}3x+2y\\3x+4y\end{bmatrix}$.

$\begin{bmatrix}1\\-1\end{bmatrix}$ y $\begin{bmatrix}2\\3\end{bmatrix}$ son vectores característicos de A correspondientes a los valores característicos $\lambda_1=1$ y $\lambda_2=6$.

Estos vectores son LI, luego forman una base B_1 de $\mathbb{R}^2$. Encontremos la matriz de la transformación referida a esta base.

$$T\begin{pmatrix}1\\-1\end{pmatrix}=\begin{bmatrix}1\\-1\end{bmatrix}=1\begin{bmatrix}1\\-1\end{bmatrix}+0\begin{bmatrix}2\\3\end{bmatrix},\quad\left(T\begin{pmatrix}1\\-1\end{pmatrix}\right)_{B_1}=\begin{pmatrix}1\\0\end{pmatrix}$$

$$T\begin{pmatrix}2\\3\end{pmatrix}=\begin{bmatrix}12\\18\end{bmatrix}=0\begin{bmatrix}1\\-1\end{bmatrix}+6\begin{bmatrix}2\\3\end{bmatrix},\quad\left(T\begin{pmatrix}2\\3\end{pmatrix}\right)_{B_1}=\begin{pmatrix}0\\6\end{pmatrix}.$$

Luego B_T, matriz de la transformación referida a la base B_1, es $B_T=\begin{bmatrix}1&0\\0&6\end{bmatrix}$, la cual es una matriz diagonal.

Geométricamente, podemos describir la situación así: $\begin{bmatrix}1\\-1\end{bmatrix}$ y $\begin{bmatrix}2\\3\end{bmatrix}$ determinan un sistema coordenado. Referida a este sistema, la transformación T tiene el efecto de dejar invariantes los vectores a lo largo de la dirección definida por $\begin{bmatrix}1\\-1\end{bmatrix}$, ya que $\lambda_1=1$, y aumentar con un factor de 6 los vectores a lo largo del eje definido por $\begin{bmatrix}2\\3\end{bmatrix}$.

Si **x** es un vector cualquiera de $\mathbb{R}^2$, el cual referido a esta base es $\begin{pmatrix}a\\b\end{pmatrix}$, entonces:

$$T(\mathbf{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ 6b \end{bmatrix} \text{ (figura 16.1).}$$

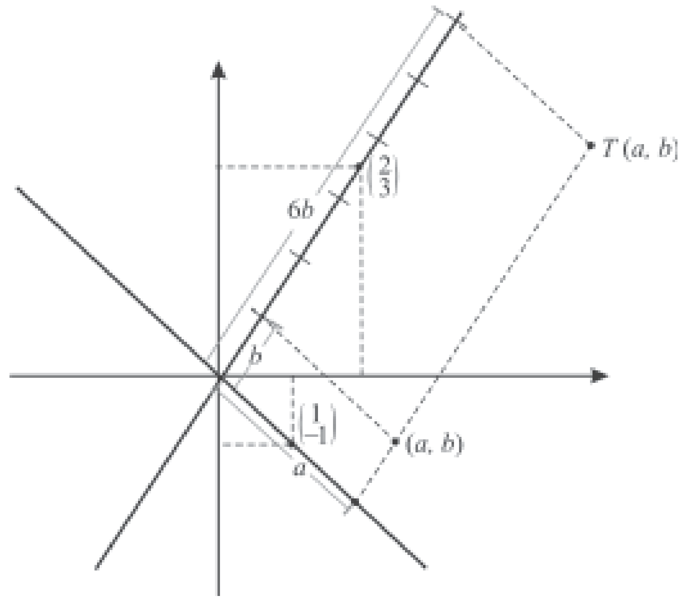


Figura 16.1

Ejemplo 2

Sea $A = I$; entonces, $I\mathbf{v} = \mathbf{v}$ para todo $\mathbf{v} \in \mathbb{R}^n$, luego todo $\mathbf{v} \neq \mathbf{0}$ es un vector característico de I con valor característico $\lambda = 1$.

Ejemplo 3

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

Determinemos los valores característicos de A y sus correspondientes vectores asociados. Debemos encontrar los escalares λ y los vectores no nulos $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ tales que:

$$\begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \lambda\begin{bmatrix} x \\ y \end{bmatrix},$$

luego:

$$\begin{aligned} 4x - 2y &= \lambda x \\ x + y &= \lambda y \end{aligned} \qquad \text{o} \qquad \begin{aligned} (4-\lambda)x - 2y &= 0 \\ x + (1-\lambda)y &= 0 \end{aligned}$$

Resolvemos el sistema homogéneo de dos ecuaciones con dos incógnitas. Este sistema tiene solución no trivial $\mathbf{x} \neq \mathbf{0}$, si y sólo si el determinante de su matriz de coeficientes es cero; esto es:

$$\begin{vmatrix} 4-\lambda & -2 \\ 1 & 1-\lambda \end{vmatrix} = 0\,.$$

Entonces:

$$(4-\lambda)(1-\lambda) + 2 = 0,$$
$$\lambda^2 - 5\lambda + 6 = 0 = (\lambda - 3)(\lambda - 2).$$

Luego, los valores característicos de A son $\lambda_1 = 3$ y $\lambda_2 = 2$.

Para encontrar los vectores característicos asociados con los valores de $\lambda_1 = 3$ y $\lambda_2 = 2$ resolvemos:

$$A\mathbf{x} = 3\mathbf{x} \qquad \text{y} \qquad A\mathbf{x} = 2\mathbf{x}$$

$$\begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 3\begin{bmatrix} x \\ y \end{bmatrix} \qquad \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 2\begin{bmatrix} x \\ y \end{bmatrix}$$

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

$$\text{o} \qquad\qquad \text{o}$$

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

$$\text{o} \qquad\qquad \text{o}$$

$$\begin{aligned} x - 2y &= 0 \\ x - 2y &= 0 \end{aligned} \qquad \begin{aligned} 2x - 2y &= 0 \\ x - y &= 0 \end{aligned}$$

cuyas soluciones están dadas por:

$$x = 2y$$

$$\mathbf{x}_1 = \begin{bmatrix} 2\gamma \\ \gamma \end{bmatrix},\ \gamma \in \left(\mathbb{R} - \{0\}\right)$$

las soluciones al sistema están dadas por:

$$x = y$$

$$\mathbf{x}_2 = \begin{bmatrix} \gamma \\ \gamma \end{bmatrix},\ \gamma \in \left(\mathbb{R} - \{0\}\right).$$

$\mathbf{x}_1$ y $\mathbf{x}_2$ están expresando, cada uno, un subespacio compuesto por todos los vectores múltiplos de $\begin{bmatrix}2\\1\end{bmatrix}$ y $\begin{bmatrix}1\\1\end{bmatrix}$, respectivamente. O, dicho de otra forma, los vectores característicos asociados con $\lambda_1 = 3$ son todos los múltiplos escalares del vector $\begin{bmatrix}2\\1\end{bmatrix}$; similarmente, los vectores característicos correspondientes a $\lambda_2 = 2$ son todos los múltiplos escalares del vector $\begin{bmatrix}1\\1\end{bmatrix}$.

Teorema 1

Sea A una matriz $n \times n$. λ es un valor característico de A si y sólo si $\det(A - \lambda I) = 0$.

Demostración

Si λ es un valor característico de A existe $\mathbf{x} \neq \mathbf{0}$ tal que:

$$A\mathbf{x} = \lambda\mathbf{x} = \lambda I\mathbf{x},$$

$$(A - \lambda I)\mathbf{x} = \mathbf{0}.$$

El sistema homogéneo $(A - \lambda I)\mathbf{x} = \mathbf{0}$ tiene solución para $\mathbf{x} \neq \mathbf{0}$ cuando $(A - \lambda I)$ es no invertible, o sea $\det(A - \lambda I) = 0$.

Recíprocamente, si $\det(A - \lambda I) = 0$, el sistema $(A - \lambda I)\mathbf{x} = \mathbf{0}$ tiene soluciones no triviales y λ es un valor característico de A. Si $\det(A - \lambda I) \neq 0$, entonces la única solución del sistema es $\mathbf{x} = \mathbf{0}$; por lo tanto, λ no es un valor característico de A.

16.2.1 El polinomio característico de $A_{n\,xn}$

Definición 2

Sea A una matriz $n \times n$. El determinante de la matriz $(A - \lambda I)$, denotado $P(\lambda)$, se denomina *polinomio característico de A*.

$$P(\lambda) = \det(A - \lambda I).$$

La ecuación $P(\lambda) = 0$ es la *ecuación característica* de A.

$$P(\lambda) = \det(A - \lambda I) = 0.$$

$P(\lambda)$ es un polinomio de grado n en λ.

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots a_{1n} \\ a_{21} & a_{22} & \dots a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots a_{nn} \end{pmatrix}, \quad \det(A-\lambda I)=\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots a_{1n} \\ a_{21} & a_{22}-\lambda & \dots a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots a_{nn}-\lambda \end{vmatrix}$$

$$P(\lambda) = \lambda^n + b_{n-1}\lambda^{n-1} + ... + b_1\lambda + b_0 = 0.$$

Esta ecuación tiene *n* raíces, varias de las cuales pueden repetirse. Si $\lambda_1, \lambda_2, ..., \lambda_m$ son las raíces diferentes de la ecuación con multiplicidades $\gamma_1, \gamma_2, ..., \gamma_m$, entonces $P(\lambda)$ puede factorizarse como:

$$P(\lambda) = (\lambda-\lambda_1)^{\gamma_1}(\lambda-\lambda_2)^{\gamma_2}...(\lambda-\lambda_m)^{\gamma_m} = 0.$$

Los números $\gamma_1, \gamma_2, ..., \gamma_m$ se llaman *multiplicidades algebraicas* de los valores característicos $\lambda_1, \lambda_2, ..., \lambda_m$, respectivamente.

O sea que si contamos multiplicidades, cada matriz $n \times n$ tiene exactamente *n* valores característicos.

16.2.2 Proceso para encontrar los valores y vectores característicos en una matriz $A_{n\times n}$

i. Halle $P(\lambda) = \det(A-\lambda I)$.

ii. Halle las raíces de $P(\lambda) = 0 \quad (\lambda_1, \lambda_2, ..., \lambda_m)$.

iii. Resuelva el sistema homogéneo $(A-\lambda_i I)\mathbf{x} = \mathbf{0}$ correspondiente a cada valor característico λ_i.

La solución del sistema homogéneo $(A-\lambda_i I)\mathbf{x} = \mathbf{0}$ no es única, ya que $\mathbf{x} \neq \mathbf{0}$. El sistema tiene infinitas soluciones; luego, al determinar los vectores **x** que satisfacen el sistema, se sacan los vectores que formen una base para el subespacio generado; es claro que todas las combinaciones lineales diferentes de cero, que se realicen con estos vectores, también son vectores característicos correspondientes al mismo valor λ.

Teorema 2

Sea *A* una matriz $n \times n$ y sea $E_\lambda = \{\mathbf{x}\colon A\mathbf{x} = \lambda\mathbf{x}\}$; entonces E_λ es un subespacio de $\mathbb{C}^n$.

Demostración

Si $A\mathbf{x}=\lambda\mathbf{x}$, entonces $(A-\lambda I)\mathbf{x}=\mathbf{0}$.

Luego E_λ es el núcleo de $(A-\lambda I)$; por lo tanto, E_λ es un subespacio de $\mathbb{C}^n$.

16.2.3 Espacio característico de *A*

Definición 3

Sea λ un valor característico de *A*. El subespacio E_λ se llama *espacio característico* de *A*, correspondiente al valor característico λ.

Ejemplo 4

Sea $A=\begin{bmatrix}1 & 2 & -1\\ 1 & 0 & 1\\ 4 & -4 & 5\end{bmatrix}$. Encontremos para A los valores y vectores característicos.

Vamos a seguir los pasos dados en el proceso:

$$\det(A-\lambda I)=\begin{vmatrix}1-\lambda & 2 & -1\\ 1 & -\lambda & 1\\ 4 & -4 & 5-\lambda\end{vmatrix}=0$$

$$=-\lambda^3+6\lambda^2-11\lambda+6=0$$

$$=(\lambda-1)(\lambda-2)(\lambda-3)=0.$$

Entonces los valores característicos de A son $\lambda_1=1,\ \lambda_2=2,\ \lambda_3=3.$

Para determinar E_1, formamos el sistema $(A-I)\mathbf{x}=\mathbf{0}$.

$$\begin{bmatrix}1-1 & 2 & -1\\ 1 & -1 & 1\\ 4 & -4 & 5-1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix} \qquad \begin{bmatrix}0 & 2 & -1\\ 1 & -1 & 1\\ 4 & -4 & 4\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

$$\begin{bmatrix}0 & 2 & -1\\ 1 & -1 & 1\\ 4 & -4 & 4\end{bmatrix}\xrightarrow{\text{operaciones elementales}}\begin{bmatrix}1 & 0 & \tfrac{1}{2}\\ 0 & 1 & -\tfrac{1}{2}\\ 0 & 0 & 0\end{bmatrix}.$$

La solución al sistema está dada por:

$$x = -\frac{1}{2}z$$
$$y = \frac{1}{2}z$$
$$z = 1z$$

$$E_1 = \text{gen}\left\{\begin{pmatrix}-1\\1\\2\end{pmatrix}\right\}; \quad \begin{pmatrix}-1\\1\\2\end{pmatrix}$$ es un vector característico asociado con $\lambda = 1$.

Determinemos E_2.

$$(A-2I)\mathbf{x} = \mathbf{0}.$$

$$\begin{bmatrix}-1 & 2 & -1\\1 & -2 & 1\\4 & -4 & 3\end{bmatrix} \xrightarrow{\text{operaciones elementales}} \begin{bmatrix}1 & 0 & \frac{1}{2}\\0 & 1 & -\frac{1}{4}\\0 & 0 & 0\end{bmatrix}.$$

La solución es:

$$x = -\frac{1}{2}z$$
$$y = \frac{1}{4}z$$
$$z = 1z$$

$$E_2 = \text{gen}\left\{\begin{pmatrix}-2\\1\\4\end{pmatrix}\right\}; \quad \begin{pmatrix}-2\\1\\4\end{pmatrix}$$ es un vector característico asociado con $\lambda = 2$.

Similarmente, encontremos E_3.

$$(A-3I)\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix}-2 & 2 & -1\\1 & -3 & 1\\4 & -4 & 2\end{bmatrix} \xrightarrow{\text{operaciones elementales}} \begin{bmatrix}1 & 0 & \frac{1}{4}\\0 & 1 & -\frac{1}{4}\\0 & 0 & 0\end{bmatrix}.$$

La solución es:

$$x = -\frac{1}{4}z$$
$$y = \frac{1}{4}z$$
$$z = 1z$$

$$E_3 = \text{gen}\left\{\begin{pmatrix}-1\\1\\4\end{pmatrix}\right\}; \quad \begin{pmatrix}-1\\1\\4\end{pmatrix} \text{ es un vector característico asociado con } \lambda = 3.$$

Ejemplo 5

Sea $A = \begin{bmatrix}1 & -1 & 0\\-1 & 2 & -1\\0 & -1 & 1\end{bmatrix}$. Determinemos los valores y vectores característicos de *A*.

$$\det(A-\lambda I) = \begin{vmatrix}1-\lambda & -1 & 0\\-1 & 2-\lambda & -1\\0 & -1 & 1-\lambda\end{vmatrix}$$

$$= -\lambda^3 + 4\lambda^2 - 3\lambda = 0$$
$$= -\lambda(\lambda-1)(\lambda-3) = 0.$$

Los valores característicos de *A* son $\lambda_1 = 0,\ \lambda_2 = 1,\ \lambda_3 = 3.$

Ahora encontremos los espacios característicos correspondientes a cada valor de λ.

$E_0, \quad (A-0I)\mathbf{x} = \mathbf{0}.$

$$\begin{bmatrix}1 & -1 & 0\\-1 & 2 & -1\\0 & -1 & 1\end{bmatrix} \longrightarrow \begin{bmatrix}1 & 0 & -1\\0 & 1 & -1\\0 & 0 & 0\end{bmatrix}, \text{ entonces } \quad \begin{aligned}x &= 1z\\y &= 1z\\z &= 1z\end{aligned}$$

$$E_0 = \text{gen}\left\{\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}.$$

$E_1, \quad (A-I)\mathbf{x} = \mathbf{0}.$

$$\begin{bmatrix}0 & -1 & 0\\-1 & 1 & -1\\0 & -1 & 0\end{bmatrix} \longrightarrow \begin{bmatrix}1 & 0 & 1\\0 & 1 & 0\\0 & 0 & 0\end{bmatrix}, \text{ entonces } \quad \begin{aligned}x &= -1z\\y &= 0z\\z &= 1z\end{aligned}$$

$$E_1 = \text{gen}\left\{\begin{pmatrix}-1\\0\\1\end{pmatrix}\right\}.$$

E_3, $\quad (A-3I)\mathbf{x} = \mathbf{0}$.

$$\begin{bmatrix}-2 & -1 & 0\\-1 & -1 & -1\\0 & -1 & -2\end{bmatrix} \longrightarrow \begin{bmatrix}1 & 0 & -1\\0 & 1 & 2\\0 & 0 & 0\end{bmatrix}, \text{ entonces } \begin{aligned} x &= \ \ 1z\\ y &= -2z\\ z &= \ \ 1z\end{aligned}$$

$$E_3 = \text{gen}\left\{\begin{pmatrix}1\\-2\\1\end{pmatrix}\right\}.$$

Ejemplo 6

Sea $A = \begin{bmatrix}0 & 1 & 0\\0 & 0 & 1\\1 & -3 & 3\end{bmatrix}$, entonces:

$$\det(A-\lambda I) = \begin{vmatrix}-\lambda & 1 & 0\\0 & -\lambda & 1\\1 & -3 & 3-\lambda\end{vmatrix} = 0$$

$$\lambda^3 - 3\lambda^2 + 3\lambda - 1 = 0,$$
$$(\lambda - 1)^3 = 0.$$

Luego $\lambda = 1$ es el único valor característico de A con multiplicidad algebraica de 3. Ahora veamos cuál es el espacio característico de $\lambda = 1$. Resolvemos $(A-I)\mathbf{x} = \mathbf{0}$.

$$E_1: \begin{bmatrix}-1 & 1 & 0\\0 & -1 & 1\\1 & -3 & 2\end{bmatrix} \longrightarrow \begin{bmatrix}1 & 0 & -1\\0 & 1 & -1\\0 & 0 & 0\end{bmatrix}, \text{ entonces } \begin{aligned} x &= z\\ y &= z\\ z &= z\end{aligned}$$

$$E_1 = \text{gen}\left\{\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}.$$

Ejemplo 7

Sea $A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$; entonces:

$$\det(A-\lambda I) = \begin{vmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{vmatrix} = 0$$

$$\det(A-\lambda I) = -\lambda^3 + 6\lambda^2 + 15\lambda + 8 = -(\lambda+1)^2(\lambda-8) = 0.$$

Luego $\lambda_1 = -1$ con multiplicidad algebraica de 2 y $\lambda_2 = 8$.

Determinemos ahora los espacios característicos.

$$E_{-1} : (A+I)\mathbf{x} = \mathbf{0}.$$

$$\begin{bmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ entonces } \begin{aligned} x &= -\frac{1}{2}y - z \\ y &= y \\ z &= z \end{aligned}$$

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

$$E_{-1} : \text{gen}\left\{\begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}.$$

$$E_8 : (A-8I)\mathbf{x} = \mathbf{0}.$$

$$\begin{bmatrix} -5 & 2 & 4 \\ 2 & -8 & 2 \\ 4 & 2 & -5 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}, \text{ entonces } \begin{aligned} x &= 1z \\ y &= \frac{1}{2}z \\ z &= 1z \end{aligned}$$

$$E_8 : \text{gen}\left\{\begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix}\right\}.$$

Observación: la factorización del polinomio característico de A, en general, no es obvia. El álgebra nos brinda dos resultados que pueden ser útiles a este respecto:

(1) El producto de todas las raíces del polinomio

$P(\lambda) = \lambda^n + b_{n-1}\lambda^{n-1} + ... + b_1\lambda + b_0 = 0$ es $(-1)^n b_0$.

(2) Si $b_{n-1},..., b_1, b_0$ son enteros, entonces $P(\lambda)$ no puede tener una raíz racional que no sea un entero. Luego las posibles raíces racionales de $P(\lambda)$ serán factores enteros de b_0. Por supuesto, $P(\lambda)$ podría tener raíces irracionales; además, $P(\lambda)$ también podrá tener raíces complejas; si esto sucede, éstas ocurren en pares conjugados.

Ejemplo 8

Sea $A = \begin{bmatrix} 2 & -1 \\ 5 & -2 \end{bmatrix}$. Determinemos los valores y vectores característicos de *A*.

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & -1 \\ 5 & -2-\lambda \end{vmatrix} = 0, \text{ entonces } \lambda^2 + 1 = 0.$$

De donde, $\lambda^2 = -1$ y $\lambda = \pm\sqrt{-1}$, o sea $\lambda = \pm\, i$, $\lambda_1 = i$ y $\lambda_2 = -i$.

Veamos los espacios característicos.

E_i : $(A - iI)\mathbf{x} = \mathbf{0}$.

$$\begin{bmatrix} 2-i & -1 \\ 5 & -2-i \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \text{ luego } \begin{array}{l} (2-i)x - y = 0 \\ (2-i)x = y \end{array}$$

Entonces, si $x = 1$, $y = 2 - i$.

Por lo tanto, $\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2-i \end{pmatrix}$ es un vector característico correspondiente a $\lambda = i$ y

$$E_i : \text{ gen}\left\{\begin{pmatrix} 1 \\ 2-i \end{pmatrix}\right\}.$$

Ahora, E_{-i} resulta de resolver $(A + iI)\mathbf{x} = \mathbf{0}$.

$$\begin{bmatrix} 2+i & -1 \\ 5 & -2+i \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \text{ lo cual lleva a } (2+i)x - y = 0.$$

Ahora, si $x = 1$, $y = (2+i)$, y entonces $\mathbf{x}_2 = \begin{pmatrix} 1 \\ 2+i \end{pmatrix}$.

Observación: note que $\lambda_2 = -i$ es el conjugado complejo de $\lambda_1 = i$, y además los componentes de $\mathbf{x}_2$ son conjugados complejos de los componentes de $\mathbf{x}_1$. Este hecho no es casual y se puede demostrar que los valores característicos de una matriz real ocurren en pares conjugados complejos y los vectores característicos correspondientes son conjugados complejos entre sí.

16.3 Propiedades

Teorema 3

Sea A una matriz $n\times n$ y sean $\lambda_1, \lambda_2, ..., \lambda_m$ valores característicos diferentes de A con sus correspondientes vectores característicos $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$. Entonces $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$ son LI.

Demostración

Razonamos por inducción sobre m.

1. $m = 2$. Suponga que $C_1\mathbf{x}_1 + C_2\mathbf{x}_2 = \mathbf{0}$. (1)

   Hagamos (1) $\times$ A: $C_1A\mathbf{x}_1 + C_2A\mathbf{x}_2 = \mathbf{0}$.

   Como $\mathbf{x}_1$ y $\mathbf{x}_2$ son vectores característicos correspondientes a valores característicos λ_1 y λ_2, entonces $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$, $i = 1, 2$.

   Luego $C_1\lambda_1\mathbf{x}_1 + C_2\lambda_2\mathbf{x}_2 = \mathbf{0}$. (2)

   Multiplicamos (1) por λ_2 y se resta de (2).

   $$C_1(\lambda_1 - \lambda_2)\mathbf{x}_1 + C_2\left(\cancelto{0}{\lambda_2 - \lambda_2}\right)\mathbf{x}_2 = \mathbf{0},$$

   $$C_1(\lambda_1 - \lambda_2)\mathbf{x}_1 = \mathbf{0}.$$

   En esta igualdad $\mathbf{x}_1 \neq \mathbf{0}$ ya que es un vector característico; $(\lambda_1 - \lambda_2) \neq 0$ porque $\lambda_1 \neq \lambda_2$, entonces $C_1 = 0$. Sustituimos este valor en (1) y se obtiene $C_2\mathbf{x}_2 = \mathbf{0}$; por lo tanto, $C_2 = 0$ y concluimos que $\mathbf{x}_1$ y $\mathbf{x}_2$ son LI.

2. Supongamos que para $m = k$ se cumple que $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ son LI y veamos para $m = k + 1$.

3. Sea $C_1\mathbf{x}_1 + C_2\mathbf{x}_2 + ... + C_k\mathbf{x}_k + C_{k+1}\mathbf{x}_{k+1} = \mathbf{0}$. (3)

   Hagamos (3) $\times$ A:

   $$C_1A\mathbf{x}_1 + C_2A\mathbf{x}_2 + ... + C_kA\mathbf{x}_k + C_{k+1}A\mathbf{x}_{k+1} = \mathbf{0},$$

y aplicando el hecho de que $A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \; i = 1, ..., k+1,$ tenemos:

$$C_1\lambda_1\mathbf{x}_1 + C_2\lambda_2\mathbf{x}_2 + ... + C_k\lambda_k\mathbf{x}_k + C_{k+1}\lambda_{k+1}\mathbf{x}_{k+1} = \mathbf{0}. \tag{4}$$

Multiplicamos (3) por λ_{k+1} y lo restamos de (4).

$$C_1(\lambda_1 - \lambda_{k+1})\mathbf{x}_1 + C_2(\lambda_2 - \lambda_{k+1})\mathbf{x}_2 + ... + C_k(\lambda_k - \lambda_{k+1})\mathbf{x}_k$$
$$+ C_{k+1}\left(\cancelto{0}{\lambda_{k+1} - \lambda_{k+1}}\right)\mathbf{x}_{k+1} = \mathbf{0}.$$
$$C_1(\lambda_1 - \lambda_{k+1})\mathbf{x}_1 + C_2(\lambda_2 - \lambda_{k+1})\mathbf{x}_2 + ... + C_k(\lambda_k - \lambda_{k+1})\mathbf{x}_k = \mathbf{0}.$$

Como según la hipótesis de inducción $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ son LI, entonces:

$$C_1(\lambda_1 - \lambda_{k+1}) = C_2(\lambda_2 - \lambda_{k+1}) = ... = C_k(\lambda_k - \lambda_{k+1}) = 0.$$

Como los valores de $\lambda_i, \; i = 1, ..., k+1$ son diferentes, entonces:

$$C_1 = C_2 = ... = C_k = 0.$$

Llevamos estos valores a (3) y obtenemos $C_{k+1}\mathbf{x}_{k+1} = \mathbf{0}$, de donde $C_{k+1} = 0$.

Concluimos entonces que $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m, \mathbf{x}_{m+1}$ son LI, completando la prueba para todo m.

El teorema anterior establece que vectores característicos correspondientes a valores característicos diferentes son linealmente independientes.

El resultado del teorema puede verificarse en los ejemplos desarrollados. Además, se observa que en algunos casos se encuentran tantos vectores característicos linealmente independientes como la multiplicidad algebraica del valor λ (ejemplos 4, 5 y 7). En el ejemplo 7 el valor $\lambda = -1$ tiene una multiplicidad algebraica 2 y se determinan dos vectores en la base del espacio característico correspondiente, esto es, dos vectores característicos LI. Sin embargo, esto no siempre es así; en el ejemplo 6 el valor caraterístico $\lambda = 1$ tiene una multiplicidad algebraica 3 y un solo vector característico LI asociado con él.

Definición 4

Sea λ un valor característico de A. Entonces la *multiplicidad geométrica de* λ es la dimensión del espacio característico correspondiente a λ (nulidad de la matriz $A - \lambda I$).

Para cada valor característico debe haber asociado al menos un vector característico. Si la multiplicidad algebraica de λ es mayor que 1, por ejemplo si A es una matriz 3×3 con dos valores característicos λ_1 y λ_2 con multiplicidades algebraicas 1 y 2, respectivamente, entonces para λ_1 habrá asociado un vector característico y para λ_2 podrá haber uno o dos vectores característicos LI. Es claro que no podrán ser más de dos, ya que en el caso de que fueran tres vectores, se completarían cuatro vectores LI en un espacio vectorial de dimensión 3, lo cual es absurdo.

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Valores y vectores característicos».

Estas consideraciones quedan expresadas en el siguiente teorema.

Teorema 4

Sea λ un valor característico de *A*. Entonces:

multiplicidad geométrica de $\lambda \leq$ multiplicidad algebraica de λ.

Teorema 5

Sea *A* una matriz $n \times n$. *A* tiene *n* vectores característicos LI si y sólo si la multiplicidad geométrica de todo valor característico es igual a la multiplicidad algebraica. En particular, *A* tiene *n* vectores característicos LI si todos sus valores característicos son diferentes.

La deducción del teorema resulta evidente de los resultados establecidos en los teoremas 3 y 4.

Teorema 6

Los valores característicos de una matriz triangular son las componentes diagonales de la matriz.

La demostración del teorema se deja como ejercicio.

En el ejemplo 5 la matriz *A* tenía un valor característico igual a 0, o sea que $\det(A-0I)=0$, luego $\det A=0$, lo cual significa que *A* no es invertible.

Teorema 7

A es invertible si y sólo si $\lambda = 0$ no es un valor característico de *A*.

La demostración del teorema se deja como ejercicio.

Finalizamos esta sección destacando algunas relaciones entre los valores característicos de *A* y los valores característicos de matrices relacionadas con *A*.

Teorema 8

Sea $A_{n\times n}$ una matriz que tiene valores característicos $\lambda_1, \lambda_2, \ldots, \lambda_k$; entonces:

a. Los valores característicos de A^T son $\lambda_1, \lambda_2, \ldots, \lambda_k$.

b. Los valores característicos de αA son $\alpha\lambda_1, \alpha\lambda_2, \ldots, \alpha\lambda_k$.

c. Los valores característicos de A^m son $\lambda_1^m, \lambda_2^m, \ldots, \lambda_k^m$ para $m = 1, 2, 3\ldots$

d. Si A^{-1} existe, los valores característicos de A^{-1} son $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_k}$.

Demostración

a. Si λ_i es valor característico de A, $\det(A-\lambda_i I) = 0$.

Ahora, $\det(A-\lambda_i I) = \det(A-\lambda_i I)^T$

$$= \det(A^T - \lambda_i I^T)$$

$$= \det(A^T - \lambda_i I).$$

Luego λ_i, $i = 1,..., k$, es un valor característico de A^T.

c. Razonemos por inducción.

Para $m = 2$. Si λ_i es un valor característico de A,

$$A\mathbf{v} = \lambda_i \mathbf{v}. \tag{1}$$

Multiplicando (1) por A,

$$A\,A\mathbf{v} = A\lambda_i \mathbf{v}, \quad A^2\mathbf{v} = \lambda_i A\mathbf{v}.$$

Aplicando (1) tenemos:

$$A^2\mathbf{v} = \lambda_i\lambda_i \mathbf{v} = \lambda_i^{\,2}\mathbf{v}.$$

Esta última ecuación significa que λ_i^2 es un valor característico de A^2.

Supongamos que para $m = k$ se cumple que $A^k\mathbf{v} = \lambda_i^{\,k}\mathbf{v}$. (2)

Veamos para $m = k + 1$.

Multiplicando (2) por A:

$$\begin{aligned} A\,A^k\mathbf{v} &= A\lambda_i^{\,k}\,\mathbf{v}, \\ A^{k+1}\mathbf{v} &= \lambda_i^{\,k} A\mathbf{v}. \end{aligned} \tag{3}$$

Aplicando (1) en (3):

$$A^{k+1}\mathbf{v} = \lambda_i^{\,k}\lambda_i\mathbf{v} = \lambda_i^{\,k+1}\mathbf{v}.$$

Luego λ_i^{k+1} es un valor característico de A^{k+1} y esto completa la prueba.
Los literales *b* y *d* se dejan como ejercicio.

Ejercicios del capítulo 4 (módulo 16)

1. Sea $A=\begin{bmatrix}2 & 2\\ 2 & 2\end{bmatrix}$. Determine cuáles de los siguientes vectores de $\mathbb{R}^2$ son vectores característicos de A; en caso de serlo, determine el valor característico asociado.

a. $(2,\ -1)$.
b. $(2,\ 2)$.
c. $(3,\ -3)$.
d. $(4,\ 4)$.
e. $(-6,\ 6)$.
f. $(-1,\ 2)$.

En los ejercicios 2 a 12 encuentre los valores y vectores característicos de la matriz dada.

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

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

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

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

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

7. $\begin{bmatrix}0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & -2 & 4\end{bmatrix}$.

8. $\begin{bmatrix}-3 & -7 & -5\\ 2 & 4 & 3\\ 1 & 2 & 2\end{bmatrix}$.

9. $\begin{bmatrix}a & 0 & 0 & 0\\ 0 & a & 0 & 0\\ 0 & 0 & a & 0\\ 0 & 0 & 0 & a\end{bmatrix}$.

10. $\begin{bmatrix}a & b & 0 & 0\\ 0 & a & 0 & 0\\ 0 & 0 & a & 0\\ 0 & 0 & 0 & a\end{bmatrix}$; $b\neq 0$.

11. $\begin{bmatrix}a & b & 0 & 0\\ 0 & a & c & 0\\ 0 & 0 & a & 0\\ 0 & 0 & 0 & a\end{bmatrix}$; $bc\neq 0$.

12. $\begin{bmatrix}a & b & 0 & 0\\ 0 & a & c & 0\\ 0 & 0 & a & d\\ 0 & 0 & 0 & a\end{bmatrix}$; $bcd\neq 0$.

13. Sea A una matriz diagonal de orden n cuyos elementos de la diagonal principal son $\lambda_1, \lambda_2, ..., \lambda_n$. Determine el polinomio característico de A y sus valores característicos.

14. Sea A una matriz triangular de orden n. Determine el polinomio característico de A, así como sus valores característicos.

15. Suponga que λ_1 es un valor característico de la matriz A, y λ_2 es un valor característico de la matriz B. ¿Es $\lambda_1 + \lambda_2$ un valor característico de $A + B$?

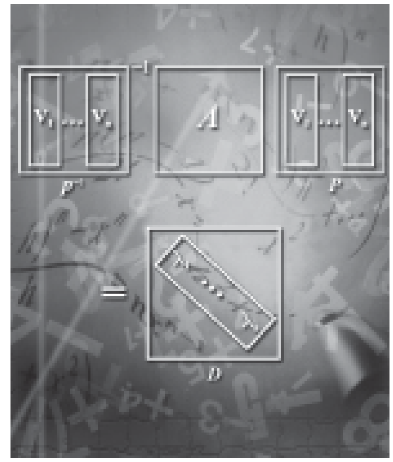
16. Realice una demostración del teorema 7.

17. Demuestre las partes *b* y *d* del teorema 8.

18. Sea A una matriz real de $n \times n$. Demuestre que si λ_1 es un valor característico complejo de A con vector característico $\mathbf{v}_1$, entonces $\overline{\lambda_1}$ es un valor característico de A con vector característico $\overline{\mathbf{v}}_1$.

Módulo 17

El problema de la diagonalización



El problema de la diagonalización consiste en encontrar una matriz D tal que $D = P^{-1}AP$.

Introducción

El problema de la diagonalización puede enunciarse de la siguiente forma: «Dada una matriz A de $n \times n$ encontrar, si es posible, una matriz D diagonal, similar o semejante a la matriz A».

Este problema está íntimamente relacionado con el problema de la determinación de los valores y vectores característicos de A, como veremos en el desarrollo siguiente.

Objetivos

1. Establecer una condición necesaria y suficiente para que una matriz sea diagonalizable.
2. Utilizar las propiedades de las matrices semejantes para trabajar algebraicamente con A a través de su matriz diagonal equivalente.

Preguntas básicas

1. ¿Cuándo una matriz $A_{n\times n}$ es diagonalizable?
2. Si A y B son matrices semejantes, ¿cómo son sus polinomios característicos y sus valores característicos?

Contenidos

17.1 Diagonalización

17.2 Condición necesaria y suficiente para que una matriz $A_{n\times n}$ sea diagonalizable

Vea el módulo 17 del programa de televisión *Álgebra lineal*.

17.1 Diagonalización

Definición 1

Decimos que la matriz $A_{n\times n}$ es *diagonalizable* si es similar a una matriz diagonal *D*. Es decir, si existe una matriz invertible *P* tal que:

$$D = P^{-1}AP.$$

Ejemplo 1

La matriz $A = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}$ dada en el ejemplo 1 del módulo 16 es diagonalizable, ya que *A* es similar a $B_T = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}$. Ambas matrices son representaciones de la transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x+2y \\ 3x+4y \end{bmatrix}$. La matriz *A* está referida a la base estándar y la matriz B_T está referida a la base formada por los vectores característicos LI $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ y $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Estos vectores forman las dos columnas de la matriz *P*.

Podemos verificar que:

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

Teorema 1

Si *A* y *B* son matrices similares de $n \times n$, entonces *A* y *B* tienen la misma ecuación característica y, por consiguiente, los mismos valores característicos.

Demostración

Si *A* y *B* son similares, entonces existe una matriz *P* tal que $B = P^{-1}AP;$ por lo tanto,

$$\det(B-\lambda I) = \det(P^{-1}AP - \lambda I) = \det\ (P^{-1}AP - P^{-1}\lambda IP) = \det\ (P^{-1}(A-\lambda I)P)$$

$$= \det P^{-1} \det\ (A-\lambda I) \det P = \det\ (P^{-1}P) \det(A-\lambda I) = \det I \det(A-\lambda I)$$

$$= \det(A-\lambda I).$$

17.2 Condición necesaria y suficiente para que una matriz $A_{n\times n}$ sea diagonalizable

Teorema 2

Una matriz A de $n \times n$ es diagonalizable si y sólo si tiene n vectores característicos LI. En este caso, A es similar a una matriz diagonal D, con $P^{-1}AP = D$, cuyos elementos en la diagonal son los valores característicos de A. P es una matriz cuyas columnas son respectivamente los n vectores característicos LI de A.

Demostración

Supongamos que A es similar a D, con $D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{bmatrix}$; entonces:

$D = P^{-1}AP$, de modo que $PD = AP$. Sea P la matriz cuyas columnas son $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.

$$P = \left[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n \right],$$

$$AP = \left[A\mathbf{x}_1, A\mathbf{x}_2, \ldots, A\mathbf{x}_j, \ldots, A\mathbf{x}_n \right].$$

Ahora:

$$PD = \left[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n \right] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{bmatrix},$$

$$= \left[\lambda_1\mathbf{x}_1, \lambda_2\mathbf{x}_2, \ldots, \lambda_j\mathbf{x}_j, \ldots, \lambda_n\mathbf{x}_n \right].$$

Luego:

$$\left[A\mathbf{x}_1, A\mathbf{x}_2, \ldots, A\mathbf{x}_j, \ldots, A\mathbf{x}_n \right] = \left[\lambda_1\mathbf{x}_1, \lambda_2\mathbf{x}_2, \ldots, \lambda_j\mathbf{x}_j, \ldots, \lambda_n\mathbf{x}_n \right], \text{ de donde}$$

$$A\mathbf{x}_j = \lambda_j\mathbf{x}_j \qquad j = 1, \ldots, n.$$

Como P es invertible, sus columnas son LI y $\mathbf{x}_j \neq \mathbf{0}$ para $j = 1, \ldots, n$; por lo tanto, λ_j es un valor característico de A y $\mathbf{x}_j$ un vector característico correspondiente. Entonces A tiene n vectores característicos LI.

Recíprocamente, supongamos que A tiene n vectores característicos LI $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ con valores característicos correspondientes $\lambda_1, \lambda_2, \ldots, \lambda_n$. Sea $P = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ la matriz cuyas columnas son los n vectores característicos LI; entonces P es invertible.

Como $A\mathbf{x}_j = \lambda_j \mathbf{x}_j$ para $j = 1, ..., n,$

$$[A\mathbf{x}_1, A\mathbf{x}_2, ..., A\mathbf{x}_n] = [\lambda_1\mathbf{x}_1, \lambda_2\mathbf{x}_2, ..., \lambda_n\mathbf{x}_n],$$
$$AP = PD.$$

Multiplicando a ambos lados por P^{-1} tenemos:

$$P^{-1}AP = D,$$

lo cual significa que *A* es diagonalizable.

Corolario

Si $\underset{n\times n}{A}$ tiene n valores característicos distintos, entonces *A* es diagonalizable.

Observación

El teorema 5 del módulo 16 establece la siguiente equivalencia:

1. *A* tiene *n* vectores característicos LI $\Leftrightarrow$ multiplicidad geométrica de λ_i es igual a la multiplicidad algebraica de λ_i, $i = 1, \ldots, m$, y el teorema anterior dice que:

2. *A* es diagonalizable $\Leftrightarrow$ A tiene *n* vectores característicos LI.

De 1 y 2 tenemos que:

A es diagonalizable $\Leftrightarrow$ multiplicidad geométrica de λ_i es igual a la multiplicidad algebraica de λ_i, $i = 1, \ldots, m$, siendo *m* el número de raíces diferentes del polinomio $P(\lambda)$.

Ejemplo 2

En el ejemplo 7 del módulo16 determinamos para $A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$ sus valores y vectores característicos así: $\det(A - \lambda I) = -(\lambda + 1)^2(\lambda - 8) = 0$ con vectores característicos $\begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ correspondientes a $\lambda = -1$ y $\begin{bmatrix} 1 \\ \frac{1}{2} \\ 1 \end{bmatrix}$ correspondiente a $\lambda = 8.$

Luego A tiene tres vectores característicos LI; por lo tanto, A es diagonalizable, siendo:

$$D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{bmatrix}.$$

La matriz P que diagonaliza la matriz A es:

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

Entonces:

$$P^{-1}AP = D.$$
$$P(P^{-1}AP)P^{-1} = PDP^{-1},$$
$$(PP^{-1})A(PP^{-1}) = A = PDP^{-1}.$$

En el teorema 6 del módulo 13 demostramos que si A es similar a B, A^n es similar a B^n. Para nuestro ejemplo, A^n es similar a D^n, o sea:

$$A^n = PD^nP^{-1},$$
$$= \begin{bmatrix} -1/2 & -1 & 1 \\ 1 & 0 & 1/2 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} (-1)^n & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & (8)^n \end{bmatrix} \times \frac{1}{9} \begin{bmatrix} -2 & 8 & -2 \\ -4 & -2 & 5 \\ 4 & 2 & 4 \end{bmatrix}.$$

En particular, si se desea obtener A^{20}, bastaría hacer el producto indicado con $n = 20$, en lugar de multiplicar 20 veces por la matriz A.

Ejemplo 3

Las matrices de los ejemplos 1, 3, 4, 5, 7 y 8 del módulo 16 son diagonalizables, ya que para cada una de ellas se determinaron n vectores característicos LI. En el ejemplo 2, módulo 16, se planteó la matriz identidad que ya es diagonal, y en el ejemplo 6 del mismo módulo se obtuvo un valor característico de multiplicidad algebraica 3, al cual sólo iba asociado un vector característico LI; luego, la matriz propuesta en el ejemplo 6 no es diagonalizable.

Cuando estudiamos las propiedades de las matrices similares (teorema 6, módulo 13) vimos que si A y B son matrices similares entonces $\det A = \det B$.

Ahora, si A es similar a una matriz diagonal D, $\det A = \det D$.

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Diagonalización de matrices».

La matriz diagonal similar a A tiene sobre la diagonal los valores característicos de A; entonces:

$$\det A = \lambda_1 \lambda_2, ..., \lambda_n.$$

También establecimos que las distintas representaciones de una misma transformación lineal son matrices similares (teorema 5, módulo 13); luego, con cualquier representación matricial que trabajemos obtendremos los mismos valores y vectores característicos.

Ejercicios del capítulo 4 (módulo 17)

En los ejercicios 1 a 6 determine si la matriz dada A es diagonalizable; en caso de serlo, determine las matrices P, D y P^{-1} tales que $A = P\,D\,P^{-1}$.

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

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

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

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

5. $\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$.

6. $\begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix}$.

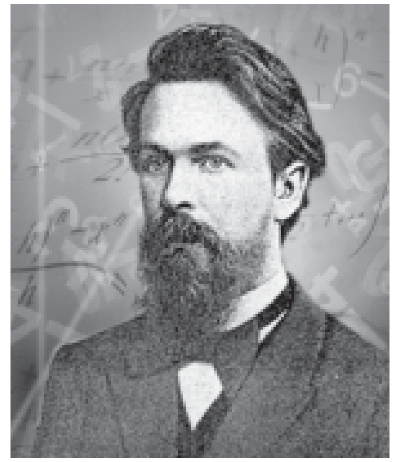
7. Diagonalice $A = \begin{pmatrix} 1 & 1 \\ 0 & \frac{1}{2} \end{pmatrix}$ y emplee la diagonalización hallada para calcular A^{12}.

8. Si A es invertible y diagonalizable, ¿es A^{-1} diagonalizable?

9. Si A y B son diagonalizables con $A = P^{-1} D_1 P$ y $B = Q^{-1} D_2 Q$, ¿es $A\,B$ una matriz diagonalizable, con $D_1\,D_2$ su matriz diagonal equivalente? ¿Es $A + B$ diagonalizable, con $D_1 + D_2$ su matriz diagonal equivalente?

Módulo 18

Aplicaciones de la teoría de valores y vectores característicos



Una de las aplicaciones más interesantes de los valores característicos es poder determinar el comportamiento de un sistema que pasa por varios estados, donde el estado siguiente sólo depende de su estado anterior en una etapa avanzada del proceso. El método para hacerlo se debe al matemático y lingüista ruso Andrei Andreyevich Markov (1856-1922) y se conoce como cadenas de Markov.

Introducción

Los valores y vectores característicos son útiles en gran parte de las matemáticas puras y aplicadas, y aparecen en contextos mucho más generales que los tratados en este tema, en el cual se estudia el comportamiento a largo plazo de un sistema dinámico discreto, descrito por medio de una ecuación en diferencias. También se emplean en ecuaciones diferenciales y sistemas dinámicos continuos. Proporcionan información crítica en el diseño de ingeniería y se presentan naturalmente en campos como la física y la química.

Las aplicaciones desarrolladas se centran en problemas de crecimiento poblacional y procesos de Markov, porque son más accesibles con los conocimientos matemáticos que el estudiante posee en el primer año de su carrera. Por tal razón, éste será el enfoque específico que le daremos al tema.

Objetivos

1. Mostrar aplicaciones a la ingeniería de la teoría de valores y vectores característicos.
2. Estudiar sistemas dinámicos en crecimientos poblacionales y procesos de Markov.

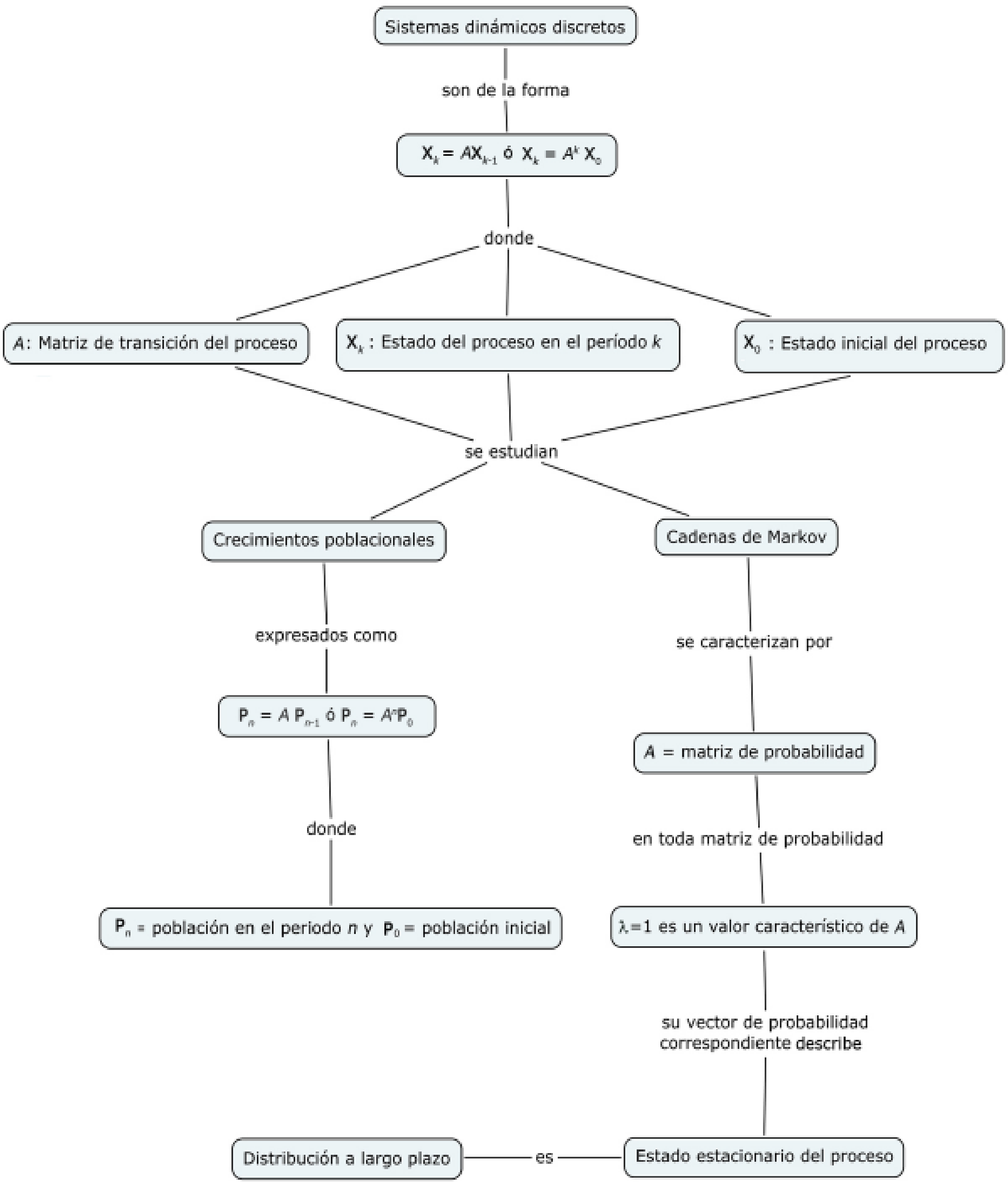
Preguntas básicas

1. ¿Cómo se conforma la matriz de transición de un proceso en un sistema dinámico?
2. Al analizar los valores característicos de la matriz de transición A en el estudio de una población, ¿cómo se sabe si la población crece o decrece?
3. ¿Qué es un proceso de Markov?
4. ¿Qué es una matriz de probabilidad?
5. ¿Cuándo un proceso de Markov alcanza su estado estacionario?

Contenidos

18.1 Ecuaciones en diferencias
    18.1.1 Un modelo de crecimiento poblacional
18.2 Procesos de Markov
    18.2.1 Modelo de funcionamiento de una máquina

Vea el módulo 18 del programa de televisión *Álgebra lineal.*



Mapa 11: módulo 18

18.1 Ecuaciones en diferencias

Un proceso que va pasando a través de diferentes estados cada cierto intervalo de tiempo, esto es, un proceso discreto en el tiempo, se describe usualmente como un sistema de ecuaciones en diferencias.

18.1.1 Un modelo de crecimiento poblacional

Estudiaremos un modelo de crecimiento poblacional para una especie de venados donde se han determinado las siguientes condiciones:

1. El número de hembras es igual al número de machos.

2. La población se considera dividida en dos grupos de edad que son:

   $P_{j,n-1}$: población juvenil (inmadura) de hembras en el año $n-1$.

   $P_{a,n-1}$: población adulta de hembras en el año $n-1$.

   La población juvenil es aquella entre 0 y 1 año, y la población adulta es la que tiene más de un año de edad.

   El crecimiento de esta población se estudia por medio de las hembras, y haciendo uso de la primera condición podemos saber cómo va la población en cualquier periodo.

3. Hay una tasa de supervivencia α de los venados jóvenes que serán adul tos en el año siguiente.

   Para los adultos también hay una tasa de supervivencia β el periodo siguiente.

4. Cada hembra adulta produce, en promedio, k hembras jóvenes para el periodo siguiente.

Haciendo uso de la información suministrada, podemos expresar las poblaciones de hembras jóvenes y adultas en el periodo n, así:

$$P_{j,n} = kP_{a,n-1}.$$

$$P_{a,n} = \alpha P_{j,n-1} + \beta P_{a,n-1}.$$

o $\mathbf{P}_n = A\mathbf{P}_{n-1}$, donde $\mathbf{P}_n = \begin{bmatrix} P_{j,n} \\ P_{a,n} \end{bmatrix}$, $A = \begin{bmatrix} 0 & k \\ \alpha & \beta \end{bmatrix}$, $\mathbf{P}_{n-1} = \begin{bmatrix} P_{j,n-1} \\ P_{a,n-1} \end{bmatrix}$.

Entonces, si $\mathbf{P_0}$ es la población inicial de hembras jóvenes y adultas, podemos expresar las poblaciones en los periodos siguientes así:

$$\mathbf{P}_1 = A\mathbf{P}_0,\quad \mathbf{P}_2 = A\mathbf{P}_1 = A(A\mathbf{P}_0) = A^2\mathbf{P}_0,\quad \mathbf{P}_3 = A\mathbf{P}_2 = A(A^2\mathbf{P}_0) = A^3\mathbf{P}_0.$$

Luego, la población en el periodo n está dada por:

$$\mathbf{P}_n = A^n \mathbf{P}_0.$$

Determinemos para A sus valores y vectores característicos:

$$\det(A-\lambda I) = \begin{vmatrix} -\lambda & k \\ \alpha & \beta-\lambda \end{vmatrix} = 0.$$

$-\lambda\beta + \lambda^2 - \alpha k = 0$, entonces $\lambda^2 - \lambda\beta - \alpha k = 0$, de donde:

$$\lambda = \frac{\beta \pm \sqrt{\beta^2 + 4\alpha k}}{2}, \text{ con } \beta,\ \alpha \text{ y } k \text{ cantidades positivas.}$$

$$\lambda_1 = \frac{\beta + \sqrt{\beta^2 + 4\alpha k}}{2} \text{ y } \lambda_2 = \frac{\beta - \sqrt{\beta^2 + 4\alpha k}}{2}.$$

Analizando estas expresiones, podemos afirmar que:

$$\lambda_1 > 0,\ \lambda_2 < 0 \text{ y } |\lambda_1| > |\lambda_2|.$$

A cada valor de λ va asociado un vector característico. Si $\mathbf{v_1}$ y $\mathbf{v_2}$ son los vectores característicos correspondientes a λ_1 y λ_2, respectivamente, entonces $\mathbf{v_1}$ y $\mathbf{v_2}$ son LI.

La población inicial $\mathbf{P}_0$ se puede expresar como combinación lineal de $\mathbf{v_1}$ y $\mathbf{v_2}$ así:

$$\mathbf{P}_0 = a_1\mathbf{v}_1 + a_2\mathbf{v}_2.$$

Como $\mathbf{P}_n = A^n\mathbf{P}_0$, entonces:

$$\mathbf{P}_n = A^n(a_1\mathbf{v}_1 + a_2\mathbf{v}_2) = a_1A^n\mathbf{v}_1 + a_2A^n\mathbf{v}_2$$

$= a_1\lambda_1{}^n\mathbf{v}_1 + a_2\lambda_2{}^n\mathbf{v}_2$, ya que $A^n\mathbf{v}_1 = \lambda_1{}^n\mathbf{v}_1$ y $A^n\mathbf{v}_2 = \lambda_2{}^n\mathbf{v}_2$ (teorema 8, parte c, módulo 16)

$$= \lambda_1{}^n\left(a_1\mathbf{v}_1 + a_2\left(\frac{\lambda_2}{\lambda_1}\right)^n \mathbf{v}_2\right).$$

Como $|\lambda_1| > |\lambda_2|$, $\left(\dfrac{\lambda_2}{\lambda_1}\right)^n$ tiende a cero cuando n es grande, entonces:

$$\mathbf{P}_n \approx \lambda_1{}^n a_1\mathbf{v}_1. \qquad (1)$$

A largo plazo la distribución de edades se estabiliza y es proporcional a $\mathbf{v}_1$. Cada grupo de edad cambiará por un factor λ_1 cada año.

Es importante notar toda la información que podemos obtener del cálculo de los valores y vectores característicos.

Una pregunta importante por resolver es: ¿la población crecerá o decrecerá? Aumenta si $\lambda_1 > 1$, y esta condición se verifica cuando:

$$\frac{\beta + \sqrt{\beta^2 + 4\alpha k}}{2} > 1, \qquad \sqrt{\beta^2 + 4\alpha k} > 2 - \beta.$$

Elevando al cuadrado:

$$\beta^2 + 4\alpha k > 4 - 4\beta + \beta^2,$$

$$4\alpha k > 4 - 4\beta,$$

$$k > \frac{1-\beta}{\alpha}.$$

Supongamos ahora que en esta población de venados se tienen:

$$k = 1, \quad \alpha = 0.6, \quad \beta = 0.8 \quad \text{y} \quad P_0 = \begin{bmatrix} 100 \\ 200 \end{bmatrix}.$$

Entonces:

$$A = \begin{bmatrix} 0 & 1 \\ 0.6 & 0.8 \end{bmatrix},$$

luego:

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ 0.6 & 0.8 - \lambda \end{vmatrix} = \lambda^2 - 0.8\lambda - 0.6 = 0,$$

$$\lambda = \frac{0.8 \pm \sqrt{(0.8)^2 + 4 \times 0.6}}{2}, \quad \lambda_1 = 1.27, \quad \lambda_2 = -0.47.$$

Sabemos que a la larga la población se estabiliza y se calcula de acuerdo con la ecuación (1). En consecuencia, la población crecerá por un factor aproximado de 1.27.

Los granjeros y otras personas del área no quieren que la población crezca. Pueden controlar la población «cosechándola» (permitiendo la caza de venados adultos); si h es la proporción de población cosechada en cada periodo, entonces la proporción de supervivencia β de la población adulta se disminuye en h y la matriz A para el modelo, será:

$$A = \begin{bmatrix} 0 & 1 \\ 0.6 & 0.8 - h \end{bmatrix}.$$

Veamos qué pasa si $h = 0.6$. En este caso:

$$A=\begin{bmatrix}0 & 1\\ 0.6 & 0.2\end{bmatrix},\quad \det(A-\lambda I)=\begin{vmatrix}-\lambda & 1\\ 0.6 & 0.2-\lambda\end{vmatrix}=\lambda^2-0.2\lambda-0.6=0$$

$$\text{y}\ \lambda_1=\frac{0.2+\sqrt{0.04+4\times 0.6}}{2}=0.88.$$

Así que $\lambda_1<1$ y la población decrecerá, luego h = 0.6 es una cosecha demasiado grande que terminará por extinguir la especie.

Si queremos que la población permanezca estable, esto es, que no crezca ni desaparezca, el valor característico mayor (λ_1) debe ser igual a 1.

Entonces:

$$\lambda_1=\frac{\beta+\sqrt{\beta^2+4\alpha K}}{2}=1,$$

$$\frac{(0.8-h)+\sqrt{(0.8-h)^2+4\times 0.6}}{2}=1,$$

$$h=0.4.$$

Así que una proporción de caza igual a 0.4 mantendrá estable la población.

Ejemplo 1

Una especie animal está clasificada en dos etapas de vida: juvenil (hasta un año de edad) y adulta. Suponga que las hembras adultas paren una vez al año un promedio de 1.6 hembras juveniles. Cada año sobrevive el 30% de los juveniles para transformarse en adultos y sobrevive el 80% de los adultos.

Veamos cómo estaría dada la población de esta especie en un periodo cualquiera *k*.

$$\mathbf{P}_k=\begin{bmatrix}P_{j,k}\\ P_{a,k}\end{bmatrix}\qquad\qquad \begin{aligned}P_{j,k}&=1.6P_{a,k-1}\\ P_{a,k}&=0.3P_{j,k-1}+0.8P_{a,k-1}\end{aligned}$$

de donde:

$$\begin{bmatrix}P_{j,k}\\ P_{a,k}\end{bmatrix}=\begin{bmatrix}0 & 1.6\\ 0.3 & 0.8\end{bmatrix}\begin{bmatrix}P_{j,k-1}\\ P_{a,k-1}\end{bmatrix},$$

$$\mathbf{P}_k=A\mathbf{P}_{k-1}.$$

Veamos cómo se comporta este sistema, al analizar sus valores y vectores característicos.

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Aplicaciones valores y vectores propios».

$$\det(A-\lambda I)=\begin{vmatrix}-\lambda & 1.6\\ 0.3 & 0.8-\lambda\end{vmatrix}=\lambda^2-0.8\lambda-0.48=0,$$

$$\lambda = \frac{0.8 \pm \sqrt{0.64 + 4 \times 0.48}}{2} = \frac{0.8 \pm 1.6}{2} .$$

Luego $\lambda_1 = 1.2$ y $\lambda_2 = -0.4.$

La población crece porque el mayor valor característico de A es 1.2, cuya magnitud es mayor que 1.

Determinemos los vectores característicos asociados con $\lambda_1 = 1.2$ y $\lambda_2 = -0.4.$

Para $\lambda_1 = 1.2,$ resolvemos $(A - 1.2I)\,\mathbf{x} = \mathbf{0}.$

$$\begin{bmatrix} -1.2 & 1.6 \\ 0.3 & -0.4 \end{bmatrix}, \text{ entonces } \quad \begin{array}{l} 0.3x - 0.4y = 0 \\ 3x = 4y \end{array}$$

Si $x = 4$ e $y = 3,$ entonces $\mathbf{x}_1 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$

Para $\lambda_2 = -0.4,$ $(A + 0.4\,I)\,\mathbf{x} = \mathbf{0}.$

$$\begin{bmatrix} 0.4 & 1.6 \\ 0.3 & 1.2 \end{bmatrix}, \text{ entonces } \quad \begin{array}{l} 0.4x + 1.6y = 0 \\ 4x = -16y \text{ o } x = -4y \end{array}$$

Si $x = 4,$ $y = -1$ y $\mathbf{x}_2 = \begin{bmatrix} 4 \\ -1 \end{bmatrix}.$

Si la población inicial $\mathbf{P}_0$ la expresamos en términos de $\mathbf{x}_1$ y $\mathbf{x}_2$, entonces:

$$\begin{aligned} \mathbf{P}_0 &= a_1\mathbf{x}_1 + a_2\mathbf{x}_2, \\ \mathbf{P}_n &= A^n\mathbf{P}_0 = A^n(a_1\mathbf{x}_1 + a_2\mathbf{x}_2), \\ &= a_1A^n\mathbf{x}_1 + a_2A^n\mathbf{x}_2 = a_1\lambda_1{}^n\mathbf{x}_1 + a_2\lambda_2{}^n\mathbf{x}_2. \end{aligned}$$

$|\lambda_2| = 0.4 < 1.$ Cuando n es grande, $\lambda_2{}^n \to 0$ y $\mathbf{P}_n \approx a_1\lambda_1{}^n\mathbf{x}_1.$

La población crece con un factor constante igual a $\lambda_1.$

La tasa de crecimiento final es de 1.2, que es un 20% anual. El vector característico $\mathbf{x}_1 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ muestra que habrá cuatro juveniles por cada tres adultos.

En los modelos estudiados se describe un proceso que está determinado por medio de la matriz A. El vector $\mathbf{P}_0$ se llama *estado inicial* del proceso y el vector $\mathbf{P}_n$ para $n \in \mathbb{N}$ se llama *n-ésimo estado del proceso.*

A la matriz A se le conoce como la *matriz de transición del proceso* y a la ecuación $\mathbf{P}_n = A\mathbf{P}_{n-1}$ se le llama *ecuación matricial en diferencias* o *sistema dinámico.*

Teorema 1

Sea A una matriz diagonalizable $n \times n$ con vectores característicos linealmente independientes $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ y sus correspondientes valores característicos $\lambda_1, \lambda_2, ..., \lambda_n$. La solución del sistema dinámico $\mathbf{X}_k = A\mathbf{X}_{k-1}$ se expresa así:

$$\mathbf{X}_k = c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + ... + c_n\lambda_n^k\mathbf{v}_n,$$

donde los coeficientes $c_1, c_2, ..., c_n$ son tales que:

$$\mathbf{X}_\text{o} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n.$$

Demostración

Los vectores $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ forman una base del espacio V ($\mathbb{R}^n$ o $\mathbb{C}^n$), luego $\mathbf{X}_0$ se puede expresar en esta base como $\mathbf{X}_\text{o} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n$.

Como $\mathbf{X}_k = A\mathbf{X}_{k-1}$, entonces:

$$\mathbf{X}_1 = A\mathbf{X}_0,\ \mathbf{X}_2 = A\mathbf{X}_1,\ \mathbf{X}_2 = A^2\mathbf{X}_0, ..., \mathbf{X}_k = A^k\mathbf{X}_0.$$

Luego:

$$\begin{aligned}\mathbf{X}_k &= A^k(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n) \\ &= c_1A^k\mathbf{v}_1 + c_2A^k\mathbf{v}_2 + ... + c_nA^k\mathbf{v}_n.\end{aligned}$$

Aplicando el teorema 8, módulo 16, donde se establece que si λ_i es un valor característico de *A*, λ_i^n es un valor característico de A^n, tenemos:

$$\mathbf{X}_k = c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + ... + c_n\lambda_n^k\mathbf{v}_n.$$

18.2 Procesos de Markov

Sean $S_1, S_2, ..., S_n$ los estados posibles de un sistema *S*. Supongamos que *S* se observa en los tiempos dados $T_1, T_2, ..., T_n$. Una *cadena de Markov* es un proceso en el cual la probabilidad empírica de que *S* se halle en un estado particular al tiempo T_k, depende solamente del estado en que se halle *S* en el tiempo T_{k-1}.

Escuche la biografía de *Andrei Andreyevich Markov* en su multimedia de *Álgebra lineal.*

La matriz de transición en un proceso de Markov es una matriz de probabilidad.

Definición 1

Sea A una matriz $n \times n$. A es una *matriz de probabilidad* si se cumple que:

i. $a_{ij} \geq 0$ para toda i y j.
ii. La suma de las componentes en cada columna es 1.

Antes de plantear modelos de procesos de Markov veremos una propiedad importante de las matrices de probabilidad.

Teorema 2

Sea $A_{n\times n}$ una matriz de probabilidad, entonces $\lambda = 1$ es un valor característico de A.

.

Demostración

$\lambda = 1$ es un valor característico de A si $\det(A-1I)=0$,

$$A - I = \begin{bmatrix} a_{11}-1 & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-1 & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-1 \end{bmatrix}.$$

Como A es una matriz de probabilidad,

$$\sum_{i=1}^{n} a_{ij} = 1, \text{ para } j = 1,\ldots,n \text{ (la suma sobre cada columna es 1).}$$

En la matriz $A-I$, la suma sobre cada columna será $\sum_{i=1}^{n} a_{ij} - 1 = 0.$

$$\begin{vmatrix} a_{11}-1 & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-1 & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-1 \end{vmatrix} \xrightarrow{R_1+(R_2+\ldots+R_n)}$$

$$\begin{vmatrix} \sum_{i=1}^{n} a_{i1}-1 & \sum_{i=1}^{n} a_{i2}-1 & \cdots & \sum_{i=1}^{n} a_{in}-1 \\ a_{21} & a_{22}-1 & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-1 \end{vmatrix} = \begin{vmatrix} 0 & 0 & \cdots & 0 \\ a_{21} & a_{22}-1 & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-1 \end{vmatrix} = 0.$$

18.2.1 Modelo de funcionamiento de una máquina

Suponga que una máquina está siempre en alguno de estos tres estados: (1) parada sin reparación (P), (2) en necesidad de ajustes (N), (3) trabajando bien (T).

Sea:

$$\begin{array}{cc} & \begin{array}{ccc} P & N & T \end{array} \\ A = & \begin{bmatrix} 1 & 1/4 & 1/18 \\ 0 & 1/2 & 8/18 \\ 0 & 1/4 & 9/18 \end{bmatrix} \end{array} \begin{array}{c} P \\ N \\ T \end{array}$$

con a_{ij} : probabilidad de que una máquina que se encuentre en el estado j pase al estado i en el periodo siguiente.

Asumimos que la probabilidad de estar en uno cualquiera de los estados al final de un periodo depende sólo del estado en que se encontraba la máquina al principio del periodo, o sea, el final del periodo anterior.

Sea $\mathbf{X}_t = \begin{bmatrix} X_{1,t} \\ X_{2,t} \\ X_{3,t} \end{bmatrix}$, donde $X_{i,t}$ es la probabilidad de que la máquina esté en el estado i al principio del periodo t.

Entonces $A\mathbf{X}_{t\text{-}1} = \mathbf{X}_t$.

Si $\mathbf{X}_0$ es la distribución de probabilidad inicial:

$$A\mathbf{X}_0 = \mathbf{X}_1,\ A\mathbf{X}_1 = \mathbf{X}_2,\ AA\mathbf{X}_0 = A^2\mathbf{X}_0 \ldots$$
$$\mathbf{X}_n = A^n\mathbf{X}_0,\quad n \geq 1.$$

Por ejemplo, si la máquina está trabajando bien al principio del periodo, entonces:

$$\mathbf{X}_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},\quad \mathbf{X}_1 = \begin{bmatrix} 1 & 1/4 & 1/18 \\ 0 & 1/2 & 8/18 \\ 0 & 1/4 & 9/18 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/18 \\ 8/18 \\ 9/18 \end{bmatrix}.$$

Si consideramos una empresa donde cada máquina se comporta con una distribución de probabilidad como muestra $\mathbf{X}_1$, podemos esperar que 1/18 de las máquinas estén paradas, 4/9 necesiten ajustes y 1/2 estén trabajando bien.

El comportamiento de la máquina está dado por la matriz *A*. La secuencia de vectores $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ se llama cadena de Markov.

Veamos, si *A* es diagonalizable, cómo calcular A^n como el producto PD^nP^{-1}.

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1/4 & 1/18 \\ 0 & 1/2-\lambda & 8/18 \\ 0 & 1/4 & 9/18-\lambda \end{vmatrix} = 0,$$

$$(1-\lambda)\left(\left(\frac{1}{2}-\lambda\right)^2-\frac{8}{18}\times\frac{1}{4}\right)=0,$$

$$(1-\lambda)\left(\lambda^2-\lambda+\frac{5}{36}\right)=0,$$

$$(1-\lambda)\left(\lambda-\frac{1}{6}\right)\left(\lambda-\frac{5}{6}\right)=0,\quad \lambda_1=1,\ \lambda_2=\frac{1}{6},\ \lambda_3=\frac{5}{6}.$$

Para $\lambda=1$, resolvemos $(A-I)\mathbf{v}=\mathbf{0}$, de donde $\mathbf{v_1}=\begin{bmatrix}1\\0\\0\end{bmatrix}$.

Para $\lambda=1/6$ se determina $\mathbf{v_2}=\begin{bmatrix}1\\-4\\3\end{bmatrix}$.

Para $\lambda=5/6$ se encuentra $\mathbf{v_3}=\begin{bmatrix}-7\\4\\3\end{bmatrix}$ (verifíquelo).

Luego:

$$P=\begin{bmatrix}1&1&-7\\0&-4&4\\0&3&3\end{bmatrix},\quad P^{-1}=\begin{bmatrix}1&1&1\\0&-1/8&1/6\\0&1/8&1/6\end{bmatrix},\quad D=\begin{bmatrix}1&0&0\\0&1/6&0\\0&0&5/6\end{bmatrix}.$$

Si se quiere saber el estado del proceso después de cinco periodos, luego de haber comenzado con las máquinas trabajando bien, entonces:

$$\mathbf{X}_5=A^5\mathbf{X}_0=A^5\begin{bmatrix}0\\0\\1\end{bmatrix},$$

$$A^5=PD^5P^{-1}=\begin{bmatrix}1&1&-7\\0&-4&4\\0&3&3\end{bmatrix}\begin{bmatrix}1&0&0\\0&(1/6)^5&0\\0&0&(5/6)^5\end{bmatrix}\begin{bmatrix}1&1&1\\0&-1/8&1/6\\0&1/8&1/6\end{bmatrix}$$

$$=\begin{bmatrix}1&0.648&0.531\\0&0.201&0.268\\0&0.151&0.201\end{bmatrix},$$

$$\text{y}\ \mathbf{X}_5=\begin{bmatrix}1&0.648&0.531\\0&0.201&0.268\\0&0.151&0.201\end{bmatrix}\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}0.531\\0.268\\0.201\end{bmatrix}.$$

Después de cinco periodos la probabilidad de que las máquinas estén trabajando bien es de, aproximadamente, 0.2, esto representa alrededor del 20%; poco menos del 27% necesita ajuste y cerca del 53% están paradas.

Se quiere saber si existe alguna distribución que se mantenga de un periodo a otro, esto es, si $\mathbf{X}_{t+1} = \mathbf{X}_t$, $\mathbf{X}_{t+1} = A\mathbf{X}_t = \mathbf{X}_t$.

Vemos que el vector de probabilidad correspondiente a $\lambda = 1$ verifica esta condición.

Cuando el proceso se comporta de acuerdo con esta distribución, se dice que ha alcanzado su *estado estacionario.* Este vector es $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$; esto es, cuando todas las máquinas están en el estado (1), o sea, paradas sin reparación, ya no ocurre ningún cambio de estado de un periodo a otro. A esta situación se llega cuando han transcurrido n periodos con n grande.

El objetivo del análisis de Markov es calcular la probabilidad de que un sistema se encuentre en un estado en particular en un tiempo futuro y determinar el comportamiento del sistema a largo plazo.

Teorema 3

Sea un proceso de Markov dado por la matriz A y la cadena $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_k, ...$. Si A es diagonalizable y todo valor característico de A distinto de 1 posee módulo menor que 1, entonces:

a. La sucesión $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_k ...$ cuando k tiende a infinito converge al $\lim_{k\to\infty} \mathbf{X}_k = \mathbf{X}_\infty$.

b. $\mathbf{X}_\infty$ es un vector del espacio característico de A asociado con $\lambda = 1$, es decir, $A\mathbf{X}_\infty = \mathbf{X}_\infty$.

Demostración

a. Como A es diagonalizable, la solución $\mathbf{X}_k$ de la ecuación en diferencias o sistema dinámico $\mathbf{X}_k = A\mathbf{X}_{k-1}$ está dada por:

$$\mathbf{X}_k = C_1\lambda_1^k \mathbf{v}_1 + C_2\lambda_2^k \mathbf{v}_2 + ... + C_r\lambda_r^k \mathbf{v}_r + C_{r+1}\lambda_{r+1}^k \mathbf{v}_{r+1} + ... + C_n\lambda_n^k \mathbf{v}_n \text{ (teorema 1).}$$

Supongamos que el valor característico $\lambda = 1$ tiene multiplicidad algebraica r, y que $|\lambda_{r+1}|$, $|\lambda_{r+2}| ... |\lambda_n|$ son menores que 1; entonces:

$$\mathbf{X}_k = C_1\mathbf{v}_1 + C_2\mathbf{v}_2 + ... + C_r\mathbf{v}_r + C_{r+1}\lambda_{r+1}^k \mathbf{v}_{r+1} + ... + C_n\lambda_n^k \mathbf{v}_n.$$

Ahora, para cada $i = r+1, \ldots, n$ tenemos que:

$$\lim_{k\to\infty} \lambda_i^k = 0 \text{ ya que } |\lambda_i| < 1.$$

Luego:

$$\lim_{k\to\infty} \mathbf{X}_k = C_1\mathbf{v}_1 + C_2\mathbf{v}_2 + \ldots + C_r\mathbf{v}_r \ldots = \mathbf{X}_\infty.$$

b. Los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ son vectores característicos asociados con $\lambda = 1$; por lo tanto, $\mathbf{X}_\infty$ es un vector del espacio característico de A asociado con $\lambda = 1$.

$\mathbf{X}_\infty$ se conoce como el estado estacionario del sistema.

Ejemplo 2

Cada año 5% de la población de la ciudad se muda a los suburbios y 3% de la población de los suburbios se muda a la ciudad. Suponga que inicialmente 60% de la población vive en la ciudad y 40% en los suburbios (figura 18.1).

Veamos cuál es la distribución de la población después un año.

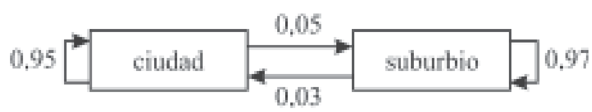


Figura 18.1

Sean $P_{c,n}$ y $P_{s,n}$ las poblaciones de la ciudad y los suburbios en el año *n*.

$$\mathbf{P}_0 = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} = \begin{bmatrix} P_{c,0} \\ P_{s,0} \end{bmatrix} \qquad \mathbf{P}_1 = \begin{bmatrix} P_{c,1} \\ P_{s,1} \end{bmatrix} \qquad \begin{aligned} P_{c,1} &= 0.95P_{c,0} + 0.03P_{s,0} \\ P_{s,1} &= 0.05P_{c,0} + 0.97P_{s,0} \end{aligned}$$

$$\mathbf{P}_1 = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix} \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 0.582 \\ 0.418 \end{bmatrix}.$$

Después de un año, aproximadamente, 58% de la población vive en la ciudad y 42% en los suburbios.

Se quiere conocer la distribución de población que permanezca a través del tiempo.

$$\mathbf{P}_{k+1} = \mathbf{P}_k = A\mathbf{P}_k.$$

$\mathbf{P}_k$ es un vector característico correspondiente a $\lambda = 1$.

$$\begin{bmatrix} 0.95-1 & 0.03 \\ 0.05 & 0.97-1 \end{bmatrix} = \begin{bmatrix} -0.05 & 0.03 \\ 0.05 & -0.03 \end{bmatrix}, \text{ entonces } \begin{array}{c} 0.05x = 0.03y \\ 5x = 3y \end{array}$$

Si $x = 3$, $y = 5$, entonces $\mathbf{v} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

Se debe determinar dentro de los vectores de E_1 un vector de probabilidad (la suma de sus componentes debe ser 1).

$\mathbf{v}_1 = \begin{bmatrix} 3/8 \\ 5/8 \end{bmatrix}$ es un vector de estado estacionario.

Se quiere conocer la población después de *k* años. Averigüemos los valores característicos de *A*.

$$\begin{vmatrix} 0.95-\lambda & 0.03 \\ 0.05 & 0.97-\lambda \end{vmatrix} = (0.95-\lambda)(0.97-\lambda) - 0.0015 = 0,$$

$$\lambda^2 - 1.92\lambda + 0.92 = 0, \qquad (\lambda-1)(\lambda-0.92) = 0, \qquad \lambda_1 = 1,\ \lambda_2 = 0.92.$$

Determinemos un vector característico correspondiente a $\lambda_2 = 0.92$.

$$\begin{bmatrix} 0.95-0.92 & 0.03 \\ 0.05 & 0.97-0.92 \end{bmatrix} = \begin{bmatrix} 0.03 & 0.03 \\ 0.05 & 0.05 \end{bmatrix}, \text{ entonces } \begin{array}{c} 0.03x = -0.03y \\ x = -y \end{array}$$

Luego:

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$\mathbf{P}_0 = a_1\mathbf{v}_1 + a_2\mathbf{v}_2, \qquad \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} = a_1\begin{bmatrix} 3 \\ 5 \end{bmatrix} + a_2\begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

$$\left[\begin{array}{cc|c} 3 & 1 & 0.6 \\ 5 & -1 & 0.4 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & 1/8 \\ 0 & 1 & 9/40 \end{array}\right].$$

$$\begin{aligned} \mathbf{P}_k &= a_1\lambda_1^k\mathbf{v}_1 + a_2\lambda_2^k\mathbf{v}_2, \\ &= \frac{1}{8}\begin{bmatrix} 3 \\ 5 \end{bmatrix} + \frac{9}{40}(0.92)^k\begin{bmatrix} 1 \\ -1 \end{bmatrix}. \end{aligned}$$

Cuando $k \to \infty$, $(0.92)^k \to 0$ y $P_k \to \begin{bmatrix} 3/8 \\ 5/8 \end{bmatrix}$.

A largo plazo la población alcanza el estado estacionario que es el vector de probabilidad correspondiente a $\lambda = 1$.

La población se estabiliza cuando 3/8 de ella vive en la ciudad y 5/8 en los suburbios.

Ejercicios del capítulo 4 (módulo 18)

1. La población de una cierta variedad de peces aumenta de manera que el crecimiento en cualquier año es el doble del crecimiento en el año anterior. Si inicialmente se tenían 50 peces y después del primer año se contabilizaron 70 peces:

   a. Encuentre el tamaño de la población de peces en cualquier año.
   b. Encuentre el tamaño de la población después del quinto año.
   c. Encuentre el año en que la población de peces alcanza 2.800 individuos.

2. Suponga que hay tres centros principales de camiones «múdese usted mismo». Cada mes, la mitad de los que están en Boston y en Los Ángeles van a Chicago; la otra mitad permanece donde está y los camiones de Chicago se dividen igualmente entre Boston y Los Ángeles. Si inicialmente la compañía tenía x_0, y_0, z_0 camiones en Boston, Chicago y Los Ángeles, respectivamente:

   a. Encuentre la distribución de camiones de la compañía en las tres ciudades, para cada mes.
   b. Determine cuál será a largo plazo la distribución de camiones de la compañía.

3. Cada año 1/10 de la gente de Estados Unidos que vive fuera de California se muda dentro y 2/10 de la gente que vive dentro de California se muda fuera. Si inicialmente y_0 y z_0 eran los tamaños de las poblaciones fuera y dentro de California:

   a. Encuentre, para el k-ésimo año, el tamaño de la población que vive fuera (dentro) de California.
   b. ¿Puede modelar este problema como un proceso de Markov?
   c. Encuentre la probabilidad de que un individuo de Estados Unidos, elegido al azar, viva fuera (dentro) de California en el k-ésimo año.
   d. Encuentre la distribución de la población americana a largo plazo.

4. Suponga que hay una epidemia en la que cada mes se enferma la mitad de los que están sanos y muere la cuarta parte de los que están enfermos. Suponga que inicialmente había s_0 y e_0 individuos sanos y enfermos, respectivamente, y ningún individuo había muerto.

   a. Encuentre la distribución de la población para el k-ésimo mes y diga cuál es la probabilidad en ese mes, para cada uno de los siguientes eventos: estar sano, estar enfermo, estar muerto.
   b. ¿Puede modelar este problema como un proceso de Markov?
   c. Encuentre la distribución de la población a largo plazo.

5. Un curso de Química se imparte en dos secciones. Si cada semana dejan el curso 1/4 de los que están en la sección A y 1/3 de los que están en la sección B, y 1/6 de cada sección se transfiere a la otra:

   a. A largo plazo, ¿cuál será la distribución de alumnos?
   b. ¿Puede modelar el problema como un proceso de Markov?

   Para resolver los literales *a* y *b* suponga que inicialmente tenían x_0 y y_0 estudiantes en las secciones A y B y ningún estudiante fuera del curso.

6. El crecimiento de un cultivo de bacterias en un medio nutritivo se observa cada dos horas y cada vez se encuentra que la población ha crecido 30% respecto de la vez anterior.

   a. Denote por P_n la población de bacterias después de $2n$ horas y describa este proceso de crecimiento por medio de una ecuación.
   b. Dado que la población inicial es de 1.000 bacterias, determine P_2 y P_4.

7. En un estudio de enfermedades infecciosas se mantiene un registro de brotes de sarampión en un colegio particular. Se estima que la población P_n infectada en la n-ésima semana está dada por la ecuación $P_{n+2} = P_{n+1} - 1/5P_n$. Si $P_0 = 0$ y $P_1 = 1.000$:

   a. Encuentre la población de infectados en la n-ésima semana.
   b. ¿Se puede modelar el problema como un proceso de Markov?
   c. ¿Cuántos infectados se tendrán después de transcurridas seis semanas?

8. Una población de conejos criados en un laboratorio tiene las siguientes características:

   - La mitad de los conejos sobrevive el primer año. De estos, la mitad sobrevive el segundo año. La duración de la máxima vida es tres años.

   - Durante el primer año los conejos no producen descendencia. El número medio de descendencia es seis durante el segundo año y ocho durante el tercer año.

Clase de primera edad	$0 \leq \text{edad} < 1$
Clase de segunda edad	$1 \leq \text{edad} < 2$
Clase de tercera edad	$2 \leq \text{edad} \leq 3$

   Si actualmente la población consta de 24 conejos en la clase de la primera edad, 24 en la clase de la segunda edad y 20 en la tercera edad, ¿cuál será la distribución de conejos cuando hayan transcurrido diez años?

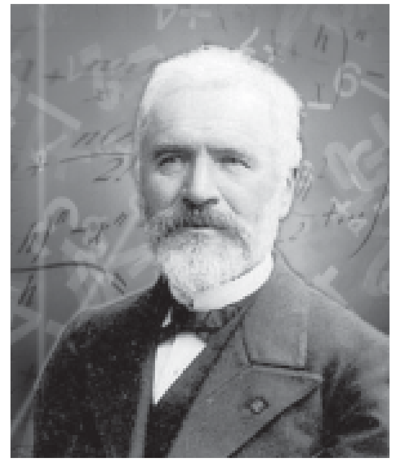
9. Una compañía de robótica quiere fabricar un brazo que deberá recoger partes de una banda transportadora para colocarlas en otra banda. Ocasionalmente el brazo falla, pero el robot está diseñado para que en caso de falla se activen circuitos secundarios. En las observaciones se descubre que si el brazo falla en una ocasión, tendrá éxito la siguiente vez 97% de las veces. Si el brazo tiene éxito en cierto intento, los circuitos secundarios se desactivarán y el brazo fallará la siguiente vez, apenas 2% de las veces. ¿Cumplirá el brazo con el requerimiento del cliente de que trabaje exitosamente 98% de las veces?

10. En un día dado, un estudiante está sano o está enfermo. De los estudiantes que están sanos hoy, 95% estará sano mañana. De los estudiantes que están enfermos hoy, 55% estará enfermo mañana.

    a. ¿Cuál será la matriz para esta situación?
    b. Suponga que el lunes, 20% de los estudiantes está enfermo. ¿Qué fracción o porcentaje de los estudiantes es probable que esté enfermo el miércoles?
    c. Si un estudiante está bien hoy, ¿cuál será la probabilidad de estar bien dentro de dos días?
    d. ¿Cuál es la probabilidad de que después de muchos días una persona dada esté enferma?

Módulo 19

Forma canónica de Jordan



Los procedimientos expuestos en este módulo aparecieron por primera vez en el trabajo del matemático francés Camille Jordan (1838-1922) titulado *Tratado sobre sustitución y ecuaciones algebraicas*, publicado en 1870.

Introducción

Hemos dicho que una transformación lineal, en un espacio vectorial de dimensión n, exhibe más claramente sus propiedades cuando se expresa con respecto a una base compuesta por vectores característicos. Cuando esto es posible, la representación matricial de la transformación es una matriz diagonal que tiene sobre la diagonal los valores característicos. El poder hacer esto tiene muchas ventajas en cuanto al manejo algebraico; sin embargo, hay casos en los cuales no se puede obtener una base del espacio compuesta por vectores propios, esto es, hay menos de n vectores característicos linealmente independientes y la matriz no es diagonalizable. Dadas las dificultadas algebraicas para operar con una matriz no diagonal, se plantea entonces el problema de encontrar una matriz 'lo más simple posible' semejante a la matriz de la transformación. La solución del problema es la *forma canónica de Jordan*, la cual engloba la teoría general de la diagonalización, dando para cualquier matriz cuadrada, diagonalizable o no, su *matriz de Jordan* equivalente. Este será precisamente el enfoque con el cual abordaremos el tema.

Objetivos

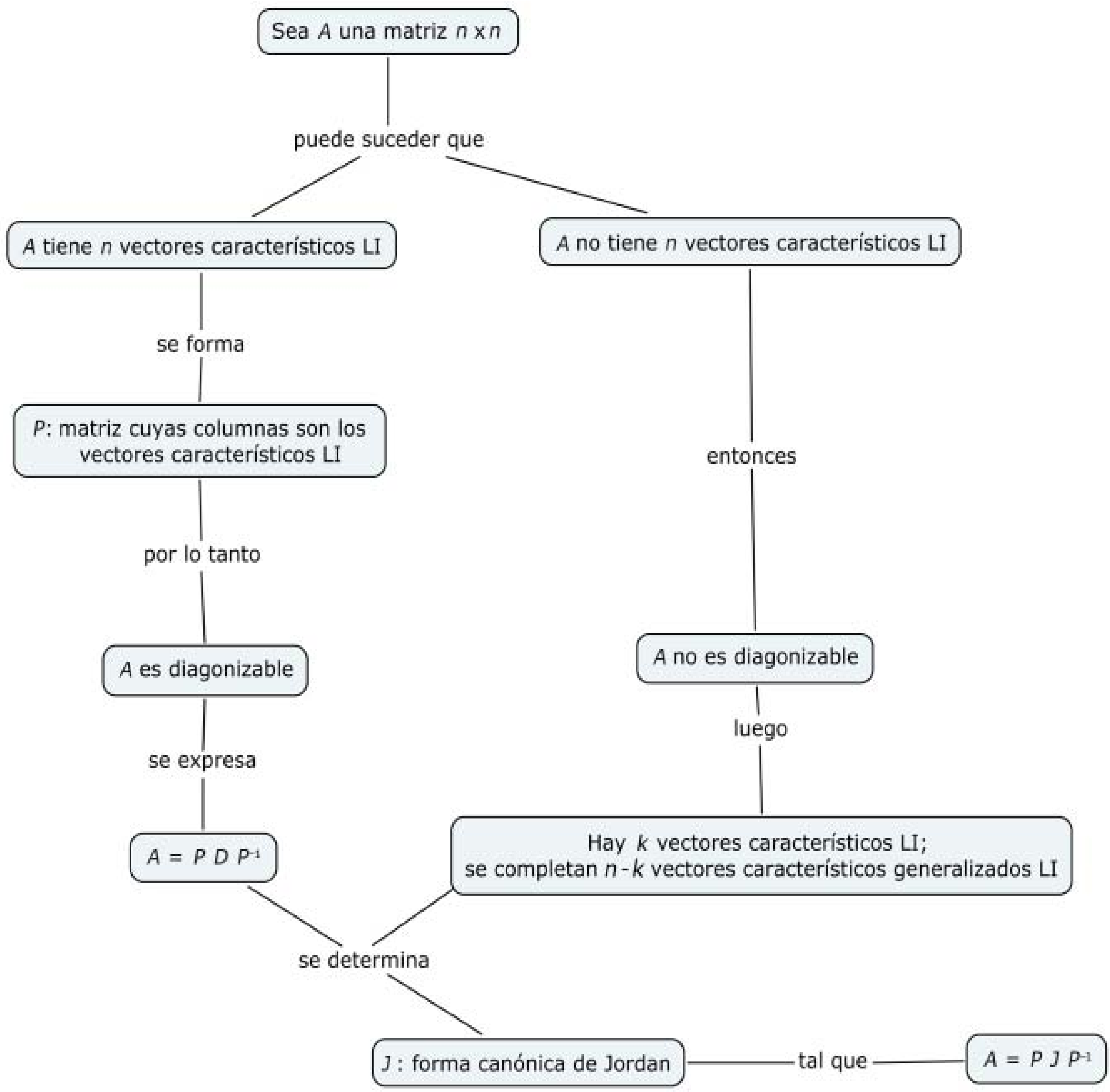
1. Determinar para cualquier matriz $A_{n\times n}$, diagonalizable o no, una matriz semejante J que es su forma canónica de Jordan.

Preguntas básicas

1. ¿Qué es una matriz bloque de Jordan?
2. ¿Qué es una matriz de Jordan?
3. ¿Cómo está formada la diagonal en una matriz de Jordan?
4. ¿Qué es un vector característico generalizado?
5. ¿Cómo se conforma la matriz P que permite llevar $A_{2\times 2}$ a su forma canónica de Jordan?
6. Si A es una matriz 3×3, ¿cómo se obtiene la matriz P que lleva a A a su forma canónica de Jordan?

Vea el módulo 19 del programa de televisión *Álgebra lineal*.

Contenidos



Mapa 12: módulo 19

19.1 Matriz de Jordan

Escuche la biografía de *Camille Jordan* en su multimedia de *Álgebra lineal.*

Comenzamos el proceso definiendo una matriz cuadrada N_k, así:

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

N_k es una matriz con unos arriba de la diagonal principal y ceros en las demás posiciones.

Ahora formamos una nueva matriz $B(\lambda) = \lambda I + N_k$; a esta matriz la llamamos *matriz de bloques de Jordan.*

$$\underset{kxk}{B(\lambda)} = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}.$$

$B(\lambda)$ es una matriz que tiene un valor λ fijo sobre la diagonal, unos encima de la diagonal y ceros en las demás posiciones.

Podemos considerar que una matriz bloque de Jordan de 1×1 (orden 1) será $B(\lambda) = (\lambda)$.

Con las matrices de bloques de Jordan se forma la matriz de Jordan *J*, la cual tiene la siguiente forma:

$$J = \begin{bmatrix} B_1(\lambda_1) & 0 & \cdots & 0 \\ 0 & B_2(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & B_r(\lambda_r) \end{bmatrix}.$$

J es una matriz que tiene en la diagonal matrices de bloques de Jordan y ceros en las demás posiciones.

Ejemplos

1. $$\begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix} \qquad \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

   Las anteriores son matrices de Jordan; los bloques de Jordan se han marcado con líneas punteadas.

2. Veamos las posibles matrices de Jordan de 3×3.

$\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$ formada por un bloque de Jordan de orden 3.

$\begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix}$ formada por un bloque de Jordan de orden 2 y uno de orden 1.

$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{bmatrix}$ formada por un bloque de Jordan de orden 1 y uno de orden 2.

$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$ formada por tres bloques de Jordan de orden 1.

$\lambda_1, \lambda_2, \lambda_3$ no necesariamente son distintos.

19.2 Forma canónica de Jordan

Para cualquier matriz A, real o compleja, se puede demostrar la existencia de una matriz de Jordan J semejante a A, tal que:

$$J = P^{-1}AP\cdot$$

Este hecho es uno de los resultados más importantes del álgebra lineal, aunque su demostración va más allá del alcance de un curso inicial.

La matriz J tiene sobre la diagonal los valores característicos de A. Esta matriz es única, excepto por el orden en que aparecen los bloques de Jordan.

Escuche el audio *El puente de Tacoma* en su multimedia de *Álgebra lineal.*

Ejemplo 3

Si A es similar a $J_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}$ también es similar a:

$$J_2 = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \quad J_3 = \begin{bmatrix} 4 & 1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix},$$

$$J_4 = \begin{bmatrix} 4 & 1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

y otras dos matrices de Jordan.

La matriz J se llama *forma canónica de Jordan de A*. Si A es diagonalizable, la forma canónica de Jordan de A será la matriz diagonal D equivalente a A, donde los bloques de Jordan serán los valores característicos de A, $\lambda_1, \lambda_2, ..., \lambda_n$ no necesariamente distintos.

A continuación veremos cómo obtener la forma canónica de Jordan para una matriz $A_{2\times 2}$. Si A es diagonalizable ya sabemos cómo obtener su matriz diagonal equivalente. Sólo resta analizar el caso en el cual A tiene un solo valor característico λ de multiplicidad algebraica 2 y multiplicidad geométrica 1. Bajo estas condiciones, si $\mathbf{v}_1$ es un vector característico correspondiente al valor característico λ, entonces existe un vector $\mathbf{v_2}$ que satisface la ecuación:

$$(A - \lambda I)\mathbf{v_2} = \mathbf{v_1}.$$

19.3 Procedimiento para obtener la forma canónica de Jordan de A_{2x2}

Definición 1

Sea A una matriz de 2×2 con un solo valor propio λ con multiplicidad algebraica 2 y multiplicidad geométrica 1. Si $\mathbf{v}_1$ es un vector característico correspondiente a λ, entonces al vector $\mathbf{v}_2$ tal que $(A-\lambda I)\mathbf{v}_2 = \mathbf{v}_1$ se le llama *vector característico generalizado de A correspondiente al valor característico* λ.

El siguiente teorema nos muestra la necesidad de encontrar el vector $\mathbf{v}_2$ como un vector característico generalizado.

Teorema 1

Sea $A_{2\times 2}$ con un único valor característico λ y un único vector característico $\mathbf{v}_1$ linealmente independiente.

Sea $\mathbf{v}_2$ un vector característico generalizado de A correspondiente al valor característico λ, y P la matriz cuyas columnas son los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$. Entonces $P^{-1}AP = J$, donde $J = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ es la forma canónica de Jordan de A.

Demostración

Como $\mathbf{v}_2$ no es un vector característico de A, $\mathbf{v}_2 \neq \alpha\mathbf{v}_1$, es decir, $\mathbf{v}_1$ y $\mathbf{v}_2$ son LI, entonces P es invertible.

$$AP = A[\mathbf{v}_1, \mathbf{v}_2] = [A\mathbf{v}_1, A\mathbf{v}_2] = [\lambda\mathbf{v}_1, A\mathbf{v}_2].$$

Como $(A-\lambda I)\mathbf{v}_2 = \mathbf{v}_1$, $A\mathbf{v}_2 - \lambda\mathbf{v}_2 = \mathbf{v}_1$, $A\mathbf{v}_2 = \mathbf{v}_1 + \lambda\mathbf{v}_2$, entonces:

$$AP = [\lambda\mathbf{v}_1, \mathbf{v}_1 + \lambda\mathbf{v}_2].$$

Ahora:

$$PJ = [\mathbf{v}_1, \mathbf{v}_2]\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} = [\lambda\mathbf{v}_1, \mathbf{v}_1 + \lambda\mathbf{v}_2].$$

Luego:

$$AP = PJ,$$

de donde $P^{-1}AP = J$.

Ejemplo 4

Transforme la matriz $A=\begin{bmatrix}-10 & -7\\ 7 & 4\end{bmatrix}$ a su forma canónica de Jordan.

Solución

$$\det(A-\lambda I)=\begin{vmatrix}-10-\lambda & -7\\ 7 & 4-\lambda\end{vmatrix}=0\,,\quad \lambda^2+6\lambda+9=0\to(\lambda+3)^2=0.$$

Luego $\lambda=-3$ es el único valor característico de A, con multiplicidad algebraica 2.

Determinemos los vectores característicos correspondientes a $\lambda=-3$.

$$(A+3I)\mathbf{x}=\mathbf{0},$$

$$\begin{bmatrix}-7 & -7\\ 7 & 7\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix},\quad \text{entonces}\quad \begin{aligned}-7x-7y&=0\\ -7x&=7y\\ -x&=y\end{aligned}$$

Por lo tanto:

$$E_{-3}=\text{gen}\left\{\begin{bmatrix}1\\ -1\end{bmatrix}\right\}.$$

Si $\mathbf{v}_1=\begin{bmatrix}1\\ -1\end{bmatrix}$, el vector $\mathbf{v_2}$ debe encontrarse como un vector característico generalizado.

Para esto resolvemos $(A+3I)\mathbf{v}_2=\mathbf{v}_1$.

$$\begin{bmatrix}-7 & -7\\ 7 & 7\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}1\\ -1\end{bmatrix},\quad \text{entonces}\quad \begin{aligned}-7x-7y&=1\\ 7x+7y&=-1\\ x+y&=-\frac{1}{7}\end{aligned}$$

Si $x=0,\quad y=-1/7,$

$$\text{luego}\ \mathbf{v}_2=\begin{bmatrix}0\\ -1/7\end{bmatrix}.$$

$$P=\begin{bmatrix}1 & 0\\ -1 & -1/7\end{bmatrix},\quad P^{-1}=-7\begin{bmatrix}-1/7 & 0\\ 1 & 1\end{bmatrix}=\begin{bmatrix}1 & 0\\ -7 & -7\end{bmatrix},$$

$$J=P^{-1}AP=\begin{bmatrix}1 & 0\\ -7 & -7\end{bmatrix}\begin{bmatrix}-10 & -7\\ 7 & 4\end{bmatrix}\begin{bmatrix}1 & 0\\ -1 & -1/7\end{bmatrix}=\begin{bmatrix}-3 & 1\\ 0 & -3\end{bmatrix}.$$

19.4 Generalización del procedimiento para obtener la forma canónica de Jordan de A_{nxn}

El método descrito puede generalizarse para obtener la forma canónica de Jordan de cualquier matriz. Es posible determinar el número de unos arriba de la diagonal en la forma canónica de Jordan de una matriz *A* de $n \times n$.

Sea λ_i un valor característico de *A* con multiplicidad algebraica r_i y multiplicidad geométrica t_i. Si $\lambda_1, \lambda_2, ..., \lambda_k$ son los valores característicos de *A*, entonces el número de unos arriba de la diagonal de la forma canónica de Jordan de *A* es:

$$(r_1 - t_1) + (r_2 - t_2) + ... + (r_k - t_k)$$

$$= \sum_{i=1}^{k} r_i - \sum_{i=1}^{k} t_i = n - \sum_{i=1}^{k} t_i.$$

Si se conoce la ecuación característica de *A*, entonces se pueden determinar las posibles formas canónicas de Jordan de *A*.

Ejemplo 5

Si el polinomio característico de *A* es $(\lambda - 3)^3(\lambda + 4)$, entonces las posibles formas canónicas de Jordan de *A*, son:

1. $\lambda_1 = 3$ con multiplicidad geométrica 3 y $\lambda_2 = -4$.

$$J = D = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -4 \end{bmatrix}.$$

2. $\lambda_1 = 3$ con multiplicidad geométrica 2 y $\lambda_2 = -4$.

$$J = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -4 \end{bmatrix}.$$

3. $\lambda_1 = 3$ con multiplicidad geométrica 1 y $\lambda_2 = -4$.

$$J = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -4 \end{bmatrix}.$$

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «La forma canónica de Jordan».

Generalizando el procedimiento para hallar el vector característico generalizado $\mathbf{v}_2$ de la matriz $A_{2\times 2}$ no diagonalizable, podemos describir el proceso para encontrar los vectores característicos generalizados en una matriz $A_{3\times 3}$ no diagonalizable.

Sea A una matriz 3×3. Suponga que λ es un valor característico de A con multiplicidad algebraica 3 y multiplicidad geométrica 1 y sea $\mathbf{v}_1$ el vector propio correspondiente. Se puede demostrar que existe un vector $\mathbf{v}_2$ tal que $(A-\lambda I)\mathbf{v}_2=\mathbf{v}_1$ con $\mathbf{v}_1$ y $\mathbf{v}_2$ linealmente independientes. Una vez obtenido $\mathbf{v}_2$ puede determinarse $\mathbf{v}_3$ resolviendo $(A-\lambda I)\mathbf{v}_3=\mathbf{v}_2$ tal que $\mathbf{v}_1, \mathbf{v}_2$ y $\mathbf{v}_3$ son linealmente independientes. Las columnas de la matriz P serán los vectores $\mathbf{v}_1, \mathbf{v}_2$ y $\mathbf{v}_3$ y la forma canónica de Jordan

será $J=\begin{bmatrix}\lambda & 1 & 0\\ 0 & \lambda & 1\\ 0 & 0 & \lambda\end{bmatrix}.$

Ejemplo 6

En el ejemplo 6 del módulo 16, para la matriz:

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

con ecuación característica $(\lambda-1)^3=0,$ determinamos un solo vector característico correspondiente a $\lambda=1,\ \mathbf{v}_1=\begin{bmatrix}1\\1\\1\end{bmatrix}.$

Encontremos los vectores característicos generalizados $\mathbf{v}_2$ y $\mathbf{v}_3,$ para determinar la matriz $P=[\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3],$ tal que $P^{-1}AP=J.$

Resolvamos $(A-I)\mathbf{v}_2=\mathbf{v}_1.$

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

$$\left[\begin{array}{ccc|c}-1 & 1 & 0 & 1\\ 0 & -1 & 1 & 1\\ 1 & -3 & 2 & 1\end{array}\right]\rightarrow\left[\begin{array}{ccc|c}1 & 0 & -1 & -2\\ 0 & 1 & -1 & -1\\ 0 & 0 & 0 & 0\end{array}\right],\qquad \begin{aligned}x&=-2+z\\ y&=-1+z\\ z&=z\end{aligned}$$

Si $z=0,\ x=-2$ e $y=-1$, entonces $\mathbf{v}_2=\begin{bmatrix}-2\\-1\\0\end{bmatrix}.$

Ahora encontremos $\mathbf{v}_3$, resolviendo $(A-\lambda I)\mathbf{v}_3=\mathbf{v}_2.$

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

$$\left[\begin{array}{ccc:c}-1&1&0&-2\\0&-1&1&-1\\1&-3&2&0\end{array}\right]\rightarrow\left[\begin{array}{ccc:c}1&0&-1&3\\0&1&-1&1\\0&0&0&0\end{array}\right],\qquad \begin{aligned}x&=3+z\\y&=1+z\\z&=\quad z\end{aligned}$$

Si $z=0$, entonces $\mathbf{v}_3=\begin{bmatrix}3\\1\\0\end{bmatrix}.$

Luego:

$$P=\begin{bmatrix}1&-2&3\\1&-1&1\\1&0&0\end{bmatrix},\qquad P^{-1}=\begin{bmatrix}0&0&1\\1&-3&2\\1&-2&1\end{bmatrix}.$$

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

Ejercicios del capítulo 4 (módulo 19)

En los ejercicios 1 a 9 determine si la matriz dada es de Jordan.

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

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

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

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

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

6. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 4 \end{bmatrix}$.

7. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$.

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

9. $\begin{bmatrix} a & 0 & 0 & 0 & 0 \\ 0 & b & 1 & 0 & 0 \\ 0 & 0 & b & 1 & 0 \\ 0 & 0 & 0 & c & 0 \\ 0 & 0 & 0 & 0 & c \end{bmatrix}$.

En los ejercicios 10 a 12 encuentre una matriz invertible P que transforme la matriz de 2×2 a su forma canónica de Jordan.

10. $\begin{bmatrix} 2 & -7 \\ 7 & -12 \end{bmatrix}$.

11. $\begin{bmatrix} 4 & 7 \\ -7 & -10 \end{bmatrix}$.

12. $\begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}$.

13. Reduzca la matriz a su forma canónica de Jordan: $\begin{bmatrix} -2 & 1 & 0 \\ -2 & 1 & -1 \\ -1 & 1 & -2 \end{bmatrix}$.

14. Haga lo mismo que en el ejercicio anterior con $A = \begin{bmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$.

15. Escriba todas las matrices de Jordan de 4×4 posibles.

En los ejercicios 16 a 19 se da el polinomio característico de una matriz A. Escriba todas las posibles formas canónicas de Jordan de A.

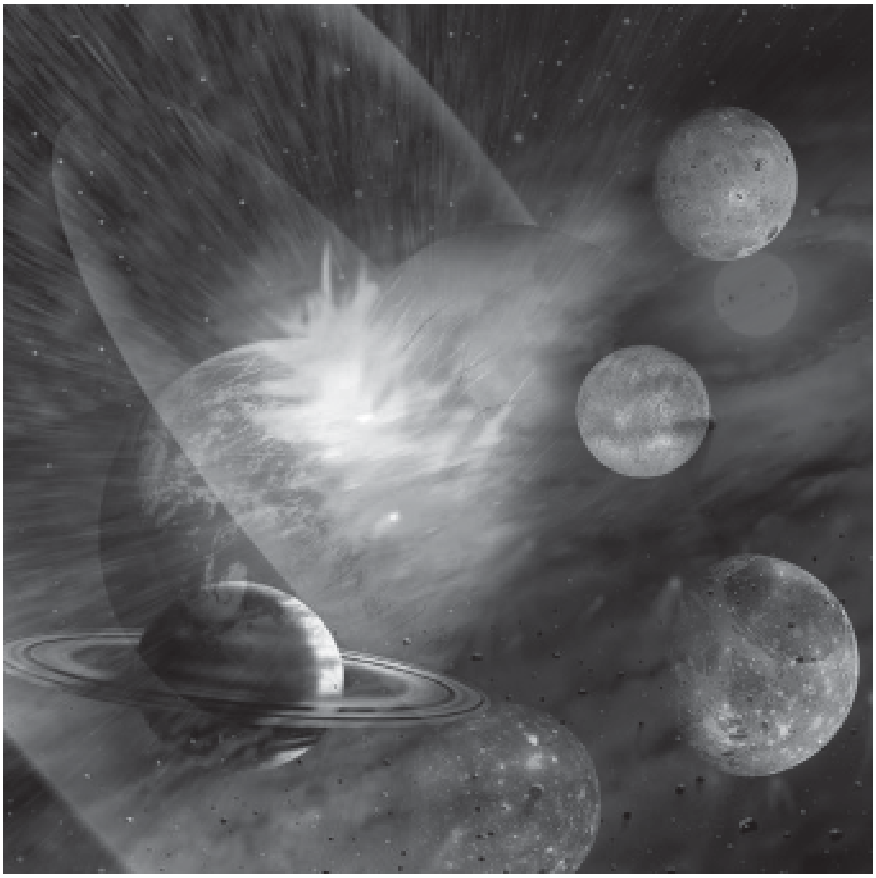
16. $\lambda^2(\lambda-1)^2$.

17. $(\lambda+3)^2(\lambda-4)^3$.

18. $(\lambda-2)(\lambda+3)^2$.

19. $(\lambda-7)^3$.

20. Usando la forma canónica de Jordan, demuestre que para cualquier matriz A de $n\times n$, $\det A=\lambda_1\lambda_2,\ldots,\lambda_n$, donde $\lambda_1,\lambda_2,\ldots,\lambda_n$ son los valores característicos de A.



Según Morris Kline[1], los valores característicos se originaron en el contexto de formas cuadráticas y en la mécanica celeste (el movimiento de los planetas), conociéndose como raíces características de la ecuación escalar.

Capítulo 5

Diagonalización ortogonal. Formas cuadráticas y aproximación de valores y vectores característicos

Contenido breve

Presentación

Las matrices simétricas surgen en las aplicaciones, con mayor frecuencia que cualquier otra clase de matrices, y esto se da por las propiedades especiales que tienen, entre ellas la más relevante es que siempre son diagonalizables y se puede formar una base ortonormal con sus vectores característicos. Estas propiedades enlazan la teoría de la diagonalización con la ortogonalidad, lo que da como resultado la *diagonalización ortogonal de las matrices simétricas.*

Una aplicación importante de la diagonalización ortogonal está en la reducción de las *formas cuadráticas* a sus ejes principales. Estas formas cuadráticas las asociamos a la ecuación general de segundo grado y cada ecuación de segundo grado, a la sección cónica correspondiente llevando la expresión de la cónica a su forma canónica.

Finalizamos el capítulo presentando el teorema de las circunferencias de Gershgorin, el cual proporciona un intervalo donde se pueden acotar los valores característicos cuando no se precisa conocerlos con exactitud. Además, se desarrolla un método iterativo para calcular el valor característico y el vector característico dominante de una matriz. Estos procedimientos se hacen necesarios debido a que cuando A es una matriz de orden grande, el cálculo de los valores característicos por medio de la ecuación característica se transforma en un problema algebraico difícil.

[1] En *Mathematical thought from ancient to modern times* (Fair Lawn, NJ.: Oxford University Press, 1972).

Módulo 20

Diagonalización ortogonal

El movimiento horizontal del sistema de masas y resortes en el cual todas las masas son iguales y todos los resortes son iguales, se puede analizar diagonalizando la matriz simétrica

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

Introducción

Estudiaremos varias propiedades importantes de las matrices simétricas, como que siempre tienen n vectores característicos LI y todos sus valores característicos son números reales. Ahora, no sólo se garantiza su diagonalización, sino que además se puede comprobar que se puede formar una base ortonormal con sus vectores característicos.

Objetivos

1. Caracterizar las matrices simétricas como aquellas que son diagonalizables ortogonalmente.
2. Desarrollar un algoritmo para determinar la matriz ortogonal que diagonaliza una matriz simétrica A.

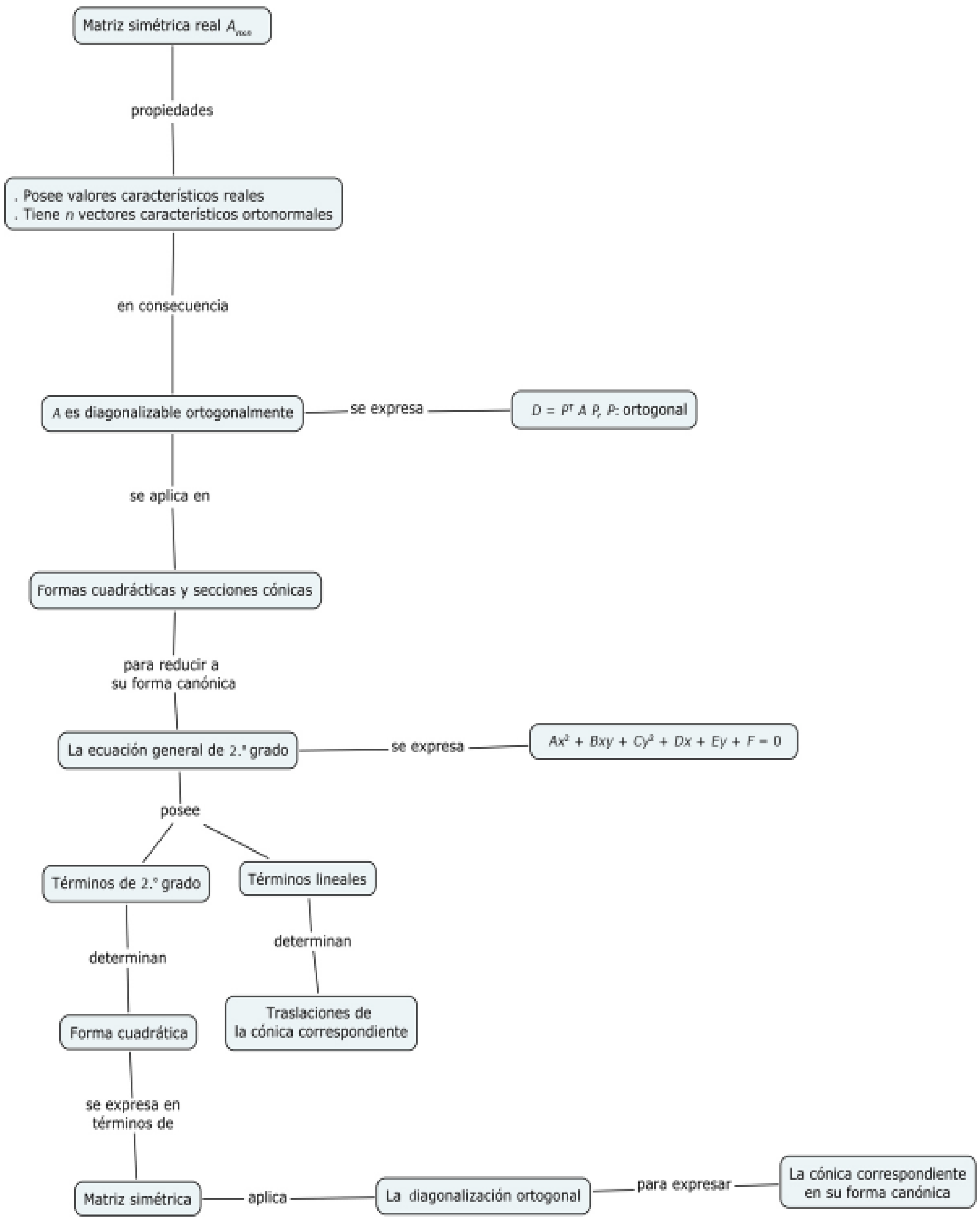
Preguntas básicas

1. ¿Cómo son los valores característicos de una matriz simétrica real A?
2. ¿Cómo son los vectores característicos que corresponden a valores característicos diferentes?
3. ¿Qué significa que una matriz es diagonalizable ortogonalmente?
4. ¿Qué se puede decir de una matriz que es diagonalizable ortogonalmente?

Contenidos

20.1 Diagonalización ortogonal
20.2 Procedimiento para encontrar n vectores característicos ortonormales de una matriz simétrica $A_{n\times n}$

Vea el módulo 20 del programa de televisión *Álgebra lineal*.



Mapa 13: módulos 20 y 21

20.1 Diagonalización ortogonal

Escuche la biografía de *Leonhard Euler* en su multimedia de *Álgebra lineal.*

Teorema 1

Si $\underset{n\times n}{A}$ es una matriz simétrica con elementos reales, entonces:

a. Sus valores característicos son reales.

b. Los vectores característicos correspondientes a distintos valores característicos son ortogonales.

Demostración

a. Sea λ un valor característico de A con vector característico $\mathbf{u}$; entonces:

$$A\mathbf{u} = \lambda\mathbf{u},\ \mathbf{u} \in \mathbb{C}^n \ \text{ y } \ \mathbf{u} \neq \mathbf{0}.$$

El producto interno entre el vector $A\mathbf{u}$ y $\mathbf{u}$ está dado por:

$(A\mathbf{u}, \mathbf{u}) = (\lambda\mathbf{u}, \mathbf{u}) = \lambda(\mathbf{u}, \mathbf{u})$ (Prop. *VI*, definición de producto interno),

o $(A\mathbf{u}, \mathbf{u}) = (\mathbf{u}, A^t\mathbf{u}) = (\mathbf{u}, A\mathbf{u})$ ya que $A = A^t$ (teorema 2, módulo 15)

$= (\mathbf{u}, \lambda\mathbf{u}) = \bar{\lambda}(\mathbf{u}, \mathbf{u})$ (Prop. *VII*, definición de producto interno),

luego $\lambda(\mathbf{u}, \mathbf{u}) = \bar{\lambda}(\mathbf{u}, \mathbf{u})$, $(\mathbf{u}, \mathbf{u}) = \|\mathbf{u}\|^2 \neq 0$, entonces $\lambda = \bar{\lambda}$.

Si $\lambda = a + bi$, $\bar{\lambda} = a - bi$,

$$a + bi = a - bi \quad \leftrightarrow \quad b = 0.$$

Luego $\lambda = a$; por lo tanto, λ es un número real.

b. Sean λ_1 y λ_2 valores característicos distintos correspondientes a vectores característicos $\mathbf{u}_1$ y $\mathbf{u}_2$. Entonces:

$$A\mathbf{u}_1 \cdot \mathbf{u}_2 = \lambda_1\mathbf{u}_1 \cdot \mathbf{u}_2 = \lambda_1(\mathbf{u}_1 \cdot \mathbf{u}_2),$$

o $$A\mathbf{u}_1 \cdot \mathbf{u}_2 = \mathbf{u}_1 \cdot A^t\mathbf{u}_2 = \mathbf{u}_1 \cdot A\mathbf{u}_2 = \mathbf{u}_1 \cdot \lambda_2\mathbf{u}_2 = \lambda_2(\mathbf{u}_1 \cdot \mathbf{u}_2),$$

luego $\lambda_1(\mathbf{u}_1 \cdot \mathbf{u}_2) = \lambda_2(\mathbf{u}_1 \cdot \mathbf{u}_2)$.

Como $\lambda_1 \neq \lambda_2$, entonces $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$, lo cual significa que $\mathbf{u}_1$ y $\mathbf{u}_2$ son ortogonales.

Observación: note que en la parte *b* del teorema se plantea el producto escalar entre $A\mathbf{u}_1$ y $\mathbf{u}_2$; esto se puede hacer así, ya que en la parte *a* del teorema se ha demostrado que los valores característicos son números reales, y como A es una matriz real, los vectores característicos serán elementos de $\mathbb{R}^n$.

Teorema 2

Sea A una matriz simétrica real de $n \times n$; entonces A tiene n vectores característicos ortonormales.

La demostración del teorema excede el alcance de este curso; sin embargo, podemos verificarlo siempre que trabajamos con matrices simétricas.

Del teorema se desprende, como consecuencia, que toda matriz simétrica real es diagonalizable y sus vectores característicos no sólo son LI, sino que son ortogonales. Si estos vectores se normalizan, se tienen n vectores característicos ortonormales de modo que la matriz P que diagonaliza la matriz simétrica A es ortogonal.

Definición 1

Una matriz $A_{n\times n}$ es *diagonalizable ortogonalmente* si existe una matriz ortogonal P tal que:

$$P^T A P = D,$$

donde D es la matriz diagonal que tiene sobre la diagonal los valores característicos de A.

Teorema 3

Sea A una matriz real de $n \times n$; entonces, A es diagonalizable ortogonalmente si y sólo si A es simétrica.

Demostración

Si A es simétrica, el teorema 2 y la definición anterior nos permiten concluir que A es diagonalizable ortogonalmente.

Recíprocamente, si A es diagonalizable ortogonalmente, existe P ortogonal, tal que: $P^T A P = D$, luego $A = P D P^T$; entonces:

$$A^T = (P D P^T)^T = (P^T)^T D^T P^T = P D P^T.$$

De donde $A = A^T$, es decir, A es simétrica.

Si $A_{n\times n}$ es una matriz simétrica con n valores característicos diferentes, sus vectores característicos serán ortogonales. Ahora, si hay valores característicos con multiplicidades mayores que 1, los vectores característicos que se obtienen al resolver el espacio característico correspondiente no necesariamente son ortogonales; si aplicamos el proceso de Gram-Schmidt, obtendremos vectores ortonormales, ¿pero serán todavía vectores característicos? La respuesta es sí, ya que al aplicar el proceso de Gram-Schmidt sólo se toman combinaciones lineales particulares de los vectores de la base del espacio característico y, por lo tanto, los vectores ortonormales obtenidos serán también vectores característicos correspondientes al mismo valor característico λ.

20.2 Procedimiento para encontrar *n* vectores característicos ortonormales de una matriz simétrica A_{nxn}

a. Encontrar los valores característicos de A.

b. Encontrar una base para cada espacio característico.

c. Aplicar el proceso de Gram-Schmidt para obtener una base ortonormal de cada espacio característico. De esta forma todos los vectores obtenidos forman un conjunto de n vectores característicos ortonormales.

Ejemplo 1

En el ejemplo 5 del módulo 16, dada $A=\begin{bmatrix}1 & -1 & 0\\ -1 & 2 & -1\\ 0 & -1 & 1\end{bmatrix}$, determinamos sus valores característicos 0, 1 y 3 con sus correspondientes vectores característicos:

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

La matriz A es simétrica. Veamos que sus vectores característicos son ortogonales.

$$\begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}-1\\0\\1\end{pmatrix}=0,\quad \begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\-2\\1\end{pmatrix}=0,\quad \begin{pmatrix}-1\\0\\1\end{pmatrix}\cdot\begin{pmatrix}1\\-2\\1\end{pmatrix}=0.$$

Para tener una base ortonormal de $\mathbb{R}^3$ sólo se necesita normalizar cada uno de estos vectores.

La base ortonormal está dada por:

$$\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix},\quad \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\1\end{pmatrix},\quad \frac{1}{\sqrt{6}}\begin{pmatrix}1\\-2\\1\end{pmatrix}.$$

Luego, la matriz P cuyas columnas son estos vectores es una matriz ortogonal, de modo que $P^T A\, P = D$.

$$\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{bmatrix}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Ejemplo 2

Sea $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. Diagonalicemos la matriz A ortogonalmente.

Resolvemos el polinomio característico de A, así:

$$\det(A-\lambda I) = \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} \qquad \begin{aligned} &= -\lambda^3+3\lambda+2 = 0 \\ &= \lambda^3-3\lambda-2 = 0 \\ &(\lambda+1)^2(\lambda-2) = 0 \end{aligned}$$

Los valores característicos de A son $\lambda_1 = -1$ con multiplicidad algebraica de 2 y $\lambda_2 = 2$.

Para determinar E_{-1}, resolvemos $(A+I)\mathbf{x} = \mathbf{0}.$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ luego } \quad \begin{aligned} x &= -y-z \\ y &= y \\ z &= \phantom{-y-{}}z \end{aligned}$$

Entonces $E_{-1} = \text{gen}\left\{\mathbf{v}_1 = \begin{pmatrix}-1\\1\\0\end{pmatrix},\ \mathbf{v}_2 = \begin{pmatrix}-1\\0\\1\end{pmatrix}\right\}.$

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Diagonalización ortogonal».

Ahora, $\mathbf{v}_1$ y $\mathbf{v}_2$ no son mutuamente ortogonales, aunque sí son linealmente independientes. Aplicamos a esta base el proceso de Gram-Schmidt.

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{|\mathbf{v}_1|} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Sea $\mathbf{v}_2{}' = \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1$

$$= \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \left[\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}\right]\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \end{bmatrix}, \quad |\mathbf{v}_2{}'| = \frac{\sqrt{6}}{2}.$$

$$\mathbf{u}_2 = \frac{\mathbf{v}_2{}'}{|\mathbf{v}_2{}'|} = \frac{2}{\sqrt{6}}\begin{pmatrix} -1/2 \\ -1/2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix}.$$

E_2 lo encontramos resolviendo $(A-2I)\mathbf{x} = \mathbf{0}$.

$$\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}; \text{ entonces:}$$

$$\begin{aligned} x &= z \\ y &= z \\ z &= z \end{aligned} \quad , \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Luego:

$$E_2 = \text{gen}\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}.$$

Para completar la base ortonormal, dividimos $\mathbf{v}_3$ por su magnitud y obtenemos:

$$\mathbf{u}_3 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}.$$

La matriz $P = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$ es la matriz ortogonal que diagonaliza la matriz simétrica *A*.

Así que:

$$P^T AP = D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Ejercicios del capítulo 5 (módulo 20)

En los ejercicios 1 a 7 diagonalice ortogonalmente las matrices simétricas dadas, determinando la matriz diagonal D y la matriz diagonalizante ortogonal P.

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

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

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

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

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

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

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

8. Suponga que $A_{n\times n}$ es una matriz simétrica real para la que todos sus valores característicos son cero. Demuestre que A es la matriz nula $n \times n$.

9. Demuestre que si una matriz real A de 2×2 tiene vectores propios ortogonales, entonces A es simétrica.

10. Muestre que si A es invertible y ortogonalmente diagonalizable, entonces A^{-1} es ortogonalmente diagonalizable.

11. Sea A una matriz real antisimétrica. Demuestre que todo valor propio de A es de la forma bi, donde $b \in \mathbb{R}$ e i es la unidad imaginaria.

Módulo 21

Formas cuadráticas y secciones cónicas

La expresión de una forma cuadrática en términos de una matriz simétrica, permite utilizar las propiedades de las matrices simétricas para reducir las ecuaciones de las secciones cónicas representadas a sus ejes principales.

Introducción

En este módulo presentamos una aplicación de la diagonalización ortogonal. Las propiedades de las matrices simétricas, estudiadas en la sección anterior, nos proporcionan una herramienta importante en la reducción de las ecuaciones de segundo grado cuando en ellas hay un término cruzado, es decir, un término en xy.

Objetivos

1. Mostrar una aplicación de la diagonalización ortogonal.
2. Completar el estudio de la ecuación general de segundo grado cuando la ecuación corresponde a una cónica rotada y/o trasladada respecto de su posición canónica.
3. Extender el estudio de las formas cuadráticas a las superficies cuadráticas en $\mathbb{R}^3$.

Preguntas básicas

1. ¿Qué es una forma cuadrática en $\mathbb{R}^2$ y cómo se relaciona con la ecuación general de segundo grado?
2. ¿Cómo se expresa matricialmente una forma cuadrática?
3. ¿Cómo se emplea la diagonalización ortogonal de la matriz A de la forma cuadrática, para llevar ésta a sus ejes principales?
4. ¿Cómo se determina el ángulo θ que deben rotar los ejes x y y para llevar la cónica a nuevos ejes $x'y'$ y expresarla en su forma canónica?
5. ¿Cómo debe ser la matriz ortogonal P que diagonaliza la matriz A para que corresponda a una matriz de rotación?
6. En la ecuación de segundo grado, eliminado el término en xy, ¿cómo se puede saber qué tipo de cónica representa?

Contenidos

21.1 Secciones cónicas

21.2 Formas cuadráticas en $\mathbb{R}^2$

21.3 Teorema 1: Teorema de los ejes principales en $\mathbb{R}^2$

21.4 Formas cuadráticas en más de dos variables

Vea el módulo 21 del programa de televisión *Álgebra lineal*.

21.1 Secciones cónicas

En el estudio de las secciones cónicas en el plano realizado en el texto de *Geometría vectorial y analítica* se han asociado estas secciones con la ecuación general de segundo grado:

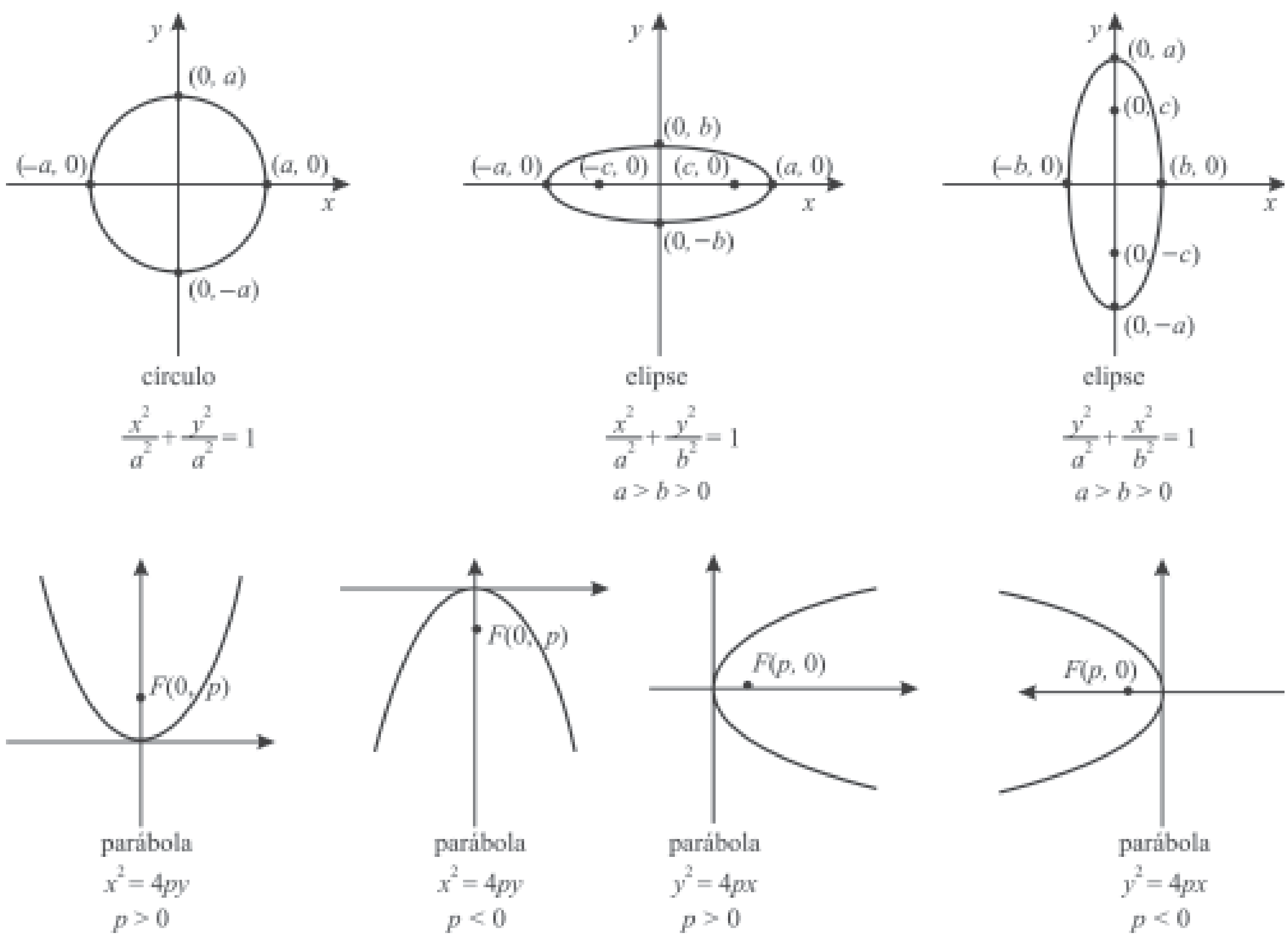
$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0,$$

donde A, B, C, D, E y F son constantes reales. (1)

La gráfica de esta ecuación es una *sección cónica* o un caso degenerado de éstas, como un punto, una recta, un par de rectas o el conjunto vacío.

Cuando la constante B es cero, o sea cuando la ecuación no tiene término en xy, los ejes de simetría de la cónica son paralelos a los ejes coordenados.

Las cónicas no degeneradas están en *posición canónica* si sus gráficas y ecuaciones son como se indica en la figura 21.1. La ecuación está en *forma canónica*.



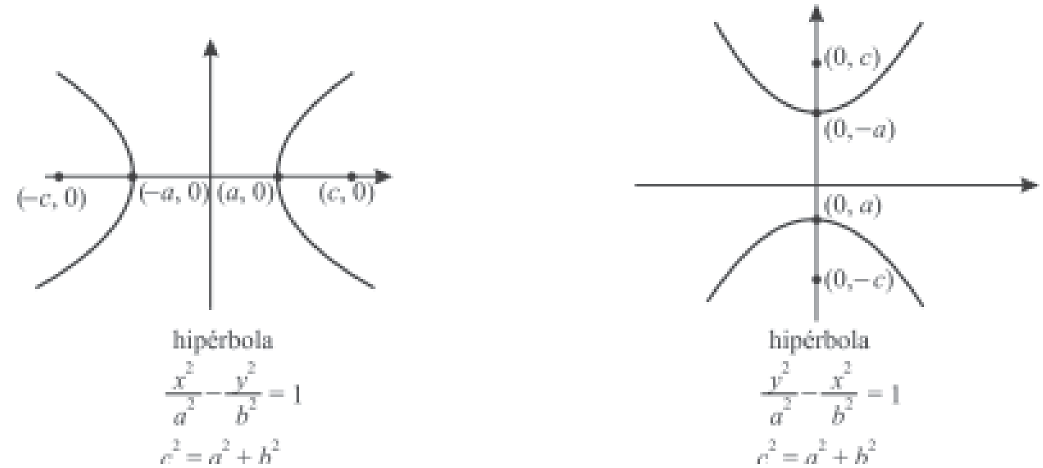


Figura 21.1

Se puede observar que las ecuaciones de las secciones cónicas cuyas gráficas están en posición canónica no tienen el término xy o *término mixto*; cuando éste aparece en la ecuación, la gráfica es una sección cónica que ha sido rotada desde su posición canónica. Además, en las ecuaciones canónicas correspondientes a círculos, elipses e hipérbolas, las variables x y y aparecen siempre al cuadrado; cuando la ecuación contiene términos cuadráticos y lineales en las variables x y y, y la ecuación no tiene término mixto, la gráfica es una sección cónica trasladada desde su posición canónica; finalmente, si aparece término en xy y términos lineales en x o en y, la gráfica es una sección cónica rotada y trasladada (figura 21.2).

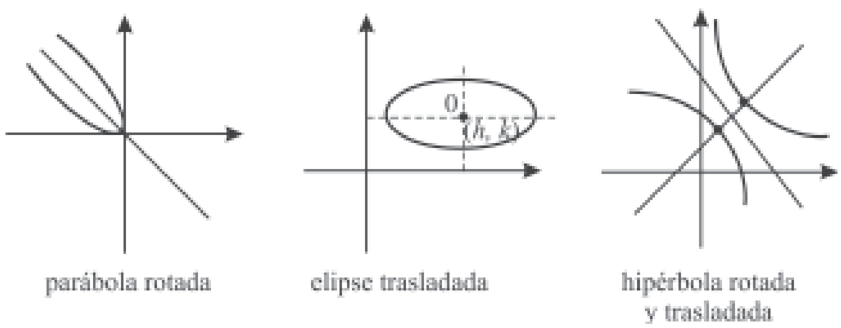


Figura 21.2

21.2 Formas cuadráticas en $\mathbb{R}^2$

En la ecuación general de segundo grado (1) consideremos únicamente los términos cuadráticos, o sea, $Ax^2 + Bxy + Cy^2$. A esta expresión la llamamos *forma cuadrática* asociada con la ecuación (1) y la denotamos por $F(x, y)$. Tenemos entonces que:

$$F(x, y) = Ax^2 + Bxy + Cy^2. \qquad (2)$$

Esta forma cuadrática la podemos representar matricialmente así:

$$F(x,y) = \mathbf{X}^T R\ \mathbf{X}, \tag{3}$$

donde $\mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}$ y R es la matriz simétrica que tiene sobre la diagonal principal los coeficientes de los términos cuadrados puros y sobre la otra diagonal se reparte simétricamente el coeficiente del término cruzado.

$$F(x,y) = (x\,y)\begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = Ax^2 + Bxy + Cy^2.$$

Por ejemplo, si $F(x,y) = 3x^2 + 2xy + y^2$, entonces:

$$\begin{aligned} F(x,y) &= [x\,y]\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \\ &= [3x+y \quad x+y]\begin{bmatrix} x \\ y \end{bmatrix} \\ &= 3x^2 + xy + xy + y^2 = 3x^2 + 2xy + y^2. \end{aligned}$$

Inversamente, si R es una matriz simétrica, entonces la ecuación (3) define una forma cuadrática.

Se puede representar $F(x,y)$ por muchas matrices, pero sólo por una matriz simétrica. En el ejemplo dado, cualquier matriz $R = \begin{bmatrix} 3 & a \\ b & 1 \end{bmatrix}$, donde $a+b=2$, verifica que $\mathbf{X}^T R\ \mathbf{X} = F(x,y)$; entonces, si $R = \begin{bmatrix} 3 & 3 \\ -1 & 1 \end{bmatrix}$, se cumple que:

$$\begin{aligned} &[x\,y]\begin{bmatrix} 3 & 3 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \\ &= [3x-y \quad 3x+y]\begin{bmatrix} x \\ y \end{bmatrix} \\ &= 3x^2 - xy + 3xy + y^2 = 3x^2 + 2xy + y^2. \end{aligned}$$

Ahora, si a la matriz R se le exige además que $a = b$, entonces sólo habrá una matriz R que satisfaga simultáneamente las dos condiciones $a+b=2$ y $a=b$.

Consideremos entonces la forma cuadrática $F(x,y) = \mathbf{X}^T R\ \mathbf{X}$ con R una matriz simétrica. Por el teorema 3 del módulo 20, R es diagonalizable ortogonalmente, esto es, existe una matriz ortogonal P tal que $R = PDP^T$, donde D es la matriz diagonal que tiene sobre la diagonal los valores característicos de R.

Entonces $\mathbf{X}^T R\,\mathbf{X}$ se puede escribir como $\mathbf{X}^T(P\,D\,P^T)\,\mathbf{X},$ y asociando, como: $(\mathbf{X}^T P)\ D\ (P^T\mathbf{X})$.

Sea $\mathbf{X}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ tal que $\mathbf{X}' = P^T\mathbf{X};$ entonces:

$$(\mathbf{X}')^T = (P^T\mathbf{X})^T = \mathbf{X}^T P.$$

Luego, la forma cuadrática expresada en las nuevas variables x', y' será:

$$F(x', y') = \mathbf{X}'^T D\ \mathbf{X}' = [x'\ y']\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} = \lambda_1 x'^2 + \lambda_2 y'^2.$$

Es decir, $F(x', y')$ es una forma cuadrática en la que no figura el término $x'y'$.

Ejemplo 1

Sea la forma cuadrática $3x^2 + 2xy + y^2$ y expresémosla en nuevas variables x' y y' donde no figure un término en $x'y'$.

$$F(x, y) = [x\ y]\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

Diagonalicemos ortogonalmente la matriz R.

$$\det(R - \lambda I) = \begin{vmatrix} 3-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = 0, \qquad \lambda^2 - 4\lambda + 2 = 0$$

$$\lambda_1 = 2 + \sqrt{2}, \quad \lambda_2 = 2 - \sqrt{2}$$

$$(R - (2+\sqrt{2})I)\mathbf{X} = \mathbf{0},$$

$$\begin{bmatrix} 1-\sqrt{2} & 1 \\ 1 & -1-\sqrt{2} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Entonces:

$$(1-\sqrt{2})x = -y \quad \text{si} \quad x = 1, \quad y = \sqrt{2} - 1,$$

luego:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ \sqrt{2}-1 \end{pmatrix} \qquad |\mathbf{v}_1| = \sqrt{1 + (\sqrt{2}-1)^2} = \sqrt{4 - 2\sqrt{2}}.$$

$$\mathbf{u}_1 = \frac{1}{\sqrt{4-2\sqrt{2}}}\begin{pmatrix} 1 \\ \sqrt{2}-1 \end{pmatrix}.$$

Después resolvemos $(R-(2-\sqrt{2})I)\mathbf{X} = \mathbf{0}$.

$$\begin{bmatrix} 1+\sqrt{2} & 1 \\ 1 & -1+\sqrt{2} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (1+\sqrt{2})x = -y;$$

entonces, si $x=-1$, $y = 1+\sqrt{2}$.

Así que:

$$\mathbf{v}_2 = \begin{pmatrix} -1 \\ 1+\sqrt{2} \end{pmatrix} \quad |\mathbf{v}_2| = \sqrt{1+(1+\sqrt{2})^2} = \sqrt{4+2\sqrt{2}}.$$

$$\mathbf{u}_2 = \frac{1}{\sqrt{4+2\sqrt{2}}}\begin{pmatrix} -1 \\ 1+\sqrt{2} \end{pmatrix}.$$

Sea $P = \begin{bmatrix} \dfrac{1}{\sqrt{4-2\sqrt{2}}} & \dfrac{-1}{\sqrt{4+2\sqrt{2}}} \\ \dfrac{\sqrt{2}-1}{\sqrt{4-2\sqrt{2}}} & \dfrac{1+\sqrt{2}}{\sqrt{4+2\sqrt{2}}} \end{bmatrix}$ la matriz ortogonal tal que $P^T R\ P = D$,

con $D = \begin{bmatrix} 2+\sqrt{2} & 0 \\ 0 & 2-\sqrt{2} \end{bmatrix}$

En las nuevas variables x' y y', la forma cuadrática se puede expresar como:

$$(2+\sqrt{2})x'^2 + (2-\sqrt{2})y'^2 = F(x', y'),$$

en la cual no hay un término en $x'y'$.

La matriz P que diagonaliza la matriz simétrica R es real y ortogonal, o sea, $P^T = P^{-1}$, entonces $1 = \det PP^{-1} = \det PP^T = \det P \det P^T = (\det P)^2$.

Por lo tanto, $\det P = \pm 1$. Si $\det P = -1$ se pueden intercambiar las columnas de P para hacer el determinante de esta nueva matriz igual a 1; esto equivale a intercambiar λ_1 y λ_2.

Veamos ahora que cuando P es real y ortogonal con determinante 1,

$P = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ para algún número θ, con $0 \le \theta \le 2\pi$, lo cual significa que P es una matriz de rotación.

Sea $P = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ real y ortogonal con $|P| = 1$.

Por el teorema 1 del módulo 15, los vectores $\begin{bmatrix} a \\ b \end{bmatrix}$ y $\begin{bmatrix} c \\ d \end{bmatrix}$ forman una base ortonormal de $\mathbb{R}^2$, es decir, son vectores unitarios y ortogonales. Luego $\begin{bmatrix} a \\ b \end{bmatrix}$ se puede expresar como $\begin{bmatrix} \cos\theta \\ \operatorname{sen}\theta \end{bmatrix}$, siendo θ el ángulo que forma el vector con el eje x positivo.

El vector $\begin{bmatrix} c \\ d \end{bmatrix}$ está formando un ángulo de $\pi/2$ con el vector $\begin{bmatrix} a \\ b \end{bmatrix}$; entonces: $\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \cos(\theta \pm \pi/2) \\ \operatorname{sen}(\theta \pm \pi/2) \end{bmatrix}$ (figura 21.3).

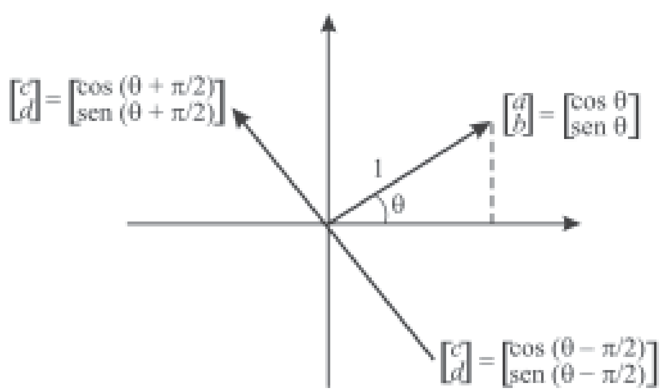


Figura 21.3

Si $\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \cos(\theta + \pi/2) \\ \operatorname{sen}(\theta + \pi/2) \end{bmatrix} = \begin{bmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{bmatrix}$, entonces:

$P = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ y $|P| = 1$.

Ahora, si $\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \cos(\theta - \pi/2) \\ \operatorname{sen}(\theta - \pi/2) \end{bmatrix} = \begin{bmatrix} \operatorname{sen}\theta \\ -\cos\theta \end{bmatrix}$, entonces $P=\begin{bmatrix}\cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta\end{bmatrix}$ y $|P|=-1$.

Luego, la matriz ortogonal P que tiene determinante 1 es una matriz de rotación; por lo tanto, se ha demostrado el siguiente teorema.

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Formas cuadráticas y secciones cónicas».

21.3 Teorema 1: Teorema de los ejes principales en $\mathbb{R}^2$

Sea $F(x,y)=Ax^2+Bxy+Cy^2$ una forma cuadrática en las variables x y y; entonces, existe un único número θ en $[0,2\pi]$ tal que $F(x,y)$ se puede escribir en la forma

$$\lambda_1 x'^2+\lambda_2 y'^2=F(x'y'),$$

donde x' y y' son los ejes obtenidos al rotar los ejes x y y un ángulo θ en sentido contrario a las manecillas del reloj.

λ_1 y λ_2 son los valores característicos de la matriz $R=\begin{bmatrix} A & B/2 \\ B/2 & C\end{bmatrix}$.

Si la forma cuadrática $F(x,y)$ está igualada a un término independiente F, tenemos una *ecuación cuadrática*, y de la ecuación $F(x',y')=F=\lambda_1 x'^2+\lambda_2 y'^2$ se dice que ha sido expresada en sus *ejes principales*.

Ejemplo 2

Tomemos la ecuación cuadrática $3x^2+2xy+y^2=4$.

En el ejemplo 1 se encontró que la forma cuadrática en sus ejes principales es $(2+\sqrt{2})x'^2+(2-\sqrt{2})y'^2$, luego la ecuación es $(2+\sqrt{2})x'^2+(2-\sqrt{2})y'^2=4$, la cual en su forma canónica es:

$$\frac{x'^2}{\dfrac{4}{2+\sqrt{2}}}+\frac{y'^2}{\dfrac{4}{2-\sqrt{2}}}=1.$$

Esta ecuación representa una elipse con $a^2=\dfrac{4}{2-\sqrt{2}}$, esto es, su eje mayor está sobre el eje y'.

$$b^2=\frac{4}{2+\sqrt{2}} \text{ y } c^2=a^2-b^2=\frac{4}{(2-\sqrt{2})}-\frac{4}{2+\sqrt{2}}=4\sqrt{2}.$$

Como $P=\begin{bmatrix} \dfrac{1}{\sqrt{4-2\sqrt{2}}} & \dfrac{-1}{\sqrt{4+2\sqrt{2}}} \\ \dfrac{\sqrt{2}-1}{\sqrt{4-2\sqrt{2}}} & \dfrac{1+\sqrt{2}}{\sqrt{4+2\sqrt{2}}} \end{bmatrix}$ tiene $\det P=1,$ entonces:

$$\cos\theta=\frac{1}{\sqrt{4-2\sqrt{2}}}, \text{ y } \operatorname{sen}\theta=\frac{\sqrt{2}-1}{\sqrt{4-2\sqrt{2}}}.$$

Como ambos son positivos, θ está en el primer cuadrante.

Usando la calculadora, determinamos que $\theta\approx 22{,}5^\circ$; por lo tanto, se trata de una elipse centrada en el origen y rotada un ángulo de 22,5º (figura 21.4).

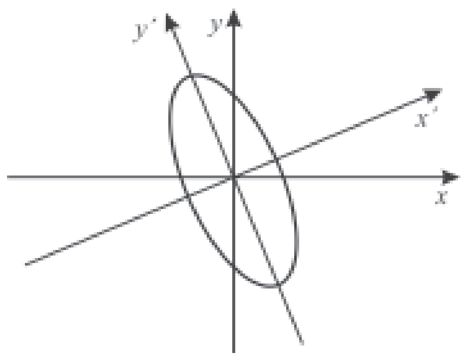


Figura 21.4

Ejemplo 3

Identifique la sección cónica cuya ecuación es $-3x^2-2xy-y^2=4.$

Solución

En el ejemplo 2 vimos que la forma cuadrática $3x^2+2xy+y^2$ en sus ejes principales se puede expresar como $(2+\sqrt{2})x'^2+(2-\sqrt{2})y'^2,$ luego la ecuación dada se puede escribir como:

$$-(2+\sqrt{2})x'^2-(2-\sqrt{2})y'^2=4.$$

Como para cualesquier números reales x' y y', $-(2+\sqrt{2})x'^2-(2-\sqrt{2})y'^2\le 0,$ no existen números reales x y y que satisfagan la ecuación dada, luego la sección cónica definida se llama *sección cónica degenerada*.

Ejemplo 4

Identifique y trace la gráfica de la ecuación $9x^2 + y^2 + 6xy - 10\sqrt{10}x + 10\sqrt{10}y + 90 = 0.$ Escriba la ecuación en forma canónica.

Solución

La forma matricial de la ecuación dada es:

$$[x\,y]\begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + [-10\sqrt{10} \quad 10\sqrt{10}]\begin{bmatrix} x \\ y \end{bmatrix} + 90 = 0.$$

Determinemos los valores característicos de $R = \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}.$

$$\det(R - \lambda I) = \begin{vmatrix} 9-\lambda & 3 \\ 3 & 1-\lambda \end{vmatrix} = \lambda^2 - 10\lambda = 0$$

$$\lambda(\lambda - 10) = 0$$

$$\lambda_1 = 0 \text{ y } \lambda_2 = 10$$

El vector característico asociado con $\lambda_1 = 0$ se obtiene resolviendo:

$$\begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \quad \text{entonces}$$

$$3x = -y, \ \mathbf{v}_1 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

Para $\lambda_2 = 10$ tenemos

$$\begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix};$$

entonces:

$$x = 3y, \ \mathbf{v}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Normalizamos estos vectores para obtener la matriz ortogonal

$$P = \begin{bmatrix} \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ -\frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix}, \ \det P = 1. \text{ Luego } P \text{ es una matriz de rotación.}$$

$$P^T R\ P = D = \begin{bmatrix} 0 & 0 \\ 0 & 10 \end{bmatrix}.$$

La forma cuadrática en los ejes $x'y'$ es $10y'^2$.

Ahora, como $\mathbf{X}' = P^T\mathbf{X},$ entonces $\mathbf{X} = P\mathbf{X}'$, luego la ecuación cuadrática en las variables $x'y'$ es:

$$10y'^2 + [-10\sqrt{10} \quad 10\sqrt{10}]\begin{bmatrix} \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ -\frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} = -90.$$

Realizando los productos se obtiene:

$$10y'^2 - 40x' - 20y' = -90,$$
$$y'^2 - 4x' - 2y' = -9,$$
$$(y'^2 - 2y' + 1) = -9 + 4x' + 1,$$
$$(y' - 1)^2 = 4(x' - 2).$$

Esta ecuación corresponde a una parábola rotada y trasladada.

Sea $y'' = y' - 1$ y $x'' = x' - 2;$ entonces la ecuación es $y''^2 = 4x''.$

El ángulo de rotación lo conocemos con la primera columna de P, así:

$$\cos\theta = \frac{1}{\sqrt{10}}, \quad \operatorname{sen}\theta = -\frac{3}{\sqrt{10}}.$$

Luego θ es un ángulo del cuarto cuadrante: $\theta = 360^\circ - 71.56^\circ = 288.44^\circ$ (figura 21.5).

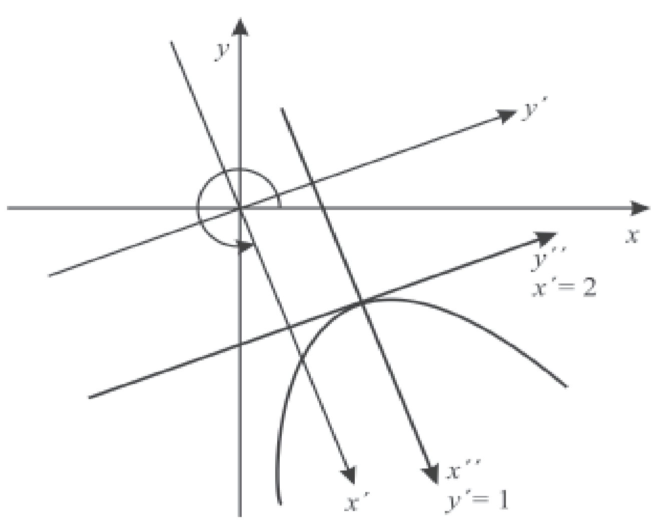


Figura 21.5

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Formas cuadráticas».

Identificamos la gráfica de una ecuación cuadrática dada en x y y mediante la ecuación obtenida después de rotar los ejes. En términos generales, esta ecuación es:

$$\lambda_1 x'^2 + \lambda_2 y'^2 + D'x' + E'y' + F' = 0.$$

Teorema 2

Dada la ecuación de segundo grado $\lambda_1 x'^2 + \lambda_2 y'^2 + D'x' + E'y' + F' = 0,$ entonces:

a. Si λ_1 y λ_2 son ambos positivos, o ambos negativos, la gráfica es una elipse, una circunferencia (si $\lambda_1 = \lambda_2$) o una sección cónica degenerada.

b. Si λ_1 y λ_2 son de signo opuesto, la gráfica es una hipérbola o dos rectas que se cortan.

c. Si $\lambda_1 = 0$ o $\lambda_2 = 0$, la gráfica es una parábola, dos rectas paralelas o una sección cónica degenerada.

La demostración se deja como ejercicio.

21.4 Formas cuadráticas en más de dos variables

Los métodos descritos se pueden usar para analizar las ecuaciones cuadráticas en más de dos variables. Veamos un ejemplo:

Ejemplo 5

Sea la ecuación cuadrática en tres variables dada por $x^2 - 2xy + 2y^2 - 2yz + z^2 = 27.$

Si $R = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$ y $\mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, entonces la ecuación se puede escribir en la forma $\mathbf{X}^T R\ \mathbf{X} = 27.$

Del ejemplo 1 del módulo 20, $P^T R\ P = D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix},$

donde $P = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/6 \end{bmatrix}.$

En consecuencia, $\mathbf{X}^T R\ \mathbf{X} = \mathbf{X}^T P\, D\, P^T \mathbf{X} = (\mathbf{X}')^T D\, \mathbf{X}',$ siendo $\mathbf{X}' = P^T \mathbf{X};$ por lo tanto, la ecuación cuadrática en sus ejes principales se puede escribir como:

$$y'^2 + 3z'^2 = 27,$$

$$\frac{y'^2}{27} + \frac{z'^2}{9} = 1.$$

La sección transversal del gráfico en el plano $x'y'$ obtenida al hacer $z' = 0$ es $y' = \pm\sqrt{27} = \pm 3\sqrt{3},$ es decir, dos rectas paralelas.

La sección transversal en el plano $x'z'$, haciendo $y' = 0$, es $z' = \pm 3$ (dos rectas paralelas).

La sección transversal en el plano $y'z'$, haciendo $x' = 0$, es la elipse $\dfrac{y'^2}{27} + \dfrac{z'^2}{9} = 1.$

El gráfico de la superficie cuadrática en $\mathbb{R}^3$ es un *cilindro elíptico* (figura 21.6).

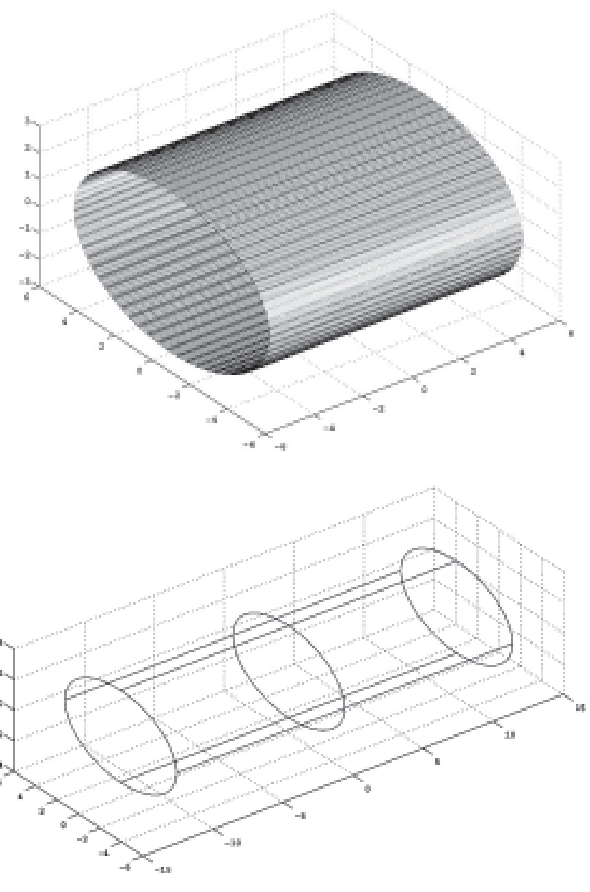


Figura 21.6

Finalmente, extendemos el concepto de forma cuadrática a cualquier número de variables.

Definición 1

Sea $\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ y R una matriz simétrica de $n \times n$.

Una *forma cuadrática* en $x_1, x_2, ..., x_n$ es una expresión de la forma $F(x_1, x_2, ..., x_n) = \mathbf{X}^T R\, \mathbf{X}$.

Ejemplo 6

Sean $A = \begin{bmatrix} 1 & 3 & 5 & 2 & 4 \\ 3 & 2 & 0 & -1 & 3 \\ 5 & 0 & 4 & 0 & 6 \\ 2 & -1 & 0 & 1 & -5 \\ 4 & 3 & 6 & -5 & 2 \end{bmatrix}$ y $\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$; entonces

$$\mathbf{X}^T R\, \mathbf{X} = [x_1\ x_2\ x_3\ x_4\ x_5] \begin{bmatrix} 1 & 3 & 5 & 2 & 4 \\ 3 & 2 & 0 & -1 & 3 \\ 5 & 0 & 4 & 0 & 6 \\ 2 & -1 & 0 & 1 & -5 \\ 4 & 3 & 6 & -5 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}.$$

Luego:

$$F(x_1, x_2, x_3, x_4, x_5) = x_1^2 + 6x_1x_2 + 10x_1x_3 + 4x_1x_4 + 8x_1x_5 + 2x_2^2 - 2x_2x_4$$
$$+6x_2x_5 + 4x_3^2 + 12x_3x_5 + x_4^2 - 10x_4x_5 + 2x_5^2.$$

Ejemplo 7

Sea la forma cuadrática:

$$5x_1^2 + 3x_1x_2 - 5x_1x_3 + 6x_1x_4 + 4x_2^2 + 9x_2x_3 + 7x_3^2 + 8x_3x_4 + 3x_4^2.$$

La matriz R que corresponde a esta forma cuadrática es:

$$R=\begin{bmatrix} 5 & 3/2 & -5/2 & 3 \\ 3/2 & 4 & 9/2 & 0 \\ -5/2 & 9/2 & 7 & 4 \\ 3 & 0 & 4 & 3 \end{bmatrix}.$$

Ejercicios del capítulo 5 (módulo 21)

En los ejercicios 1 a 5 escriba cada forma cuadrática como $\mathbf{X}^T\ R\,\mathbf{X}$, donde R es una matriz simétrica.

1. $2x^2 + 5xy - 9y^2.$
2. $3x^2 - 3xy + 6xz - 9y^2 + 7yz - z^2.$
3. $-4x^2 + 6xy - 3y^2 + 7xz.$
4. $xy = a; \quad a > 0.$
5. $6x^2 + 5xy - 6y^2.$

En los ejercicios 6 a 12 escriba la ecuación cuadrática en la forma $\mathbf{X}^T\ R\,\mathbf{X} = d$ y elimine el término xy rotando los ejes un ángulo θ. Escriba la ecuación en términos de las nuevas variables, identifique la sección cónica obtenida y realice una representación gráfica cuando esto sea posible.

6. $4x^2 + 4xy + y^2 = 9.$
7. $4x^2 + 4xy - y^2 = 9.$
8. $xy = a; \quad a > 0.$
9. $-x^2 + 2xy - y^2 = 0.$
10. $x^2 - 3xy + 4y^2 = 1.$
11. $6x^2 + 5xy - 6y^2 = -7.$
12. $9x^2 + y^2 + 6xy = 4.$

En los ejercicios 13 a 15 identifique la gráfica de la ecuación y escriba ésta en forma canónica.

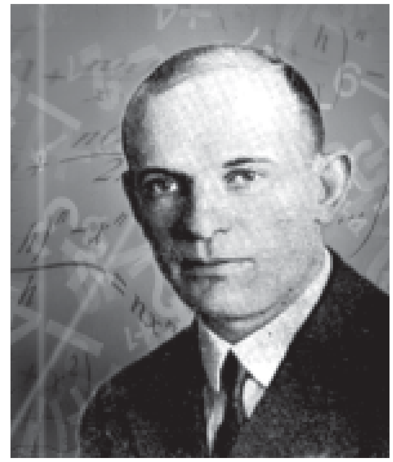
13. $5x^2 + 12xy - 12\sqrt{13}x = 36.$
14. $5x^2 + 5y^2 - 6xy - 30\sqrt{2}x + 18\sqrt{2}y + 82 = 0.$
15. $x^2 - y^2 + 2\sqrt{3}xy + 6x = 0.$

En los ejercicios 16 y 17 escriba la forma cuadrática en términos de las nuevas variables x', y', z' de manera que se eliminen los términos de productos cruzados (xy, xz, yz).

16. $x^2 - 2xy + 2y^2 - 2yz + z^2.$
17. $x^2 - 2xy + y^2 - 2xz - 2yz + z^2.$

Módulo 22

Aproximación de valores y vectores característicos



Al matemático ruso Semyon Gershgorin (1901-1933) se debe el teorema del círculo, publicado en 1931, el cual proporciona una aproximación de los valores característicos cuando no se requiere su cálculo exacto o cuando la ecuación característica es un polinomio difícil de factorizar.

Introducción

En el capítulo anterior desarrollamos un algoritmo para calcular el polinomio característico de A, y a partir de él, extrayendo sus raíces, obtener los valores característicos de A. Esta forma de solución presenta algunos problemas. Uno de ellos tiene que ver con el condicionamiento del polinomio característico, y es que cuando éste está mal condicionado, errores pequeños de redondeo en los coeficientes del polinomio pueden conducir a errores grandes en las raíces. Otro problema es que incluso si los coeficientes del polinomio fueran exactos, es difícil encontrar todas las raíces del polinomio. Estas dificultades han motivado el diseño de métodos numéricos para el cálculo directo de los valores y vectores característicos.

Objetivos

1. Determinar un intervalo de valores donde se puedan acotar los valores característicos de una matriz sin necesidad de hacer muchos cálculos.
2. Proporcionar un método numérico para el cálculo del vector característico dominante de $A_{n\times n}$, sin resolver el polinomio característico.

Preguntas básicas

1. ¿Cómo se construye un círculo de Gershgorin?
2. ¿Qué dice el teorema del círculo de Gershgorin?
3. ¿Cuándo una matriz $A_{n\times n}$ tiene un valor característico dominante?
4. ¿Cómo se define la secuencia de iteraciones para calcular el valor y el vector característico dominante de A?
5. ¿En qué consiste el método de la potencia con normalización?
6. ¿Cómo se calcula el error relativo que se comete al calcular el valor característico dominante?

Vea el módulo 22 del programa de televisión *Álgebra lineal*.

Contenidos

22.1 El teorema del círculo de Gershgorin
22.2 Cálculo numérico de valores y vectores característicos
    22.2.1 El método de la potencia
    22.2.2 El método de la potencia con normalización

Escuche la biografía de *Semyon Gershgorin* en su multimedia de *Álgebra lineal.*

22.1 El teorema del círculo de Gershgorin

El siguiente teorema nos proporciona una aproximación de los valores característicos de A con un mínimo de operaciones algebraicas.

Antes de presentar el teorema, definiremos algunos términos que figuran en su enunciado.

Sea A una matriz $n \times n$ de componentes reales o complejas.

$$\underset{n\times n}{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Se define el número:

$$\gamma_1 = |a_{12}| + |a_{13}| + \ldots + |a_{1n}| = \sum_{j=2}^{n} |a_{1j}|,$$

$$\gamma_2 = |a_{21}| + |a_{23}| + \ldots + |a_{2n}| = \sum_{\substack{j=1 \\ j\neq 2}}^{n} |a_{2j}|,$$

y en general:

$$\gamma_i = |a_{i1}| + |a_{i2}| + \ldots + |a_{i,i-1}| + |a_{i,i+1}| + \ldots + |a_{in}| = \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|.$$

Esto es, γ_i es la suma de los valores absolutos de las componentes del i-ésimo renglón de A, excepto la componente que está sobre la diagonal principal.

Sea $D_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq \gamma_i\}$. D_i es un círculo en el plano complejo con centro en a_{ii} y radio γ_i, con $i = 1,..., n$.

El círculo D_i está compuesto por todos los puntos del plano complejo, sobre y dentro de las circunferencias $C_i = \{z \in \mathbb{C} : |z - a_{ii}| = \gamma_i\}$. A estas circunferencias se les llama *circunferencias de Gershgorin.*

Teorema 1: Teorema del círculo de Gershgorin

Sea A un matriz de $n \times n$ y D_i el conjunto definido anteriormente; entonces, cada valor característico de A está contenido en al menos uno de los D_i. Esto es, si los valores característicos de A son $\lambda_1, \lambda_2, ..., \lambda_k$, entonces:

$$\{\lambda_1, \lambda_2, \dots, \lambda_k\} \subset \bigcup_{i=1}^{n} D_i.$$

Demostración

Sea λ un valor propio de A con $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ un vector característico correspondiente. Sea x_i la componente de $\mathbf{x}$ de mayor valor absoluto, $|x_i| \geq |x_j|,\ j = 1, \dots, n,\ |x_i| > 0,$ ya que $\mathbf{x} \neq \mathbf{0}$.

Ahora, $A\mathbf{x} = \lambda\mathbf{x}$.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_i \\ \vdots \\ \lambda x_n \end{bmatrix}.$$

La componente i de esta ecuación, es:

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ii}x_i + \dots + a_{in}x_n = \lambda x_i.$$

Restando $a_{ii}\, x_i$ a ambos lados de la ecuación, tenemos:

$$\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}x_j = \lambda x_i - a_{ii}x_i = (\lambda - a_{ii})x_i.$$

Tomando valor absoluto a ambos lados, y aplicando la desigualdad de Cauchy Schwarz $\left(|a_{ij}x_j| \leq |a_{ij}||x_j|\right)$ se obtiene:

$$|(\lambda - a_{ii})x_i| = \left|\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}x_j\right| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}||x_j|,$$

y dividiendo a ambos lados por $|x_i|$ obtenemos:

$$|(\lambda - a_{ii})| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|\frac{|x_j|}{|x_i|} \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| = \gamma_i, \text{ ya que } |x_j| \leq |x_i|.$$

Vea en su multimedia de *Álgebra lineal* el código fuente en MATLAB para ilustrar «Circunferencia de Gershgorin».

De la última expresión concluimos que $\lambda \in D_i$, y esto concluye la prueba.

Ejemplo 1

Sea $A = \begin{bmatrix} 1 & 3 & -1 & 4 \\ 2 & 5 & 0 & -7 \\ 3 & -1 & 6 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix}$; entonces, $a_{11} = 1$, $a_{22} = 5$, $a_{33} = 6$, $a_{44} = 4$.

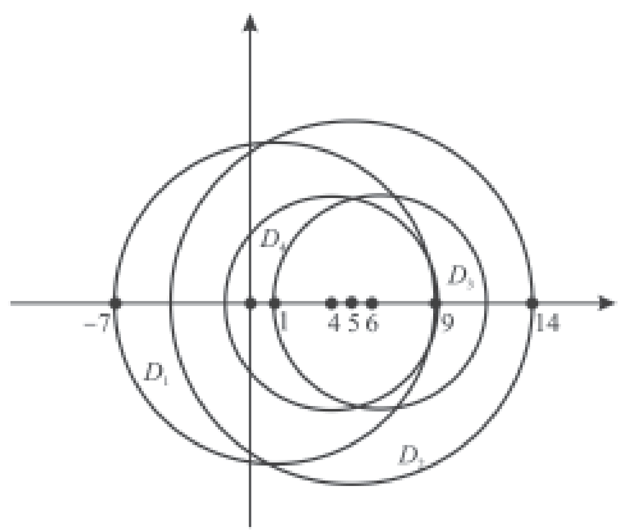
$$
\begin{aligned}
\gamma_1 &= |3| + |-1| + |4| = 8, \\
\gamma_2 &= |2| + |-7| = 9, \\
\gamma_3 &= |3| + |-1| + |1| = 5, \\
\gamma_4 &= |2| + |3| = 5.
\end{aligned}
$$

Los valores característicos de A se encuentran dentro de los círculos descritos por:

$$
D_1 = \{ z \in \mathbb{C} : |z - 1| \leq 8 \}, \quad D_2 = \{ z \in \mathbb{C} : |z - 5| \leq 9 \},
$$

$$
D_3 = \{ z \in \mathbb{C} : |z - 6| \leq 5 \}, \quad D_4 = \{ z \in \mathbb{C} : |z - 4| \leq 5 \}.
$$

Gráficamente podemos ilustrar esta situación dentro del plano complejo, así (figura 22.1):



$C_1 = (x-1)^2 + y^2 = 64$ $\quad$ $C_3 = (x-6)^2 + y^2 = 25$

$C_2 = (x-5)^2 + y^2 = 81$ $\quad$ $C_4 = (x-4)^2 + y^2 = 25$

Figura 22.1

Todos los valores característicos de *A* se encuentran dentro de estas cuatro circunferencias.

Examinando la gráfica anterior, resulta evidente que si λ es un valor característico de *A*, entonces $|\lambda| \leq 14$ y $-7 \leq \mathbb{R}e\lambda \leq 14$ ($\mathbb{R}e\lambda$: parte real de λ).

Ejemplo 2

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

Aplicando el teorema del círculo de Gershgorin podemos demostrar que los valores característicos de *A* son números reales positivos.

Demostración

Como *A* es una matriz simétrica, los valores característicos de *A* son reales (teorema 1, módulo 20).

Acotemos los valores característicos de *A* por medio de los círculos de Gershgorin (figura 22.2).

$$D_1 = \{ z \in \mathbb{C} : |z-2| \leq 13/12 \}.$$
$$D_2 = \{ z \in \mathbb{C} : |z-3| \leq 2 \}.$$
$$D_3 = \{ z \in \mathbb{C} : |z-5| \leq 17/6 \}.$$
$$D_4 = \{ z \in \mathbb{C} : |z-4| \leq 13/4 \}.$$

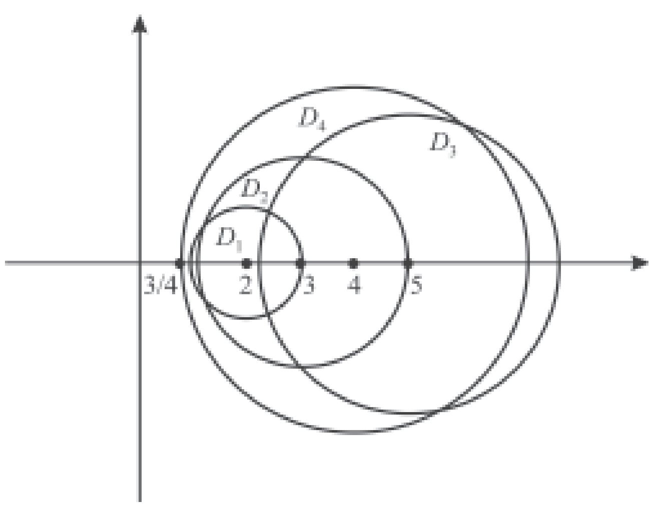


Figura 22.2

El extremo izquierdo de D_4 sobre el eje x es $4 - 13/4 = 3/4$.

Luego $\lambda \geq 3/4$; por lo tanto, los valores característicos de A son reales y positivos.

22.2 Cálculo numérico de valores y vectores característicos

22.2.1 El método de la potencia

Definición 1

Sean $\lambda_1, \lambda_2, ..., \lambda_n$ los valores característicos de una matriz A de $n \times n$ y suponga que $|\lambda_1| > |\lambda_i|$ para $i = 2, ..., n$.

Entonces se dice que λ_1 es el *valor característico dominante* de A. Si $\mathbf{v}_1$ es un vector característico de A correspondiente a λ_1, entonces $\mathbf{v}_1$ se denomina *vector característico dominante.*

Sea A una matriz diagonalizable de $n \times n$, tal que λ_1 es el valor característico dominante y sean $\mathbf{u_1}, \mathbf{u_2}, ..., \mathbf{u_n}$ n vectores característicos linealmente independientes de A, correspondientes a $\lambda_1, \lambda_2, ..., \lambda_n$. Si $\mathbf{x}_0$ es un vector de $\mathbb{R}^n$, entonces $\mathbf{x}_0 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n$ con $c_i \in \mathbb{R}$.

Se puede, sin pérdida de generalidad, hacer $c_1 \neq 0$.

Definimos una secuencia de iteraciones, tales que:

$$\mathbf{x}_{n+1} = A\mathbf{x}_n .$$

Entonces:

$$\begin{aligned}
\mathbf{x}_1 &= A\mathbf{x}_0 \\
\mathbf{x}_2 &= A\mathbf{x}_1 = A^2\mathbf{x}_0 \\
\vdots \quad & \quad \vdots \quad\quad \vdots \\
\mathbf{x}_k &= A\mathbf{x}_{k\text{-}1} = A^k\mathbf{x}_0
\end{aligned}$$

$$\begin{aligned}
\mathbf{x}_k &= A^k(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_n\mathbf{u}_n) \\
&= c_1A^k\mathbf{u}_1 + c_2A^k\mathbf{u}_2 + ... + c_nA^k\mathbf{u}_n \\
&= c_1\lambda_1^{\,k}\mathbf{u}_1 + c_2\lambda_2^{\,k}\mathbf{u}_2 + ... + c_n\lambda_n^{\,k}\mathbf{u}_n \\
&= \lambda_1^{\,k}\left(c_1\mathbf{u}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{u}_2 + ... + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{u}_n\right).
\end{aligned}$$

Como $|\lambda_1| > |\lambda_i|$ para $i = 2, ..., n$, $\left|\dfrac{\lambda_i}{\lambda_1}\right|^k$ se acerca a cero a medida que k aumenta; por lo tanto, $\mathbf{x}_k = A^k \mathbf{x}_0 \approx \lambda_1{}^k c_1 \mathbf{u}_1$.

$\mathbf{x}_k$ es entonces un múltiplo escalar de $\mathbf{u}_1$ y será un vector característico correspondiente a λ_1.

Si $\mathbf{u}_1 = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, entonces $\lambda_1{}^k c_1 \mathbf{u}_1 = \begin{bmatrix} \lambda_1{}^k c_1 a_1 \\ \lambda_1{}^k c_1 a_2 \\ \vdots \\ \lambda_1{}^k c_1 a_n \end{bmatrix}$.

Escogemos dentro de las componentes de $\mathbf{x}_k$, $a_j \neq 0$ y formamos el cociente:

$$\alpha_j^{k+1} = \frac{j\text{-ésima componente de } A^{k+1} x_0}{j\text{-ésima componente de } A^k x_0} = \frac{\lambda_1{}^{k+1} c_1 a_j}{\lambda_1{}^k c_1 a_j} = \lambda_1.$$

De este modo calculamos λ_1; es decir, calculamos el cociente de la j-ésima componente de $\mathbf{x}_{k+1}$ y $\mathbf{x}_k$ y dejamos que k crezca. Una vez determinado λ_1, se tiene $\mathbf{x}_k$ como el vector característico correspondiente a λ_1.

Ejemplo 3

Use el método de la potencia para encontrar el valor característico y el vector característico dominante de:

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

Solución

Sea $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Calculemos:

$$\mathbf{x}_1 = A\mathbf{x}_0 = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \qquad \alpha_1^{(1)} = \frac{1}{1} = 1$$

$$\mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 8 \end{bmatrix} \qquad \alpha_1^{(2)} = \frac{10}{1} = 10 \qquad \alpha_2^{(2)} = \frac{8}{3}$$

$$\mathbf{x}_3 = A\mathbf{x}_2 = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 10 \\ 8 \end{bmatrix} = \begin{bmatrix} 34 \\ 36 \end{bmatrix} \qquad \alpha_1^{(3)} = \frac{34}{10} = 3.4$$

$$\alpha_2^{(3)} = \frac{36}{8} = 4.5$$

Si continuamos de esta manera, obtenemos la tabla 22.1, donde los resultados están redondeados a cuatro cifras decimales.

Tabla 22.1

Iteración	$\mathbf{x}_k$ (como vector renglón)	α_1^k	α_2^k
0	[1, 0]	_	_
1	[1, 3]	1	_
2	[10, 8]	10	2.6667
3	[34, 36]	3.4	4.5
4	[142, 140]	4.17	3.8889
5	[562, 564]	3.9578	4.0286
6	[2.254, 2.252]	4.0107	3.993
7	[9.010, 9.012]	3.9973	4.0018
8	[36.046, 36.044]	4.0007	3.9996

Los valores de $\alpha_1^{(k)}$ y $\alpha_2^{(k)}$ convergen a 4, que es el valor característico de *A* que puede obtenerse por desarrollo del polinomio característico. Ahora el vector $\mathbf{v}_1 = \begin{bmatrix} 36.046 \\ 36.044 \end{bmatrix}$ es aproximadamente igual a un vector característico correspondiente a $\lambda = 4$. Para simplificar este vector, lo dividimos por la componente de mayor magnitud y obtenemos $\mathbf{v}_1' = \begin{bmatrix} 1 \\ 0.9999 \end{bmatrix}$, el cual es muy aproximado a $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ que es el vector característico correspondiente a $\lambda = 4$.

En este caso el método ha funcionado bien, pero vale la pena decir algo acerca de los casos en los cuales el método falla. Puede suceder que:

1. Se aplica el método a una matriz no diagonalizable.
2. *A* no tiene un valor característico dominante o el valor característico dominante λ_1 es apenas un poco mayor que λ_2 en términos absolutos, esto es, $|\lambda_1 / \lambda_2|$ es apenas menor que 1 y las potencias de $|\lambda_1 / \lambda_2|$ no tienden a cero rápidamente.

3. A veces, cuando los elementos de A han tenido que calcularse, hay errores de redondeo que cuando se calcula A^m se convierten en errores considerables.

22.2.2 El método de la potencia con normalización

En el ejemplo anterior podemos observar que el tamaño de las componentes del vector $\mathbf{x}_k$ crece rápidamente; esto lo podemos evitar *normalizando* o *graduando* el vector, o sea, dividiéndolo por la componente de mayor magnitud cada vez que se hace una iteración. Esta modificación sobre el método anterior se llama *método de potencias con normalización.*

Al aplicar este procedimiento se obtiene un vector característico $\mathbf{u}$ cuya mayor componente es 1 y entonces es posible obtener el valor característico dominante resolviendo la ecuación $A\mathbf{u} = \lambda_1\mathbf{u}$ para λ_1.

Ejemplo 4

Determinemos el vector y el valor característico dominante para la matriz del ejemplo anterior, aplicando el método de la potencia con normalización.

Solución

Si $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, entonces:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \text{y} \quad \mathbf{x}_1' = \begin{bmatrix} 1/3 \\ 1 \end{bmatrix}.$$

$$\mathbf{x}_2 = A\mathbf{x}_1' = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1/3 \\ 1 \end{bmatrix} = \begin{bmatrix} 10/3 \\ 4/3 \end{bmatrix}, \qquad \mathbf{x}_2' = \frac{3}{10}\begin{bmatrix} 10/3 \\ 4/3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4/10 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.4 \end{bmatrix}$$

La tabla 22.2 muestra más iteraciones.

El vector característico se puede tomar como $\mathbf{x}_k = [1,\ 0.9998]$, el cual es una buena aproximación del vector [1 1], y el valor característico λ_1 se obtiene de:

$$A\mathbf{u} = A\begin{bmatrix} 0.9991 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.9991 \\ 3.9982 \end{bmatrix} = \lambda_1\begin{bmatrix} 0.9991 \\ 1 \end{bmatrix},$$

de donde $\alpha_1^8 = 4.0027$ y $\alpha_2^8 = 3.9982$ son las aproximaciones para $\lambda_1 = 4$.

Al resolver problemas mediante métodos iterativos, siempre nos hacemos la pregunta: ¿cuándo hay que parar? Una buena forma de resolver esta duda es parar cuando el *error relativo* es suficientemente pequeño.

Tabla 22.2

Iteración	$\mathbf{x}_k$	$\mathbf{x}'_k$ (normalizado)
0	[1, 0]	[1, 0]
1	[1, 3]	[0.3333, 1]
2	[3.3333, 1.3333]	[1, 0.4]
3	[2.2, 2.8]	[0.7857, 1]
4	[3.7857, 3.5714]	[1, 0.9434]
5	[3.8302, 3.8868]	[0.9854, 1]
6	[3.9854, 3.9709]	[1, 0.9964]
7	[3.9892, 3.9928]	[0.9991, 1]
8	[3.9991, 3.9982]	[1, 0.9998]

Si x es la solución exacta y x' la aproximación, el error relativo, E_r, está dado por:

$$E_r = \left|\frac{x' - x}{x}\right|.$$

Ahora bien, no conocemos la solución exacta, pero si el método converge a ella, el valor calculado en la n-ésima iteración está más cerca al valor exacto que el calculado en la iteración $n-1$. Podemos hacer una estimación del error relativo, así:

$$E_r = \left|\frac{x^{(n)} - x^{(n-1)}}{x^{(n)}}\right|.$$

En el ejemplo dado, el valor de λ calculado al hacer la séptima iteración era 3.9892 y al hacer la octava, sería:

$$\frac{3.9991}{0.9991} = 4.0027,$$

$$E_r = \left|\frac{4.0027 - 3.9892}{4.0027}\right| = 0.0033.$$

Esto es, $E_r = 0.33\%$.

Ejercicios del capítulo 5 (módulo 22)

En los ejercicios 1 a 3 dibuje las circunferencias de Gershgorin para la matriz dada A y encuentre una cota para $|\lambda|$ si λ es un valor característico de A.

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

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

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

4. Sea $A = \begin{bmatrix} 2 & \frac{1}{2} & -\frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & 4 & \frac{1}{2} & 1 \\ -\frac{1}{3} & \frac{1}{2} & 3 & 1 \\ \frac{1}{4} & 1 & 1 & 5 \end{bmatrix}$. Demuestre que los valores propios de A son números reales positivos.

5. Sea $A = \begin{bmatrix} -3 & \frac{1}{2} & \frac{1}{4} & 1 \\ \frac{1}{2} & -5 & 1 & 2 \\ \frac{1}{4} & 1 & -4 & 1 \\ 1 & 2 & 1 & -6 \end{bmatrix}$. Demuestre que los valores propios de A son números reales negativos.

6. Se dice que la matriz A de $n \times n$ tiene diagonal estrictamente dominante si $|a_{ii}| > \gamma_i$ para $i = 1, 2, \ldots, n$, donde $\gamma_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$. Demuestre que si $A_{n\times n}$ es una matriz con diagonal estrictamente dominante, entonces A es invertible.

En los problemas 7 a 9 calcule el valor característico dominante y el vector característico correspondiente mediante el método de potencias con normalización.

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

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

9. $\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

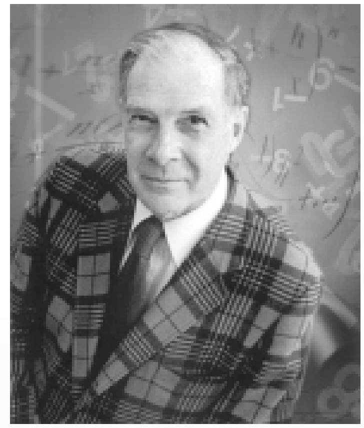
10. Use el método de la potencia para estimar el valor característico dominante de $A = \begin{bmatrix} 1 & 7 \\ 6 & 3 \end{bmatrix}$.
    a. Redondee a cinco cifras significativas y continúe las iteraciones hasta que el error relativo estimado $Er^{(n)} < 0.001$.
    b. Calcule el valor característico dominante en forma exacta. ¿Cuál es el valor exacto de Er?

11. Demuestre que las iteraciones del método de potencias no convergen para la matriz $A = \begin{bmatrix} -3 & 5 \\ -2 & 3 \end{bmatrix}$. Explique por qué.

Apéndice

Una aplicación interesante de los espacios vectoriales

1. Teoría de la codificación



Aplicando conceptos elementales de los espacios vectoriales, Richard Hamming creó un ingenioso método para codificar y decodificar mensajes detectando la ocurrencia de un posible error en su transmisión.

En la sociedad moderna, la comunicación digital, mediante las computadoras o vía satélite, ha penetrado los más variados ámbitos de la actividad humana. La comunicación digital se refiere a transmisión de la información por medio de ceros y unos. Estas cadenas se llaman mensajes binarios y estos mensajes están codificados de modo que se pueda recibir la información tal como fue enviada. Estos mensajes que se reciben de un satélite están sujetos a errores que pueden ser producidos por la estática o cualquier otro tipo de interferencia; por consiguiente, es importante poder codificar un mensaje de manera que, después de haberse sometido a interferencia, pueda decodificarse correctamente.

Se han desarrollado muchas formas de codificar los mensajes. Una de ellas consiste en aumentar un dígito extra, 0 ó 1, dependiendo de que el número de unos del mensaje sea par o impar; esta técnica se llama control de paridad. Con este procedimiento se puede detectar cuándo hay un error en el mensaje, pero no se puede saber dónde; tampoco puede saberse si han ocurrido uno, tres o cinco errores; además, cuando el número de errores es par, estos no se detectan.

Los códigos de corrección de errores generalizan el control de paridad de modo que se puedan ubicar los errores y así poder corregirlos. Richard Hamming introdujo estas teorías a principios de 1950 cuando trabajaba en laboratorios Bell.

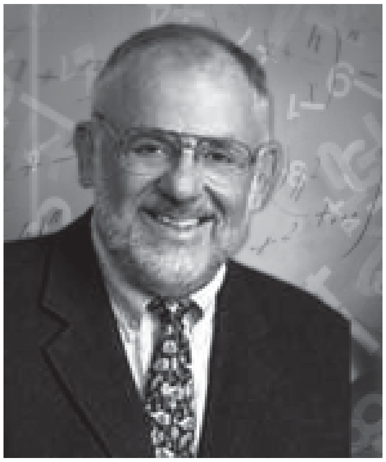
Describiremos un código de Hamming que corrige errores únicos en mensajes formados por cuatro ceros y unos. Pero antes de desarrollar este ejemplo, introduciremos algunos conceptos.

Una *palabra* es una *n*-tupla de ceros y unos, a la que también se le llama *cadena de longitud n.*

Definimos un espacio vectorial denotado por Z_2^n, compuesto por todas las palabras de longitud *n*. La suma y la multiplicación por un escalar se definen en la misma forma que se hace en $\mathbb{R}^n$, sólo que acá los escalares se definen en $Z_2 = \{0, 1\}$, que son los enteros módulo 2.

Las operaciones de suma y multiplicación en Z_2 están dadas por las siguientes tablas:

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad\qquad \begin{array}{c|cc} \times & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$



Con 66 años y múltiples reconocimientos en el mundo de las ciencias y las matemáticas, Cleve Moler es el presidente y cofundador de MathWorks, compañía fundada en 1984 y dedicada a desarrollar y suministrar software para el trabajo científico y de ingeniería.

Antes de unirse de tiempo completo a la compañía en 1989, Moler se dedicó a enseñar matemáticas y ciencias de la computación en universidades de Michigan, Stanford y Nuevo México durante casi veinte años y permaneció cinco años en dos compañías de hardware (la organización Intel Hypercube y Ardent Computer).

Además de ser el autor de la primera versión de MATLAB, Moler es uno de los autores de las bibliotecas científicas de los programas LINPACK y EISPACK y es coautor de tres libros de texto de métodos numéricos.

Su último libro, *El cómputo numérico con MATLAB*, es un texto de introducción a métodos numéricos, MATLAB y técnicas de computación.

Con estas dos operaciones, y tomando los escalares en Z_2, el conjunto Z_2^n satisface todos los axiomas de un espacio vectorial. En conclusión, Z_2^n es un espacio vectorial sobre Z_2.

La base canónica de

$$\mathbb{R}^n: \ \mathbf{e}_1 = (1, 0, 0... 0), \ \mathbf{e}_2 = (0, 1, 0... 0), ..., \ \mathbf{e}_n = (0, 0, ..., 0, 1)$$

también es una base para Z_2^n ; por lo tanto, Z_2^n tiene dimensión *n*.

Todos los conceptos, como subespacios, base, dependencia lineal, conjunto generado **v**, espacio renglón, espacio columna, núcleo, rango y nulidad se aplican a espacios vectoriales sobre Z_2 y a matrices cuyos elementos son de Z_2.

Una base de Z_2^n tiene n vectores, $B=\{\mathbf{v}_1, \mathbf{v}_2 ... \mathbf{v}_n\}$.

Un vector **v** de Z_2^n es una combinación lineal $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n$ con $c_i = 0$ ó 1, de modo que se pueden producir 2^n combinaciones distintas, esto es, el espacio vectorial Z_2^n tiene 2^n elementos.

¿Cómo codificar un mensaje de modo que si ocurre un solo error en la transmisión, éste pueda ser detectado y corregido en la recepción final? Veamos una forma de hacerlo:

Se toman todas las palabras de longitud 4 y se añaden tres controles de paridad. De esta forma se producen palabras de longitud 7, es decir, nos situamos en Z_2^7.

Definimos sobre Z_2^7 un código de corrección de un solo error de Hamming, denotado $C_{7,4}$, así:

Sea H la matriz sobre Z_2 dada por:

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Las columnas de H son la representación binaria de los números del 1 al 8, los cuales son los vectores de Z_2^3 distintos de cero.

Las columnas 1, 2 y 4 de H son linealmente independientes, luego $\rho(H) = 3$, y como $\rho(H) + \nu(h) = 7$, $\nu(h) = 4$.

El espacio nulo de H se llama código $(7,4)$ de Hamming. Una base B para el espacio nulo de H es:

$$\{\mathbf{u}_1 = (1, 0, 0, 0, 0, 1, 1),\ \mathbf{u}_2 = (0, 1, 0, 0, 1, 0, 1), \mathbf{u}_3 = (0, 0, 1, 0, 1, 1, 0), \mathbf{u}_4 = (0, 0, 0, 1, 1, 1, 1)\}.$$

nu H es un subespacio de dimensión 4 de Z_2^7 que tiene $2^4 = 16$ vectores.

Como $H(e_i) = h_i$ (columna i de H, $i = 1, 2, ..., 7$), entonces los vectores de la base estándar Z_2^7 no pertenecen al núcleo de H. Sea $\mathbf{v} \in$ nu H; entonces $(\mathbf{v}+e_i) \notin$ nu H para $i = 1,..., 7$. Si $Hv = h_j$, entonces $(\mathbf{v}+e_j) \in \text{nu}\,H$, además $(\mathbf{v}+e_i) \notin \text{nu}\,H$, para $i \neq j$. Es decir, si se cambia alguna coordenada a un vector de $\text{nu}\,H$, el vector resultante no pertenece a $\text{nu}\,H$. Ahora, si $H\mathbf{v}$ es la j-ésima columna de H, al cambiar la j-ésima componente de **v,** el vector resultante estará en $\text{nu}\,H$.

2. Codificación y decodificación

Veamos cómo codificar un mensaje y decodificar su recepción distorsionada. Suponga que la palabra por codificar corresponde a la cadena binaria 0101 y que hubo una alteración de un dígito.

Para codificar 0101 se expresa la combinación lineal **v** en la base del código de Hamming (7, 4), así:

$$\begin{aligned}\mathbf{v} = 0\mathbf{u}_1 + 1\mathbf{u}_2 + 0\mathbf{u}_3 + 1\mathbf{u}_4 &= \mathbf{u}_2 + \mathbf{u}_4\\ &= (0, 1, 0, 0, 1, 0, 1) + (0, 0, 0, 1, 1, 1, 1)\\ &= (0, 1, 0, 1, 0, 1, 0).\end{aligned}$$

La palabra codificada **v** está en $\text{nu}\,H$.

Suponga que el mensaje recibido es 1101010. Sea $\mathbf{v}' = (1, 1, 0, 1, 0, 1, 0)$. Para decodificar, calculamos $H\mathbf{v}'$:

$$H\mathbf{v}' = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Como $H\mathbf{v}'$ es la primera columna de H, entonces $(\mathbf{v}'+e_1) \in$ nu H y ningún $\mathbf{v}'+e_i$, $i \neq 1$, está en nu H; por lo tanto, $(\mathbf{v}'+e_1) = \mathbf{v}$ es el mensaje corregido que coincide con el originalmente codificado, y así se recupera el mensaje correcto 0101.

Ejemplo

Suponga que $\mathbf{v} = 0111011$ es un mensaje codificado con $C_{7,4}$. Si a lo sumo hay un error en la transmisión, ¿cuál fue el mensaje original?

Solución

Calculemos $H\mathbf{v}$:

$$H\mathbf{v} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

$H\mathbf{v}$ es la cuarta columna de H, entonces $\mathbf{v}+e_4$ será el mensaje corregido y $\mathbf{v}+e_4$ se obtiene cambiando la cuarta componente de $\mathbf{v}$. El vector corregido es entonces 0110011 y el mensaje correcto es 0110.

Ejercicios

1. Haga una lista de los 16 vectores de $C_{7,\,4}$.

2. Sea $\mathbf{v} = (1, 1, 0, 0, 1, 1, 1)$ un mensaje codificado en $C_{7,\,4}$, donde a lo sumo se ha cometido un error. Determine si $\mathbf{v}$ está en $C_{7,\,4}$. Si lo está, decodifíquelo; si no, corríjalo y decodifique el mensaje correcto.

3. Sean $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Efectúe las operaciones indicadas Z_2^3.

   a. $\mathbf{u}+\mathbf{v}$.

   b. $\mathbf{u}-\mathbf{w}$.

   c. $\mathbf{u}+\mathbf{v}-\mathbf{w}$.

   d. Sea $A = [\mathbf{u},\ \mathbf{v},\ \mathbf{w}] = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Calcule A^2 y A^3 sobre $\mathbb{Z}_2$.

   e. ¿Es $\{\mathbf{u},\mathbf{v}\}$ LI sobre $\mathbb{Z}_2$? ¿Y cómo son $\{\mathbf{u},\mathbf{w}\}$ y $\{\mathbf{u},\mathbf{v},\mathbf{w}\}$?

   f. Determine una base y los vectores del espacio nulo de A sobre $\mathbb{Z}_2$. ($A = [\mathbf{u},\mathbf{v},\mathbf{w}]$).

Respuestas

Respuestas de los ejercicios impares

Capítulo 1

Módulo 1

7. Sí.

9. Sí.

11. No, no se cumple el axioma *i*.

13. No, no se cumple el axioma *vi*.

15. No, no se cumple el axioma *vii*.

Módulo 2

1. Sí.

3. Sí.

5. No.

7. Sí.

9. No.

11. Sí.

13. $W_1 \cap W_2 = \left\{ A \in M_{23} : A = \begin{bmatrix} 0 & b & c \\ d & e & 0 \end{bmatrix} \text{ con } b = c \right\}.$

Módulo 3

1. a. Sí, -3, 0. b. No. c. Sí, 0, 0.

3. a. No. b. Sí, 3, -2. c. Sí, 2, -5.

5. $\text{gen}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\} = \{(x, y) \in \mathbb{R}^2 : x = y\}$. Bisectriz del 1.º y 3.er cuadrantes.

7. El plano *xz* en $\mathbb{R}^3 = \left\{\begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3 \; y = 0\right\}$.

9. $\mathbb{P}_1$: conjunto de polinomios de grado menor o igual que 1.

11. Sí.

13. Sí.

Módulo 4

5. LI.

7. LI.

9. LD, $P_3 = -2p_1 + 3p_2$.

11. LI.

19. $x_3 = \begin{pmatrix}13\\-6\\1\end{pmatrix}$.

25. $\alpha = 1$.

27. $\{(1, 2, 4), (-1, 0, 2), (1, 0, 0)\}$.

Módulo 5

5. Sí.

7. Sí.

9. Sí.

11. Sí.

13. $\left\{\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}, \begin{bmatrix}0&1&0\\1&0&0\\0&0&0\end{bmatrix}, \begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix}, \text{etc.}\right\}, 6.$

15. $\left\{\begin{pmatrix}0\\1\\-1\end{pmatrix}, \begin{pmatrix}1\\0\\2\end{pmatrix}\right\}.$

17. $\{-1+x,\ 1+x^2\}.$

19. $\{(1, 0, 1, 0),\ (0, 1, -1, 0),\ (1, 0, 0, 0),\ (0, 0, 0, 1)\}.$

21. El conjunto de todos los vectores de la forma $(a,\ a-b,\ 2a+b,\ -a+3b)$, con a y $b \in \mathbb{R}$.

23. $\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}.$

25. $\left\{\begin{pmatrix}3\\1\\0\end{pmatrix}, \begin{pmatrix}-2\\0\\1\end{pmatrix}\right\}.$

Módulo 6

1. $\{(1, 5, 3, 0, 1),\ (0, 0, 1, 7, 3),\ (0, 0, 0, 0, 1)\}$; $\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}3\\1\\0\end{pmatrix}, \begin{pmatrix}1\\3\\1\end{pmatrix}\right\}$; $\rho = 3.$

3. Todos los renglones, todas las columnas, $\rho = 4$.

5. $\left\{\begin{pmatrix}-2\\1\\0\end{pmatrix}\right\}$, $\{(1, 2, 0),\ (0, 0, 1)\}$, $\left\{\begin{pmatrix}1\\2\\-3\end{pmatrix}, \begin{pmatrix}-3\\5\\0\end{pmatrix}\right\}$, $\rho = 2,\ \nu = 1$.

7. $\left\{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\ 3/2\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\ -5/8\\0\\ -3/4\\1\end{pmatrix}\right\}$, $\{(0,\ 2,\ -3,\ 1,\ 2),\ (0,\ 0,\ 0,\ 4,\ 3)\}$; $\left\{\begin{pmatrix}2\\-2\\4\end{pmatrix}, \begin{pmatrix}1\\3\\6\end{pmatrix}\right\}$, $\rho = 2,\ \nu = 3.$

9. $\left\{\begin{pmatrix}-1\\1\\ -3/2\\1\end{pmatrix}\right\}$, $\{(1,\ 0,\ 0,\ 1),\ (0,\ 1,\ 0,\ -1),\ (0,\ 0,\ 2,\ 3)\}$; $\left\{\begin{pmatrix}-1\\0\\4\\3\end{pmatrix}, \begin{pmatrix}-1\\0\\0\\-1\end{pmatrix}, \begin{pmatrix}0\\2\\-2\\0\end{pmatrix}\right\}$, $\rho = 3,\ \nu = 1.$

11. $\left\{\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\end{pmatrix}\right\}.$

13. No.

15. Sí.

21. a. $n,\ \ m \geq n.$ b. $m,\ \ m \leq n.$ c. $m = n.$

Módulo 7

1. $\begin{bmatrix}3\\1\end{bmatrix}.$ 3. $\begin{bmatrix}5/2\\ 11/2\\ -1/2\end{bmatrix}.$ 5. $\begin{bmatrix}2\\2\\-1\end{bmatrix}.$ 7. $\begin{bmatrix}2\\1\\-4\\3\end{bmatrix}.$ 9. $\begin{bmatrix}3/7\\ -2/7\end{bmatrix}, \begin{bmatrix}0\\ \sqrt{2}\end{bmatrix}, \begin{bmatrix}-2/7\\ 3/7\end{bmatrix}.$

11. $\begin{pmatrix}0\\3\end{pmatrix}.$

13. $-1 + 9x^2 + 7x^3.$

15. a. $(\mathbf{v})_{B_2} = \begin{pmatrix}3\\2\\-7\end{pmatrix}, \quad (\mathbf{w})_{B_2} = \begin{pmatrix}2\\3\\-3\end{pmatrix}.$

b. $\begin{bmatrix} 2 & 1 & 0 \\ 1 & -2/5 & 3/5 \\ 0 & 2/5 & 2/5 \end{bmatrix}$.

c. $(\mathbf{v})_{B_1} = \begin{pmatrix} 8 \\ -2 \\ -2 \end{pmatrix}$, $\quad (\mathbf{w})_{B_1} = \begin{pmatrix} 7 \\ -1 \\ 0 \end{pmatrix}$.

d. Lo mismo que en *c*.

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

f. Lo mismo que en *a*.

17. $\left\{ \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 3 \end{pmatrix} \right\}$.

19. $\vec{i}_r = \begin{pmatrix} \cos\theta \\ \operatorname{sen}\theta \end{pmatrix}$, $\quad \vec{j}_r = \begin{pmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{pmatrix}$.

21. Linealmente dependiente (LD).

23. Linealmente dependiente (LD).

Capítulo 2

Módulo 8

1. a. $\left\{(0, 0, 1), (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)\right\}$. b. $\left\{(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}), (\frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}), (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})\right\}$.

3. $\left\{(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0), (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})\right\}$.

5. $\frac{2}{\sqrt{3}}(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}) + \frac{4}{\sqrt{6}}(-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}) - \frac{2}{\sqrt{2}}(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$.

7. $\dfrac{1}{\sqrt{14}}$.

9. $\left\{\left(\frac{1}{\sqrt{5}}, 0, \frac{2}{\sqrt{5}}\right), (0, 1, 0), \left(-\frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}\right)\right\}$.

11. a. $\frac{1}{49}\begin{pmatrix} -186 \\ 75 \\ 118 \end{pmatrix}$. b. $\frac{1}{7}\begin{pmatrix} 3 \\ -2 \\ 6 \end{pmatrix}$. c. $\mathbf{v} = \frac{1}{49}\begin{pmatrix} -186 \\ 75 \\ 118 \end{pmatrix} + \frac{13}{49}\begin{pmatrix} 3 \\ -2 \\ 6 \end{pmatrix}$.

Módulo 9

1. $y = -0.2 + 1.1x$.

3. $y = \frac{2}{5}x - \frac{4}{5}$.

5. $y \approx 2.61 + 2.08x + 0.31x^2$.

7. i. $\begin{pmatrix} 4/3 \\ -1/3 \end{pmatrix}$. ii. $2\sqrt{5}$.

9. $y = 3.4627 - 3.1314x + 0.5718x^2$.

Módulo 10

3. T_2.

5. $$\left\{\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \ldots \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}\right\}.$$

7. $\left\{\left(\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}\right), \left(\frac{i}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right\}$.

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

Capítulo 3

Módulo 11

1. Sí.
3. Sí.
5. Sí.
7. No.
9. Sí.
11. No.
13. Sí.
15. No.
17. Sí.
19. No.
21. $T(x\,y) = (-x, y)$.
25. $(T_1 + T_2)(\mathbf{x}) = (A+B)\mathbf{x}, \quad (T_1 \text{ o } T_2)(\mathbf{x}) = AB\mathbf{x}$.
27. $(T_1 \text{ o } T_2)(x, y) = (x+4y,\ 3x-y)$.
$(T_2 \text{ o } T_1)(x, y, z) = (2x-y+3z,\ -3x-y-7z,\ -y-z)$.

Módulo 12

1. $\text{nu}\,T = \{0\}$, imagen $T = \mathbb{R}^2$; $\rho_{(t)} = 2$; $v_t = 0$.
3. $\text{nu}\,T = \{0\}$, imagen $T = \mathbb{R}^3$; $\rho_{(t)} = 3$; $v_t = 0$.
5. $\text{nu}\,T = \{0\}$, imagen $T = \text{gen}\{x, x^2, x^3\}$; $\rho_{(t)} = 3$; $v_t = 0$.
7. $\text{nu}\,T = \{f \in C[0, 1] : f(1) = 0\}$, imagen $T = \mathbb{R}$; $\rho_{(t)} = 1$; el núcleo es un espacio de dimensión infinita, luego $v_T = \infty$.
9. $\text{nu}\,T =$ imagen $T = \{(x, y) : x = y\}$.

11. El núcleo de T es el espacio solución del sistema homogéneo de ecuaciones lineales $A\mathbf{x} = \mathbf{0}$.

13. $T(\mathbf{x}) = A(\mathbf{x})$ con $A = \begin{bmatrix} 2 & -1 & 1 \\ 2 & -1 & 1 \\ 2 & -1 & 1 \end{bmatrix}$.

Módulo 13

1. a. $\begin{bmatrix} 1 & -2 \\ 1 & 2 \end{bmatrix}$. b. $\begin{bmatrix} 1 & -2 \\ 2 & 0 \end{bmatrix}$. c. $\begin{bmatrix} 3 & -2 \\ -1 & 2 \end{bmatrix}$. d. $\begin{bmatrix} 3 & -2 \\ 2 & 0 \end{bmatrix}$. e. $(4, 0)$.

3. a. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$. b. $\begin{bmatrix} -1 & -1/3 & 0 \\ 1 & 2/3 & 0 \end{bmatrix}$. c. $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$.

5. a. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$. b. $\begin{bmatrix} 0 & 1 \\ 0 & -1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$.

7. a. $\begin{bmatrix} 10 \\ 5 \\ 5 \end{bmatrix}$. b. $\begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}$.

9. a. $\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -3 \\ 4 \end{bmatrix}$. b. $\begin{bmatrix} 1 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ -10 \end{bmatrix}$. c. $\begin{bmatrix} -2 \\ 25 \end{bmatrix}$.

11. En 1. $\operatorname{nu} T = \left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right\}$; imagen $T = \mathbb{R}^2$.

En 3. $\operatorname{nu} T = \operatorname{gen}\left\{\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}\right\}$; imagen $T = \mathbb{R}^2$.

En 5. $\operatorname{nu} T = \{0\}$; imagen $T = \operatorname{gen}\{x^2, x^3\}$.

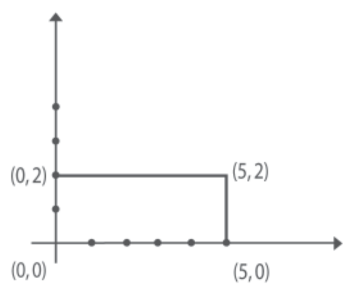
En 7. $\text{nu}\, T = \text{gen}\left\{\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}\right\}$; imagen $T = \text{gen}\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}\right\}$.

15. a. $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ cy \end{pmatrix}$, $c > 0$.

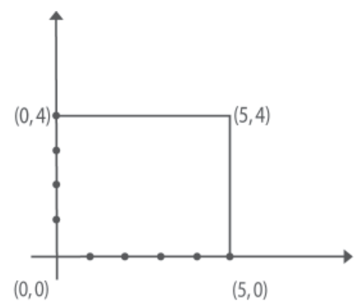
Expansión: $c > 1$.
Contracción: $0 < c < 1$.

b. Matriz de la transformación: $\begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}$.

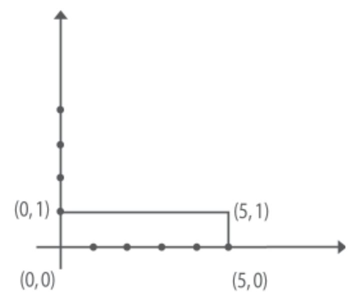
c.



Rectángulo inicial



Expansión a lo largo del eje y con $c = 2$

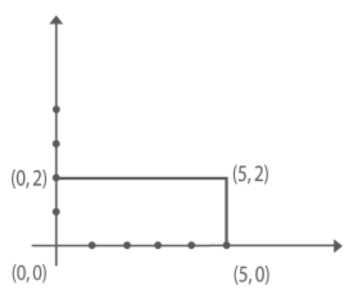


Contracción a lo largo del eje y con $c = 1/2$

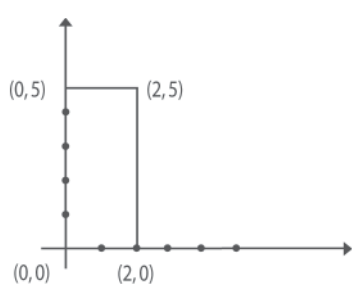
17. a. $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$.

b. Matriz de la transformación: $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

c.



Rectángulo inicial



Reflexión con respecto a la recta $y = x$

Módulo 14

1. Solución posible: $T:\mathbb{R}^2\to\mathbb{R}^3,\ T(x,y)=(x-y,\ 2x+y,\ y).$

$T:\mathbb{R}^3\to\mathbb{R}^2,\ T(x,y,z)=(x+y-z,\ y+z).$

$T:\mathbb{R}^2\to\mathbb{R}^2,\ T(x,y)=\ (2x+y,\ x-y).$

3. $T:D_n\to\mathbb{R}^n,\ T\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & a_{nn} \end{bmatrix}=\begin{bmatrix} a_{11} \\ a_{22} \\ \vdots \\ a_{nn} \end{bmatrix}.$

5. a. Sí. b. No. c. No. d. Sí.

7. No.

11. $T^{-1}(A)\ =B^{-1}A.$

Módulo 15

3. $\left(\dfrac{1}{\sqrt{2}},\ \dfrac{1}{\sqrt{2}}\right).$

5. $T:\mathbb{R}^2\to\mathbb{R}^2$ tal que $T\begin{pmatrix} x \\ y \end{pmatrix}=\alpha\begin{pmatrix} x \\ y \end{pmatrix},\ \ \alpha\neq 1.$

Capítulo 4

Módulo 16

1. a. No. b. Sí, $\lambda=4$. c. Sí, $\lambda=0$. d. Sí, $\lambda=4$. e. Sí, $\lambda=0$. f. No.

3. $\lambda=1,\quad \mathbf{v}=\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$

5. $\lambda_1=1$ con $\mathbf{v}_1=\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ \ \lambda_2=-1$ con $\mathbf{v}_2=\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix},\ \ \lambda_3=2$ con $\mathbf{v}_3=\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$

7. $\lambda_1 = 1$ con $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $\lambda_2 = -1$ con $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$, $\lambda_3 = 3$ con $\mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix}$, $\lambda_4 = 2$ con $\mathbf{v}_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$.

9. $\lambda = a$ mult. alg 4, $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{v}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$.

11. $\lambda = a$ mult. alg 4, $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$.

13. $P(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_n)$, valores característicos: $\lambda_1, \lambda_2, ..., \lambda_n$.

15. No. Para que $\lambda_1 + \lambda_2$ sea un valor característico de $A + B$, es necesario que el vector característico de A correspondiente a λ_1 sea el mismo que el de B correspondiente a λ_2, lo que, en general, no se cumple.

Módulo 17

1. $D = \begin{pmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 2 \end{pmatrix}$, $P = \begin{pmatrix} -i & i & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $P^{-1} = \begin{pmatrix} -i/2 & 1/2 & 0 \\ i/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

3. $D = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{pmatrix}$, $P^{-1} = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 2 & -1 & 0 & 0 \\ 0 & 0 & 1/2 & -1/2 \end{pmatrix}$.

5. $D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$, $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, $p^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

7. $A = PDP^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$.

$$A^{12} = PD^{12}P^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & (1/2)^{12} \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 - 1/2^{11} \\ 0 & 1/2^{12} \end{pmatrix}.$$

9. D_1D_2 no es necesariamente una diagonalización de AB, $AB = P^{-1}D_1PQ^{-1}D_2Q$. Esto se daría en el caso en que $P = Q$ y así $AB = P^{-1}D_1PP^{-1}D_2P$; entonces, $AB = PD_1D_2P^{-1}$.

Si A y B son matrices diagonalizables, $A+B$ no necesariamente es diagonalizable; por ejemplo: $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$ y $\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ son diagonalizables, pero su suma $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ no es diagonalizable.

Módulo 18

1. a. $P_n = 20(2^n) + 30$ cualquiera sea $n \in \mathbb{N}$.

b. El número de peces después de cinco años es $P_5 = 20(2^5) + 30 = 640 + 30 = 670$.

c. $n \approx 7.11$ años.

3. a. Dentro: $\left(\dfrac{2z_0 - y_0}{3}\right)\left(\dfrac{7}{10}\right)^k + \left(\dfrac{y_0 + z_0}{3}\right)$.

Fuera: $\left(\dfrac{y_0 - 2z_0}{3}\right)\left(\dfrac{7}{10}\right)^k + 2\left(\dfrac{y_0 + z_0}{3}\right)$.

b. Sí. Matriz de transición: $\dfrac{1}{10}\begin{bmatrix} 8 & 1 \\ 2 & 9 \end{bmatrix}$.

c. Probabilidad de vivir dentro: $\dfrac{(2z_0 - y_0)}{3(y_0 + z_0)}\left(\dfrac{7}{10}\right)^k + \dfrac{1}{3}$.

Probabilidad de vivir fuera: $\dfrac{(y_0 - 2z_0)}{3(y_0 + z_0)}\left(\dfrac{7}{10}\right)^k + \dfrac{2}{3}$.

d. Dentro: $\dfrac{1}{3}(y_0 + z_0)$; fuera: $\dfrac{2}{3}(y_0 + z_0)$.

5. a. A largo plazo la distribución de alumnos es $\begin{bmatrix} 0 \\ 0 \\ x_0 + y_0 \end{bmatrix}$, el cual es un vector propio correspondiente a $\lambda = 1$. El vector de probabilidad correpondiente será $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

b. Sí. Matriz de transición: $\dfrac{1}{12}\begin{bmatrix} 7 & 2 & 0 \\ 2 & 6 & 0 \\ 3 & 4 & 12 \end{bmatrix}$.

7. a. $P_n = 1.000\sqrt{5}\left[\left(\dfrac{5+\sqrt{5}}{10}\right)^n - \left(\dfrac{5-\sqrt{5}}{10}\right)^n\right]; \quad n \in \mathbb{N}.$

b. No.

c. $P_6 = 1.000\sqrt{5}\left[\left(\dfrac{5+\sqrt{5}}{10}\right)^6 - \left(\dfrac{5-\sqrt{5}}{10}\right)^6\right].$

9. El brazo trabaja exitosamente 97.97% de las veces, muy cerca al 98% exigido por el cliente, así que es muy probable que se acepte.

Módulo 19

1. No.

3. Sí.

5. No.

7. Sí.

9. No.

11. $P = \begin{bmatrix} 1 & 0 \\ -1 & 1/7 \end{bmatrix}.$

13. $P = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \end{bmatrix}, \ J = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}.$

15. $\begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix}, \quad \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{bmatrix}, \quad \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{bmatrix}, \quad \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{bmatrix},$

$$\begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix}.$$

Los λ_i no necesariamente son distintos; además, los bloques de Jordan pueden permutarse sobre la diagonal.

17. $\begin{bmatrix} -3 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \begin{bmatrix} -3 & 1 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \begin{bmatrix} -3 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix},$

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

19. $\begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \begin{bmatrix} 7 & 1 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \begin{bmatrix} 7 & 1 & 0 \\ 0 & 7 & 1 \\ 0 & 0 & 7 \end{bmatrix}.$

Capítulo 5

Módulo 20

1. $\begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}; \quad P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$

3. $\begin{bmatrix} -3 & 0 & 0 \\ 0 & 1+2\sqrt{2} & 0 \\ 0 & 0 & 1-2\sqrt{2} \end{bmatrix}; \quad P = \begin{bmatrix} 1/\sqrt{2} & 1/2 & 1/2 \\ -1/\sqrt{2} & 1/2 & 1/2 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}.$

5. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{bmatrix}; \quad P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$

7. $\begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad P = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix}.$

Módulo 21

1. $[x \ y]\begin{bmatrix} 2 & 5/2 \\ 5/2 & -9 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$

3. $[x\ y\ z]\begin{bmatrix} -4 & 3 & 7/2 \\ 3 & -3 & 0 \\ 7/2 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}.$

5. $[x\ y]\begin{bmatrix} 6 & 5/2 \\ 5/2 & -6 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$

7. $[x\ y]\begin{bmatrix} 4 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 9;\qquad \dfrac{x'^2}{\left(\dfrac{18}{\sqrt{41}+3}\right)} - \dfrac{y'^2}{\left(\dfrac{18}{\sqrt{41}-3}\right)} = 1;$ hipérbola; $\theta \approx 19.33^{\text{o}}$.

9. $[x\ y]\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 0;\qquad y'^2 = 0;$ recta que pasa por el origen; $\theta = \pi/4 = 45^{\text{o}}$.

11. $[x\ y]\begin{bmatrix} 6 & 5/2 \\ 5/2 & -6 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = -7;\qquad \dfrac{y'^2}{14/13} - \dfrac{x'^2}{14/13} = 1;$ hipérbola; $\theta \approx 11.31^{\text{o}}$.

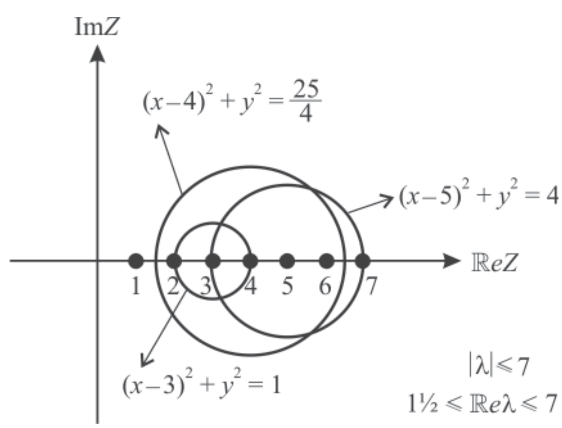
13. Hipérbola; $\dfrac{x''^2}{4} - \dfrac{y''^2}{9} = 1.$

15. Hipérbola; $\dfrac{x''^2}{\frac{9}{8}} - \dfrac{y''^2}{\frac{9}{8}} = 1.$

17. $(x\ y\ z)\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix},\quad -x'^2 + 2y'^2 + 2z'^2.$

Módulo 22

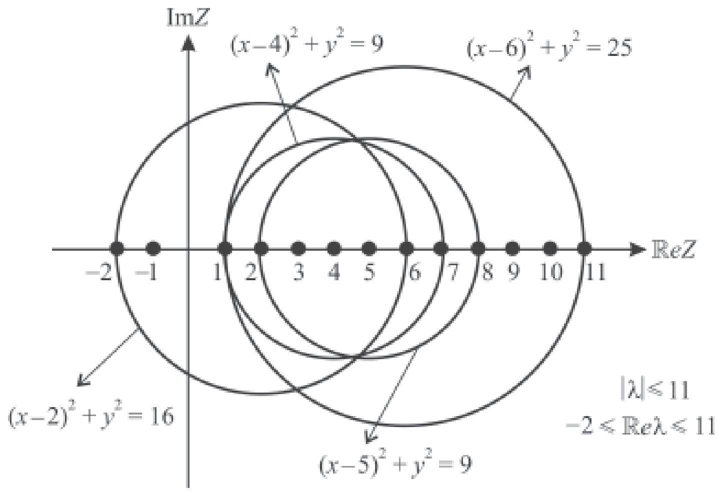
1.



$|\lambda| \leq 7$

$1\frac{1}{2} \leq \mathbb{R} \in \lambda \leq 7$

3.



$$|\lambda| \le 11$$
$$-2 \le \mathbb{R} \in \lambda \le 11$$

5. Como A es una matriz simétrica, sus valores característicos son números reales. Al trazar las circunferencias de Gershgorin éstas quedan completamente contenidas a la izquierda del eje y, con lo cual se establece que los valores característicos son negativos.

7. $-4, \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$

9. $8, \begin{pmatrix} 1 \\ 0.5 \\ 1 \end{pmatrix}.$

11. Comenzando con $x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ y sin normalización, se obtiene $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \ldots$; por lo tanto, el método no converge. La razón de esto es que A no posee un valor característico dominante.

Bibliografía

Florey, F. (1980), *Fundamentos de álgebra lineal y aplicaciones*, México, Prentice Hall.

Grossman, S. *Álgebra lineal*, Bogotá, McGraw-Hill.

Hill, R. (1997), *Álgebra lineal elemental con aplicaciones*, México, Prentice Hall Hispanoamericana.

Kolman, B. (1999), *Álgebra lineal con aplicaciones y MATLAB*, México, Prentice Hall.

Kolman, B. y D. Hill (2006), *Álgebra lineal*, México, Prentice Hall.

Lay, D. C. (1999), *Álgebra lineal y sus aplicaciones*, México, Addison Wesley Longman.

Nakos, G. y D. Joyner (1999), *Álgebra lineal con aplicaciones*, México, International Thomson Editors.

Perry, W. (1980), *Álgebra lineal con aplicaciones*, México, McGraw-Hill.

Pita Ruiz, C. (1991), *Álgebra lineal*, México, McGraw-Hill.

Restrepo, P. *et al*. (1997), *Álgebra lineal con aplicaciones*, Medellín, Universidad Nacional de Colombia, sede Medellín.

Strang, G. (1982), *Álgebra lineal y sus aplicaciones*, México, Fondo Educativo Interamericano.

Villamayor, O. (1982), *Álgebra lineal*, Washington, Secretaría General de la OEA.

Ude@

Educación no presencial