Notas de álgebra lineal

Alfredo Gómez Rodríguez
Instituto de Física y Facultad de Ingeniería, U.N.A.M

octubre de 2019

Índice general

Introducción xi

I Operaciones y estructuras algebraicas. 1

1. Operaciones binarias **3**
1.0.1. un ejemplo interesante 6
1.0.2. otro ejemplo 7
1.0.3. Operaciones parciales 7

2. Semigrupos y monoides **9**

3. Grupos **11**
3.1. ejemplos de grupos 11
3.2. propiedades elementales de los grupos 12

4. Anillos **15**
4.1. Ejemplos de anillos 16
4.1.1. divisores de cero 16

5. Campos **17**
5.1. Ejemplos de campos 18

II Espacios vectoriales. 19

6. Concepto de espacio vectorial **21**
6.1. Algunos ejemplos de espacios vectoriales 23
6.1.1. Los espacios $\mathbb{R}^n$ 23
6.1.2. Los espacios $\mathbb{C}^n$ 26
6.1.3. Los espacios P_n 27
6.1.4. Los espacios $M(m,n)$ 28
6.1.5. Los espacios $M^*(m,n)$ 29
6.2. Algunas propiedades básicas de los espacios vectoriales. 30
6.3. Subespacios 33

6.4. Sistemas de ecuaciones lineales homogéneas. 35
6.5. combinaciones lineales 36
6.5.1. Des-sumando y des-multiplicando 38
6.6. generadores 38
6.6.1. cómo encontrar un generador para el espacio solución de un sistema lineal homogéneo. 40
6.7. Independencia lineal y dependencia lineal 41
6.7.1. ¿Cómo determinar la dependencia o independencia? . . . 43
6.8. bases y dimensión 46
6.8.1. Algunos resultados adicionales 53
6.9. coordenadas 53
6.10. cambio de base y matriz de transición 56
6.10.1. Encontrando la matriz de transición. 57
6.11. Isomorfismos y el principio del isomorfismo 60
6.11.1. Principio del isomorfismo 62
6.11.2. cambio de base de nuevo 62
6.12. Espacios renglón y columna 63
6.13. El espacio nulo de una matriz 68
6.14. La estructura de los sistemas inhomogéneos. 68
6.15. Espacios de funciones, Wronskianos 69
6.15.1. dependencia e independencia lineal de funciones. 70

III Transformaciones lineales 75

7. Transformaciones lineales 79
7.0.2. algunas propiedades de las transformaciones lineales. . . . 80
7.1. algunas transformaciones importantes 81
7.1.1. ejemplos de transformaciones lineales. 81
7.2. núcleo e imagen 84
7.2.1. imagen y suprayectividad. 87
7.2.2. Núcleo e inyectividad. 87
7.3. teorema de la dimensión. 88
7.4. representaciones matriciales. 88
7.4.1. la matriz de transición de nuevo 89
7.4.2. calculando la matriz asociada a una transformación lineal. 90
7.4.3. Dada la matriz calcular la regla de correspondencia 92

8. Álgebra de transformaciones 93
8.1. Suma de transformaciones lineales. 93
8.2. Producto por números. 94
8.3. La suma de dos transformaciones lineales es lineal 94
8.4. Linealidad del producto por un número 95
8.5. El conjunto de todas las transformaciones lineales de V en W es un espacio vectorial. 95

8.5.1. ¿hay un zero (elemento idéntico) para la suma de transformaciones? . . . 95
8.5.2. ¿hay un negativo (inverso aditivo) para la suma de transformaciones ? . . . 96
8.6. Matriz asociada a la suma y producto por números. . . . 96
8.7. Composición de transformaciones. . . . 98
8.7.1. definición de composición. . . . 98
8.7.2. La composición de dos transformaciones lineales es lineal. 98
8.7.3. Matriz de la composición de dos transformaciones lineales. 99
8.7.4. Las propiedades más importantes de estas operaciones . . 99

9. La transformación inversa. 103
9.1. La inversa de una transformación. . . . 103
9.1.1. la inversa es única . . . 103
9.1.2. la inversa de una transformación lineal es lineal . . . 104
9.1.3. La matriz de la inversa. . . . 105
9.1.4. Criterios de invertibilidad. . . . 105
9.1.5. La inversa de una composición. . . . 106
9.1.6. un ejemplo completo . . . 107
9.2. Cambio de base y matrices . . . 110
9.3. Matrices "vacías" . . . 111

10. Geometría de las transformaciones (I). 113
10.1. Haciendo geometría en $\mathbb{R}^2$ . . . 113
10.2. Efecto de transformaciones lineales sobre rectas y segmentos . . 113
10.3. Areas de paralelogramos y de imágenes de paralelogramos bajo transformaciones lineales. . . . 114
10.4. Quiralidad. . . . 115
10.4.1. ejemplos . . . 115

11. eigenvalores y eigenvectores 117
11.1. algunas definiciones . . . 117
11.1.1. en términos de las matrices. . . . 118
11.2. Hallando los eigenvalores . . . 118
11.3. hallando los eigenvectores. . . . 119
11.3.1. primeros ejemplos . . . 119
11.4. Un ejemplo sencillo. . . . 120
11.5. Para obtener eigenvalores y eigenvectores. . . . 122
11.6. Matriz de un operador en una eigenbase. . . . 122
11.6.1. independencia de eigenvectores. . . . 122
11.6.2. Matriz diagonalizadora . . . 123
11.7. Un ejemplo completo . . . 124

12. Interpretación geométrica de las transformaciones lineales (II). 131

IV Espacios con producto interior 133

13. Espacios con producto interior 135
13.1. Motivación . . . 135
13.2. Concepto de producto interior . . . 135
13.3. Ejemplos de productos interiores . . . 136
13.4. Algunas propiedades de los productos interiores . . . 140
13.5. La desigualdad de Cauchy-Schwarz . . . 140
13.6. Normas . . . 142
13.6.1. motivación. . . . 142
13.6.2. Caso general. . . . 142
13.7. Distancias (métricas) . . . 146
13.8. Ángulos. . . . 148
13.9. Conjuntos ortogonales, normales y ortonormales . . . 148
13.9.1. Los conjuntos ortogonales son independientes . . . 149
13.9.2. expansión con respecto a una base ortonormal . . . 149
13.10Matriz con respecto a una base ortonormal. . . . 150
13.11Productos y cordenadas . . . 150
13.12complementos ortogonales . . . 151
13.13Proyecciones . . . 152
13.13.1Motivación . . . 152
13.13.2Caso general . . . 153
13.14Mejor aproximación. . . . 155
13.15Método de Gram-Schmidt . . . 156
13.15.1Formulación del problema. . . . 156
13.15.2Idea básica . . . 156
13.15.3Formulación explícita . . . 157

14. Cuadrados mínimos. 159
14.0.4. motivación . . . 159
14.0.5. Teoría . . . 159
14.1. Aplicaciones . . . 161
14.1.1. usando un poco de cálculo . . . 163
14.1.2. otros comentarios . . . 163
14.1.3. un caso general y típico . . . 164
14.1.4. otro enfoque . . . 165

V Operadores lineales en espacios con producto interior. 167

15. productos interiores y funciones lineales 169

16. Operadores adjuntos **171**
16.1. Matriz del operador adjunto con respecto a una base ortonormal. 172
16.2. Propiedades del adjunto 173
16.3. regla de correspondencia del adjunto 174
16.3.1. ¿y si la base no era ortonormal? 176
16.4. Un ejemplo 176

17. Operadores y matrices Hermiteanos, antihermiteanos, simétricos, antisimétricos, unitarios, ortogonales y normales. **179**
17.1. Algunas definiciones básicas 179
17.2. Algunos teoremas 180
17.3. Teorema espectral. 184
17.4. resolución espectral 186
17.4.1. productos exteriores 186
17.4.2. resolución espectral del operador y de la identidad 187
17.4.3. Proyectores de nuevo 191

18. Cuádricas. **193**
18.1. Introducción a las ideas principales 193
18.1.1. propiedades de A 194
18.2. diagonalización de formas cuadráticas. 194
18.3. cónicas 195
18.4. superficies cuádricas. 196

A. Dependencia e independencia lineales. **197**

B. Isomorfismos **199**

C. Comentarios sobre las referencias **201**

Prefacio

Estas son las notas para un curso de álgebra lineal. Estas notas están en un proceso de continua revisión, para versiones más recientes puede Usted consultar http://www.paginaspersonales.unam.mx/archivos/index/alias:alfredogomez

Introducción

El plan 2015 para la asignatura “álgebra lineal” impartida en la División de Ciencias Básicas de la Facultad de Ingeniería de la UNAM nos ha obligado a revisar nuestras notas de clase. El resultado de dicha revisión es el presente juego de notas, que no pretenden ser sino eso, notas. El lector indudablemente notará algunos de los inumerables errores en este texto, y los errores no son nada más de ortografía o redacción. Sin embargo, esta obra es, como toda obra humana, perfectible y jamás perfecta. Agradeceré si me informan de los errores bajo el principio de “error detectado, error eliminado.”

Con la revisión 2015 de los planes de estudio, se vuelve a incluir .^{es}tructuras algebraicas.^{en} el curso. En 2016 desaparecieron los "Wronskianos"que, sin embargo, he dejado en estas notas.

Hace ya algunos años los físicos prometieron darnos “computadoras cuánticas”. Al momento de escribir estas notas aun no cumplen su promesa y, por desgracia, en otras ocasiones (los cuasicristales, los superconductores) no las han cumplido. Sin embargo de lograrse la computación cuántica nuestro modesto curso se convertiría en uno de los más importantes; sugiero al lector que le eche un ojo a un texto de información cuántica, verá en él álgebra lineal y más álgebra lineal.

Más recientemente, los campos de .^{a}rtificial inteligence"(inteligencia artificial), "machine learning"(aprendizaje por máquinas) y "deep learning"(aprendizaje profundo) hacen uso abundante de la estadística y del álgebra lineal. No dudo que algunos de mis estudiantes lleguen a trabajar en "big data"(grandes datos) como "data scientists"(científicos de datos).

Parte I

Operaciones y estructuras algebraicas.

Capítulo 1

Operaciones binarias

Definición 1 *Sea S un conjunto. Una operación binaria en S es una función*

$$F : S \times S \to S$$

Si $\otimes$ es una operación binaria, frecuentemente escribimos

$$\otimes(a, b) = a \otimes b$$

y decimos que en $\otimes(a, b)$ hemos usado la notación de prefijo y en $a \otimes b$ la de infijo.

Definición 2 *A una función*

$$G : S \to S$$

se le llama operación unaria y, en general.

$$H : S \times S \times ...S \to S$$

donde S aparece n veces en el dominio es una operación $n-$aria.

Definición 3 *Sea $B \subset S$ y $\otimes$ una operación binaria en S. Se dice que B es cerrado bajo la operación $\otimes$ si*

$$x, y \in B \Rightarrow x \otimes y \in B$$

Definición 4 *Sea $\otimes$ una operación binaria en S. Se dice que S es cerrado bajo la operación $\otimes$ si*

$$x, y \in S \Rightarrow x \otimes y \in S$$

Esta propiedad recibe también el nombre de cerradura, se dice que la operación es cerrada o satisface la cerradura. Esto es redundante pues la definición misma de operación binaria implica que $x \otimes y \in S$ para todo $x, y \in S$.

Definición 5 *Se dice que $\otimes$ es conmutativa si*

$$a \otimes b = b \otimes a$$

para cualesquiera a,$b \in S$.

Definición 6 *Se dice que $\otimes$ es asociativa si para cualesquiera a,b,$c \in S$ se cumple que*

$$a \otimes (b \otimes c) = (a \otimes b) \otimes c$$

Definición 7 *Se dice que l es una identidad izquierda (o elementio idéntico izquierdo) para $\otimes$ si*

$$l \otimes a = a$$

para toda $a \in S$.

Definición 8 *Se dice que r es una identidad derecha (o elemento idéntico derecho) para $\otimes$ si*

$$a \otimes r = a$$

para toda $a \in S$.

Definición 9 *Se dice que e es una identidad (o elemento idéntico) para $\otimes$ si*

$$a \otimes e = e \otimes a = a$$

para toda $a \in S$.

Comentario 10 *e es una identidad para $\otimes$ si y sólo si e es una identidad derecha y es una identidad izquierda.*

Ejemplo 11 *La suma y la diferencia son operaciones en los enteros, en los racionales y en los reales. La suma es conmutativa, la diferencia no lo es. La suma es asociativa en tanto que la diferencia no. La división no es una operación binaria pues no se permite división por cero.*

Comentario 12 *A las operaciones se les llama, a veces, leyes de composición internas. Esto se hace para contrastarlas con funciones del tipo $B \times S \to S$ en donde se usan elementos de un conjunto "externo" B, estas funciones se llaman leyes de composición externa.*

Más ejemplos de operaciones serían:

- La suma de números naturales, enteros, racionales, reales o complejos.
- El producto de números naturales, enteros, racionales, reales o complejos.
- La suma de polinomios.
- El producto de polinomios.

- La suma de matrices del mismo tamaño.
- En el conjunto de las matrices de $n \times n$ el producto es una operación, note que en el conjunto de todas las matrices el producto no es una operación pues no todo par de matrices puede ser multiplicado (las matrices deben de ser conformables, es decir, el primer factor debe de tener tantas columnas como renglones tenga el segundo factor).
- La unión e intersección de conjuntos son operaciones entre conjuntos.

Una manera de definir la operación es dando, explícitamente, el conjunto S y la regla de correspondencia de la operación $\otimes$. Otra manera, útil cuando S es un conjunto finito, es dando la "tabla" de la operación, análoga a las tablas de multiplicar.

Ejemplo 13 *Sea $S = \{e, c, c^2, c^3, m_1, m_2, m_3, m_4\}$ con la tabla de multiplicar*

$*$	e	c	c^2	c^3	m_1	m_2	m_3	m_4
e	e	c	c^2	c^3	m_1	m_2	m_3	m_4
c	c	c^2	c^3	e	m_2	m_3	m_4	m_1
c^2	c^2	c^3	e	c	m_3	m_4	m_1	m_2
c^3	c^3	e	c	c^2	m_4	m_1	m_2	m_3
m_1	m_1	m_4	m_3	m_2	e	c^3	c^2	c
m_2	m_2	m_1	m_4	m_3	c	e	c^3	c^2
m_3	m_3	m_2	m_1	m_4	c^2	c	e	c^3
m_4	m_4	m_3	m_2	m_1	c^3	c^2	c	e

lo que quiere decir, a guisa de ejemplo, que $m_3 * c^2 = m_1$ etc. Dejo al estudiante convencerse de que $*$ es asociativa y no es conmutativa.

Definición 14 *Sean $\otimes$ y $\circledast$ dos operaciones binarias en S. Se dice que $\otimes$ es distributiva sobre $\circledast$ si, para cualesquiera a,b,c $\in S$ tenemos que*

$$a \otimes (b \circledast c) = (a \otimes b) \circledast (a \otimes c)$$

Teorema 15 *Sea $\otimes$ una operación binaria en S. Si l es una identidad izquierda y r es una identidad derecha, entonces $l = r$. En este caso $e = l = r$ es una identidad.*

Demostración. $l \otimes r = r$ pues l es una identidad izquierda, pero $l \otimes r = l$ pues r es una identidad derecha; por lo tanto $l = r$ es una identidad. ∎

Teorema 16 *La identidad respecto a una operación binaria es única.*

Demostración. sean e y e' dos identidades. Entonces $e \otimes e' = e$ por ser e' una identidad pero $e \otimes e' = e'$ por ser e una identidad. Luego $e = e'$. ∎

Definición 17 *Sea S un conjunto con una operación binaria $\otimes$ y suponga que e es una identidad. Se dice que L es una inversa izquierda de $a \in S$ si*

$$L \otimes a = e$$

Se dice que R es una inversa derecha de $a \in S$ si

$$a \otimes R = e$$

y que a' es una inversa (bilátera) de $a \in S$ si

$$a' \otimes a = a \otimes a' = e$$

Teorema 18 *Si la operación $\otimes$ es asociativa y tiene un elemento idéntico, la inversa de un elemento $a \in S$ es única.*

Demostración. Sean b y c inversas de a. Entonces $c \otimes a = a \otimes c = e$ y $b \otimes a = a \otimes b = e$ (donde e es el elemento idéntico). Por ello $b = b \otimes e = b \otimes (a \otimes c) = (b \otimes a) \otimes c = e \otimes c = c$ ■

Teorema 19 *Sea $\otimes$ una operación binaria asociativa en S,si $a \in S$ tiene una inversa izquierda L y una inversa derecha R, entonces $L = R$ y a tiene una inversa.*

Demostración. Como $L \otimes a = e$ y $a \otimes R = e$ tenemos que $L \otimes (a \otimes R) = L \otimes e = L$ pero, por la asociatividad $L \otimes (a \otimes R) = (L \otimes a) \otimes R = e \otimes R = R$ de modo que $L = R$. ■

Definición 20 *Se dice que un elemento $a \in S$ es cancelable con respecto a una operación binaria $\otimes$ si para cualesquiera x,$y \in S$*

$$(a \otimes x = a \otimes y) \ \vee (x \otimes a = y \otimes a) \Rightarrow x = y$$

Comentario 21 *Ciertamente si $\otimes$ es asociativa y a tiene inverso, a es cancelable.*

Dejo al lector verificar que para la operación cuya tabla hemos dado más arriba, hay un elemento idéntico bajo $*$ y todo elemento de S tiene un inverso bajo $*$. El elemento idéntico es e. El inverso de c es c^3, el de m_1 es m_1 etc.

1.0.1. un ejemplo interesante

Sea S un conjunto y defina una operación binaria en S mediante

$$a * b = b$$

para cualesquiera $a, b \in S$. Entonces:

- cualquier elemento de S es un idéntico izquierdo.

- si z fuera un idéntico derecho entonces para todo $a \in S$ tendríamos que $a * z = a = z$ de modo que $z = a$ para todo a, por ello, no hay idénticos derechos (a menos, claro, que el conjunto S tenga un sólo elemento).
- dejo al lector ver que la operación no es conmutativa pero que es asociativa.
- además $a * b = a * c \Rightarrow b = c$ (se vale una de las posibles leyes de cancelación). Sin embargo $b * a = c * a \Rightarrow a = a$ y no se aplica la otra cancelación.

1.0.2. otro ejemplo

Sea X un conjunto y

$$\mathcal{F} = \{F \mid F : X \to X\}$$

el conjunto de todas las funciones de X en X. Entonces la composición de funciones $\circ$ es una operación binaria en $\mathcal{F}$, pues la composicion de dos funciones en $\mathcal{F}$ siempre puede hacerse y da como resultado otra función en $\mathcal{F}$.

La función identidad en X

$$I_X : X \to X$$

definida mediante

$$I_X(y) = y$$

$\forall y \in X$ es la identidad en X pues

$$I_X \circ F = F \circ I_X = F$$

$\forall F \in \mathcal{F}$.

Una función $G \in \mathcal{F}$ tendrá inversa izquierda si y sólo si G es inyectiva.

Una función $G \in \mathcal{F}$ tendrá inversa derecha si y sólo si G es suprayectiva y una función $G \in \mathcal{F}$ tendrá inversa (bilátera) si y sólo si G es biyectiva.

1.0.3. Operaciones parciales

A veces es necesario relajar la condición de que la operación se pueda aplicar a todos los elementos de $S \times S$ y basta con estipular que se puede aplicar a algunos de sus elementos. Un buen ejemplo sería el conjunto de todas las matrices con la operación de multiplicación. No todas las parejas de matrices se pueden multiplicar, algunas sí (se dice entonces que los factores son conformables). Por ello la multiplicación matricial no es una operación binaria en el conjunto de todas las matrices, pero se dice que es una operación parcial (o parcialmente definida).

En términos más formales, una operación parcialmente definida en S es una función de la forma

$$F : B \times B \to S$$

donde $B \subset S$.

Cuando se tiene una operación parcial, se puede hablar también de identidades e inversas derechas e izquierdas. Sin embargo puede verse facilmente que ya no es cierto el teorema que dice que si hay identidad derecha e izquierda entonces éstas son iguales. Por ejemplo, las matrices de 3×2 tienen como identidad izquierda a la matriz identidad de 3×3 pero como identidad derecha a la matriz identidad de 2×2

Capítulo 2

Semigrupos y monoides

Definición 22 *Se llama estructura algebraica, o sistema algebraico, a un conjunto S con una o más operaciones (unarias, binarias etc.).*

Definición 23 *Un grupoide (o según, algunos autores, magma) es un conjunto S con una operación binaria.*

Definición 24 *Se llama semigrupo al sistema algebraico $\{S, \otimes\}$ consistente de un conjunto S y una operación binaria $\otimes$ asociativa.*

Definición 25 *Se llama monoide al sistema algebraico $\{S, \otimes, e\}$ consistente de un conjunto S y una operación binaria $\otimes$ asociativa con elemento identidad e.*

Comentario 26 *Algunos autores llaman semigrupo a lo que nosotros hemos llamado monoide. No hay uniformidad en la literatura por lo que el lector deberá estár pendiente de posibles usos alternativos.*

Ejemplo 27 *El ejemplo que sigue es probablemente el más paradigmático. Considere un conjunto X y el conjunto $\hom(X, X) = \{F : X \to X\}$ de todas las funciones de X a X (este conjunto también se denota como X^X). La operación a considerar en $\hom(X, X)$ es la composición de funciones y el elemento idéntico es la función identidad $I_X : X \to X$ cuya regla de correspondencia es $I_X(y) = y$ para cualquier $y \in X$. Dejamos al lector verificar que $\{\hom(X, X), \circ, I_X\}$ es un monoide.*

Ejemplo 28 *Hay muchos ejemplos de monoides. Los enteros (y los números naturales) son un monoide respecto a la adición, siendo el número cero el elemento idéntico. También son un monoide con respecto a la multiplicación pero el elemento idéntico es el número uno. El lector debería encontrar por si mismo muchos otros ejemplos en los racionales, reales, complejos, polinomios, matrices etc.*

Capítulo 3

Grupos

Un grupo G es un conjunto (también llamado G) con una operación binaria $*$ y que satisface las siguientes propiedades:

- Para todo $a, b \in G \quad a * b \in G$ (cerradura, axioma realmente redundante puesto que hemos supuesto que $*$ es una operación binaria en G)

- Para todo $a, b, c \in G$ tenemos que $a * (b * c) = (a * b) * c$ (propiedad asociativa)

- Existe un elemento, llamado $e \in G$ y tal que para todo $a \in G$ tenemos que $a * e = e * a = a$ (idéntico bajo $*$)

- Para todo elemento $a \in G$ existe otro, llamado inverso y denotado mediante a^{-1} y tal que $a * a^{-1} = a^{-1} * a = e$

- Si además para todo $a, b \in G$ tenemos que $a * b = b * a$ (propiedad conmutativa) decimos que G es un grupo conmutativo o abeliano.

Comentario 29 *un grupo es un monoide en el que todo elemento tiene inverso.*

3.1. ejemplos de grupos

- El conjunto S con la operación $*$ cuya tabla de multiplicar hemos dado arriba, es un grupo. Es el llamado grupo dihedral D_4.

- Los números enteros, los números racionales, los números reales y los números complejos forman un grupo bajo adición. Los números naturales no forman grupo bajo adición porque no contienen a los negativos.

- Los números racionales diferentes de cero, los números reales diferentes de cero y los números complejos diferentes de cero forman grupo con respecto a la multiplicación.

- Los números complejos de módulo igual a uno forman grupo bajo multiplicación.
- Los números reales de valor absoluto uno ($= \{1, -1\}$) forman grupo bajo multiplicación.
- Las matrices reales o complejas de $m \times n$ forman un grupo bajo adición.
- Las matrices reales o complejas de $n \times n$ invertibles forman un grupo bajo multiplicación (llamado grupo lineal general).
- Las matrices reales ortogonales de $n \times n$ forman un grupo bajo multiplicación.
- Las matrices complejas unitarias de $n \times n$ forman un grupo bajo multiplicación.
- Las matrices reales o complejas de $n \times n$ invertibles y triangulares superiores forman un grupo bajo multiplicación.
- Las matrices reales o complejas de $n \times n$ invertibles y determinante igual a 1 forman un grupo bajo multiplicación (grupo especial lineal).

3.2. propiedades elementales de los grupos

Teorema 30 *en un grupo el elemento idéntico es único:*

Demostración. sea z otro idéntico de modo que para todo $x \in G$

$$ex = xe = x$$
$$zx = xz = x$$

entonces

$$z = ze = e$$

■

Teorema 31 *en un grupo el elemento inverso de $x \in G$ es único.*

Demostración. sean x^{-1} y z dos inversos de x de modo que

$$x^{-1}x = xx^{-1} = e$$
$$zx = xz = e$$

entonces

$$z = ez = (x^{-1}x)z = x^{-1}(xz) = x^{-1}e = x^{-1}$$

■

Teorema 32 *Si en un grupo G $xa = xb$ entonces $a = b$ (es decir, en un grupo todo elemento es cancelable).*

Demostración. : $xa = xb \Rightarrow x^{-1}(xa) = x^{-1}(xb) \Rightarrow (x^{-1}x)a = (x^{-1}x)b \Rightarrow ea = eb \Rightarrow a = b$. ■

Teorema 33 *: Si en un grupo G $ax = bx$ entonces $a = b$.*

Demostración. : $ax = bx \Rightarrow (ax)x^{-1} = (bx)x^{-1} \Rightarrow a(xx^{-1}) = b(xx^{-1}) \Rightarrow ae = be \Rightarrow a = b$. ■

Los dos teoremas anteriores reciben el nombre de leyes de cancelación.

Teorema 34 *: si $a, b \in G$ entonces $(ab)^{-1} = b^{-1}a^{-1}$*

Demostración. Por la unicidad de inversas basta con probar que $b^{-1}a^{-1}$ es una inversa de ab.

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})(a^{-1}) = aea^{-1} = aa^{-1} = e$$
$$(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}(e)b = b^{-1}b = e$$

■

Teorema 35 *: $(a^{-1})^{-1} = a$*

Dejamos la demostración al lector.

Definición 36 *: Sea G un grupo y H un subconjunto de G. Decimos que H es un subgrupo de G si $e \in H$, para todo $x, y \in H$ $xy \in H$ (cerradura de H bajo multiplicación) y para todo $x \in H$ $x^{-1} \in H$.*

Teorema 37 *: Sea G un grupo y H un subgrupo, entonces H es un grupo con respecto a la misma multiplicación de G.*

Demostración. : Como H es cerrado bajo la multiplicación $*$ cumple el primer axioma de los grupos. Si $a, b, c \in H$ entonces $a, b, c \in G$ y $a * (b * c) = (a * b) * c$, cumpliendo el segundo axioma (propiedad asociativa). Como $e \in H$, si $a \in H$ entonces $a \in G$ y $ea = ae = a$ de modo que el idéntico bajo $*$ en H es el mismo que el elemento idéntico de G. Por último, si $a \in H$ entonces $a \in G$ y existe $a^{-1} \in G$ tal que $a * a^{-1} = a^{-1} * a = e$, pero $a^{-1} \in H$ por la definición de subgrupo. ■

Note, pues, que un subgrupo es un grupo que es subconjunto de otro grupo y tiene la misma operación.

Como ejemplo, el grupo de las matrices complejas invertibles de $n \times n$ tiene como subgrupo el de las matrices complejas unitarias de $n \times n$.

Comentario 38 *Es fácil demostrar que en la tabla de multiplicaión de un grupo cada elemento aparece una y sólo una vez en cada renglón y cada columna.*

Capítulo 4

Anillos

Un anillo es un conjunto A que tiene definidas dos operaciones binarias $+$ y $\times$ tales que:

- Para todo $a, b \in A \quad a + b \in A$ (cerradura, axioma realmente redundante puesto que hemos supuesto que $+$ es una operación binaria en A)

- Para todo $a, b \in A$ tenemos que $a + b = b + a$ (propiedad conmutativa)

- Para todo $a, b, c \in A$ tenemos que $a + (b + c) = (a + b) + c$ (propiedad asociativa)

- Existe un elemento, $0 \in A$ (llamado cero) y tal que para todo $a \in A$ tenemos que $a + 0 = 0 + a = a$ (idéntico bajo $+$)

- Para todo elemento $a \in A$ existe otro, llamado negativo de a (o inverso aditivo) y denotado mediante $-a$ y tal que $a + (-a) = (-a) + a = 0$

- Para todo $a, b \in A \quad a \times b \in A$ (cerradura, axioma realmente redundante puesto que hemos supuesto que $*$ es una operación binaria en A)

- Para todo $a, b, c \in A$ tenemos que $a \times (b \times c) = (a \times b) \times c$ (propiedad asociativa)

- Para cualesquiera a,b y $c \in A$ $a \times (b + c) = a \times b + a \times c$ (propiedad distributiva)

- Para cualesquiera a,b y $c \in A$ $(a + b) \times c = a \times c + b \times c$ (propiedad distributiva)

Si adicionalmente

- Para todo $a, b \in A$ tenemos que $a \times b = b \times a$ (propiedad conmutativa) decimos que A es un anillo conmutativo.

- Si existe $1 \in A$ (llamado unidad) tal que para todo $a \in A$ $1 \times a = a \times 1 = a$ decimos que A es un anillo con unidad.
- En un anillo conmutativo y con unidad A, si todo elemento $a \in A$ y diferente de cero es cancelable,entonces se dice que A es un dominio entero o de integridad.

Note que los anillos son grupos abelianos con respecto a la suma.

4.1. Ejemplos de anillos

- Los enteros son un anillo conmutativo y con unidad con respecto a la suma y producto de enteros. Son, de hecho, un dominio entero.
- Los polinomios con coeficientes reales o complejos forman un anillo con respecto a la suma y producto de polinomios (un dominio entero).
- Los números reales y los complejos forman un anillo (un dominio entero).
- Las matrices reales o complejas de $n \times n$ forman un anillo. Son un anillo con unidad I_n (la matriz identidad de $n \times n$). No forman un dominio entero, esto se puede ver con un contraejemplo, las matrices

$$C = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$
$$D = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$$

  tienen la propiedad de que

$$CD = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$$

  y sin embargo $C \neq 0$ y $D \neq 0$, pero $0 = 0D$ de modo que $CD = 0D$ y si D fuera cancelable tendríamos que $C = 0$ contradiciendo lo que sabemos de C. Entonces en el conjunto de las matrices de 2×2 hay elementos no cero que no son cancelables, no es un dominio entero.

4.1.1. divisores de cero

En un anillo se dice que a y b son divisores de cero si $a \neq 0$, $b \neq 0$ y $ab = 0$.

Teorema 39 *Sea A un anillo conmutativo y con unidad. Entonces A es un dominio entero si y sólo si no hay divisores de cero.*

Dejo la prueba al lector. Note que el último ejemplo hace, de hecho, uso de esta propiedad. Ahí se vio que $C \neq 0$ y $D \neq 0$, pero $0 = 0D$.

Capítulo 5

Campos

Un campo es un conjunto F que tiene definidas dos operaciones binarias $+$ y $\times$ tales que:

- Para todo $a, b \in F \quad a + b \in F$ (cerradura, axioma realmente redundante puesto que hemos supuesto que $+$ es una operación binaria en A)
- Para todo $a, b \in F$ tenemos que $a + b = b + a$ (propiedad conmutativa)
- Para todo $a, b, c \in F$ tenemos que $a + (b + c) = (a + b) + c$ (propiedad asociativa)
- Existe un elemento, llamado $0 \in F$ y tal que para todo $a \in F$ tenemos que $a + 0 = 0 + a = a$ (idéntico bajo $+$)
- Para todo elemento $a \in F$ existe otro, llamado negativo de a (o inverso aditivo) y denotado mediante $-a$ y tal que $a + (-a) = (-a) + a = 0$
- Para todo $a, b \in F \quad a \times b \in F$ (cerradura, axioma realmente redundante puesto que hemos supuesto que $*$ es una operación binaria en F)
- Para todo $a, b, c \in F$ tenemos que $a \times (b \times c) = (a \times b) \times c$ (propiedad asociativa)
- Para cualesquiera a,b y $c \in F$ $a \times (b + c) = a \times b + a \times c$ (propiedad distributiva)
- Para cualesquiera a,b y $c \in F$ $(a + b) \times c = a \times c + b \times c$ (propiedad distributiva)
- Para todo $a, b \in F$ tenemos que $a \times b = b \times a$ (propiedad conmutativa)
- Existe $1 \in A$ tal que para todo $a \in A$ $1 \times a = a \times 1 = a$
- Para todo $a \in F$ tal que $a \neq 0$ existe $a^{-1} \in F$ tal que $a \times a^{-1} = a^{-1} \times a = 1$

Un campo es, pues, un anillo conmutativo con identidad en el que todo elemento no cero tiene un inverso multiplicativo (esto implica que todo elemento no cero es cancelable: $ac = bc$ con $c \neq 0$ implica que $acc^{-1} = bcc^{-1}$ o sea $a = b$ y el campo es también un dominio entero).

Comentario 40 *Algunos autores (sobre todo los españoles) llaman "cuerpo" a lo que hemos llamado campo.*

5.1. Ejemplos de campos

- El campo de los números racionales
- El campo de los números reales
- El campo de los números complejos
- Los enteros no son un campo pues (excepto ± 1) los enteros no tienen inverso multiplicativo.

Parte II

Espacios vectoriales.

Capítulo 6

Concepto de espacio vectorial

Un espacio vectorial sobre el campo F es un conjunto V cuyos elementos llamamos vectores y con dos operaciones, la primera es una suma

$$+ : V \times V \to V$$

y la segunda es un producto de vectores por escalares (números del campo F)

$$\cdot : F{\times}V \to V$$

que satisfacen las siguientes propiedades.

- 1)Si $a, b \in V$ entonces $a + b \in V$ (cerradura)
- 2)Si $a, b \in V$ entonces $a + b = b + a$ (propiedad conmutativa)
- 3)Si $a, b, c \in V$ entonces $(a + b) + c = a + (b + c)$ (propiedad asociativa)
- 4) Hay un elemento $0 \in V$ con la propiedad de que si $a \in V$ entonces $0 + a = a + 0 = a$.
- 5) Para todo elemento $a \in V$ hay otro elemento $(-a) \in V$ (llamado negativo de a) y que tiene la propiedad de que $a + (-a) = (-a) + a = 0$.
- 6) Si $\alpha \in F$ y $a \in V$ entonces $\alpha \cdot a \in V$
- 7) Si $a \in V$ entonces $1 \cdot a = a$
- 8)Si $\alpha, \beta \in F$ y $a \in V$ entonces $(\alpha + \beta) \cdot a = \alpha \cdot a + \beta \cdot a$
- 9) Si $\alpha \in F$ y $a, b \in V$ entonces $\alpha \cdot (a + b) = \alpha \cdot a + \alpha \cdot b$
- 10)Si $\alpha, \beta \in F$ y $a \in V$ entonces $(\alpha\beta) \cdot a = \alpha(\beta \cdot a) = \beta(\alpha a)$

Notación 41 *ordinariamente usaremos la notación* v *(y el contexto indica si* v *es vector o escalar) para indicar un vector en un espacio vectorial. Otras posibilidades son* $\underline{v}$, $\overrightarrow{v}$, $\widehat{v}$ *etc. En mecánica cuántica se usa la llamada "notación de Dirac" en donde los vectores se denotan como* $\mid v\rangle$ *(y a esto se le llama "ket", esto es una broma pues en ingles el símbolo* $\langle\rangle$ *es un "bracket" por lo que su mitad derecha* $|\rangle$ *es un "ket"). Una notación frecuente para los vectores es del tipo* $\bar{v}$ *pero note usted que la barra también se usa en matemáticas para la conjugación de números complejos (si* $z = x + iy$ *entonces* $\bar{z} = x - iy$*, a veces para la conjugación se usa el símbolo* $*$ *y* $z^* = x - iy$*).*

Comentario 42 *La operación* $+$ *es, por lo tanto, una ley de composición interna en tanto que la operación* $\cdot$ *es una ley de composición externa.*

Comentario 43 *Note usted que el símbolo* $+$ *es ambiguo. A veces es la suma de números, a veces es la suma de vectores. El contexto debe remover las ambiguedades. Esto es similar a lo que ocurre en programación orientada a objetos cuando tenemos una "sobrecarga" (overloading). Si por cualquier razón necesitamos distinguir notacionalmente las operaciones podemos usar* $+$ *y* $\oplus$ *.*

Comentario 44 *Algunos autores llaman a* $\cdot$*" una acción del campo* F *sobre los vectores".*

Cuando $F = \mathbb{R}$ decimos que el espacio vectorial es real, cuando $F = \mathbb{C}$ decimos que se trata de un espacio vectorial complejo. A veces a F se le llama "campo base (ground field)".

Los axiomas 1 y 6 son redundantes dado que se ha supuesto que la suma de vectores es una operación binaria y que el producto de vectores por números es una operacion externa. De hecho el axioma 2 también sale sobrando en el sentido de que un sistema que satisfaga los restantes axiomas también debe necesariamente satisfacer el axioma 2 (¿le gustaría intentar la demostración de esto?).

El ejemplo más sencillo de espacio vectorial es el espacio de todas las "flechas" (técnicamente llamadas "vectores geométricos" o "segmentos dirigidos") en dos o tres dimensiones, esto es, el espacio de los vectores que ya conocemos, entendidos como objetos con magnitud, dirección y sentido. Como ejemplos de flechas podemos citar las fuerzas, el vector de posicion de una partícula, la velocidad, la aceleración, el campo eléctrico y el campo magnético. Lo que tienen en común es que pueden ser representados pictóricamente mediante flechas. Pero lo que hace que estas flechas sean vectores es el hecho de que podemos combinar dos flechas v y w para formar una tercera flecha $v + w$ (mediante la "regla del paralelogramo) y que podemos combinar un número real α y una flecha v para formar una nueva flecha (aquella cuyo tamaño es $\mid \alpha \mid$ veces el de v y tiene el mismo sentido que v si $\alpha > 0$ y el sentido opuesto si $\alpha < 0$). Dejo al lector el checar que el conjunto de las flechas con esta adición y multiplicación por números es un espacio vectorial real.

Ejercicio 45 *Muestre que el conjunto A de todas las flechas en el espacio bidimensional ordinario, con las operaciones descritas anteriormente, forma un espacio vectorial real.*

Comentario 46 *Suele llamarse "espacio" a un conjunto con algo de estructura. Así hablamos de espacios vectoriales, espacios normados, espacios métricos y espacios topológicos.*

Comentario 47 *A veces en vez de "espacio vectorial" se usa la expresión "espacio lineal".*

Comentario 48 *Algunos autores usan en vez de un campo F un anillo R. En este caso (con los mismos axiomas) decimos que se trata de un "módulo sobre el anillo R" en vez de un "espacio vectorial sobre el campo F". Nosotros no trataremos los módulos en lo absoluto.*

Comentario 49 *En general usaremos el mismo símbolo (por ejemplo V) para referirnos al conjunto (de vectores) o a la estructura (espacio vectorial). Si fuera necesario distinguirlos podríamos usar $V = \{\mid V \mid, +, \cdot\}$ para denotar el espacio vectorial V construido con el conjunto $\mid V \mid$. Pero en este curso no tendremos que preocuparnos por esto.*

6.1. Algunos ejemplos de espacios vectoriales

Además del espacio vectorial real de las flechas (vectores geométricos), tenemos algunos ejemplos importantes de espacios vectoriales.

6.1.1. Los espacios $\mathbb{R}^n$

Sea n un entero positivo y

$$\mathbb{R}^n = \{(x_1, x_2, ...x_n) \mid x_i \in \mathbb{R} \ \forall i = 1, 2...n\}$$

el conjunto de todas las eneadas ordenadas (es decir, no es lo mismo (a, b) que (b, a)) de números reales. Aquí se entiende que dos elementos x y y de $\mathbb{R}^n$ dados por

$$x = (x_1, x_2, ...x_n)$$
$$y = (y_1, y_2, ...y_n)$$

son iguales si y sólo si $\forall i = 1, 2...n$

$$x_i = y_i$$

En $\mathbb{R}^n$ definimos la suma de $x = (x_1, x_2, ...x_n) \in \mathbb{R}^n$ y $y = (y_1, y_2, ...y_n) \in \mathbb{R}^n$ como aquel elemento $x + y \in \mathbb{R}^n$ dado por

$$x + y = (x_1 + y_1, x_2 + y_2, ...x_n + y_n)$$

y definimos el producto de $\alpha \in \mathbb{R}$ y $x = (x_1, x_2, ...x_n) \in \mathbb{R}^n$ como el elemento $\alpha x \in \mathbb{R}^n$ dado por

$$\alpha x = (\alpha x_1, \alpha x_2, ...\alpha x_n)$$

Puede demostrarse que $\mathbb{R}^n$ con la suma y el producto por escalares recien definidos es un espacio vectorial real, procediendo de la manera siguiente:

- Si $a, b \in \mathbb{R}^n$ entonces $a = (x_1, x_2, ...x_n)$ y $b = (y_1, y_2, ...y_n)$ con todas las x_i y las y_i reales. Entonces $a + b = (x_1 + y_1, x_2 + y_2, ...x_n + y_n)$ con lo que se ve que $a + b$ tiene exáctamente n componentes y todas ellas son reales pues, por la cerradura en los reales, $x_i + y_i$ es real para toda i. Esto prueba la cerradura.

- Si $a, b \in \mathbb{R}^n$ entonces $a = (x_1, x_2, ...x_n)$ y $b = (y_1, y_2, ...y_n)$ con todas las x_i y las y_i reales. Por ello $a+b = (x_1+y_1, x_2+y_2, ...x_n+y_n)$ en tanto que $b+a = (y_1+x_1, y_2+x_2, ...y_n+x_n) = (x_1+y_1, x_2+y_2, ...x_n+y_n)$ y donde el último paso es válido por la conmutatividad de la suma de números reales. Esto prueba que $a + b = b + a$ y la operación $+$ es conmutativa.

- Si $a, b, c \in \mathbb{R}^n$ entonces $a = (x_1, x_2, ...x_n)$ $b = (y_1, y_2, ...y_n)$ y $c = (z_1, z_2, ...z_n)$ con todas las x_i , las y_i y las z_i reales. Entonces $b + c = (y_1 + z_1, y_2 + z_2, ...y_n + z_n)$ y $a+(b+c) = (x_1+(y_1+z_1), x_2+(y_2+z_2), ...x_n+(y_n+z_n))$. Por otra parte $a + b = (x_1 + y_1, x_2 + y_2, ...x_n + y_n)$ y $(a + b) + c = ((x_1 + y_1) + z_1, (x_2 + y_2) + z_2, ...(x_n + y_n) + z_n) = (x_1 + (y_1 + z_1), x_2 + (y_2 + z_2), ...x_n + (y_n + z_n))$ donde el último paso se desprende de la asociatividad de números reales. Por ello $a + (b + c) = (a + b) + c$ y se vale la asociatividad.

- Para verificar el axioma cuatro necesitamos primero indagar quién es el cero. Sabemos que $0 \in \mathbb{R}^n$ y que para cualquier $a \in \mathbb{R}^n$ se debe cumplir el que $a + 0 = 0 + a = a$ por lo que, con la misma a usada anteriormente si $0 = (y_1, y_2, ...y_n)$ tenemos que $a + 0 = (x_1 + y_1, x_2 + y_2, ...x_n + y_n) = a = (x_1, x_2, ...x_n)$ de modo que

$$x_1 + y_1 = x_1$$
$$x_2 + y_2 = x_2$$
$$...$$
$$x_n + y_n = x_n$$

  y resulta que $y_1 = y_2...y_n = 0$ (el número cero). Por ello $0 = (0, 0, .., 0)$ (el 0 del lado izquierdo es el cero en $\mathbb{R}^n$ en tanto que los n ceros en el lado derecho representan el número cero de los reales). Sabiendo quién debe de ser el cero, es fácil ver que, en realidad, $0 = (0, 0.., 0)$ es el cero de $\mathbb{R}^n$. En efecto, si $a = (x_1, x_2, ...x_n)$ entonces $a + 0 = (x_1 + 0, x_2 + 0, ...x_n + 0) = (x_1, x_2, ...x_n) = a$.

- Igual que en el caso anterior, dado un elemento a tenemos primero que deducir quién es su inverso $-a$ y luego demostrar que en efecto es el inverso.

Como $-a \in \mathbb{R}^n$ $-a = (y_1, y_2, ...y_n)$ y si $a = (x_1, x_2, ...x_n)$ entonces $a+(-a)$ será $(x_1 + y_1, x_2 + y_2, ...x_n + y_n) = 0 = (0, 0, .., 0)$ y

$$\begin{aligned} x_1 + y_1 &= 0 \\ x_2 + y_2 &= 0 \\ &... \\ x_n + y_n &= 0 \end{aligned}$$

y

$$\begin{aligned} y_1 &= -x_1 \\ y_2 &= -x_2 \\ &... \\ y_n &= -x_n \end{aligned}$$

y concluimos que $(-a) = (-x_1, -x_2, ... - x_n)$. Que éste es realmente el negativo se desprende del hecho de que $a+(-a) = (x_1 - x_1, x_2 - x_2, ...x_n - x_n) = (0, 0, .., 0) = 0$.

- Si $\alpha \in \mathbb{R}$ y $a = (x_1, x_2, ...x_n) \in \mathbb{R}^n$ entonces $\alpha a = (\alpha x_1, \alpha x_2, ...\alpha x_n)$ y vemos que αa es una eneada de números reales, demostrando que $\alpha a \in \mathbb{R}^n$.

- Si $\alpha, \beta \in \mathbb{R}$, y $a \in \mathbb{R}^n$ entonces, para $a = (x_1, x_2, ...x_n)$ tenemos que

$$\begin{aligned} (\alpha + \beta)(a) &= ((\alpha + \beta)x_1, (\alpha + \beta)x_2, ...(\alpha + \beta)x_n) \\ &= (\alpha x_1 + \beta x_1, \alpha x_2 + \beta x_2, ...\alpha x_n + \beta x_n) \end{aligned}$$

por la distributividad de la multiplicación de números reales. Por otra parte

$$\alpha a + \beta a = (\alpha x_1 + \beta x_1, \alpha x_2 + \beta x_2, ...\alpha x_n + \beta x_n)$$

mostrando que $(\alpha + \beta)(a) = \alpha a + \beta a$.

- Si $a, b \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$ entonces

$$\begin{aligned} \alpha(a + b) &= \alpha(x_1 + y_1, x_2 + y_2, ...x_n + y_n) \\ &= (\alpha(x_1 + y_1), \alpha(x_2 + y_2), ...\alpha(x_n + y_n)) \\ &= (\alpha x_1 + \alpha y_1, \alpha x_2 + \alpha y_2, ...\alpha x_n + \alpha y_n) \end{aligned}$$

valiéndose el último paso por la distributividad entre números reales. Por otra parte

$$\alpha a + \alpha b = (\alpha x_1 + \alpha y_1, \alpha x_2 + \alpha y_2, ...\alpha x_n + \alpha y_n)$$

de donde se ve que $\alpha(a + b) = \alpha a + \alpha b$.

- Si $\alpha, \beta \in \mathbb{R}$ y $a \in \mathbb{R}^n$ entonces

  $$(\alpha\beta)a = ((\alpha\beta)x_1, (\alpha\beta)x_2, ...(\alpha\beta)x_n)$$

  pero

  $$\begin{aligned}\alpha(\beta a) &= \alpha(\beta x_1, \beta x_2, ...\beta x_n) = (\alpha(\beta x_1), \alpha(\beta x_2), ...\alpha(\beta x_n))\\ &= ((\alpha\beta)x_1, (\alpha\beta)x_2, ...(\alpha\beta)x_n)\end{aligned}$$

  por la asociatividad del producto de números reales. Por ello $(\alpha\beta)a = \alpha(\beta a)$. Similarmente $(\alpha\beta)a = \beta(\alpha a)$.

- $1a = 1(x_1, x_2, ...x_n) = (1x_1, 1x_2, .., 1x_n) = (x_1, x_2, ...x_n)$ con lo que $1a = a$.

Concluimos que $\mathbb{R}^n$, con las operaciones arriba definidas, es un espacio vectorial real.

6.1.2. Los espacios $\mathbb{C}^n$

De una manera análoga sea n un entero positivo y

$$\mathbb{C}^n = \{(x_1, x_2, ...x_n) \mid x_i \in \mathbb{C}\ \forall i = 1, 2...n\}$$

el conjunto de eneadas ordenadas de números complejos. Se entiende aquí que dos elementos x y y de $\mathbb{C}^n$ dados por

$$\begin{aligned}x &= (x_1, x_2, ...x_n)\\ y &= (y_1, y_2, ...y_n)\end{aligned}$$

son iguales si y sólo si $\forall i = 1, 2...n$

$$x_i = y_i$$

En $\mathbb{C}^n$ definimos la suma de $x = (x_1, x_2, ...x_n) \in \mathbb{C}^n$ y $y = (y_1, y_2, ...y_n) \in \mathbb{C}^n$ como aquel elemento $x + y \in \mathbb{C}^n$ dado por

$$x + y = (x_1 + y_1, x_2 + y_2, ...x_n + y_n)$$

y definimos el producto de $\alpha \in \mathbb{C}$ y $x = (x_1, x_2, ...x_n) \in \mathbb{C}^n$ como el elemento $\alpha x \in \mathbb{C}^n$ dado por

$$\alpha x = (\alpha x_1, \alpha x_2, ...\alpha x_n)$$

Puede demostrarse que $\mathbb{C}^n$ con la suma y el producto por números complejos recién definidos es un espacio vectorial complejo.

Ejercicio 50 *Demuestre que $\mathbb{C}^n$ con la suma y el producto por números complejos definidos arriba es un espacio vectorial complejo.*

Ejercicio 51 *Si $\mathbb{C}^n$ se equipa con la misma suma que se definió en el ejercicio anterior pero la multiplicación por escalares se restringe a números reales, ¿es la estructura resultante un espacio vectorial real?*

Notación 52 *Los vectores de $\mathbb{C}^n$ y $\mathbb{R}^n$ pueden representarse en las siguientes formas que son equivalentes:*

$$(x_1, x_2, ...x_n)$$
$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$
$$\begin{pmatrix} x_1 & x_2 & ... & x_n \end{pmatrix}^T$$

y frecuentemente consideramos estos vectores como si fueran matrices de $1 \times n$ o $n \times 1$. El punto central es que los vectores de $\mathbb{C}^n$ (o de .$\mathbb{R}^n$) son colecciones de exactamente n números complejos (o reales). De la misma manera, consideramos a los separadores entre los números como matematicamente irrelevantes, de manera que un espacio vacío, una coma, un punto y coma etc. (el llamado espacio blanco de la computación) podrían ser usados para separar los números.

Notación 53 *Hay una notación útil que restringimos a $\mathbb{C}^n$. Si denotamos por v al vector*

$$v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

entonces, por el símbolo v^* denotamos el transpuesto conjugado (adjunto) del vector y

$$v^* = (\overline{x}_1 \ \overline{x}_2 ... \overline{x}_n)$$

que vemos como una matriz de $1 \times n$

6.1.3. Los espacios P_n

Sea $n$un entero positivo y

$$P_n = \{a_0 + a_1x + a_2x^2 + ...a_nx^n \mid a_i \in \mathbb{R} \ \forall i = 0, 1...n\}$$

el conjunto de todos los polinomios con coeficientes reales y grado menor o igual a n. Aquí se entiende que dos miembros $P(x)$y $Q(x)$de $P_n$dados por

$$P(x) = a_0 + a_1x + a_2x^2 + ...a_nx^n$$
$$Q(x) = b_0 + b_1x + b_2x^2 + ...b_nx^n$$

son iguales si y sólo si $\forall i = 0, 1...n$

$$a_i = b_i$$

En P_n definimos la suma de $P(x) = a_0 + a_1x + a_2x^2 + ...a_nx^n \in P_n$ y $Q(x) = b_0 + b_1x + b_2x^2 + ...b_nx^n \in P_n$ como el elemento $(P+Q)(x) \in P_n$ dado por

$$(P+Q)(x) = (a_0+b_0) + (a_1+b_1)x + (a_2+b_2)x^2 + ...(a_n+b_n)x^n$$

y definimos el producto de $\alpha \in \mathbb{R}$ y $P(x) = a_0 + a_1x + a_2x^2 + ...a_nx^n \in P_n$ como el elemento $\alpha x \in P_n$ dado por

$$(\alpha P)(x) = (\alpha a_0) + (\alpha a_1)x + (\alpha a_2)x^2 + ...(\alpha a_n)x^n$$

Puede demostrarse que P_n con la suma y el producto por números reales recién definidos es un espacio vectorial real.

Ejercicio 54 *Demuestre que P_n con la suma y el producto por números reales definidos anteriormente es un espacio vectorial real.*

6.1.4. Los espacios $M(m,n)$

Sean my $n$enteros positivos y

$$M(m,n) = \left\{ \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix} | a_{ij} \in \mathbb{R}\ \forall i = 0,1...m \forall j = 0,1...n \right\}$$

el conjunto de todas las matrices reales de $m \times n$.

Para referirnos a los elementos de una matriz dada $A \in M(m,n)$ usaremos la notación $a_{ij} = [A]_{ij}$ y diremos que $A = B$ (con $A, B \in M(m,n)$) si y sólo si $\forall i = 0,1...m \forall j = 0,1...n$

$$[A]_{ij} = [B]_{ij}$$

Definimos una suma en $M(m,n)$ mediante

$$[A+B]_{ij} = [A]_{ij} + [B]_{ij}$$

$(\forall i = 0,1...m \forall j = 0,1...n)$ donde se sobreentiende que $A, B \in M(m,n)$.

La multiplicación de elementos de $M(m,n)$ por números reales se define mediante

$$[\alpha A]_{ij} = \alpha [A]_{ij}$$

$(\forall i = 0,1...m\ \forall j = 0,1...n)$ con $A \in M(m,n)$.

$M(m,n)$ con las operaciones anteriormente definidas es un espacio vectorial real.

Ejercicio 55 *Demuestre que $M(m,n)$ con la suma y el producto por números reales definidos anteriormente es un espacio vectorial real.*

Comentario 56 *Ya hemos visto que $M(n,n)$ es un espacio vectorial y, anteriormente, que es un anillo con unidad. Se dice, en este caso, que se trata de un álgebra.*

6.1.5. Los espacios $M^*(m,n)$

Similarmente definimos (para enteros positivos my n)

$$M^*(m,n) = \left\{ \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix} | a_{ij} \in \mathbb{C}\ \forall i = 0,1...m \forall j = 0,1...n \right\}$$

como el conjunto de todas las matrices complejas de $m \times n$.

Para referirnos a los elementos de una matriz dada $A \in M^*(m,n)$ usaremos la notación

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

y diremos que $A = B$ (con $A, B \in M^*(m,n)$) si y sólo si $\forall i = 0,1...m \forall j = 0,1...n$

$$[A]_{ij} = [B]_{ij}$$

Definimos una suma en $M^*(m,n)$ mediante

$$[A+B]_{ij} = [A]_{ij} + [B]_{ij}$$

($\forall i = 0,1...m \forall j = 0,1...n$) donde se entiende que $A, B \in M^*(m,n)$.

La multiplicación de elementos de $M(m,n)$ por números complejos se define mediante

$$[\alpha A]_{ij} = \alpha [A]_{ij}$$

($\forall i = 0,1...m \forall j = 0,1...n$) y con $A \in M^*(m,n)$.

$M^*(m,n)$ con estas operaciones es un espacio vectorial complejo.

Ejercicio 57 *Demuestre que $M^*(m,n)$ con la suma y el producto por números complejos definidos anteriormente es un espacio vectorial complejo.*

Comentario 58 *También se usan las notaciones $M_{mn}(\mathbb{R})$ o $M_{mn}(\mathbb{C})$ en vez de $M(m,n)$ y $M^*(m,n)$.*

Comentario 59 *Todo campo puede ser visto como un espacio vectorial sobre él mismo (tome $n = 1$ en las definiciones de $\mathbb{C}^n$ y de $\mathbb{R}^n$). Por ello el campo de los números complejos puede verse como espacio vectorial complejo, esto contesta la pregunta que frecuentemente se hace de si hay espacios complejos de "flechas" (vectores geométricos). De acuerdo con la interpretación geométrica de los complejos (diagrama de Argand), los complejos son flechas que pueden sumarse y multiplicarse por números. Pero en este caso al multiplicar un vector por un número se le cambia el tamaño y se le rota o gira (si el vector x se multiplica por el número $z = \rho e^{i\theta}$ entonces su tamaño se altera por un factor ρ y su ángulo de fase se incrementa por θ).*

Comentario 60 *Si F es un campo y G un subcampo, entonces también F puede pensarse como un espacio vectorial sobre G .*

Comentario 61 *En el caso de los espacios de polinomios podríamos definir espacios de polinomios con coeficientes sobre otros campos aparte del de los números reales.*

Ejemplo 62 *Otros ejemplos de espacios vectoriales serían:*
a) el conjunto de todas las secuencias x_i ($i = 1...\infty$) de números reales o complejos. La suma se define como $(x + y)_i = x_i + y_i$ y el producto por números como $(\alpha x)_i = \alpha x_i$.
b) si U y V son espacios vectoriales sobre un mismo campo, el producto cartesiano de U y V se define como

$$U \times V = \{(u, v) \mid u \in U, v \in V\}$$

definiendo la suma mediante

$$(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2)$$

y el producto como

$$\alpha(u, v) = (\alpha u, \alpha v)$$

Esto permitiría, por ejemplo, tener un espacio donde los vectores sean parejas de polinomios, o parejas en donde el primer elemento sea un polinomio y el segundo una matriz. A esta construcción se le llama "suma directa externa".
c) Si S es un conjunto, el conjunto de todas las funciones de S a los reales (o los complejos o a cualquier otro campo)

$$F = \{f : S \to \mathbb{R}\}$$

es un espacio vectorial si definimos

$$(f + g)(x)) = f(x) + g(x)$$

y

$$(\alpha f)(x) = \alpha f(x)$$

El conjunto F (y también el correspondiente espacio) también se denotan como $\mathbb{R}^S$.

6.2. Algunas propiedades básicas de los espacios vectoriales.

Hay varias propiedades importantes de todos los espacios vectoriales.

El axioma 4 indica que hay un vector especial 0 que tiene la propiedad de que $0 + v = v + 0 = v$ para todo $v \in V$, demostramos a continuación que sólo hay un vector con estas propiedades:

Teorema 63 *Sea V un espacio vectorial sobre un campo F. Entonces si $z \in V$ tiene la propiedad de que $z + v = v + z = v$ para todo $v \in V$ entonces $z = 0$. En palabras: el idéntico aditivo es único.*

Demostración. $z = z + 0$ (por el axioma 4)
$= 0$ (porque $z + 0 = 0$) ■
Analogamente, dado v, el vector $-v$ cuya existencia se asevera en el axioma 5 es único y tenemos el teorema:

Teorema 64 *Sea V un espacio vectorial sobre el campo F y sea $v \in V$. Si $w \in V$ tiene la propiedad de que $v+w = w+v = 0$ entonces $w = -v$ (el inverso es único).*

Demostración. $w = w + 0$ (por el axioma 4)
$= w + (v + (-v))$ (por el axioma 5)
$= (w + v) + (-v)$ (por el axioma 3
$= 0 + (-v)$ (por las propiedades que supusimos para w)
$= -v$ (por el axioma 4) ■
De hecho tenemos una forma débil del teorema 63 que a menudo es útil:

Teorema 65 *Sea V un espacio vectorial sobre el campo F y sea $v \in V$. Si $w \in V$ tiene la propiedad de que $v + w = v$ entonces $w = 0$.*

Demostración. $(-v) + (v + w) = (-v) + v$ (por las propiedades asumidas para w)
$(-v + v) + w = 0$ (por los axiomas 3 y 4)
$0 + w = 0$ (por el axioma 5)
$w = 0$ (por el axioma 4) ■

Teorema 66 *Sea V un espacio vectorial (real o complejo). Entonces $0v = 0$ $\forall v \in V$.*

Demostración. Dado que $0v = 0v + 0$(por el axioma 4)
$= 0v + (v + (-v))$ (por el axioma 5)
$= (0v + v) + (-v)$(por el axioma 3)
$= (0v + 1v) + (-v)$(por el axioma 7)
$= (0 + 1)v + (-v)$ (por el axioma 8)
$= (1)v + (-v)$ (pues en todo campo $1 + 0 = 1$)
$= v + (-v)$ (por el axioma 7)
$= 0$ (por el axioma 4)
Note que no se usó la propiedad conmutativa de la suma. ■

Teorema 67 *Sea V un espacio vectorial (real o complejo). Entonces $(-1)v = -v$*

Demostración. Dado que $(-1)v = (-1)v + 0$ (axioma 4)
$= (-1)v + (v + (-v))$ (axioma 5)
$= ((-1)v + v) + (-v)$ (axioma 3)
$= ((-1)v + 1v) + (-v)$ (axioma 7)
$= ((-1 + 1)v) + (-v)$ (axioma 8)
$= ((0)v) + (-v)$ (en todo campo $1 + (-1) = 0$)
$= 0 + (-v)$ (teorema anterior)
$= -v$ (axioma 4)
De nuevo, note que no se usó la propiedad conmutativa de la suma. ∎

Teorema 68 *Sea V un espacio vectorial(real or complejo). Entonces si λ es un número (real or complejo dependiendo de si el espacio V es real o complejo) tenemos que $\lambda\overline{0} = \overline{0}$ (momentáneamente llamaré $\overline{0}$ al vector cero y 0 al número cero).*

Demostración. $\lambda\overline{0} = \lambda(0\overline{0})$ (por el teorema 66)$= (\lambda 0)\overline{0}$ (por el axioma 10)
$= 0\overline{0}$ (propiedades básicas de los números)
$= \overline{0}$ (por el teorema 66) ∎

Teorema 69 *Sea V un espacio vectorial (real o complejo). Entonces si $v \in V$ entonces $-(-v) = v$.*

Demostración. Puesto que $v + (-v) = 0$ (axioma 5)
$-(-v) = v$ (por el teorema 64) ∎

Teorema 70 *Sea V un espacio vectorial (real o complejo) λ un número y $v \in V$. Entonces si $\lambda v = 0$ entonces $\lambda = 0$ o $v = 0$.*

Demostración. Si $\lambda = 0$ el teorema es verdadero. Suponga que $\lambda \neq 0$ de modo que $\exists \lambda^{-1}$.
Entonces $\lambda^{-1}(\lambda v) = \lambda^{-1}0 = 0$ (por el teorema 68)
pero $\lambda^{-1}(\lambda v) = (\lambda^{-1}\lambda)v$ (por el axioma 10)
$= 1v$ (propiedades básicas de los números reales o complejos)
$= v$ (por el axioma 7)
Consecuentemente $v = 0$. ∎

Ejercicio 71 *Demuestre que si V es un espacio vectorial, $v \in V$ y λ es un número, entonces $(-\lambda)v = -(\lambda v) = \lambda(-v)$.*

Ejercicio 72 *Pruebe que si V es un espacio vectorial, $v \in V$ y λ es un número, entonces $(-\lambda)(-v) = \lambda v$.*

Ejercicio 73 *Pruebe que si V es un espacio vectorial y $u + w = v + w$ entonces $u = v$*

Ejercicio 74 *Pruebe que si V es un espacio vectorial y $\lambda u = \lambda v$ para algún número $\lambda \neq 0$ entonces $u = v$.*

6.3. Subespacios

Definición 75 *Sea V un espacio vectorial. Decimos que W es un subespacio de V si: a) $W \subset V$, b) W es un espacio vectorial con las mismas operaciones que V.*

Comentario 76 *Por"las mismas operaciones que V" entendemos lo siguiente: Si a,b $\in W$, puesto que $W \subset V$ tenemos que a,b $\in V$, y consecuentemente podemos formar su suma $a + b$; decimos que esta suma es la que es inducida en W por la suma en V. Similarmente, si $a \in W$,puesto que $W \subset V$ tenemos que $a \in V$, consecuentemente podemos formar el producto λa (donde λ es un número); decimos que este producto es el que es inducido en W por el producto en V.*

Un subespacio es, entonces, un subconjunto que también es un espacio vectorial. Note que todo espacio V es un subespacio de sí mismo (un subespacio trivial). A fin de verificar si un conjunto W es un subespacio de un espacio vectorial V tendríamos, en principio, que verificar que $W \subset V$ y que W (con la suma y producto inducidos por V) es un espacio vectorial. En la práctica es mucho más sencillo usar el siguiente teorema:

Teorema 77 *Sea V un espacio vectorial. Entonces W es un subespacio de V si y sólo si las siguientes cuatro condiciones se cumplen: a) $W \subset V$, b) $W \neq \phi$, c) Si $a, b \in W$ entonces $a + b \in W$ (cerradura) y d) Si k es un número y $a \in W$ entonces $k \cdot a \in W$.*

Demostración. $\Rightarrow$

Si W es un subespacio de V, entonces $W \subset V$

Si W es un subespacio de V, entonces es un espacio vectorial, consecuentemente $W \neq \phi$ pues $0 \in W$ por el axioma 4

Si W es un subespacio de V,entonces es un espacio vectorial, consecuentemente por el axioma 1 si $a, b \in W$ entonces $a + b \in W$.

Si W es un subespacio de V, entonces es un espacio vectorial, consecuentemente por el axioma 6 si k es un número y $a \in W$ entonces $k \cdot a \in W$.

$\Leftarrow$

Si W cumple las cuatro condiciones, entonces el hecho de que cumple los axiomas 1,2,3,5,6,7,8,9 y 10 es una consecuencia simple del hecho de que V es un espacio vectorial y la demostración se deja como ejercicio. El axioma 4 es más sutil:

Puesto que (por la hipótesis b) $W \neq \phi$, W tiene algún elemento, llámelo x, entonces (por la hipótesis d) $0x \in W$, pero (por la hipótesis a) si $0x \in W$ then $0x \in V$, pero por el teorema 4 $0x = 0$. Entonces $0 \in W$. Para cualquier $a \in W$ dado que también es cierto que $a \in V$ $a + 0 = 0 + a = a$ y el axioma 4 se cumple. ∎

Ejercicio 78 *Llene los huecos en la prueba del teorema 77.*

Ejemplo 79 *Si V es un espacio vectorial y $W = \{0\}$ entonces W es, trivialmente, un subespacio de V, se llama "espacio nulo". Los subespacios triviales de V son, entonces, V mismo y $\{0\}$.*

Ejemplo 80 *Claramente ϕ no puede ser un subespacio de ningún espacio pues $0 \notin \phi$.*

Ejercicio 81 *Sea V un espacio vectorial y $a \in V$ tal que $a \neq 0$. ¿ es $\{0, a\}$ un subespacio de V? Vemos que $\{0, a\} \neq \phi$ y que $\{0, a\} \subset V$. Comencemos con la cerradura con respecto a la multiplicación por escalares. Si $\alpha \in F$ y $x \in \{0, a\}$ entonces αx debe de ser uno de $\alpha 0$, αa o sea 0,αa. Con el caso $\alpha x = 0$ no hay problema, pero cuando $x = a$ tenemos que αa debe de ser 0 o a. El primer caso implica que α o a sea cero, pero a no es cero de modo que α debe de ser cero. El segundo caso implica que $\alpha a = a$ y $(\alpha - 1)a = 0$ implicando que $\alpha = 1$. Por ello la cerradura no puede cumplirse si α no es cero o uno. Por ello concluimos que $\{0, a\}$ no es cerrado bajo multiplicación escalar y no es un subespacio. De hecho tampoco es cerrado bajo la suma, pues al sumar $x + y$ con ambos sumandos en $\{0, a\}$ las posibilidades son las de la tabla siguiente*

x	y	$x+y$
0	0	0
0	a	a
a	0	a
a	a	$2a$

y vemos que cuando tenemos $a+a = 2a$ para que valga la cerradura $2a \in \{0, a\}$ o sea $2a = 0$ o $2a = a$. Ambos casos implicarían que $a = 0$ contrario a las hipótesis. $\{0, a\}$ no es cerrado bajo suma tampoco.

Teorema 82 *Si U y V son subespacios de W, entonces $U \cap V$ es también es subespacio de W.*

Demostración. 1) Claramente el $0 \in W$ también satisface $0 \in U$ y $0 \in V$ con lo que $0 \in U \cap V$
2) Claramente $U \cap V \subset W$
3) si $x, y \in U \cap V$ entonces $x, y \in U$ y $x, y \in V$ con lo que $x + y \in U$ y $x + y \in V$ (pues U y V son subespacios) y $x+y \in U \cap V$ probando la cerradura bajo adición.
4) si $x \in U \cap V$ y $\alpha \in F$ entonces $x \in U$ y $x \in V$ con lo que $\alpha x \in U$ y $\alpha x \in V$ (por la cerradura bajo multiplicación escalar de los subespacios) y $\alpha x \in U \cap V$ con lo que $U \cap V$ también es cerrado bajo multiplicación.
Consecuentemente $U \cap V$ es un subespacio de W. ■

Sin embargo $U \cup V$ no es un subespacio en general, aunque $U \cup V \neq \phi$ $(0 \in U \cup V)$ y $U \cup V \subset W$ resulta que las cerraduras pueden fallar. Esto se ve claramente con un ejemplo: considere $W = \mathbb{R}^2$, $U = \{\alpha(1, 0) \mid \alpha \in \mathbb{R}\}$ y $V = \{\beta(0, 1) \mid \beta \in \mathbb{R}\}$, dejo al lector verificar que tanto U como V son subespacios de W. Pero $(1, 0) + (0, 1) = (1, 1) \notin U \cup V$ pues $(1, 1) \notin U$ y $(1, 1) \notin V$. También deberá el lector mostrar que $U \cup V$ sí es cerrado bajo la multiplicación por escalares. Es la cerradura bajo $+$ la que causa problemas.

Comentario 83 *Si vemos al espacio vectorial como una estructura algebraica, podemos pensar al subespacio como una subestructura. En matemáticas es muy común estudiar estructuras y luego sus posibles subestructuras.*

Ejemplo 84 *Si U y V son subespacios de W, entonces podemos definir la suma de U y V como*

$$U + V = \{u + v \mid u \in U, v \in V\}$$

y dejamos como ejercicio para el lector el verificar que $U + V$ es un subespacio de W.

6.4. Sistemas de ecuaciones lineales homogéneas.

Una clase especial pero importante de subespacios es la formada por los conjuntos solución de sistemas de ecuaciones lineales homogeneas.

El sistema más general tiene la forma:

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

y decimos que es un sistemas de ecuaciones lineales homogeneas de m equaciones en n incógnitas.

Si definimos una matriz A de $m \times n$ como

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

y la matriz X de $n \times 1$(que consideramos como un elemento de $\mathbb{R}^n$ o de $\mathbb{C}^n$) mediante

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

entonces podemos escribir el sistema como

$$AX = 0$$

donde

$$0 = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

es el vector cero de $\mathbb{R}^m$ (o $\mathbb{C}^n$ según sea el caso).

Es conveniente definir el conjunto solución de este sistema como

$$S = \{X \in F^n \mid AX = 0\}$$

donde F^n quiere decir $\mathbb{R}^n$ o $\mathbb{C}^n$ según sea el caso.

Teorema 85 *S es un subespacio de F^n*

Demostración. *a*) claramente $S \subset F^n$

b) $S \neq \phi$ pues $A0 = 0$ (aquí el 0 en la izquierda es el cero de F^n, en tanto que el cero en la derecha es el cero de F^m) y $0 \in S$.

c) si x,$y \in S$ entonces $Ax = 0$ y $Ay = 0$, por lo que $A(x+y) = Ax + Ay = 0 + 0 = 0$ y $x + y \in S$.

d) si $x \in S$ y $\lambda \in F$ entonces $A(\lambda x) = \lambda A(x) = \lambda 0 = 0$ de modo que $\lambda x \in S$.

La conclusión se desprende del teorema 9. ■

6.5. combinaciones lineales

En todo lo que sigue sea V un espacio vectorial.

Sea $\alpha = \{v_1, v_2, ...v_n\} \subset V$ un conjunto de n vectores en V. Una combinación lineal de los vectores en α es una expresión de la forma

$$c_1 v_1 + c_2 v_2 + ...c_n v_n$$

donde c_1, c_2 ...c_n son números.

Ejemplo 86 *¿es $x = (9,2,7)$ combinación lineal de $u = (1,2,-1)$ y $v = (6,4,2)$?*
Lo que hay que preguntarse es si existen escalares α y β tales que $x = \alpha u + \beta v$. Si existieran

$$(9,2,7) = \alpha(1,2,-1) + \beta(6,4,2) = (\alpha + 6\beta, 2\alpha + 4\beta, -\alpha + 2\beta)$$

por lo que tenemos el sistema lineal inhomogéneo

$$\begin{aligned} \alpha + 6\beta &= 9 \\ 2\alpha + 4\beta &= 2 \\ -\alpha + 2\beta &= 7 \end{aligned}$$

cuya matriz aumentada es

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

indicando que

$$\begin{aligned}
\alpha &= -3 \\
\beta &= 2
\end{aligned}$$

y x es, en efecto, combinación lineal de u y v.

Ejemplo 87 *¿es $x = (9,3,7)$ combinación lineal de $u = (1,2,-1)$ y $v = (6,4,2)$?*
Lo que hay que preguntarse es si existen escalares α y β tales que $x = \alpha u + \beta v$. Si existieran

$$(9,3,7) = \alpha(1,2,-1) + \beta(6,4,2) = (\alpha + 6\beta, 2\alpha + 4\beta, -\alpha + 2\beta)$$

por lo que tenemos el sistema lineal inhomogéneo

$$\begin{aligned}
\alpha + 6\beta &= 9 \\
2\alpha + 4\beta &= 3 \\
-\alpha + 2\beta &= 7
\end{aligned}$$

cuya matriz aumentada es

$$\begin{aligned}\begin{pmatrix} 1 & 6 & 9 \\ 2 & 4 & 3 \\ -1 & 2 & 7 \end{pmatrix} &\sim \begin{pmatrix} 1 & 6 & 9 \\ 0 & -8 & -15 \\ 0 & 8 & 16 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 6 & 9 \\ 0 & 8 & 15 \\ 0 & 1 & 2 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 0 & -3 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 0 & -3 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\end{aligned}$$

que no tiene soluciones. Luego x no es combinación lineal de u y v.

6.5.1. Des-sumando y des-multiplicando

Hay un "truco" que usaremos fecuentemente y que consiste en tomar vectores de $\mathbb{R}^n$ o de $\mathbb{C}^n$ y descomponerlos combinaciones lineales. Un ejemplo vale un millón de palabras:

Considere el vector $(2x + 3y, -x + 8y)$, usando las regla para la suma y producto por escalares vemos que

$$\begin{aligned}(2x + 3y, -x + 8y) &= (2x, -x) + (3y, 8y) \\ &= x(2, -1) + y(3, 8)\end{aligned}$$

en el primer renglón hemos descompuesto la expresión en dos sumandos, uno que sólo incluye la variable x y otro que sólo incluye la variable y; a continuación hemos factorizado los escalares x y y.

6.6. generadores

Teorema 88 *Sea V un espacio vectorial y $\alpha = \{v_1, v_2, ...v_n\} \subset V$. El conjunto $Gen(\alpha) = Gen(v_1, v_2, ...v_n)$ de todas las combinaciones lineales de elementos en α y definido por*

$$Gen(\alpha) = \{c_1v_1 + c_2v_2 + ...c_nv_n \mid c_i \in F \ \forall i = 1, 2...n\}$$

es un subespacio de V.

Demostración. *a*) $Gen(\alpha)$ claramente es un subconjunto de V .

b) $0 = 0v_1 + 0v_2 + ..,0v_n$, de modo que $0 \in Gen(\alpha)$ y $Gen(\alpha) \neq \phi$

c) Si $x, y \in Gen(\alpha)$ entonces $x = c_1v_1 + c_2v_2 + ...c_nv_n$ y $y = d_1v_1 + d_2v_2 + ...d_nv_n$ donde los c_i y los d_i son números. Entonces $x + y = (c_1 + d_1)v_1 + (c_2 + d_2)v_2 + ...(c_n + d_n)v_n$ y $x + y \in Gen(\alpha)$

d) Si $x \in Gen(\alpha)$ entonces $x = c_1v_1 + c_2v_2 + ...c_nv_n$ donde los c_i son números. Entonces para $\lambda \in F$

$\lambda x = (\lambda c_1)v_1 + (\lambda c_2)v_2 + ...(\lambda c_n)v_n$ y $\lambda x \in Gen(\alpha)$.

La conclusión se sigue del teorema 9. ■

Algunos autores escriben en vez de $Gen(v_1, v_2, ...v_n)$ $L(v_1, v_2, ...v_n)$ o incluso $\langle(v_1, v_2, ...v_n)\rangle$.

Decimos que estos vectores generan un subespacio $Gen(\alpha)$ de V, o que los vectores son un conjunto generador para $Gen(\alpha)$. En general, decimos que α genera W si $W = Gen(\alpha)$. En otras palabras, un conjunto α es un generador de un espacio W si cada vector de W es una combinación lineal de vectores en α.

Por los ejemplos anteriores, podemos ver que $u = (1, 2, -1)$ y $v = (6, 4, 2)$ no generan $\mathbb{R}^3$ pues hay vectores que no son combinaciones lineales de u y v.

En general, un conjunto generador para $\mathbb{R}^n$ o $\mathbb{C}^n$ está dado por

$$\{(1, 0, 0.., 0), (0, 1, 0, .., 0), ...(0, 0, 0.., 1)\}$$

un conjunto generador para P_n está dado por

$$\{1, x, x^2...x^n\}$$

en tanto que un conjunto generador para $M(m, n)$ y $M^*(m, n)$ es

$$\left\{\begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & 0 & ... & 0 \\ & & \vdots & \\ 0 & 0 & ... & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & ... & 0 \\ 0 & 0 & ... & 0 \\ & & \vdots & \\ 0 & 0 & ... & 0 \end{pmatrix}, ..., \begin{pmatrix} 0 & 0 & ... & 0 \\ 0 & 0 & ... & 0 \\ & & \vdots & \\ 0 & 0 & ... & 1 \end{pmatrix}\right\}$$

Todo espacio tiene un número infinito de posibles generadores.

Más adelante mostraremos cómo obtener generadores de conjuntos solución de sistemas lineales homogéneos.

Ejemplo 89 $\mathbb{R}^3$ *contiene los vectores* $(1, 0, 0)$ *y* $(0, 1, 0)$ *y* $S = Gen((1, 0, 0), (0, 1, 0))$ *es un subespacio de* $\mathbb{R}^3$*. pero* S *no coincide con* $\mathbb{R}^3$ *pues* $\{(1, 0, 0), (0, 1, 0)\}$ *no genera* $\mathbb{R}^3$ *(por ejemplo* $(0, 0, 1)$ *no está en* S*).*

En álgebra lineal se consideran dos clases importantes de problemas relativos a generadores:

- Dado un conjunto $\{v_1, v_2, ...v_n\}$ determinar el espacio generado $Gen(v_1, v_2, ...v_n)$.
- Dado un espacio, determinar un generador.

6.6.1. cómo encontrar un generador para el espacio solución de un sistema lineal homogéneo.

Veamos primero, mediante un ejemplo, la estrategia básica.
Considere el sistema

$$\begin{aligned} 3x + 5y - 4z &= 0 \\ -3x - 2y + 4z &= 0 \\ 6x + y - 8z &= 0 \end{aligned}$$

La matriz de coeficientes es

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

cuya forma escalonada reducida por renglones es:

$$\begin{bmatrix} 1 & 0 & -\frac{4}{3} \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

por lo que un sistema equivalente (en el sentido de tener exactamente las mismas soluciones) es

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

En la matriz reducida debemos identificar las columnas pivote, es decir las columnas con un uno delantero (o principal). Éstas sosn la primera y la segunda columna. Las variables correspondientes (x y y) se llaman variables ligadas y las restantes (z en este caso) se llaman variables libres . La parte crucial consiste en tomar el vector genérico de $\mathbb{R}^3$ (pues el conjunto solución es un subespacio de $\mathbb{R}^3$) y reexpresarlo en términos de las variables libres usando las ecuaciones anteriores de modo que

$$(x, y, z) = (\frac{4}{3}z, 0, z) = z(\frac{4}{3}, 0, 1)$$

Factorizando los escalares, obtenemos las combinaciones lineales que dan todas las soluciones, de ellas podemos inferir el generador (los vectores cuyas combinaciones lineales dan las soluciones).

Por ello un generador para el espacio solución del sistema propuesto es

$$\{(\frac{4}{3}, 0, 1)\}$$

En el caso general lo que hay que hacer es:

- del sistema S inferimos la matriz de coeficientes M
- calculamos la forma escalonada reducida por renglones N de la matriz M
- de N inferimos el sistema de ecuaciones S' que es equivalente a S.
- de N inferimos cuáles son las variables ligadas (las que corresponden a colunmas conteniendo los unos principales, delanteros o pivotes).
- de N inferimos cuáles son las variables libres (las que no son ligadas).
- tomamos el vector genérico $v = (x_1, x_2 ... x_n)$ de F^n (n es el número de incógnitas)
- reexpresamos v usando las ecuaciones de S' poniendo las variables ligadas en términos de las libres.
- reescribimos v factorizando los escalares de manera que se vea que las soluciones son combinaciones lineales de ciertos vectores de F^n.
- estos vectores constituyen un generador.

6.7. Independencia lineal y dependencia lineal

Considere un espacio vectorial V, un conjunto $\alpha = \{v_1, v_2, ... v_n\} \subset V$ y la ecuación lineal

$$x_1 v_1 + x_2 v_2 + ... x_n v_n = 0$$

donde los números $x_1, x_2 .. x_n$ son incógnitas cuyos valores deseamos determinar. En otras palabras, deseamos ver cuáles combinaciones lineales de vectores en α dan el vector cero.

Obviamente, si $x_1 = x_2 = ... x_n = 0$ la combinación lineal correspondiente da 0 (se dice que la combinación lineal es trivial). Pero nos preguntamos si hay soluciones en las que no todos los coeficientes sean cero.

Definición 90 *Dado un espacio vectorial V y un conjunto $\alpha = \{v_1, v_2, ... v_n\} \subset V$ decimos que α es linealmente independiente (o que los vectores en α son linealmente independientes) si*

$$x_1 v_1 + x_2 v_2 + ... x_n v_n = 0 \Rightarrow x_1 = x_2 = ... x_n = 0$$

Por otro lado, decimos que α es linealmente dependiente (o que los vectores en α son linealmente dependientes) si hay números $x_1, x_2 .. x_n$, no todos cero, y tales que

$$x_1 v_1 + x_2 v_2 + ... x_n v_n = 0$$

Ejercicio 91 *Demuestre que cualquier conjunto que contenga al vector cero* 0 *es linealmente dependiente. Sea el conjunto* $\{v_1, v_2, ...v_n, 0\}$*, entonces*

$$0v_1 + 0v_2 + ..,0v_n + 1\overrightarrow{0} = \overrightarrow{0}$$

y tenemos una combinación lineal de los vectores que da cero y no todos los coeficientes son cero (el coeficiente de $\overrightarrow{0}$ *es uno)*

Teorema 92 *Sea* V *un espacio vectorial,* $S = \{v_1, v_2, ...v_n\}$ *un subconjunto de* V *que es linealmente independiente y* T *un subconjunto de* S*. Entonces* T *también es independiente.*

Demostración. Suponga (renumerando los vectores si es necesario) que $T = \{v_1, v_2, ...v_k\}$ con $k \leq n$, es dependiente, entonces existen $x_1, x_2..x_k$, no todos cero, y tales que $x_1v_1 + x_2v_2 + ...x_kv_k = 0$ pero esto implica que $x_1v_1 + x_2v_2 + ...x_kv_k + 0v_{k+1} + .,0v_n = 0$ y tenemos una combinación lineal de $\{v_1, v_2, ...v_n\}$ que da cero y no todos los coeficientes son cero, lo cual contradice la independencia de S. ■

Corolario 93 *Sea* V *un espacio vectorial,* $S = \{v_1, v_2, ...v_n\}$ *un subconjunto de* V *y* T *un subconjunto de* S *linealmente dependiente. Entonces* S *también es dependiente.*

Pero no es cierto que un subconjunto de un conjunto linealmente dependiente sea necesariamente dependiente ni que si S tiene un subconjunto independiente S tenga que ser independiente.

En resumidas cuentas

$$S \text{ independiente } \Rightarrow T \text{ independiente}$$
$$T \text{ dependiente} \Rightarrow S \text{ dependiente}$$

y note por favor que las proposiciones son contrapositivas la una de la otra. En general la proposición $P \Rightarrow Q$ es lógicamente equivalente a $\sim Q \Rightarrow \sim P$. La aseveración falsa "$S$ dependiente$\Rightarrow$ T dependiente" es la conversa de "T dependiente$\Rightarrow$ S dependiente" y la contrapositiva de "T independiente$\Rightarrow$ S independiente".

Si se nos dice que $\{a, b\}$ es dependiente, entonces $xa + yb = 0$ para números x y y no ambos cero. Suponga que $x \neq 0$, entonces $a = (\frac{-y}{x})b$ y los vectores son colineales (o paralelos).

Si se nos dice que $\{a, b, c\}$ es dependiente, entonces $xa + yb + zc = 0$ para números x , y y z no todos cero. Suponga que $x \neq 0$, entonces $a = (\frac{-y}{x})b + (\frac{-z}{x})c$ y los vectores son coplanares.

Teorema 94 *Sea* V *un espacio vectorial y* $S = \{v_1, v_2, ...v_n\}$ *un subconjunto de* V*. Si* v_n *es combinación lineal de* $\{v_1, v_2, ...v_{n-1}\}$ *entonces* S *es linealmente dependiente.*

Demostración. Existen, por hipótesis, números $x_1, x_2..x_{n-1}$ tales que $v_n = x_1v_1 + x_2v_2 + ...x_{n-1}v_{n-1}$ de modo que $x_1v_1 + x_2v_2 + ...x_{n-1}v_{n-1} + (-1)v_n = 0$ con lo que tenemos una combinación lineal de vectores de $\{v_1, v_2, ...v_n\}$ que no es trivial y da cero. Por ello S es dependiente. ■

Y hay un converso,

Teorema 95 *Sea V un espacio vectorial y $S = \{v_1, v_2, ...v_n\}$ un subconjunto de V. Si S es linealmente dependiente, entonces uno de los vectores de S es combinación lineal de los restantes.*

Demostración. Si S es linealmente dependiente entonces hay números no todos cero y tales que $x_1v_1 + x_2v_2 + ...x_{n-1}v_{n-1} + x_nv_n = 0$. Suponga que $x_1 \neq 0$ (y lo mismo vale si es otro el que no es cero) entonces $v_1 = (\frac{-1}{x_1})(x_2v_2 + ...x_{n-1}v_{n-1} + x_nv_n)$ demostrando el teorema. ■

6.7.1. ¿Cómo determinar la dependencia o independencia?

Si le dan a Usted un espacio vectorial V y un conjunto $\alpha = \{v_1, v_2, ...v_n\} \subset V$ y le preguntan si α es linealmente independiente lo que debe de hacer es plantear la ecuación $x_1v_1+x_2v_2+...x_nv_n = 0$ y resolverla (se convertirá eventualmente en un sistema homogéneo de ecuaciones lineales). Si la única solución es la trivial, el conjunto es linealmente independiente. Si hay otras soluciones, α es linealmente dependiente.

Ejemplo 96 *¿son los vectores de $\mathbb{R}^3$ $(1, 3, 2)$, $(-1, 2, 5)$ y $(0, 1, 4)$ independientes o dependientes?.La ecuación a resolver es*

$$x_1(1, 3, 2) + x_2(-1, 2, 5) + x_3(0, 1, 4) = (0, 0, 0)$$

o

$$(x_1 - x_2, 3x_1 + 2x_2 + x_3, 2x_1 + 5x_2 + 4x_3) = (0, 0, 0)$$

que da lugar al sistema lineal homogéneo

$$\begin{aligned} x_1 - x_2 &= 0 \\ 3x_1 + 2x_2 + x_3 &= 0 \\ 2x_1 + 5x_2 + 4x_3 &= 0 \end{aligned}$$

cuya matriz de coeficientes es

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

y al reducir

$$
\begin{aligned}
A &\sim \begin{pmatrix} 1 & -1 & 0 \\ 0 & 5 & 1 \\ 0 & 7 & 4 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & \frac{1}{5} \\ 0 & 7 & 4 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 0 & \frac{1}{5} \\ 0 & 1 & \frac{1}{5} \\ 0 & 0 & \frac{13}{5} \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 0 & \frac{1}{5} \\ 0 & 1 & \frac{1}{5} \\ 0 & 0 & 1 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}
$$

por lo que el sistema original es equivalente al sistema

$$
\begin{aligned}
x_1 &= 0 \\
x_2 &= 0 \\
x_3 &= 0
\end{aligned}
$$

cuya solución es la trivial y los vectores son independientes.

Ejemplo 97 *Considere el subconjunto*

$$\beta = \left\{ \begin{pmatrix} -1 & 16 \\ -4 & 13 \end{pmatrix}, \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ -2 & 1 \end{pmatrix} \right\}$$

de $M(2,2)$ *. ¿es* β *linealmente dependiente o linealmente independiente?.Planteamos la ecuación básica que es la de una combinación lineal de vectores de* β *que de cero:*

$$x \begin{pmatrix} -1 & 16 \\ -4 & 13 \end{pmatrix} + y \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} + z \begin{pmatrix} -1 & 2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

y

$$\begin{pmatrix} -x + 2y - z & 16x + 3y + 2z \\ -4x + 3y - 2z & 13x + 4y + z \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

de modo que

$$
\begin{aligned}
-x + 2y - z &= 0 \\
16x + 3y + 2z &= 0 \\
-4x + 3y - 2z &= 0 \\
13x + 4y + z &= 0
\end{aligned}
$$

que es un sistema lineal homogéneo cuya matriz de coeficientes es

$$A = \begin{pmatrix} -1 & 2 & -1 \\ 16 & 3 & 2 \\ -4 & 3 & -2 \\ 13 & 4 & 1 \end{pmatrix}$$

y al reducir

$$\begin{aligned} A &\sim \begin{pmatrix} 1 & -2 & 1 \\ 16 & 3 & 2 \\ -4 & 3 & -2 \\ 13 & 4 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & 35 & -14 \\ 0 & -5 & 2 \\ 0 & 30 & -12 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & 5 & -2 \\ 0 & -5 & 2 \\ 0 & 5 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & 5 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \frac{1}{5} \\ 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

y un sistema equivalente al original es

$$\begin{aligned} x + \frac{z}{5} &= 0 \\ y - \frac{2}{5}z &= 0 \end{aligned}$$

Las variables "ligadas" de este sistema, que son las variables correspondientes a las columnas con los unos delanteros (unos principales, pivotes), o sea x y y. Las variables libres son las restantes, o z en nuestro caso. A partir de aquí la receta es: ponga todo en función de las variables libres.

Por ello un vector en el conjunto solución del sistema será

$$(x, y, z) = (-\frac{z}{5}, \frac{2}{5}z, z) = z(\frac{-1}{5}, \frac{2}{5}, 1)$$

indicando que todos los múltiplos de $(\frac{-1}{5}, \frac{2}{5}, 1)$ dan soluciones. Por ello para cualquier $z \neq 0$ tendremos soluciones no triviales. En conclusión, los vectores de β son linealmente dependientes.

Comentario 98 *De los ejemplos anteriores se puede ver que, al reducir la matriz de coeficientes, si hay al menos una variable "libre" el sistema tendrá soluciones no triviales. Si no hay variables libres y todas son "ligadas" la única solución posible al sistema lineal homogéneo será la trivial.*

6.8. bases y dimensión

Definición 99 *Sea V un espacio vectorial. Una base $\alpha = \{v_1, v_2..v_n\}$ es un conjunto que: 1) genera V. 2) es linealmente independiente.*

Si V es un espacio vectorial y α es una base, entonces dado $v \in V$ hay números c_1, $c_2 ... c_n$ con la propiedad de que

$$v = \sum_{i=1}^{n} c_i v_i$$

pues α genera a V. Estos números son únicos, pues si tuviéramos que

$$v = \sum_{i=1}^{n} d_i v_i$$

entonces (restando ambas ecuaciones)

$$0 = \sum_{i=1}^{n} (c_i - d_i) v_i$$

lo cual implica que $\forall i$

$$c_i = d_i$$

pues el conjunto α es linealmente independiente.

Tenemos el siguiente teorema básico:

Teorema 100 *Sea V un espacio vectorial sobre el campo F. Sea $\alpha = \{v_1, v_2..v_n\}$ una base de V. Sea $\{w_1, w_2..w_m\} \subset V$ y suponga que $m > n$. Entonces $\beta = \{w_1, w_2..w_m\}$ es linealmente dependiente.*

Demostración. Ofrecemos aquí una prueba sencilla usando algunos hechos elementales sobre sistemas de ecuaciones, una demostración alternativa puede hallarse en el libro de Lang [1]

Dado que α es una base, todo elemento de β puede ser expresado como una combinación lineal de vectores de α, consecuentemente

$$\begin{aligned} w_1 &= c_{11} v_1 + c_{12} v_2 + ... c_{1n} v_n \\ w_2 &= c_{21} v_1 + c_{22} v_2 + ... c_{2n} v_n \\ &\vdots \\ w_m &= c_{m1} v_1 + c_{m2} v_2 + ... c_{mn} v_n \end{aligned}$$

y una combinación lineal de los vectores de β que de 0 debe satisfacer

$$0 = \gamma_1 w_1 + \gamma_2 w_2 + ... \gamma_n w_n$$

de modo que

$$\begin{aligned} 0 &= \sum_{i=1}^{m} \gamma_i w_i = \sum_{i=1}^{m} \gamma_i \sum_{j=1}^{n} c_{ij} v_j \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} c_{ij} \gamma_i \right) v_j \end{aligned}$$

pero como α es linealmente independiente

$$\left(\sum_{i=1}^{m} c_{ij} \gamma_i \right) = 0$$

que es un conjunto de n ecuaciones con m incógnitas y $m > n$. Es bien sabido que un sistema lineal homogéneo en el cual el número de incógnitas excede al de ecuaciones siempre tiene una solución no trivial. Por ello podemos encontrar números $\{\gamma_1, \gamma_2 ... \gamma_m\}$ no todos cero, y tales que una combinación lineal de vectores de α con estos números como coeficientes da cero. Luego β es linealmente dependiente. ■

Este teorema tiene una consecuencia importante:

Teorema 101 *Sea V un espacio vectorial y sean α y β bases de V. Entonces α y β tienen el mismo número de elementos.*

Demostración. Sean $\alpha = \{v_1, v_2..v_n\}$ y $\beta = \{w_1, w_2..w_m\}$ dos bases. Si $m > n$ entonces, por el teorema 101 tenemos que β es linealmente dependiente, contradiciendo el que β sea una base. Si $n > m$ entonces, de nuevo, por el teorema 101 tenemos que w α es linealmente dependiente, contradiciendo el que α sea una base. Por lo tanto $n = m$. ■

Definición 102 *Sea V un espacio vectorial. Si V tiene una base con un número finito de elementos, se dice que V es de dimensión finita. De lo contrario decimos que V es de dimensión infinita.*

Definición 103 *En vista del hecho de que cualesquiera dos bases en un espacio de dimensión finita tienen el mismo número de elementos (gracias al teorema 101) podemos hacer la siguiente definición:*

Definición 104 *Sea V un espacio vectorial de dimensión finita.El número n de elementos en cualquier base se llama la dimensión del espacio y se denota mediante* $\dim(V)$.

Comentario 105 *En lo que sigue se supondrá que todos los espacios vectoriales son de dimensión finita.*

Ejemplo 106 *En $\mathbb{R}^n$ el conjunto $\alpha = \{(1,0,0..,0),(0,1,0,..,0),...(0,0,.,1))\}$ es una base, se llama base canónica de $\mathbb{R}^n$. Por ello* $\dim(\mathbb{R}^n) = n$.

Ejemplo 107 *En* $\mathbb{C}^n$ *la base* $\alpha = \{(1,0,0..,0),(0,1,0,..,0),...(0,0,.,1))\}$ *es la base canónica. Por ello* $\dim(\mathbb{C}^n) = n$

Ejemplo 108 *En* P_n *la base canónica es* $\{1, x, x^2...x^n\}$ *y* $\dim(P_n) = n+1$.

Ejemplo 109 *En* $M(m,n)$ *y en* $M^*(m,n)$ *la base canónica está dada por*

$$\left\{ \begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & 0 & ... & 0 \\ \vdots & \vdots & ... & \vdots \\ 0 & 0 & ... & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & ... & 0 \\ 0 & 0 & ... & 0 \\ \vdots & \vdots & ... & \vdots \\ 0 & 0 & ... & 0 \end{pmatrix}, ... \begin{pmatrix} 0 & 0 & ... & 0 \\ 0 & 0 & ... & 0 \\ \vdots & \vdots & ... & \vdots \\ 0 & 0 & ... & 1 \end{pmatrix} \right\}$$

y $\dim(M(m,n)) = \dim(M^*(m,n)) = m \times n$

Ejemplo 110 *¿Es* $A = \{(1,0)\}$ *una base de* $\mathbb{R}^2$*?.*
Más adelante veremos que dado que el conjunto A *tiene un sólo elemento y* $\dim(\mathbb{R}^2) = 2$*,* A *no puede ser una base. Pero analicemos, como ejemplo de los cálculos habituales en este tema, el asunto en detalle. Como* $(1,0) \neq (0,0)$ *el conjunto es linealmente independiente. Pero no genera* $\mathbb{R}^2$ *pues*

$$(x,y) = \alpha(1,0)$$

implica que

$$\begin{aligned} x &= \alpha \\ y &= 0 \end{aligned}$$

y ningún vector de $\mathbb{R}^2$ *con* $y \neq 0$ *puede ser combinación lineal de vectores en* A*. Confirmamos que* A *no es una base de* $\mathbb{R}^2$*.*

Ejemplo 111 *¿Es* $A = \{(1,0),(0,1)\}$ *una base de* $\mathbb{R}^2$*?. Claramente* A *genera* $\mathbb{R}^2$ *pues*

$$(x,y) = (x,0) + (0,y) = x(1,0) + y(0,1)$$

y es independiente pues si

$$x(1,0) + y(0,1) = (0,0)$$

tendremos que

$$(x,y) = (0,0)$$

y $x = 0$ *y* $y = 0$*.* A *es una base de* $\mathbb{R}^2$*.*

Ejemplo 112 *¿Es* $A = \{(1,0),(0,1),(1,1)\}$ *una base de* $\mathbb{R}^2$*?. Más adelante veremos que dado que* $\dim(\mathbb{R}^2) = 2$*,* A *no puede ser una base. Pero, si procedemos directamente, vemos que* A *genera* $\mathbb{R}^2$ *pues*

$$(x,y) = \alpha(1,0) + \beta(0,1) + \gamma(1,1)$$

implica que

$$(x,y) = (\alpha+\gamma, \beta+\gamma)$$

y

$$\begin{aligned} x &= \alpha+\gamma \\ y &= \beta+\gamma \end{aligned}$$

cuya matriz aumentada es

$$\begin{pmatrix} 1 & 0 & 1 & x \\ 0 & 1 & 1 & y \end{pmatrix}$$

que ya está en forma escalonada reducida por renglones. Las columnas pivote son la primera y la segunda, las variables ligadas son, por lo tanto, α *y* β *en tanto que hay sólo una variable libre:* γ*. Por ello el sistema tiene soluciones, de hecho un número infinito de ellas.* A *genera a* $\mathbb{R}^2$*. Sin embargo* A *no es linealmente independiente pues*

$$(0,0) = (\alpha+\gamma, \beta+\gamma)$$

implica

$$\begin{aligned} \alpha+\gamma &= 0 \\ \beta+\gamma &= 0 \end{aligned}$$

que tiene soluciones no triviales (por ejemplo $\gamma = -1$*,* $\alpha = \beta = 1$*.* A *no es base de* $\mathbb{R}^2$*.*

Hemos visto que para que un conjunto sea una base es preciso que sea un generador y que sea linealmente independiente. Sin embargo si conocemos la dimensión del espacio obtenemos criterios más sencillos.

Teorema 113 *Sea* V *un espacio vectorial con* $\dim(V) = n$*. Si* $\alpha = \{v_1, v_2..v_n\}$ *genera a* V *entonces es linealmente independiente y, consecuentemente, es una base de* V*.*

Teorema 114 *Sea* V *un espacio vectorial con* $\dim(V) = n$*. Si* $\alpha = \{v_1, v_2..v_n\}$ *es linealmente independiente entonces genera a* V *y, consecuentemente, es una base de* V*.*

Cuando el espacio es $V = \{0\}$ se dice que es de dimensión cero. No hemos tenido oportunidad de demostrar que si U es subespacio de V entonces $\dim(U) \leq \dim(V)$ pero más adelante usaremos este teorema.

Ejemplo 115 *Demuestre que* $\{(1,2,1),(2,9,0),(3,3,4)\}$ *es una base de* $\mathbb{R}^3$*.*
Primero veamos la independencia, suponga que

$$x(1,2,1) + y(2,9,0) + z(3,3,4) = (0,0,0)$$

por lo que

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

cuya matriz de coeficientes es

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{pmatrix}$$

que tiene como forma escalonada reducida por renglones

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

de modo que un sistema equivalente es

$$\begin{aligned} x &= 0 \\ y &= 0 \\ z &= 0 \end{aligned}$$

probando que los vectores son independientes.
Como tenemos tres vectores independientes en un espacio de dimensión tres, resulta que forman base. Sin embargo, para ilustrar el punto, veamos que los vectores también generan a $\mathbb{R}^3$*. Tome cualquier* $(a, b, c) \in \mathbb{R}^3$ *y preguntémonos si es combinación lineal de* $(1, 2, 1), (2, 9, 0)$ *y* $(3, 3, 4)$

$$x(1, 2, 1) + y(2, 9, 0) + z(3, 3, 4) = (a, b, c)$$

por lo que

$$\begin{aligned} x + 2y + 3z &= a \\ 2x + 9y + 3z &= b \\ x + 4z &= c \end{aligned}$$

de modo que el problema puede ser parafraseado preguntando si este sistema lineal inhomogéneo es consistente para cualquir elección de (a, b, c)*. Hay varias formas de contestar esto, la primera es notar que*

$$det \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{pmatrix} = -1$$

con lo que $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{pmatrix}$ *es una matriz invertible y el sistema inhomogéneo* $Ax = y$ *tiene siempre solución única. Otra opción es ver que, como la forma*

escalonada reducida por renglones de A es la identidad, la matriz es invertible. Más adelante veremos mejores métodos pero ahora hagámoslo por los métodos Gaussianos que ya conocemos:
El sistema

$$\begin{aligned} x + 2y + 3z &= a \\ 2x + 9y + 3z &= b \\ x + 4z &= c \end{aligned}$$

tiene como matriz aumentada

$$\begin{bmatrix} 1 & 2 & 3 & a \\ 2 & 9 & 3 & b \\ 1 & 0 & 4 & c \end{bmatrix}$$

y al reducirla queda

$$\begin{bmatrix} 1 & 0 & 0 & 8b - 36a + 21c \\ 0 & 1 & 0 & 5a - b - 3c \\ 0 & 0 & 1 & 9a - 2b - 5c \end{bmatrix}$$

indicando que el sistema es compatible para todos los valores de a , b y c .

El siguiente ejemplo es más complicado e ilustra la manera de extraer una base para espacios solución de sistemas lineales homogéneos.

Ejemplo 116 *Determine una base y la dimensión del espacio de soluciones del sistema homogéneo*

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

Para resolverlo escribimos la correspondiente matriz de coeficientes

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

que puede ser reducida

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

y las variables ligadas (las que corresponden a las columnas con unos principales) son x_1, x_3 *y* x_4 *por lo que las variables libres son* x_2 *y* x_5. *Un sistema equivalente al original es*

$$\begin{aligned} x_1 + x_2 + x_5 &= 0 \\ x_3 + x_5 &= 0 \\ x_4 &= 0 \end{aligned}$$

y podemos extraer una base escribiendo el vector genérico de $\mathbb{R}^5$ *como* $(x_1, x_2, x_3, x_4, x_5)$

y reexpresándolo en términos de las variables libres como

$$(x_1, x_2, x_3, x_4, x_5) = (-x_2 - x_5, x_2, -x_5, 0, x_5) =$$
$$x_2(-1, 1, 0, 0, 0) + x_5(-1, 0, -1, 0, 1)$$

de donde vemos que un generador para el espacio solución es

$$\{(-1, 1, 0, 0, 0), (-1, 0, -1, 0, 1)\}$$

pero una gracia de la reducción Gaussiana es que este método siempre da vectores independientes (decimos esto sin demostración) por lo que es también una base. La dimensión del conjunto solución es dos.

Comentario 117 *A veces es conveniente pensar en los elementos de una base en un orden dado. Entonces hablamos de una "base ordenada". Así pues* $\alpha = \{v, w\}$ *y* $\beta = \{w, v\}$ *son la misma base pero diferentes bases ordenadas.*

Comentario 118 *que el procedimiento Gaussiano dé siempre vectores independientes requiere, para su demostración, el teorema de la dimensión que se verá más adelante junto con el concepto de rango.*

6.8.1. Algunos resultados adicionales

Hay algunos resultados importantes pero que, por no aparecer en nuestro programa, serán simplemente presentados de manera breve:

- Sea V un espacio vectorial con $\dim(V) = n$. Si $\alpha = \{v_1, v_2, ...v_m\}$ con $m < n$ es un conjunto linealmente independiente, entonces podemos hallar vectores $v_{m+1}, v_{m+2}, ...v_n$ tales que $\beta = \{v_1, v_2, ...v_n\}$ sea una base de V. Esti a veces se enuncia diciendo que un conjunto linealmente independiente puede "completarse a una base".
- Si W es un subespacio de V entonces $\dim(W) \leq \dim(V)$. Aún más, $\dim(W) = \dim(V)$ si y sólo si $V = W$.

6.9. coordenadas

Dado un espacio vectorial V equipado con una base $\alpha = \{v_1, v_2..v_n\}$ ya hemos visto que cualquier vector $v \in V$ puede escribirse de manera única como

$$v = \sum_{i=1}^{n} c_i v_i$$

Los números c_1, $c_2...c_n$ se llaman coordenadas de v con respecto a la base α. Claramente estos números cambian si cambiamos de la base α a otra base β. El vector de $\mathbb{R}^n$ (o $\mathbb{C}^n$)

$$(v)_\alpha = (c_1, c_2...c_n)$$

se llama vector de coordenadas de v con respecto a la base α.

La matriz de $n \times 1$ (que puede ser pensada como un vector de $\mathbb{R}^n$ o de $\mathbb{C}^n$) dada por

$$[v]_\alpha = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

se llama matriz de coordenadas de v con respecto a la base α.

Ejemplo 119 *Las coordenadas de un vector $(x_1, x_2...x_n)$ en $\mathbb{R}^n$ o $\mathbb{C}^n$ con respecto a la base canónica es la misma eneada $(x_1, x_2...x_n)$, es decir, $((x_1, x_2...x_n))_c$ $= (x_1, x_2...x_n)$ donde c denota la base canónica.*

Ejemplo 120 *Las coordenadas de un vector $P(x) = a_0 + a_1x + ...a_nx^n$ en P_n con respecto a la base canónica son $(p(x))_c = (a_0, a_1, ...a_n)$.*

Ejemplo 121 *En $V = P_4$ considere la base $A = \{v_1, v_2, v_3, v_4, v_5\}$ dada por*

$$\begin{aligned} v_1 &= 1 \\ v_2 &= 1 + x \\ v_3 &= 1 + x + x^2 \\ v_4 &= 1 + x + x^2 + x^3 \\ v_5 &= 1 + x + x^2 + x^3 + x^4 \end{aligned}$$

Obtener las coordenadas en la base A del vector $v = 2 + 3x + 2x^2 + 5x^3 - 2x^4$. Hay que encontrar los escalares c_1, c_2, c_3, c_4 y c_5 tales que

$$v = c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 + c_5v_5$$

o sea

$$2 + 3x + 2x^2 + 5x^3 - 2x^4$$
$$= c_1(1) + c_2(1 + x) + c_3(1 + x + x^2) + c_4(1 + x + x^2 + x^3) + c_5(1 + x + x^2 + x^3 + x^4)$$

de modo que

$$\begin{aligned} &2 + 3x + 2x^2 + 5x^3 - 2x^4 \\ &= (c_1 + c_2 + c_3 + c_4 + c_5)(1)+ \\ &(c_2 + c_3 + c_4 + c_5)(x)+ \\ &(c_3 + c_4 + c_5)(x^2)+ \\ &(c_4 + c_5)(x^3)+ \\ &c_5x^4 \end{aligned}$$

o sea

$$
\begin{aligned}
c_1 + c_2 + c_3 + c_4 + c_5 &= 2 \\
c_2 + c_3 + c_4 + c_5 &= 3 \\
c_3 + c_4 + c_5 &= 2 \\
c_4 + c_5 &= 5 \\
c_5 &= -2
\end{aligned}
$$

cuya matriz aumentada es

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

que se reduce como

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

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

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

$$
\sim \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & -1 \\
0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & -3 \\
0 & 0 & 0 & 1 & 0 & 7 \\
0 & 0 & 0 & 0 & 1 & -2
\end{pmatrix}
$$

y el sistema equivalente es

$$
\begin{aligned}
c_1 &= -1 \\
c_2 &= 1 \\
c_3 &= -3 \\
c_4 &= 7 \\
c_5 &= -2
\end{aligned}
$$

Note que el sistema también podría haberse resuelto por sustitución regresiva, pero nosotros seguimos la política de usar Gauss-Jordan.

6.10. cambio de base y matriz de transición

En esta sección consideramos el siguiente escenario: Sea V un espacio vectorial con $\dim(V) = n$, sean $\alpha = \{v_1, v_2..v_n\}$ y $\beta = \{w_1, w_2..w_n\}$ dos bases de V. Sean $v \in V$ y

$$[v]_\alpha = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

$$[v]_\beta = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}$$

las matrices de coordenadas v con respecto a las bases α y β respectivamente. Queremos ver cuál es la conexión entre $[v]_\alpha$ y $[v]_\beta$, esto es, queremos ver cuál es la relación que hay entre las coordenadas de un vector dado con respecto a dos bases diferentes.

Dado que β es una base, dados v_i (con $i = 1...n$) deben existir números P_{ji} tales que

$$v_i = \sum_{j=1}^{n} P_{ji} w_j$$

en otras palabras, P_{ji} es la $j-$ésima coordenada en la base β del $i-$ésimo vector de la base α. Entonces, puesto que

$$v = \sum_{i=1}^{n} c_i v_i$$

tenemos que

$$v = \sum_{i=1}^{n} c_i \left(\sum_{j=1}^{n} P_{ji} w_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} P_{ji} c_i \right) w_j$$

pero como

$$v = \sum_{j=1}^{n} d_j w_j$$

resulta que, por la unicidad de las coordenadas, que

$$d_j = \sum_{i=1}^{n} P_{ji} c_i$$

La matriz P con elementos $[P]_{ji} = P_{ji}$ se llama matriz de transición de la base α a la base β y se denota como M_β^α, las ecuaciones anteriores claramente implican que

$$[v]_\beta = M_\beta^\alpha [v]_\alpha$$

y M_β^α es una matriz tal que da las coordenadas de v con respecto a la base β mediante una sencilla multiplicación de M_β^α por las coordenadas de v con respecto a la base α. Es de importancia práctica el notar que M_β^α es la matriz cuya $j-$ésima columna contiene las coordenadas en la base β del $j-$ésimo vector de la base α.

Igualmente, si queremos pasar de la base β a la base α tendremos que

$$[v]_\alpha = M_\alpha^\beta [v]_\beta$$

de modo que, combinando las últimas ecuaciones,

$$\begin{aligned} [v]_\beta &= M_\beta^\alpha [v]_\alpha \\ &= M_\beta^\alpha M_\alpha^\beta [v]_\beta \end{aligned}$$

y como ésto vale para cualquier $[v]_\beta$ tendremos que

$$M_\beta^\alpha M_\alpha^\beta = I_{n\times n}$$

donde $I_{n\times n}$ es la identidad de $n \times n$.

Esto implica que las matrices de transcición son no-singulares y que

$$M_\beta^\alpha = \left(M_\alpha^\beta\right)^{-1}$$

6.10.1. Encontrando la matriz de transición.

Considere el ejemplo:

En $\mathbb{R}^3$ tenemos las bases

$$\begin{aligned} \alpha &= \{(1,1,1),(0,1,1),(0,0,1)\} \\ \beta &= \{(1,0,1),(1,2,1),(1,2,2)\} \end{aligned}$$

Calcular la matriz M_α^β de transición de la base β a la base α .

La solución “oficial” se obtiene hallando las coordenadas de los vectores de la base β en términos de la base α. Por ello

$$\begin{aligned} (1,0,1) &= m_{11}(1,1,1) + m_{21}(0,1,1) + m_{31}(0,0,1) \\ (1,2,1) &= m_{12}(1,1,1) + m_{22}(0,1,1) + m_{32}(0,0,1) \\ (1,2,2) &= m_{13}(1,1,1) + m_{23}(0,1,1) + m_{33}(0,0,1) \end{aligned}$$

con lo que

$$\begin{aligned} m_{11} &= 1 \\ m_{11} + m_{21} &= 0 \\ m_{11} + m_{21} + m_{31} &= 1 \end{aligned}$$

y

$$\begin{aligned} m_{11} &= 1 \\ m_{21} &= -1 \\ m_{31} &= 1 \end{aligned}$$

Similarmente

$$\begin{aligned} m_{12} &= 1 \\ m_{12} + m_{22} &= 2 \\ m_{12} + m_{22} + m_{32} &= 1 \end{aligned}$$

con lo que

$$\begin{aligned} m_{12} &= 1 \\ m_{22} &= 1 \\ m_{32} &= -1 \end{aligned}$$

y finalmente

$$\begin{aligned} m_{13} &= 1 \\ m_{13} + m_{23} &= 2 \\ m_{13} + m_{23} + m_{33} &= 2 \end{aligned}$$

por lo que

$$\begin{aligned} m_{13} &= 1 \\ m_{23} &= 1 \\ m_{33} &= 0 \end{aligned}$$

y

$$M_\alpha^\beta = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

Ahora vamos a ver cómo resolver el problema de manera económica:

- Formamos matrices A y B que tienen en sus <u>columnas</u> las coordenadas (respecto a la base canónica de $\mathbb{R}^n$ o $\mathbb{C}^n$ según sea el caso) de los vectores de las bases α y β respectivamente.

- Forme una nueva matriz D que se obtiene yuxtaponiendo A y B

$$D = [A, B]$$

- hagamos la reducción Gaussiana de esta matriz (forma escalonada reducida por renglones)

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

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

$$C = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 & 2 \end{bmatrix}$$

  y haciendo la reducción

$$D = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{bmatrix}$$

- La matriz buscada es simplemente el bloque derecho

$$M_{\alpha}^{\beta} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

Lo presenté con un ejemplo pero funciona siempre en espacios $\mathbb{R}^n$ o $\mathbb{C}^n$, que son todos los espacios de dimensión finita, en vista del principio del isomorfismo (que es el tema que sigue).

Para recordar todo el tema conviene usar el siguiente esquema:

1. Sean $\alpha = \{v_1, ...v_n\}$ y $\beta = \{w_1, ...w_n\}$ las bases y v un vector.
2. entonces $[v]_\beta = M_\beta^\alpha [v]_\alpha$
3. si tomamos $v = v_j$ resulta que

$$[v_j]_\alpha = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

   donde el 1 aparece en el lugar j-ésimo.
4. pero multiplicar M por este vector nos da la j–ésima columna de M.
5. luego las columnas de M tienen las coordenadas de los vectores de la base α expresados en términos de la base β .

6.11. Isomorfismos y el principio del isomorfismo

Recuerde de su curso de álgebra que si A y B son conjuntos y

$$f : A \to B$$

es un mapeo entre A y B, A se llama dominio de f y B es el codominio de f.

Decimos que f es inyectiva si $\forall x, y \in A$

$$f(x) = f(y) \Rightarrow x = y$$

o, lo que es lo mismo

$$x \neq y \Rightarrow f(x) \neq f(y)$$

y decimos que f es suprayectiva si $\forall y \in B\ \exists x \in A$ tal que $f(x) = y$.

Definición 122 *Un mapeo $f : A \to B$ es una biyección (o es una función biyectiva) si y sólo si es inyectiva y suprayectiva.*

La importancia de las biyecciones reside en el hecho elemental de que un mapeo tiene un mapeo inverso si y sólo si es una biyección.

Se dice que dos espacios vectoriales son isomórficos (o isomorfos) cuando son básicamente el mismo espacio en lo que a estructura algebraica se refiere. Esto ocurre cuando están relacionados por una biyección que también preserva la estructura algebraica; más rigurosamente, tenemos la definición:

Definición 123 *Decimos que dos espacios V y W son isomórficos (y escribimos $V \sim W$) si hay un mapeo $\phi : V \to W$ tal que: 1) ϕ es una biyección, 2) para cualesquiera u, v en V tenemos que $\phi(u+v) = \phi(u) + \phi(v)$ y 3) para cualquier número α y cualquier vector u tenemos que $\phi(\alpha u) = \alpha\phi(u)$. En este caso el mapeo ϕ se llama "isomorfismo entre V y W".*

Teorema 124 *La relación de isomorfismo $\sim$ entre espacios es una relación de equivalencia.*

Demostración. $a)$ $V \sim V$ puesto que $\phi : V \to V$ y con regla de correspondencia $\phi(x) = x\ \forall x$ (el mapeo identidad) claramente satisface las condiciones para ser un isomorfismo.

$b)$ Si $V \sim W$ entonces hay un isomorfismo $\phi : V \to W$ y el mapeo $\phi^{-1} : W \to V$ es también un isomorfismo (la demostración es mejor posponerla hasta que veamos que la inversa de una función que cumple las propiedades 2 y 3 también las cumple).

$c)$ Si $U \sim V$ y $V \sim W$ entonces hay isomorfismos $\phi : U \to V$ y $\psi : V \to W$ y el mapeo compuesto $\psi \circ \phi : U \to W$ es también un isomorfismo (tambiénesto se verá más adelante). Consecuentemente $U \sim W$. ■

Teorema 125 *Si V es un espacio vectorial n dimensional con una base α entonces el mapeo $\phi : V \to \mathbb{R}^n$ (o $\phi : V \to \mathbb{C}^n$ si V es complejo) y dado por*

$$\phi(v) = [v]_\alpha$$

es un isomorfismo, se llama isomorfismo de coordenadas.

Demostración. *a*) Sea $\alpha = \{v_1, v_2..v_n\}$, entonces para cualesquiera v, $w \in V$ tenemos que

$$v = \sum_{i=1}^{n} c_i v_i$$

$$w = \sum_{i=1}^{n} d_i v_i$$

entonces

$$v + w = \sum_{i=1}^{n} (c_i + d_i) v_i$$

lo cual quiere decir que

$$[v + w]_\alpha = [v]_\alpha + [w]_\alpha$$

b) si γ es un número y $v = \sum_{i=1}^{n} c_i v_i$ entonces

$$\gamma v = \sum_{i=1}^{n} (\gamma c_i) v_i$$

lo cual quiere decir que

$$[\gamma v]_\alpha = \gamma [v]_\alpha$$

c) El mapeo es inyectivo puesto que $[v]_\alpha = [w]_\alpha$ implica que

$$v = \sum_{i=1}^{n} c_i v_i = w$$

d) el mapeo es suprayectivo, puesto que dado $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$ hay un vector $v \in V$ tal que $[v]_\alpha = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$, a saber, $v = \sum_{i=1}^{n} c_i v_i$. ■

Corolario 126 *Cualesquiera dos espacios V y W de la misma dimensión finita son isomórficos.*

Demostración. Sea $n = \dim(V) = \dim(W)$ y sea F $\mathbb{R}^n$ o $\mathbb{C}^n$ dependiendo de si el espacio es real o complejo. Entonces $V \sim F$ y $W \sim F$, pero como $\sim$ es una relación de equivalencia, por la simetría, $V \sim F$ y $F \sim W$ y por la transitividad $V \sim W$. ■

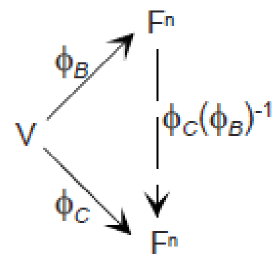


Figura 6.1: Ilustración del cambio de base.

6.11.1. Principio del isomorfismo

Dos espacios isomórficos son "el mismo espacio" en lo que a estructura algebraica se refiere. Por esta razón, cuando resolvemos cualquier problema en un espacio dado, podemos, si queremos, mapear el espacio a otro espacio isomórfico, resolver el problema en el nuevo espacio y regresar al espacio original con la respuesta. Esta aseveración pragmática se conoce como "principio del isomorfismo".

Debe enfatizarse que dos espacios isomórficos son lo mismo sólo en lo que a estructura algebraica (de espacio vectorial) se refiere. Por ejemplo, P_2 y $\mathbb{R}^3$ son isomorfos pero en P_2 podemos diferenciar e integrar los vectores en tanto que en $\mathbb{R}^3$ no (excepto en un sentido trivial). Son similares sólo como espacios vectoriales, por ejemplo en $M(2,2)$ podemos multiplicar los vectores y obtener un anillo, pero en $\mathbb{R}^4$ (que es isomorfo a $M(2,2)$) normalmente nos restringimos a la estructura de espacio vectorial.

Es también interesante notar que un isomorfismo mapea combinaciones lineales en combinaciones lineales (y con los mismos coeficientes), vectores linealmnete independientes en vectores linealmente independientes y vectores linealmente dependientes en vectores linealmente dependientes.

6.11.2. cambio de base de nuevo

Como ya hemos visto, al elegir una base para V (digamos, la base B), automáticamente tenemos un isomorfismo (de coordenadas) $\phi_B : V \to F^n$ (aquí F representa al campo sobre el cual tenemos el espacio V). Similarmente, para otra base C tendremos un isomorfismo $\phi_C : V \to F^n$. El mapeo compuesto $\phi_C \circ \phi_B^{-1} : F^n \to F^n$ es también un isomorfismo entre F^n y él mismo. La situación se representa de manera esquemática en el diagrama siguiente:

Más adelante veremos cuál es el parentesco entre $\phi_C \circ \phi_B^{-1}$ y la matríz de transición.

6.12. Espacios renglón y columna

Considere una matriz real o compleja A de $m \times n$ donde

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

El $i-$ésimo renglón de A es el vector de $\mathbb{R}^n$ o $\mathbb{C}^n$ definido por

$$R_i(A) = (a_{i1}, a_{i2}, ...a_{in})$$

en tanto que la $j-$ésima columna de A es el vector de $\mathbb{R}^m$ o $\mathbb{C}^m$ definido por

$$C_j(A) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

de modo que A puede ser vista de las siguiente tres manera equivalentes:

- Como un arreglo rectangular de $m \times n$ números.
- Como una colección de m vectores en $\mathbb{R}^n$ o $\mathbb{C}^n$.
- Como una colección de n vectores en $\mathbb{R}^m$ o $\mathbb{C}^m$.

En adición introducimos la siguiente definición:

Definición 127 *El espacio renglón de la matriz real o compleja A de $m \times n$ ($Ren(A)$) es el subespacio de $\mathbb{R}^n$ o $\mathbb{C}^n$ generado por los m renglones de A.*

Definición 128 *El espacio columna de la matriz real o compleja A de $m \times n$ ($Col(A)$) es el subespacio de $\mathbb{R}^m$ o $\mathbb{C}^m$ generado por las n columnas de A.*

Comentario 129 *Debe de notarse que $Ren(A) = Col(A^T)$ y que $Ren(A^T) = Col(A)$ pues la operación de transposición intercambia renglones y columnas.*

Los conceptos de espacio renglón y espacio columna tienen una estrecha conexión con las operaciones elementales sobre las matrices y los métodos de Gauss y Gauss-Jordan para resolver sistemas de ecuaciones lineales. recuerde que las operaciones elementales sobre una matriz A son de los siguientes tipos:

- intercambiar cualesquiera dos renglones
- multiplicar cualquier renglón (visto como si fuera un vector de $\mathbb{R}^n$ o $\mathbb{C}^n$) por cualquier número α.que no sea cero

- agregar a cualquier renglón cualquier múltiplo de otro renglon.

Tenemos un teorema importante:

Teorema 130 *Las operaciones elementales sobre una matriz A no cambian su espacio renglón $Ren(A)$.*

Demostración. a)Al intercambiar dos renglones tenemos el mismo generador.

b) Al multiplicar un vector (digamos que el v_j) por $\alpha \neq 0$ vemos que toda combinación lineal $v = \sum_{i=1}^{n} c_i v_i$ se puede escribir con un término $(\frac{c_j}{\alpha})(\alpha v)$

c) La dejamos al lector. ■

Sorprendentemente, las dimensiones de los espacios renglón y columna de una matriz siempre son iguales:

Teorema 131 *Sea A una matriz cualquiera, real o compleja. La dimensión de su espacio renglón y de su espacio columna son iguales. Este número común se llama rango de la matriz.*

Ejemplo 132 *Hallar una base y la dimensión del espacio renglón de*

$$A = \begin{pmatrix} 1 & 3 & 4 & 2 \\ 2 & 1 & 3 & -1 \\ 3 & 4 & 7 & 1 \end{pmatrix}$$

Como

$$\begin{aligned} A &\sim \begin{pmatrix} 1 & 3 & 4 & 2 \\ 0 & -5 & -5 & -5 \\ 3 & 4 & 7 & 1 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 3 & 4 & 2 \\ 0 & -5 & -5 & -5 \\ 0 & -5 & -5 & -5 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 3 & 4 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 3 & 4 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

tenemos que el espacio renglón de A también es generado por

$$\{(1,0,1,-1),(0,1,1,1)\}$$

(el vector cero puede omitirse siempre de una lista de generadores). Además, a causa del esclonamiento los vectores no cero que quedan en la forma escalonada reducida por renglones son independientes, pues

$$\alpha(1,0,1,-1)+\beta(0,1,1,1)=(0,0,0,0)$$

implica que

$$(\alpha,\beta,\alpha+\beta,-\alpha+\beta)=(0,0,0,0)$$

y $\alpha=\beta=0$ *con lo que concluimos que los vectores son independientes. Luego la base pedida es* $\{(1,0,1,-1),(0,1,1,1)\}$ *y el espacio es de dimensión dos.*

Comentario 133 *Los vectores renglón no cero de una forma escalonada reducida por renglones son, siempre, linealmente independientes gracias al escalamiento. Lo mismo ocurre con los vectores no cero de cualquier forma escalonada.*

Ejemplo 134 *Determine una base para el subespacio* V *de* $\mathbb{R}^5$ *generado por*

$$\begin{aligned}v_1&=(1,2,-1,4,3)\\v_2&=(3,1,2,2,4)\\v_3&=(2,1,1,-1,0)\\v_4&=(2,-4,6,-11,-5)\end{aligned}$$

Como V *es también el espacio renglón de*

$$A=\begin{pmatrix}1&2&-1&4&3\\3&1&2&2&4\\2&1&1&-1&0\\2&-4&6&-11&-5\end{pmatrix}$$

y

$$
\begin{aligned}
A &\sim \begin{pmatrix} 1 & 2 & -1 & 4 & 3 \\ 0 & -5 & 5 & -10 & -5 \\ 0 & -3 & 3 & -9 & -6 \\ 0 & -8 & 8 & -19 & -11 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 2 & -1 & 4 & 3 \\ 0 & 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 3 & 2 \\ 0 & -8 & 8 & -19 & -11 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -3 & -3 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}
\end{aligned}
$$

de modo que una base de V es

$$\{(1,0,1,0,1),(0,1,-1,0,-1),(0,0,0,1,1)\}$$

y $\dim(V) = 3$.

Comentario 135 *Por todo lo anterior podemos decir que al ir haciendo la reducción de una matriz vamos cambiando de generador (en cada paso intermedio los renglones siguen siendo generadores del mismo espacio renglón). Al final acabamos con una base, la base puede tener menos elementos que el conjunto generador inicial, esto se refleja en el hecho de que la forma escalonada tiene algunos renglones que constan exclusivamente de ceros.*

Comentario 136 *Nosotros hemos seguido la política de hacer reducciones a forma escalonada reducida por renglones pero lo anteriormente dicho sobre espacios renglón y sus bases también se aplica a las formas escalonadas. Los renglones no cero de cualquier forma escalonada (reducida o no) forman una base del espacio renglón.*

Considere el espacio $M_{2\times 2}(\mathbb{R})$ de las matrices cuadradas de 2×2 con entradas reales y el subespacio W generado por

$$S = \left\{ \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 7 & 4 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 4 & 3 \\ 1 & 2 \end{pmatrix} \right\}$$

proporcione una base para W.

Solución 137 *Seleccionamos una base de* $M_{2\times 2}(\mathbb{R})$*, a saber, la base canónica*

$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

Usamos el isomorfismo de coordenadas $\phi : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}^4$ definido mediante

$$\phi(m) = [m]_\beta$$

Bajo el isomorfismo, la imagen de S es

$$S' = \{(1, 2, 3, 4), (7, 4, -1, 0), (4, 3, 1, 2)\}$$

Por ello el espacio generado por S' es el espacio renglón de la matriz

$$M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 7 & 4 & -1 & 0 \\ 4 & 3 & 1 & 2 \end{bmatrix}$$

Usando operaciones elementales reducimos M para obtener
$\begin{bmatrix} 1 & 0 & -\frac{7}{5} & -\frac{8}{5} \\ 0 & 1 & \frac{11}{5} & \frac{14}{5} \\ 0 & 0 & 0 & 0 \end{bmatrix}$
Por ello el espacio generado por S' tiene la base

$$\gamma = \left\{ (1, 0, \frac{-7}{5}, \frac{-8}{5}), (0, 1, \frac{11}{5}, \frac{14}{5}) \right\}$$

Usando el isomorfismo inverso resulta que una base de W es

$$\left\{ \begin{pmatrix} 1 & 0 \\ \frac{-7}{5} & \frac{-8}{5} \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ \frac{11}{5} & \frac{14}{5} \end{pmatrix} \right\}$$

Comentario 138 *Ya hemos visto que la dimensión del espacio columna de una matriz es igual a la dimensión del espacio renglón y que éste número es el rango de la matriz. Por otra parte, al hacer la reducción Gaussiana de la matriz, vemos que el rango es igual al número de pivotes (y al número de renglones no-cero en su forma reducida).*

6.13. El espacio nulo de una matriz

Sea A una matriz (real o compleja) de $m \times n$. El espacio nulo de A es el espacio solución del sistema lineal homogéneo correspondiente, es decir

$$null(A) = \{x \in F^n \mid Ax = 0\}$$

La dimensión del espacio nulo está dada por el número de variables libres en el correspondiente sistema. Más adelante veremos que siempre $n = rango(A) + \dim(null(A))$.

6.14. La estructura de los sistemas inhomogéneos.

Considere el sistema lineal inhomogéneo

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

donde no todos los números b_i son cero, y re-escribámoslo usando las matrices

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

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

como

$$AX = b$$

El conjunto solución de este sistema es el conjunto

$$S = \{X \in F^n \mid AX = b\}$$

donde F es $\mathbb{R}^n$ o $\mathbb{C}^n$ según sea el caso.

Teorema 139 *S no es un subespacio de F^n.*

Demostración. a) claramente $S \subset F^n$

b)es bien sabido que algunos sistemas de ecuaciones lineales inhomogéneos no tienen soluciones (sistemas inconsistentes). Consecuentemente no siempre es el caso que $S \neq \phi$.

c) Sean $x, y \in S$, entonces $Ax = b$ y $Ay = b$ de modo que $A(x + y) = Ax + Ay = b + b = 2b$ y no puede ocurrir que $A(x + y) = b$ a menos que $b = 0$, en cuyo caso el sistema era homogéneo. Concluimos que S no es cerrado bajo adición de vectores.

d) Sean $x \in S$ y γ un número. Entonces $Ax = b$ por lo que $A(\gamma x) = \gamma Ax = \gamma b$ y no puede ocurrir que $A(\gamma x) = b$ a menos que $b = \gamma b$, esto es, a menos que $b = 0$ (en cuyo caso el sistema era homogéneo) o $\gamma = 1$. En este último caso S no es cerrado bajo multiplicación por escalares. ■

6.15. Espacios de funciones, Wronskianos

Considere el conjunto de todas las funciones que tienen como dominio el intervalo cerrado $[a, b] \subset \mathbb{R}$ y codominio $\mathbb{R}$

$$\mathcal{F} = \{g : [a, b] \to \mathbb{R}\}$$

Es posible definir una suma en $\mathcal{F}$. Para $f, g \in \mathcal{F}$ defina $(f + g) : [a, b] \to \mathbb{R}$ por medio de

$$(f + g)(x) = f(x) + g(x)$$

para toda $x \in [a, b]$.

También se puede definir una multiplicación de números reales por elementos de $\mathcal{F}$. Para $f \in \mathcal{F}$ y $\alpha \in \mathbb{R}$ defina $(\alpha f) : [a, b] \to \mathbb{R}$ mediante

$$(\alpha f)(x) = \alpha(f(x))$$

para toda $x \in [a, b]$.

Resulta que, $\mathcal{F}$ con estas operaciones, es un espacio vectorial real.

Teorema 140 *$\mathcal{F}$, con las operaciones recién definidas, es un espacio vectorial real.*

Demostración. *1)por definición, si $f, g \in \mathcal{F}$ entonces $f + g \in \mathcal{F}$*

2)puesto que $(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$ se sigue que $f + g = g + f$ para cualesquiera $f, g \in \mathcal{F}$.

3) dado que $(f + (g + h))(x) = f(x) + (g + h)(x) = f(x) + g(x) + h(x) = (f + g)(x) + h(x) = ((f + g) + h)(x)$ tenemos que $(f + (g + h)) = ((f + g) + h)$ para cualesquiera $f, g, h \in \mathcal{F}$.

4) defina $0 : [a, b] \to \mathbb{R}$ mediante $0(x) = 0$. Entonces, claramente, $f + 0 = 0 + f = f$ para toda $f \in \mathcal{F}$.

5) dada $f \in \mathcal{F}$ defina $(-f) : [a, b] \to \mathbb{R}$ mediante $(-f)(x) = -f(x)$. Entonces es fácil ver que $f + (-f) = (-f) + f = 0$ para toda $f \in \mathcal{F}$.

6) por definición, si $f \in \mathcal{F}$ *y* $\alpha \in \mathbb{R}$ *entonces* $\alpha f \in \mathcal{F}$
7) sean $\alpha, \beta \in \mathbb{R}$, *y* $f \in \mathcal{F}$. *Como* $((\alpha+\beta)f)(x) = (\alpha+\beta)f(x) = \alpha f(x)+\beta f(x) = (\alpha f + \beta f)(x)$ *se desprende que* $(\alpha + \beta)f = \alpha f + \beta f$ *para cualesquiera* $\alpha, \beta \in \mathbb{R}$ *y* $f \in \mathcal{F}$.
8)Como

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

, se desprende que $\alpha(f+g) = \alpha f + \alpha g$ *para toda* $\alpha \in \mathbb{R}$ *y cualesquiera* $f, g \in \mathcal{F}$.
9)Puesto que $(\alpha(\beta f))(x) = \alpha((\beta f)(x)) = \alpha(\beta f(x)) = (\alpha\beta)f(x) = ((\alpha\beta)f)(x)$ *tenemos que* $\alpha(\beta f) = (\alpha\beta)f$ *para cualesquiera* $\alpha, \beta \in \mathbb{R}$ *y* $f \in \mathcal{F}$.
10)Como $(1f)(x) = 1(f(x)) = f(x)$ *se sigue que* $1f = f$ *para toda* $f \in \mathcal{F}$. ■

Algunos de los subespacios importantes de $\mathcal{F}$ son los subespacios de todas las funciones de $[a, b] \to \mathbb{R}$ que son contínuas, diferenciables, integrables, etc.

El espacio $\mathcal{F}$ no es de dimensión finita, esto se puede ver considerando la secuencia infinita de funciones linealmente independientes $\{1, x, x^2, ...\}$.

6.15.1. dependencia e independencia lineal de funciones.

Considere el espacio $\mathcal{F}$ o cualquiera de sus subespacios y $\alpha = \{f_1, f_2..f_n\} \subset \mathcal{F}$ .El conjunto α es linealmente dependiente si y sólo si hay números $\alpha_1, \alpha_2..\alpha_n$, no todos cero, tales que

$$\alpha_1 f_1 + \alpha_2 f_2 + ...\alpha_n f_n = 0$$

pero esta ecuación implica que para toda $x \in [a, b]$

$$\alpha_1 f_1(x) + \alpha_2 f_2(x) + ...\alpha_n f_n(x) = 0$$

dando, de facto, un número infinito de condiciones sobre los coeficientes.

método 1 (por inspección)

Este método se puede explicar mejor con un ejemplo.

Ejemplo 141 *Determine si el conjunto* $\{\cos^2(x), \sin^2(x), 5\}$ *es linealmente dependiente o linealmente independiente.*
Como $sen^2(x)+\cos^2(x) = 1$ *tenemos que* $5sen^2(x)+5\cos^2(x) = 5$ *o* $5sen^2(x)+5\cos^2(x) - 5 = 0$ *probando que el conjunto es dependiente.*

El método depende de la posibilidad de conocer una relación entre las funciones. Por ello las siguientes relaciones son a menudo útiles:

- $sen^2(x) + \cos^2(x) = 1$

- $sen(x+y) = sen(x)\cos(y) + \cos(x)sen(y)$
- $cos(x+y) = \cos(x)\cos(y) - sen(x)sen(y)$
- $sen(2x) = 2sen(x)\cos(x)$
- $\cos(2x) = \cos^2(x) - sen^2(x) = 1 - 2sen^2(x) = 2\cos^2(x) - 1$
- $\log(xy) = \log(x) + \log(y)$
- $\exp(x+y) = \exp(x) \times \exp(y)$

etc. Casi cualquier identidad trigonométrica, exponencial o logarítmica puede ser útil.

Note que este método puede servirnos para descubrir que el conjunto es dependiente pero que no puede ser usado para establecer que el conjunto es independiente.

método 2 (escoja valores para las funciones).

De nuevo, esto se explica mejor con un ejemplo.

Ejemplo 142 *Encuentre si el conjunto* $\{x\ sen(x), \cos(x)\}$ *es linealmente dependiente o linealmente independiente.*
Tome una combinación lineal $\alpha\ x\ sen(x) + \beta\cos(x) = 0$ *. Como esto debe de valer para cualquier número real* x *se vale en particular para* $x = 0$ *por lo que* $\beta = 0$ *. Entonces* $\alpha\ x\ sen(x) = 0$ *y tomando esta vez* $x = \frac{\pi}{2}$ *tenemos que* $\frac{\alpha\pi}{2} = 0$ *y* $\alpha = 0$. *Consecuentemente, el conjunto es linealmente independiente.*

Este método consiste en escojer valores arbitrarios (¡pero escogidos inteligentemente!) para la variable independiente de modo que podamos concluir que todos los coeficientes de la combinación lineal son cero. Note que este método puede ser usado para descubrir que el conjunto es linealmente independiente pero no puede ser usado para demostrar que es linealmente dependiente. Este método, al igual que el precedente, no siempre funciona.

Vale la pena ver en algún detalle porqué el método no siempre sirve. Suponga que el conjunto es

$$S = \{f_1(x), f_2(x), ...f_n(x)\}$$

y que usamos los valores de x dados por $x_1, x_2...x_m$.

Si las ecuaciones de dependencia

$$\begin{aligned}
\alpha_1 f_1(x_1) + \alpha_2 f_2(x_1) + ...\alpha_n f_n(x_1) &= 0 \\
\alpha_1 f_1(x_2) + \alpha_2 f_2(x_2) + ...\alpha_n f_n(x_2) &= 0 \\
&\vdots \\
\alpha_1 f_1(x_m) + \alpha_2 f_2(x_m) + ...\alpha_n f_n(x_m) &= 0
\end{aligned}$$

implican que todas las α_i son cero, entonces las funciones son independientes. Aunque otra elección de las x_j tuviera soluciones no triviales, como las ecuaciones deben de valer para toda x, la conclusión no varía, son independientes.

Pero si con la elección original hay soluciones no triviales, eso no garantiza la dependencia pues existe la posibilidad de que otra elección de las x_j implique que la única solución sea la trivial.

Así que jamás use este método para probar dependencia.

Si $m = n$ entonces tendremos tantas incógnitas como ecuaciones, en este caso podemos usar un teorema de álgebra elemental que dice que dicho sistema tiene solución no trivial si y solo si el determinante de la matriz de coeficientes se anula. Por ello si el determinante no vale cero, podemos concluir que solo tendrá la solución trivial.

el método del Wronskiano.

Hay un tercer método de gran importancia en el contexto de las ecuaciones diferenciales.

Suponga que $\alpha = \{f_1, f_2..f_n\} \subset \mathcal{F}$ consiste de funciones que son diferenciables al menos $n-1$ veces en $[a, b]$. Entonces, tomando derivadas sucesivamente de la combinación lineal $\alpha_1 f_1 + \alpha_2 f_2 + ...\alpha_n f_n = 0$ tendremos

$$\begin{aligned}
\alpha_1 f_1 + \alpha_2 f_2 + ...\alpha_n f_n &= 0 \\
\alpha_1 \frac{df_1}{dx} + \alpha_2 \frac{df_2}{dx} + ...\alpha_n \frac{df_n}{dx} &= 0 \\
\alpha_1 \frac{d^2 f_1}{dx^2} + \alpha_2 \frac{d^2 f_2}{dx^2} + ...\alpha_n \frac{d^2 f_n}{dx^2} &= 0 \\
&\vdots \\
\alpha_1 \frac{d^{n-1} f_1}{dx^{n-1}} + \alpha_2 \frac{d^{n-1} f_2}{dx^{n-1}} + ...\alpha_n \frac{d^{n-1} f_n}{dx^{n-1}} &= 0
\end{aligned}$$

dando un sistema lineal homogéneo de n ecuaciones en n incógnitas $\alpha_1, \alpha_2...\alpha_n$. Pero semejante sistema tiene una solución no trivial si y sólo si el determinante de su matriz de coeficientes es cero, por ello tenemos que si para alguna $x \in [a, b]$

$$\det \begin{pmatrix}
f_1 & f_2 & ... & f_n \\
\frac{df_1}{dx} & \frac{df_2}{dx} & ... & \frac{df_n}{dx} \\
\frac{d^2 f_1}{dx^2} & \frac{d^2 f_2}{dx^2} & ... & \frac{d^2 f_n}{dx^2} \\
\vdots & \vdots & ... & \vdots \\
\frac{d^{n-1} f_1}{dx^{n-1}} & \frac{d^{n-1} f_2}{dx^{n-1}} & ... & \frac{d^{n-1} f_n}{dx^{n-1}}
\end{pmatrix} \neq 0$$

entonces la única solución posible es la trivial y las funciones son linealmente independientes. Esto es equivalente a decir que si las funciones son dependientes entonces el determinante es cero para todas las x en el dominio de las funciones.

El Wronskiano de las funciones, denotado por $W(f_1(x), f_2(x)...f_n(x))$ se define como

$$W(f_1(x), f_2(x)...f_n(x)) = \det \begin{pmatrix} f_1 & f_2 & ... & f_n \\ \frac{df_1}{dx} & \frac{df_2}{dx} & ... & \frac{df_n}{dx} \\ \frac{d^2f_1}{dx^2} & \frac{d^2f_2}{dx^2} & ... & \frac{d^2f_n}{dx^2} \\ \vdots & \vdots & ... & \vdots \\ \frac{d^{n-1}f_1}{dx^{n-1}} & \frac{d^{n-1}f_2}{dx^{n-1}} & ... & \frac{d^{n-1}f_n}{dx^{n-1}} \end{pmatrix}$$

y si para alguna $x \in [a,b]$ $W(f_1(x), f_2(x)...f_n(x) \neq 0$ entonces las funciones son linealmente independientes (si las funciones son linealmente dependientes entonces $W(f_1(x), f_2(x)...f_n(x) = 0 \; \forall x)$.

El converso no es, en general, verdadero. Si $W(f_1(x), f_2(x)...f_n(x) = 0 \; \forall x$ las ecuaciones ciertamente tienen soluciones no triviales pero éstas podrían ser diferentes para diferentes valores de x como se ilustra en el siguiente ejemplo.

Ejemplo 143 *Considere las funciones $f : (-\infty, \infty) \to \mathbb{R}$ y $g : (-\infty, \infty) \to \mathbb{R}$ definidas mediante*

$$f(x) = x^2$$
$$g(x) = \mid x \mid x$$

Entonces

$$W(f(x), g(x)) = \det \begin{pmatrix} x^2 & \mid x \mid x \\ 2x & 2 \mid x \mid \end{pmatrix}$$
$$2 \mid x \mid x^2 - 2 \mid x \mid x^2 = 0$$

pero si las funciones satisficieran $x_1 f(x) + x_2 g(x) = 0$, entonces usando nuestro método dos con $x = -1$ y $x = 1$ tendríamos

$$x_1 - x_2 = 0$$
$$x_1 + x_2 = 0$$

de modo que $x_1 = x_2 = 0$ probando que las funciones f y g son linealmente independientes. De hecho $f = g$ en $[0, \infty)$ y $f = -g$ in $(-\infty, 0]$.

Sin embargo si hay otras restriccciones impuestas sobre f_1 y f_2 (como la de ser soluciones de ciertas ecuaciones diferenciales), la condición del Wronskiano puede ser tanto necesaria como suficiente.

Como ejemplo sencillo, suponga que para las funciones f y g

$$W(f(x), g(x)) = \det \begin{pmatrix} f(x) & g(x) \\ \frac{df}{dx} & \frac{dg}{dx} \end{pmatrix}$$
$$= f(x)\frac{dg}{dx} - g(x)\frac{df}{dx} = 0$$

y que $f(x) \neq 0$ para toda x en su dominio, entonces

$$\frac{d\frac{g(x)}{f(x)}}{dx} = \frac{f(x)\frac{dg(x)}{dx} - g(x)\frac{df(x)}{dx}}{f(x)^2} = \frac{0}{f(x)^2} = 0$$

e integrando

$$\frac{g(x)}{f(x)} = A$$

para alguna constante real A. Entonces $1g(x) - Af(x) = 0$ y las funciones son dependientes. En el ejemplo de más arriba $f(0) = g(0) = 0$ y esta demostración no se aplica.

Comentario 144 *Si tenemos una función $f : A \to B$ y $C \subset A$ entonces podemos definir una función $g : C \to B$ mediante la regla de correspondencia $g(x) = f(x)$ para toda $x \in C$. La función g recibe el nombre de restricción de f al conjunto C y suele denotarse como $g = f\mid_C$. A veces los libros dicen cosas como "las funciones son independientes en $(-10, 10)$ pero dependientes en $(0, 10)$", e incluso barbaridades como "independencia o dependencia en un punto". Lo correcto podría ser "las funciones f y h son independientes pero sus restricciones $f\mid_C$ y $g\mid_C$son dependientes".*

Ejemplo 145 *Considere el conjunto $\{sen(x), \cos(x)\}$. El Wronskiano será*

$$\begin{aligned} w(sen(x), \cos(x)) &= \det \begin{bmatrix} sen(x) & \cos(x) \\ \cos(x) & -sen(x) \end{bmatrix} \\ &= -sen^2(x) - \cos^2(x) \\ &= -1 \neq 0 \end{aligned}$$

Por lo tanto las funciones son independientes.

Parte III

Transformaciones lineales

Cuando uno ya ha estudiado una estructura (como la de espacios vectoriales) incluyendo sus subestructuras (como los subespacios) ya es hora de estudiar los "morfismos" que en este caso son las transformaciones lineales.

En buena medida el álgebra lineal estudia los espacios vectoriales y las transformaciones lineales entre espacios, tema al que nos dedicamos en esta parte.

La importancia de las transformaciones lineales radica en que "preservan" estructura, en el sentido que más adelante se verá.

Capítulo 7

Transformaciones lineales

Sean V y W espacios vectoriales sobre un mismo campo F. Una transformación de V a W es cualquier función

$$T : V \to W$$

(lo cual también suele escribirse como

$$V \underset{\to}{T} W$$

)

Las transformaciones son, pues, funciones entre espacios vectoriales.

Cuando $V = W$ se dice que T es un operador.

Comentario 146 *Algunos sinónimos de "función":mapeo, aplicación (este último vocablo es usado por los autores españoles). Toda transformación es una función pero no toda función es una transformación.*

Sean V y W espacios vectoriales. Una transformación lineal de V a W es una transformación $T : V \to W$ que satisface las siguientes dos propiedades:

- Para todo v, w en V tenemos que $T(v + w) = T(v) + T(w)$
- Para todo v en V y todo α en F tenemos que $T(\alpha v) = \alpha T(v)$

Cuando $V = W$ se suele llamar a T "operador lineal en V".

Comentario 147 *He aquí un ejemplo de transformación que satisface la primera condición pero no la segunda:*

$$T : \mathbb{C} \to \mathbb{C}$$
$$T(z) = \bar{z}$$

pues

$$T(a + b) = \overline{(a + b)} = \overline{a} + \overline{b} = T(a) + T(b)$$
$$T(\beta b) = \overline{\beta b} = \overline{\beta}\ \overline{b} = \overline{\beta} T(b) \neq \beta T(b)$$

Comentario 148 *Un ejemplo de transformación que cumple la segunda condición mas no la primera sería:*

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$
$$T(x,y) = \begin{cases} (x,y) \text{ si } x = 0 \text{ o } y = 0 \\ (0,0) \text{ en los restantes casos} \end{cases}$$

Comentario 149 *Un isomorfismo es una transformación lineal biyectiva.*

Comentario 150 *Se dice que las transformaciones lineales "preservan estructura" porque mapean una suma en otra suma y un producto por números en otro producto por números.*

Comentario 151 *Ya hemos mencionado que en matemáticas cuando uno estudia estructuras algebraicas considera luego las subestructuras. También ocurre que uno indaga los mapeos que preservan estructura, a ellos se les llama, en general, morfismos. Si el morfismo tiene el mismo dominio y codominio puede ser llamado operador lineal o endomorfismo.*

Comentario 152 *Un endomorfismo que también es isomorfismo se llama automorfismo.*

7.0.2. algunas propiedades de las transformaciones lineales.

Teorema 153 *Sean V y W espacios vectoriales y $T : V \to W$ una transformación lineal. Entonces:*

- $T(\hat{0}) = \hat{0}$ (aquí $\hat{0}$ es el vector cero para distinguirlo del escalar cero)
- $T(-a) = -T(a)$
- $T(a-b) = T(a) - T(b)$
- $T(\sum_{i=1}^{n} c_i v_i) = \sum_{i=1}^{n} c_i T(v_i)$ donde los $c_i \in F$ y $v_i \in V$ $(i = 1, 2...n)$

Pruebas:

- $T(\hat{0}) = T(0\hat{0}) = 0T(\hat{0}) = \hat{0}$
- $T(-a) = T((-1)a) = (-1)T(a) = -T(a)$
- $T(a-b) = T(a+(-1)b) = T(a) + T((-1)b) = T(a) + (-1)T(b) = T(a) - T(b)$
- (demostración por inducción) si $n = 1$ el teorema es cierto pues, en este caso $T(\sum_{i=1}^{1} c_i v_i) = T(c_1 v_1) = c_1 T(v_1) = \sum_{i=1}^{1} c_i T(v_i)$.
  Suponga ahora que el teorema es cierto para $n = k$, entonces $T(\sum_{i=1}^{k+1} c_i v_i) = T(\sum_{i=1}^{k} c_i v_i + c_{k+1} v_{k+1}) = T(\sum_{i=1}^{k} c_i v_i) + T(c_{k+1} v_{k+1}) = \sum_{i=1}^{k} c_i T(v_i) + c_{k+1} T(v_{k+1})$
  $= \sum_{i=1}^{k+1} c_i T(v_i)$ completando el paso inductivo.

7.1. algunas transformaciones importantes

La identidad $I_V : V \to V$ en V se define como

$$I_V v = v \ \forall v \in V$$

y la transformación cero $\hat{0} : V \to W$ se define mediante

$$\hat{0}v = 0 \ \forall v \in V$$

Como

$$\begin{aligned} I_V(a+b) &= a+b = I_V(a) + I_V(b) \\ I_V(\beta a) &= \beta a = \beta(I_V(a)) \end{aligned}$$

y

$$\begin{aligned} \hat{0}(a+b) &= 0 = \hat{0}a + \hat{0}b \\ \hat{0}(\beta a) &= 0 = \beta 0 = \beta\hat{0}(a) \end{aligned}$$

tenemos que tanto $\hat{0}$ como I_V son transformaciones lineales.

Hay otro tipo de transformaciones cuya importancia será evidente más tarde. Considere los espacios $\mathbb{R}^n$ y $\mathbb{R}^m$ (o $\mathbb{C}^n$ y $\mathbb{C}^m$) y una matriz real (o compleja) de $m \times n$. Definamos la transformación

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

(o $T : \mathbb{C}^n \to \mathbb{C}^m$ según el caso) mediante

$$T(x) = Ax$$

en donde el elemento $x \in \mathbb{R}^n$ se visualiza como un vector columna, es decir, como matriz de $n \times 1$ de modo que Ax es el producto de dos matrices, una de $m \times n$ y otra de $n \times 1$ dando por rersultado una matriz de $m \times 1$ (o sea, un vector de $\mathbb{R}^m$). Esta transformación, llamada "matriz por vector", es lineal pues por las propiedades de la multiplicación matricial tenemos que

$$\begin{aligned} T(x+y) &= A(x+y) = Ax + Ay = T(x) + T(y) \\ T(\beta x) &= A(\beta x) = \beta Ax = \beta T(x) \end{aligned}$$

7.1.1. ejemplos de transformaciones lineales.

Ejemplo 154 *Considere $T : \mathbb{R}^3 \to \mathbb{R}^2$ definida mediante*

$$T(x,y,z) = (2x+3y+4z, x+2y+z)$$

¿es T una transformación lineal?
a) Considere $a = (x,y,z)$ y $b = (x',y',z')$. Entonces $a+b = (x+x', y+y', z+z')$

y

$$
\begin{aligned}
T(a) &= (2x+3y+4z, x+2y+z) \\
T(b) &= (2x'+3y'+4z', x'+2y'+z') \\
&T(a)+T(b) \\
&= (2x+3y+4z+2x'+3y'+4z', x+2y+z+x'+2y'+z') \\
&T(a+b) \\
&= (2(x+x')+3(y+y')+4(z+z'), (x+x')+2(y+y')+(z+z')) \\
&= (2x+3y+4z+2x'+3y'+4z', x+2y+z+x'+2y'+z')
\end{aligned}
$$

con lo que $T(a+b) = T(a)+T(b)$
b)Sea $\alpha \in \mathbb{R}$ *. Entonces si* $a = (x,y,z)$ *tendremos que* $\alpha a = (\alpha x, \alpha y, \alpha z)$ *y*

$$
\begin{aligned}
T(\alpha a) &= (2\alpha x + 3\alpha y + 4\alpha z, \alpha x + 2\alpha y + \alpha z) \\
&= \alpha((2x+3y+4z, x+2y+z)) \\
&= \alpha T(a)
\end{aligned}
$$

con lo que $T(\alpha a) = \alpha T(a)$ *y la transformación es lineal.*

Ejemplo 155 *Del ejemplo anterior, si* $T : \mathbb{R}^3 \to \mathbb{R}^2$ *está dada por* $T(x,y,z) = (2x+3y-4z, x+2y+z)$ *podemos decir que, como*

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

la transformación lineal T *es de la forma* Ax *con* $A = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 1 \end{pmatrix}$.

Comentario 156 *Espero que el lector no se sienta confundido por el hecho de que en algún momento un vector* v *de (digamos)* $\mathbb{R}^3$ *se escriba como* $v = (x,y,z)$ *o como* $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ *o incluso* $v = (x\ y\ z)$*. Lo que define al vector de* $\mathbb{R}^3$ *es el tener exactamente tres componentes y ser éstas números del campo* F*. Los signos de agrupación, los separadores o la disposición de los elementos (vertical, horizontal etc.) son matemáticamente irrelevantes pero, según el contexto, una notación puede ser más conveniente que otra.*

Ejemplo 157 *En* P_n *considere la transformación* $T : P_n \to P_n$ *definida mediante*

$$T(P(x)) = \frac{dP(x)}{dx}$$

Como
a) $T(P(x)+Q(x)) = \frac{d(P(x)+Q(x))}{dx} = \frac{dP(x)}{dx} + \frac{dQ(x)}{dx} = T(P(x)) + T(Q(x))$ *y*
b) $T(\beta P(x)) = \frac{d\beta P(x)}{dx} = \beta \frac{dP(x)}{dx} = \beta T(P(x))$ *tenemos que* T *es lineal. Hemos*

usado varias propiedades comunes de la derivada, tales como el hecho de que la derivada de una suma de funciones es la suma de las derivadas y que la derivada del producto de un número por una función es el número por la derivada de la función.

Ejercicio 158 *Considere $T : P_n \to \mathbb{R}$ definida mediante*

$$T(P(x)) = \int_0^1 P(x)dx$$

¿es T lineal?

Ya hemos visto que una transformación lineal mapea una combinación lineal en otra.

Teorema 159 *Sean V y W espacios vectoriales y $T : V \to W$ una transformación lineal. Sea $\{v_1, v_2...v_n\} \subset V$ un conjunto linealmente dependiente. Entonces $\{T(v_1), T(v_2)...T(v_n)\}$ también es linealmente dependiente.*

Demostración. Como $\{v_1, v_2...v_n\}$ es linealmente dependiente, existen números γ_1, $\gamma_2, ...\gamma_n$, no todos cero, tales que

$$\sum_{i=1}^{n} \gamma_i v_i = 0$$

y por lo tanto

$$T\left[\sum_{i=1}^{n} \gamma_i v_i\right] = 0$$
$$\left[\sum_{i=1}^{n} \gamma_i T(v_i)\right] = 0$$

probando que $\{T(v_1), T(v_2)...T(v_n)\}$ es linealmente dependiente. ■

De aquí se sigue que si $\{T(v_1), T(v_2)...T(v_n)\}$ es linealmente independiente, entonces $\{v_1, v_2...v_n\}$ tiene que ser linealmente independiente.

El converso del teorema no es cierto como se puede ver con un contraejemplo:

Ejemplo 160 *Considere $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (x, 0)$, entonces el conjunto $\{(1, 0), (0, 1)\}$ es independiente pero $\{T(1, 0), T(0, 1)\} = \{(1, 0), (0, 0)\}$ es dependiente.*

Para especificar completamente una transformación lineal basta con dar sus valores en una base cualquiera, como se muestra en el siguiente teorema:

Teorema 161 *Una transformación lineal queda determinada por sus valores en una base cualquiera.*

Demostración. Sean V y W espacios vectoriales y $T : V \to W$ una transformación lineal. Sea

$$A = \{v_1, v_2, ...v_n\}$$

una base de V. Llamemos $w_i = T(v_i)$. Entonces si $v \in V$ tendremos que

$$v = \sum_{i=1}^{n} c_i v_i$$

donde $c_1, c_2, ..c_n$ son las coordenadas de v en la base A. Por lo tanto

$$\begin{aligned} T(v) &= T\left(\sum_{i=1}^{n} c_i v_i\right) = \sum_{i=1}^{n} T(c_i v_i) \\ &= \sum_{i=1}^{n} c_i T(v_i) = \sum_{i=1}^{n} c_i w_i \end{aligned}$$

lo cual quiere decir que el valor de $T(v)$ está determinado completamente por las coordenadas de v y los valores de T en la base. ■

Ejemplo 162 *¿cómo son las transformaciones lineales η de un campo F a un espacio vectorial V (sobre el campo F)? Es fácil ver que*

$$\eta(x) = \eta(x1) = x\eta(1)$$

por lo que las transformaciones lineales de F a V son multiplicación por un vector $\eta(1)$. Conversamente, si $\phi(x) = xv$ con $v \in V$ ϕ es lineal. Hay una correspondencia biunívoca entre los mapeos lineales $\eta : F \to V$ y los vectores de V.

7.2. núcleo e imagen

Definición 163 *Sean V y W espacios vectoriales y $T : V \to W$ una transformación lineal*
El núcleo (kernel) de T se define como

$$N(T) = \ker(T) = \{x \in V \mid Tx = 0\}$$

La imagen (recorrido,rango) de T se define como

$$\mathrm{Im}(T) = \{Tx \mid x \in V\}$$

Teorema 164 *Sean V y W espacios vectoriales y $T : V \to W$ una transformación lineal. Entonces $\ker(T)$ es un subespacio de V.*

Demostración. a) de la definición vemos que $\ker(T) \subset V$
b) Como $T(0) = 0$ tenemos que $0 \in \ker(T)$ y $\ker(T) \neq \phi$
c) Si $a, b \in \ker(T)$ entonces $T(a) = 0$ y $T(b) = 0$ por lo que $T(a+b) = T(a) + T(b) = 0 + 0 = 0$ y $a + b \in \ker(T)$
d) Si $a \in \ker(T)$ y α es un escalar, entonces $T(a) = 0$ por lo que $T(\alpha a) = \alpha T(a) = \alpha 0 = 0$ y $\alpha a \in \ker(T)$ Q.E.D. ■

Teorema 165 *Sean V y W espacios vectoriales y $T : V \to W$ una transformación lineal. Entonces $\mathrm{Im}(T)$ es un subespacio de W.*

Demostración. a) De la definición vemos que $\mathrm{Im}(T) \subset W$
b) Como $T(0) = 0$ tenemos que $0 \in$ y $\mathrm{Im}(T) \neq \phi$
c) Si $a, b \in \mathrm{Im}(T)$ entonces existen $c, d \in V$ tales que $T(c) = a$ y $T(d) = b$ con lo que $T(c+d) = T(c) + T(d) = a + b$ y, por ello, existe $c + d \in V$ tal que $T(c+d) = a + b$ y $a + b \in \mathrm{Im}(T)$
d) Si $a \in \mathrm{Im}(T)$ y γ es un escalar, entonces existe $c \in V$ tal que $T(c) = a$, por ello $T(\gamma c) = \gamma T(c) = \gamma a$ por lo
que existe $\gamma c \in V$ tal que $T(\gamma c) = \gamma a$ y $\gamma a \in \mathrm{Im}(T)$. Q.E.D. ■

Ejemplo 166 *El núcleo de $\hat{0}$ es todo V pues $\hat{0}(x) = 0$ para todo $x \in V$ y la imagen de $\hat{0}$ es $\{0\}$. El núcleo de I_V es $\{0\}$ en tanto que la imagen es V.*

Ejemplo 167 *Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal dada por*

$$T(x, y, z) = (x + 2y + z, 3x + y - 2z, -x - 7y - 6z)$$

Dar bases para $\ker(T)$ y $\mathrm{Im}(T)$.
a) un vector $(x, y, z) \in \ker(T)$ si y sólo si $T(x, y, z) = 0$ por lo que

$$(x + 2y + z, 3x + y - 2z, -x - 7y - 6z) = (0, 0, 0)$$

o

$$\begin{aligned} x + 2y + z &= 0 \\ 3x + y - 2z &= 0 \\ -x - 7y - 6z &= 0 \end{aligned}$$

y el núcleo buscado es el conjunto solución de este sistema lineal homogéneo. La matriz de coeficientes es

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & -2 \\ -1 & -7 & -6 \end{pmatrix}$$

y

$$
\begin{aligned}
A &\sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & -5 & -5 \\ 0 & -5 & -5 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}
\end{aligned}
$$

y un sistema equivalente al original (en el sentido de tener las mismas soluciones) es

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

en el cual las variables ligadas son x y y en tanto que z es la única variable libre. Por ello, los vectores en el conjunto solución están dados por

$$(x, y, z) = (z, -z, z) = z(1, -1, 1)$$

y una base del núcleo es $B = \{(1, -1, 1))\}$.
b) para hallar la imagen note que

$$
\begin{aligned}
T(x, y, z) &= (x + 2y + z, 3x + y - 2z, -x - 7y - 6z) \\
&= (x, 3x, -x) + (2y, y, -7y) + (z, -2z, -6z) \\
&= x(1, 3, -1) + y(2, 1, -7) + z(1, -2, -6)
\end{aligned}
$$

por lo que un generador para la imagen es el conjunto

$$\{(1, 3, -1), (2, 1, -7), (1, -2, -6)\}$$

y de aquí hay que extraer una base. Para ello notemos que la imagen es el espacio renglón de

$$B = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 1 & -7 \\ 1 & -2 & -6 \end{pmatrix}$$

y como

$$
\begin{aligned}
B &\sim \begin{pmatrix} 1 & 3 & -1 \\ 0 & -5 & -5 \\ 0 & -5 & -5 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 3 & -1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 0 & -4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}
\end{aligned}
$$

de modo que una base de la imagen es

$$\{(1,0,-4),(0,1,1)\}$$

Nótese que $B = A^T$. Esto siempre es cierto, la imagen es el espacio columna de la matriz A cuando la transformación es del tipo $T(x) = Ax$.

7.2.1. imagen y suprayectividad.

La imagen es, entre otras cosas, una herramienta para valorar la suprayectividad. Para ello disponemos del siguiente:

Teorema 168 *Sean V y W espacios vectoriales, $T : V \to W$ una transformación lineal. Entonces T es suprayectiva si y sólo si*

$$\mathrm{Im}(T) = W$$

Demostración. Sea T suprayectiva. Claramente $\mathrm{Im}(T) \subset W$, ahora bien, si $x \in W$ entonces, como T es suprayectiva, existe $a \in V$ tal que $T(a) = x$ pero esto implica que $x \in \mathrm{Im}(T)$ por lo que $W \subset \mathrm{Im}(T)$ y $\mathrm{Im}(T) = W$. ■

Conversamente, sea $\mathrm{Im}(T) = W$. Si $x \in W$ entonces como $W = \mathrm{Im}(T)$ $x \in \mathrm{Im}(T)$ .con lo que $x = T(a)$ para alguna $a \in V$ y la función es suprayectiva.

7.2.2. Núcleo e inyectividad.

El núcleo es básicamente una herramienta para valorar inyectividad, para ello resulta muy útil el siguiente:

Teorema 169 *Sean V y W espacios vectoriales, $T : V \to W$ una transformación lineal. Entonces T es inyectiva si y sólo si*

$$N(T) = \{0\}$$

Demostración. Sea T inyectiva. Si $x \in N(T)$ entonces $T(x) = 0 = T(0)$ lo cual, por la inyectividad, implica que $x = 0$, por otra parte ya sabemos que $0 \in N(T)$ por lo que $N(T) = \{0\}$. ■

Conversamente, sea $N(T) = \{0\}$. Si $T(x) = T(y)$ entonces $T(x) - T(y) = 0$ y $T(x - y) = 0$ con lo que $(x - y) \in N(T)$ y $x - y = 0$ (pues $N(T) = \{0\}$) y $x = y$. Luego T es inyectiva.

7.3. teorema de la dimensión.

Teorema 170 *(de la dimensión). Sean V y W espacios vectoriales, $T : V \to W$ una transformación lineal, $\alpha = \{v_1, v_2..v_n\}$ y $\beta = \{w_1, w_2..w_m\}$ bases de V y W respectivamente. Entonces*

$$\dim(V) = \dim(N(T)) + \dim(\mathrm{Im}(T))$$

Si T es biyectiva, $\dim(N(T)) = 0$ y $\dim(\mathrm{Im}(T)) = \dim(W) = m$ con lo que $\dim(V) = \dim(W)$. Es decir, si T es biyectiva, las dimensiones del dominio y codominio coinciden, pero el converso no es cierto en general.

Comentario 171 *Muchos autores llaman a $\dim(\mathrm{Im}(T))$ "rango de T", pero cuidado con este término que es usado con diversos significador. A $\dim(N(T))$ se le llama "nulidad de T".*

Comentario 172 *Cuando la transformación $T : V \to W$ es del tipo $T(x) = Ax$ donde A es una matriz y x una matriz de una sola columna, $\mathrm{Im}(T)$ es simplemente el espacio columna $\mathrm{col}(A)$.*

Comentario 173 *Cuando la transformación $T : V \to W$ es del tipo $T(x) = Ax$ donde A es una matriz y x una matriz de una sola columna, $N(T)$ se llama "espacio nulo" de A .*

Comentario 174 *Por lo anterior el teorema de la dimensión también se expresa como n=rango+nulidad (A es de $m \times n$)*

Comentario 175 *El núcleo (=espacio nulo) es, por lo tanto, el espacio solución de $Ax = 0$.*

7.4. representaciones matriciales.

Sean V y W espacios vectoriales, y $\alpha = \{v_1, v_2..v_n\}$ y $\beta = \{w_1, w_2..w_m\}$ bases de V y W respectivamente.

Si $T : V \to W$ es una transformación lineal, como $Tv_j \in W$ existen números M_{ij} tales que

$$Tv_j = \sum_{i=1}^{\dim(W)} M_{ij} w_i$$

La matriz definida de esta manera recibe el nombre de matriz de T relativa a las bases α y β y se representa como $M_\beta^\alpha(T)$.

La importancia de esta matriz puede verse tomando $v \in V$, que puede escribirse en términos de la base α y

$$v = \sum_{h=1}^{\dim(V)} c_h v_h$$

(aquí los números c_i son las coordenadas de v en la base α) y entonces

$$\begin{aligned} T(v) = T\left(\sum_{h=1}^{\dim(V)} c_h v_h\right) &= \sum_{h=1}^{\dim(V)} c_h T(v_h) = \sum_{h=1}^{\dim(V)} c_h \sum_{i=1}^{\dim(W)} M_{ih} w_i \\ &= \sum_{i=1}^{\dim(W)} \left(\sum_{h=1}^{\dim(V)} M_{ih} c_h\right) w_i \end{aligned}$$

(aquí hemos usado libremente el hecho de que en un espacio vectorial podemos reordenar las sumas como deseemos). De aquí se ve que $\sum_{h=1}^{\dim(V)} M_{ih} c_h$ es la i-ésima coordenada de $T(v)$ en la base β. También vemos que $\sum_{h=1}^{\dim(V)} M_{ih} c_h$ es la i-ésima componente de $M[v]_\alpha$ por lo que obtenemos el resultado fundamental de que

$$[Tv]_\beta = M_\beta^\alpha(T)[v]_\alpha$$

donde $[v]_\alpha$ y $[Tv]_\beta$ son las coordenadas de v y Tv en sus respectivas bases.

Más adelante veremos una manera fácil de calcular $M_\beta^\alpha(T)$ en ciertos casos (Si la base β es ortonormal $M_{hj} = \langle w_h, Tv_j \rangle$ donde $\langle\rangle$ es el producto interior que será discutido en el próximo capítulo).

Note que si $\dim(V) = n$ y $\dim(W) = m$ entonces $M_\beta^\alpha(T)$ es de $m \times n$.

Esquemáticamente

$$\begin{array}{ccc} V & \overset{T}{\rightarrow} & W \\ \phi_\alpha \downarrow & & \phi_\beta \downarrow \\ F^n & \overset{M_\beta^\alpha(T)}{\rightarrow} & F^m \end{array}$$

donde ϕ_α y ϕ_β son los isomorfismos de coordenadas. En la literatura se dice que el diagrama conmuta o que es un diagrama conmutativo, esto quiere decir que da lo mismo ir directamente de V a W mediante T que ir primero de V a F^n mediante el isomorfismo de coordenadas, de F^n a F^m mediante multiplicación por $M_\beta^\alpha(T)$ y finalmente llegar a W mediante ϕ_β^{-1}.

7.4.1. la matriz de transición de nuevo

Sea I_V el operador identidad en un espacio V. Sean $\alpha = \{v_1, v_2..v_n\}$ y $\beta = \{w_1, w_2..w_n\}$ dos bases de V y M_β^α la matriz de transición de α a β. Sea $n = \dim(V)$.

Entonces:

1. $M_{\alpha}^{\alpha}(I_V) = I_{n\times n}$.
2. $M_{\beta}^{\beta}(I_V) = I_{n\times n}$
3. $M_{\beta}^{\alpha}(I_V) = M_{\beta}^{\alpha}$ pues

$$[I_V v]_\beta = [v]_\beta = M_{\beta}^{\alpha}(I_V)[v]_\alpha$$

4. $M_{\alpha}^{\beta}(I_V) = M_{\alpha}^{\beta} = \left(M_{\beta}^{\alpha}\right)^{-1}$ pues

$$[I_V v]_\alpha = [v]_\alpha = M_{\alpha}^{\beta}(I_V)[v]_\beta$$

Además, sea $L : V \to V$ el operador lineal definido en las bases mediante

$$L(v_i) = w_i$$

entonces:

1. $$M_{\beta}^{\alpha}(L) = I_{n\times n}$$

   por la mera definición de matriz asociada.

2. $$M_{\alpha}^{\alpha}(L) = \left(M_{\beta}^{\alpha}\right)^{-1} = M_{\alpha}^{\beta}$$

7.4.2. calculando la matriz asociada a una transformación lineal.

El esquema que proponemos para sistematizar el cálculo de la matriz asociada a una transformación lineal es el siguiente:

1. evalúe la transformación en cada uno de los vectores de la base α, es decir, $T(v_1)$, $T(v_2)...T(v_n)$.
2. calcule las coordenadas de cada vector $T(v_j)$ en la base β.
3. las coordenadas de $T(v_j)$ forman la j-ésima columna de $M_{\beta}^{\alpha}(T)$
4. Note que si $n = \dim(V)$ y $m = \dim(W)$ entonces $M_{\beta}^{\alpha}(T)$ será una matriz de $m \times n$.

Ejemplo 176 *Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal definida mediante*

$$T(x, y, z) = (2x - 3y + z, x - y - z, x + y + 5z)$$

y equipe a $\mathbb{R}^3$ con su base canónica $\alpha = \{(1,0,0),(0,1,0),(0,0,1)\}$ tanto para el dominio como para el codominio. Calcule la matríz $M_\alpha^\alpha(T)$.
Solución: De acuerdo con el esquema recomendado calculamos

$$\begin{aligned}
T(1,0,0) &= (2,1,1) = (2)(1,0,0) + (1)(0,1,0) + (1)(0,0,1) \\
T(0,1,0) &= (-3,-1,1) = (-3)(1,0,0) + (-1)(0,1,0) + (1)(0,0,1) \\
T(0,0,1) &= (1,-1,5) = (1)(1,0,0) + (-1)(0,1,0) + (5)(0,0,1)
\end{aligned}$$

en donde hemos re-escrito el resultado como combinación lineal de los vectores de la base (en este caso es muy fácil pues es la base canónica).
Por ello la matriz solicitada es

$$M_\alpha^\alpha(T) = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & 5 \end{bmatrix}$$

Note que

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

Ejemplo 177 *Considere la transformación lineal $S : P_3 \to P_2$ definida mediante*

$$S(q) = \frac{dq}{dx}$$

y considere las bases

$$\begin{aligned}
\alpha &= \{1, x, x^2, x^3\} \\
\beta &= \{1, x, x^2\}
\end{aligned}$$

para P_3 y P_2 respectivamente. Calcule la matriz de S en las bases dadas.
Solución: Por el algoritmo recomendado, calculamos

$$\begin{aligned}
S(1) &= 0 = (0)(1) + (0)(x) + (0)(x^2) \\
S(x) &= 1 = (1)(1) + (0)(x) + (0)(x^2) \\
S(x^2) &= 2x = (0)(1) + (2)(x) + (0)(x^2) \\
S(x^3) &= 3x^2 = (0)(1) + (0)(x) + (3)(x^2)
\end{aligned}$$

y por ello

$$M_\beta^\alpha(S) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

7.4.3. Dada la matriz calcular la regla de correspondencia

A veces tenemos el problema inverso. Si se nos dá $M_{\beta}^{\alpha}(T)$ ¿podemos inferir la regla de correspondencia de T ? La respuesta es afirmativa, la clave es, una vez más,

$$[Tv]_{\beta} = M_{\beta}^{\alpha}(T)[v]_{\alpha}$$

Ejemplo 178 *Si $T : \mathbb{R}^2 \to \mathbb{R}^3$ tiene como matriz asociada a las bases canónicas de $\mathbb{R}^2$ y $\mathbb{R}^3$*

$$M_{\beta}^{\alpha} = \begin{bmatrix} 2 & 3 \\ -1 & 0 \\ 4 & 3 \end{bmatrix}$$

entonces para $v \in \mathbb{R}^2$ si $[v]_{\alpha} = \begin{bmatrix} x \\ y \end{bmatrix}$

$$\begin{aligned} [Tv]_{\beta} &= M_{\beta}^{\alpha}(T)[v]_{\alpha} \\ &= \begin{bmatrix} 2 & 3 \\ -1 & 0 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} 2x + 3y \\ -x \\ 4x + 3y \end{bmatrix} \end{aligned}$$

por lo que

$$\begin{aligned} Tv &= (2x + 3y)(1, 0, 0) + (-x)(0, 1, 0) + (4x + 3y)(0, 0, 1) \\ &= (2x + 3y, -x, 4x + 3y) \end{aligned}$$

Capítulo 8

Álgebra de transformaciones

Hemos visto cómo en un espacio vectorial hay dos operaciones de tal manera que los vectores pueden sumarse y multiplicarse por números. Ahora vamos a ver que las transformaciones también pueden sumarse y multiplicarse por numeros.

8.1. Suma de transformaciones lineales.

En este capítulo sean V y W dos espacios vectoriales sobre el mismo campo F y $S : V \to W$ y $T : V \to W$ dos transformaciones.

Vamos a definir una nueva transformación llamada suma de S y T que denotaremos como $S + T$. Recuerde usted que para definir cualquier función (transformación o no) se requiere:

- especificar un dominio
- especificar un codominio
- especificar una regla de correspondencia

Por ello definimos $S + T$ como aquella transformación tal que:

- su dominio es V
- su codominio es W
- su regla de correspondencia está definida para cualquier $x \in V$ mediante

$$(S + T)(x) = S(x) + T(x)$$

Nóte que el signo $+$ del lado izquierdo se refiere a la suma de transformaciones, en tanto que el signo $+$ del lado derecho es la suma en W.

En resumidas cuentas

$$S + T : V \to W$$
$$(S + T)(x) = S(x) + T(x)$$

8.2. Producto por números.

Dado un número $\alpha \in F$ y una transformación vamos a definir una nueva transformación $S : V \to W$ llamada producto de α y de S y que denotaremos como αS.

αT se define con las siguientes estipulaciones:

- su dominio es V
- su codominio es W
- su regla de correspondencia, para cualquier $x \in V$ y cualquier $\alpha \in F$ es $(\alpha S)(x) = \alpha(S(x))$

8.3. La suma de dos transformaciones lineales es lineal

Suponga, de momento, que tanto S como T son lineales, tome $x, y \in V$, entonces

$$\begin{aligned}(S + T)(x + y) &= S(x + y) + T(x + y)\\ &= S(x) + S(y) + T(x) + T(y)\\ &= (S(x) + T(x)) + (S(y) + T(y))\\ &= (S + T)(x) + (S + T)(y)\end{aligned}$$

donde hemos usado tanto la definición de $S + T$ como la linealidad de S y de T .

Análogamente, para $x \in V$ y $\alpha \in F$ tendremos

$$\begin{aligned}(S + T)(\alpha x) &= S(\alpha x) + T(\alpha x)\\ &= \alpha S(x) + \alpha T(x)\\ &= \alpha(S(x) + T(x))\\ &= \alpha((S + T)(x))\end{aligned}$$

completando la prueba de la linealidad.

8.4. Linealidad del producto por un número

Suponga, de momento, que T es lineal, tome $x, y \in V$, entonces y $\alpha \in F$, entonces

$$\begin{aligned}(\alpha T)(x+y) &= \alpha(T(x+y)) \\ &= \alpha(T(x)+T(y)) \\ &= \alpha T(x)+\alpha T(y) \\ &= (\alpha T)(x)+(\alpha T)(y)\end{aligned}$$

donde hemos usado tanto la definición de αT como la linealidad de T .

Análogamente, para $x \in V$ y $\beta \in F$ tendremos

$$\begin{aligned}(\alpha T)(\beta x) &= \alpha(T(\beta x)) \\ &= \alpha(\beta T(x)) \\ &= \alpha\beta(T(x)) \\ &= \beta(\alpha T(x)) \\ &= \beta((\alpha T)(x))\end{aligned}$$

completando la prueba de la linealidad.

8.5. El conjunto de todas las transformaciones lineales de V en W es un espacio vectorial.

Al conjunto de todas las transformaciones lineales de V en W

$$\mathcal{F} = \{T : V \to W \mid T \text{ es lineal}\}$$

se le denota por

$$Lin(V, W)$$

o

$$mor(V, W)$$

o

$$\hom(V, W)$$

Vamos a ver que $\mathcal{F}$ es un espacio vectorial.

8.5.1. ¿hay un zero (elemento idéntico) para la suma de transformaciones?

Primeramente vamos a suponer que sí lo hay a fin de descubrir quién es. Suponga pues que existe $\widehat{0} : V \to W$ y tal que para todo $S \in \mathcal{F}$ tenemos que $S + \widehat{0} = \widehat{0} + S = S$. Entonces para cualquier $x \in V$

$$(S+\widehat{0})(x) = S(x) + \widehat{0}(x) = S(x)$$

con lo que

$$\widehat{0}(x) = 0$$

para todo $x \in V$ y $\widehat{0}$ es la transformación cero anteriormente discutida.

Conversamente, hay que verificar que $\widehat{0}$ hace bien su trabajo. Evaluemos pues para $x \in V$

$$\begin{aligned}(S + \widehat{0})(x) &= S(x) + \widehat{0}(x) = S(x) + 0 = S(x) \\ &= 0 + S(x) = \widehat{0}(x) + S(x) = (\widehat{0} + S)(x)\end{aligned}$$

con lo que $S + \widehat{0} = \widehat{0} + S = S$.

8.5.2. ¿hay un negativo (inverso aditivo) para la suma de transformaciones ?

Sea $S \in \mathcal{F}$. Primero supondremos que existe $H \in \mathcal{F}$ y tal que $S + H = H + S = \widehat{0}$. Entonces para todo $x \in V$

$$(S + H)(x) = S(x) + H(x) = \widehat{0}(x) = 0$$

y tenemos que

$$H(x) = -S(x) = ((-1)S)(x)$$

con lo que el candidato a negativo es $-S = (-1)S$.

Verificamos que en efecto

$$(S + (-1)S)(x) = S(x) + ((-1)S)(x) = S(x) - S(x) = 0$$

De hecho $Lin(V, W)$ es un espacio vectorial sobre F con las operaciones de suma y producto por elementos de F que hemos definido, dejamos la demostración (usando los axiomas para los espacios vectoriales) al lector, ya habiendo hecho la parte más difícil (la existencia del cero y del inverso).

8.6. Matriz asociada a la suma y producto por números.

De nuevo sean V y W dos espacios vectoriales sobre el mismo campo F y $S : V \to W$ y $T : V \to W$ dos transformaciones lineales. Sea $\alpha \in F$. Equipemos ahora a V y W con las bases $A = \{v_1, v_2, ...v_n\}$ y $B = \{w_1, w_2, ...w_m\}$ respectivamente.

¿Cuál será la matriz $M_B^A(S + T)$?

Ciertamente esperamos que, para cualquier $x \in V$

$$[(S + T)(x)]_B = M_B^A(S + T)\,[x]_A$$

pero como

$$[(S + T)(x)]_B = [S(x) + T(x)]_B = [S(x)]_B + [T(x)]_B$$

y además

$$\begin{aligned}[S(x)]_B &= M_B^A(S)\,[x]_A \\ [T(x)]_B &= M_B^A(T)\,[x]_A\end{aligned}$$

resulta que

$$[(S+T)(x)]_B = M_B^A(S)\,[x]_A + M_B^A(T)\,[x]_A = \left(M_B^A(S) + M_B^A(T)\right)[x]_A$$

con lo que obtenemos que

$$M_B^A(S+T)\,[x]_A = \left(M_B^A(S) + M_B^A(T)\right)[x]_A$$

y como esto vale para cualquier $x \in V$

$$M_B^A(S+T) = M_B^A(S) + M_B^A(T)$$

de modo que la matriz de la suma de dos transformaciones lineales es la suma de las matrices de las transformaciones.

De una manera análoga, ¿Cuál será la matriz $M_B^A(\alpha S)$?

Ciertamente esperamos que, para cualquier $x \in V$

$$[(\alpha S)(x)]_B = M_B^A(\alpha S)\,[x]_A$$

pero como

$$[(\alpha S)(x)]_B = [\alpha(S(x))]_B = \alpha\,[S(x)]_B$$

y además

$$[S(x)]_B = M_B^A(S)\,[x]_A$$

resulta que

$$[(\alpha S)(x)]_B = \alpha M_B^A(S)\,[x]_A = \left(\alpha M_B^A(S)\right)[x]_A$$

con lo que obtenemos que

$$M_B^A(\alpha S)\,[x]_A = \left(\alpha M_B^A(S)\right)[x]_A$$

y como esto vale para cualquier $x \in V$

$$M_B^A(\alpha S) = \alpha M_B^A(S)$$

de modo que la matriz del producto de un número por una transformación es el producto del número por la matriz de la transformación.

Comentario 179 *De hecho los espacios* $Lin(V,W)$ *y* $M(m,n)$ *(o* $M^*(m,n)$ *según el caso) son isomorfos. El isomorfismo es* $S \mapsto M_B^A(S)$.

8.7. Composición de transformaciones.

Sabemos que matrices del mismo tamaño pueden sumarse y hemos visto que esto corresponde a la suma de transformaciones. También sabemos que las matrices pueden multiplicarse por números, y que esto corresponde al producto de números por transformaciones.

Pero las matrices también pueden multiplicarse entre sí cuando son conformables (es decir, el primer factor tiene tantas columnas como renglones tiene el segundo factor). ¿a qué operación entre transformaciones corresponde esto?.

8.7.1. definición de composición.

Sean V, W, Z espacios vectoriales sobre un mismo campo F. Si $S : V \to W$ y $T : W \to Z$ son transformaciones entonces podemos definir una nueva transformación $T \circ S$ de la siguiente manera:

- su dominio es V
- su codominio es Z
- su regla de correspondencia es

$$(T \circ S)(x) = T(S(x))$$

Lo anterior suele ponerse en forma de diagrama

$$\begin{tikzcd} V \arrow[r, "S"] \arrow[dr, "T \circ S"'] & W \arrow[d, "T"] \\ & Z \end{tikzcd}$$

y se dice que el diagrama es conmutativo, con ello se quiere decir que da lo mismo ir primero de V a W mediante S y luego de W a Z mediante T que ir directamente de V a Z mediante $T \circ S$.

8.7.2. La composición de dos transformaciones lineales es lineal.

Sean F, S, V, W, T, W y Z como en la subsección anterior pero suponga que tanto S como T son lineales. Sean $x, y \in V$, entonces

$$\begin{aligned} (T \circ S)(x + y) &= T(S(x + y)) \text{ por la definición de composición} \\ &= T(S(x) + S(y)) \text{ por la linealidad de } S \\ &= T(S(x)) + T(S(y)) \text{ por la linealidad de } T \\ &= (T \circ S)(x) + (T \circ S)(y) \text{ por la definición de composición} \end{aligned}$$

análogamente, sea $x \in V$ y $\alpha \in F$, entonces

$$\begin{aligned}(T \circ S)(\alpha x) &= T(S(\alpha x)) \text{ por la definición de composición} \\ &= T(\alpha S(x))\text{por la linealidad de } S \\ &= \alpha T(S(x))\text{por la linealidad de } T \\ &= \alpha\left[(T \circ S)(x)\right] \text{por la definición de composición}\end{aligned}$$

8.7.3. Matriz de la composición de dos transformaciones lineales.

Equipemos ahora a V, W y Z con las bases $A = \{v_1, v_2, ...v_n\}$ y $B = \{w_1, w_2, ...w_m\}$ y $C = \{z_1, z_2, ...z_p\}$ respectivamente.

Entonces el "trabajo" de $M_C^A(T \circ S)$ es (para cualquier $x \in V$)

$$\left[(T \circ S)(x)\right]_C = M_C^A(T \circ S)\left[x\right]_A$$

pero por una parte

$$\begin{aligned}\left[(S)(x)\right]_B &= M_B^A(S)\left[x\right]_A \\ \left[(T)(x)\right]_C &= M_C^B(T)\left[x\right]_B\end{aligned}$$

y por otra

$$(T \circ S)(x) = T(S(x))$$

con lo que

$$\begin{aligned}\left[(T \circ S)(x)\right]_C &= \left[(T(S(x))\right]_C \\ &= M_C^B(T)\left[S(x)\right]_B \\ &= (M_C^B(T))M_B^A(S)\left[x\right]_A \\ &= (M_C^B(T)M_B^A(S))\left[x\right]_A \\ &= M_C^A(T \circ S)\left[x\right]_A\end{aligned}$$

y como esto vale para cualquier x tendremos que

$$M_C^A(T \circ S) = M_C^B(T)M_B^A(S)$$

con lo que la matriz de una composición es el producto de las matrices de las transformaciones. Note, como mnemotecnia, que en el lado izquierdo no aparece el nombre de la base que a la derecha aparece dos veces, una vez arriba y otra vez abajo.

8.7.4. Las propiedades más importantes de estas operaciones .

Ahora el escenario será un poco más complicado.

Sean U, V, W y X espacios vectoriales y F, G, H, S, T transformaciones lineales

$$\begin{gathered}F : U \to V\\ G : U \to V\\ H : W \to X\\ S : V \to W\\ T : V \to W\end{gathered}$$

entonces:

- $S \circ (F + G) = S \circ F + S \circ G$ (distributividad)
- $(S + T) \circ F = S \circ F + T \circ F$ (otra distributividad)
- $\alpha(S \circ F) = (\alpha S) \circ F = S \circ (\alpha F)$ $(\alpha \in F)$
- $H \circ (S \circ F) = (H \circ S) \circ F$ (asociatividad)
- $T \circ I_V = T$ (I_V es la identidad en V)
- $I_W \circ T = T$ (I_W es la identidad en W)

Recuerde usted que dos funciones son iguales si y sólo si:

- tienen el mismo dominio
- tienen el mismo codominio
- tienen la misma regla de correspondencia.

Tome, por ejemplo, la primera igualdad. Las expresiones a ambos lados tienen el mismo dominio (que es U, el dominio de F y G) y el mismo codominio, que es W (el codominio de S). Si $x \in U$ tendremos que

$$\begin{aligned}(S \circ (F + G))(x) &= S((F + G)(x))\text{por la definición de composición}\\ &= S(F(x) + G(x))\text{por la definición de suma}\\ &= S(F(x)) + S(G(x)) \text{ por la linealidad de } S\\ &= (S \circ F)(x) + (S \circ G)(x) \text{ por la definición de composición}\\ &= ((S \circ F) + (S \circ G))\,(x) \text{ por la definición de suma}\end{aligned}$$

Para la segunda igualdad, notemos que ambos miembros tienen dominio U y codominio W, y

$$\begin{aligned}((S + T) \circ F)(x) &= (S + T)(F(x)) \text{ por la definición de composición}\\ &= S(F(x)) + T(F(x)) \text{ por la definición de suma}\\ &= (S \circ F)(x) + (T \circ F)(x) \text{ por la definición de composición}\\ &((S \circ F) + (T \circ F))\,(x) \text{ por la definición de suma}\end{aligned}$$

Para la tercera parte ambos miembros tienen dominio U y codominio W, y

$$\begin{aligned}\alpha(S\circ F)(x) &= \alpha((S\circ F)(x)) \text{ por la definición de producto}\\ &= \alpha S(F(x)) \text{ por la definición de composición}\\ &= (\alpha S)(F(x)) \text{ por la definición de producto}\\ &= ((\alpha S)\circ F)(x) \text{ por la definición de composición}\end{aligned}$$

y además

$$\begin{aligned}\alpha(S\circ F)(x) &= \alpha((S\circ F)(x)) \text{ por la definición de producto}\\ &= \alpha S(F(x)) \text{ por la definición de composición}\\ &= S(\alpha F(x)) \text{ por la linealidad de } S\\ &= S((\alpha F)(x)) \text{ por la definición de producto}\\ &= (S\circ(\alpha F))(x) \text{ por la definición de composición}\end{aligned}$$

Para la cuarta identidad, notamos que ambos lados tienen dominio U y codominio X. Entonces

$$\begin{aligned}(H\circ(S\circ F))(x) &= H((S\circ F)(x)) \text{ por la definición de composición}\\ &= H(S(F(x))) \text{ por la definición de composición}\\ &= (H\circ S)(F(x)) \text{ por la definición de composición}\\ &= ((H\circ S)\circ F)(x) \text{ por la definición de composición}\end{aligned}$$

con lo que vemos que la operación de composición es asociativa. Dejo al lector convencerse de que esta operación no es conmutativa.

La quinta identidad involucra a V como dominio en ambos lados y a W como codominio. Tenemos, pues

$$\begin{aligned}(T\circ I_V)(x) &= T(I_V(x)) \text{ por la definición de composición}\\ &= T(x) \text{ por la definición de identidad en } V\end{aligned}$$

en tanto que la sexta también involucra a V como dominio en ambos lados y a W como codominio y un sencillo cómputo revela que

$$\begin{aligned}(I_W\circ T)(x) &= I_w(T(x)) \text{ por la definición de composición}\\ &= T(x) \text{ por la definición de identidad en } W\end{aligned}$$

Por ello las matrices identidad (del tamaño adecuado) se comportan como identidades derecha o izquierda con respecto a la multiplicación matricial. Recuerde que la composición (o la multiplicación de matrices) no son operaciones binarias (en el conjunto de todas las transformaciones lineales o de todas las matrices) pues no siempre se pueden realizar.

Comentario 180 *¿es única esta identidad derecha o izquierda? Considere*

$$a = \begin{bmatrix} 1 & 3 & 5\\ 2 & 4 & 6\end{bmatrix}$$

y

$$u = \begin{bmatrix} \frac{5}{6} & \frac{1}{3} & \frac{-1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{6} & \frac{1}{3} & \frac{5}{6} \end{bmatrix}$$

y usted puede verificar facilmente que

$$au = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} \frac{5}{6} & \frac{1}{3} & \frac{-1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{6} & \frac{1}{3} & \frac{5}{6} \end{bmatrix} = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = a$$

pero u no es la matriz identidad.

Capítulo 9

La transformación inversa.

9.1. La inversa de una transformación.

Sean V y W dos espacios vectoriales sobre un campo F y $T : V \to W$ una transformación.

Decimos que T tiene una inversa (denotada por T^{-1}) o que es invertible (o no singular) si existe una transformación $T^{-1} : W \to V$ tal que:

- $T^{-1} \circ T = I_V$
- $T \circ T^{-1} = I_W$

9.1.1. la inversa es única

Suponga que $T : V \to W$ tuviera dos inversas, denotemos una por T^{-1} y la otra por G, entonces

$$\begin{aligned} T^{-1} &: W \to V \\ G &: W \to V \\ T^{-1} \circ T &= I_V \\ T \circ T^{-1} &= I_w \\ G \circ T &= I_V \\ T \circ G &= I_w \end{aligned}$$

por ello

$$\begin{aligned} G &= G \circ I_w \\ &= G \circ (T \circ T^{-1}) \\ &= (G \circ T) \circ T^{-1} \\ &= I_V \circ T^{-1} \\ &= T^{-1} \end{aligned}$$

9.1.2. la inversa de una transformación lineal es lineal

Suponga que $T : V \to W$ es una transformación lineal invertible.
Entonces si $T(x) = y$ tendremos que

$$\begin{aligned} T(x) &= y \Rightarrow \\ T^{-1}(T(x)) &= T^{-1}(y) \Rightarrow \\ \left(T^{-1} \circ T\right)(x) &= T^{-1}(y) \Rightarrow \\ I_V(x) &= T^{-1}(y) \Rightarrow \\ x &= T^{-1}(y) \end{aligned}$$

y también que

$$\begin{aligned} x &= T^{-1}(y) \Rightarrow \\ T(x) &= T(T^{-1}(y)) \Rightarrow \\ T(x) &= \left(T \circ T^{-1}\right)(y) \Rightarrow \\ T(x) &= (I_W)(y) \Rightarrow \\ T(x) &= (y) \end{aligned}$$

por lo que podemos concluir que

$$T(x) = (y) \Longleftrightarrow x = T^{-1}(y)$$

Por ello una manera sencilla de demostrar que la inversa de una transformación lineales lineal sería la siguiente:

$$\begin{aligned} T^{-1}(a+b) = T^{-1}(a) + T^{-1}(b) &\Longleftrightarrow \\ (a+b) = T(T^{-1}(a) + T^{-1}(b)) &\Longleftrightarrow \\ (a+b) = T(T^{-1}(a)) + T(T^{-1}(b)) &\Longleftrightarrow \\ (a+b) = \left(T \circ T^{-1}\right)(a) + \left(T \circ T^{-1}\right)(b) &\Longleftrightarrow \\ (a+b) = I_W(a) + I_W(b) &\Longleftrightarrow \\ (a+b) = a+b & \end{aligned}$$

y como la última aseveración es siempre cierta

$$T^{-1}(a+b) = T^{-1}(a) + T^{-1}(b)$$

Análogamente

$$\begin{aligned} T^{-1}(\beta a) &= \beta T^{-1}(a) \Longleftrightarrow \\ T(T^{-1}(\beta a)) &= T(\beta T^{-1}(a)) \Longleftrightarrow \\ \left(T \circ T^{-1}\right)(\beta a) &= \beta(T(T^{-1}(a))) \Longleftrightarrow \\ I_W(\beta a) &= \beta\left(T \circ T^{-1}\right)(a) \Longleftrightarrow \\ \beta a &= \beta I_W(a) \Longleftrightarrow \\ \beta a &= \beta a \end{aligned}$$

y como la última igualdad siempre es cierta tenemos que

$$T^{-1}(\beta a) = \beta T^{-1}(a)$$

9.1.3. La matriz de la inversa.

Sea $T : V \to W$ lineal e invertible, $A = \{v_1, v_2, ...v_n\}$ una base de V y $B = \{w_1, w_2, ...w_m\}$ una base de W.

Como

$$\begin{aligned} T^{-1} \circ T &= I_V \\ M_A^A(T^{-1} \circ T) &= M_A^B(T^{-1}) M_B^A(T) \\ &= M_A^A(I_V) = I \end{aligned}$$

tenemos que

$$M_A^B(T^{-1}) = \left[M_B^A(T)\right]^{-1}$$

de modo que la matriz de la inversa de T es la inversa de la matriz de T.

Por esto una manera de ver si una transformación es invertible es ver si su matriz (respecto a cualquier base que escojamos) es invertible.

9.1.4. Criterios de invertibilidad.

Una transformación lineal, al igual que cualquier otra función, tendrá inversa si y sólo si es biyectiva. Ahora bien, una función es biyectiva si y sólo si es inyectiva y suprayectiva. Pero ya hemos visto que una transformación lineal $T : V \to W$ es inyectiva si y sólo si

$$N(T) = 0$$

o, de manera equivalente, si y sólo si

$$\dim(N(T)) = 0$$

y que es suprayectiva si y sólo si

$$\mathrm{Im}(T) = W$$

o, lo que es lo mismo, si y sólo si

$$\dim(W) = \dim(\mathrm{Im}(T))$$

Además recuerde el teorema de la dimensión:

$$\dim(V) = \dim(N(T)) + \dim(\mathrm{Im}(T))$$

De aquí podemos sacar algunas conclusiones:

- Si T es invertible, entonces $\dim(V) = \dim(W)$. Esto es útil sobre todo en la forma: si $\dim(V) \neq \dim(W)$ entonces T no puede ser invertible.
- Cuando $\dim(V) = \dim(W)$, la inyectividad implica $\dim(V) = \dim(\mathrm{Im}(T))$, pero como $\mathrm{Im}(T)$ es un subespacio de W, tendremos que $\mathrm{Im}(T) = W$ y la transformación es necesariamente también suprayectiva (y biyectiva, consecuentemente).
- Cuando $\dim(V) = \dim(W)$, la suprayectividad implica $\dim(N(T)) = \dim(V) - \dim(\mathrm{Im}(T)) = \dim(V) - \dim(W) = 0$, y, por lo tanto, la inyectividad de T (y la biyectividad de T, consecuentemente).

9.1.5. La inversa de una composición.

Sean V, W, Z espacios vectoriales sobre un mismo campo F. Si $S : V \to W$ y $T : W \to Z$ son transformaciones invertibles, entonces $T \circ S : V \to Z$ es invertible y

$$(T \circ S)^{-1} = S^{-1} \circ T^{-1}$$

La demostración se basa en ver que $T^{-1} \circ S^{-1}$ "se comporta como inversa" y por la unicidad debe, entonces, ser la inversa.

Así que considere

$$\begin{aligned}
(T \circ S) \circ \left(S^{-1} \circ T^{-1}\right) &= T \circ (S \circ S^{-1}) \circ T^{-1} \\
&= T \circ (I_W) \circ T^{-1} \\
&= (T \circ I_W) \circ T^{-1} \\
&= T \circ T^{-1} = I_Z
\end{aligned}$$

y también

$$\begin{aligned}
\left(S^{-1} \circ T^{-1}\right) \circ (T \circ S) &= S^{-1} \circ (T^{-1} \circ T) \circ S \\
&= S^{-1} \circ (I_W) \circ S \\
&= (S^{-1} \circ I_W) \circ S \\
&= S^{-1} \circ S \\
&= I_V
\end{aligned}$$

con lo que se demuestra que $(T \circ S)^{-1} = S^{-1} \circ T^{-1}$ (note el cambio en el orden de aparición de S y de T, es importante).

Este "teorema" tiene su contrapartida entre las matrices, si A y B son matrices conformables (es decir, que se pueden multiplicar) y tienen inversos, ocurre que

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

Otras propiedades elementales de las inversas que el lector puede verificar fácilmente son:

- Si $\alpha \in F$, $\alpha \neq 0$ y $S : V \to W$ es invertible, entonces αS también es invertible y $(\alpha S)^{-1} = \alpha^{-1} S^{-1}$.
- Si $S : V \to W$ es invertible, entonces $(S^{-1})^{-1} = S$.

9.1.6. un ejemplo completo

Sean

$$\begin{aligned} V &= \{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\} \\ W &= \left\{ \begin{pmatrix} a & b \\ b & c \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\} \\ T &: V \to W \\ T(ax^2 + bx + c) &= \begin{pmatrix} a+c & 3b \\ 3b & a-c \end{pmatrix} \end{aligned}$$

y equipemos a V y W con las bases

$$\begin{aligned} A &= \{x^2, x, 1\} \\ B &= \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \end{aligned}$$

respectivamente.

Primero vamos a calcular $M_B^A(T)$ y para ello calculamos el efecto de T en la base A usando la regla de correspondencia

$$\begin{aligned} T(x^2) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = (1)\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + (0)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + (1)\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ T(x) &= \begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix} = (0)\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + (3)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + (0)\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ T(1) &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = (1)\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + (0)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + (-1)\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

con lo que vemos que

$$M_B^A(T) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

A continuación vamos a determinar si T tiene inversa. Ya que tenemos la matriz de T el mejor camino será indagar si esta matriz es invertible. Así pues

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

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

de donde vemos que la matriz es invertible y, por lo tanto, T es invertible. Además

$$M_A^B(T^{-1}) = M_B^A(T)^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & \frac{-1}{2} \end{pmatrix}$$

Notas:

- A veces puede usted determinar si la transformación es invertible verificando si es (o no es) una biyección. También ayuda tener en mente el teorema de la dimensión.
- En casos como el del ejemplo, en el que la matriz de la transformación es cuadrada, puede usted saber si la matriz (y por lo tanto la transformación) es invertible calculando el determinante. En el presente ejemplo

$$\det \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & -1 \end{pmatrix} = -6 \neq 0$$

  confirmando que la matriz es invertible.

Nuestro siguiente paso será calcular la regla de correspondencia de T^{-1}. Dado $w = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ tenemos que $[w] = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ y como

$$\begin{aligned} \left[T^{-1}(w)\right]_A &= M_A^B(T^{-1})\,[w]_B \\ &= \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & \frac{-1}{2} \end{pmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ &= \begin{bmatrix} \frac{a+c}{2} \\ \frac{b}{3} \\ \frac{a-c}{2} \end{bmatrix} \end{aligned}$$

de modo que

$$T^{-1}(w) = (\frac{a+c}{2})x^2 + (\frac{b}{3})x + (\frac{a-c}{2})1$$

En resumidas cuentas:

- calculamos la matriz de la transformación usando su regla de correspondencia
- calculamos la inversa de dicha matriz
- con esta inversa inferimos la regla de correspondencia de la transformación inversa.

Nota: lo hicimos de una manera complicada. Un alumno me señaló lo siguiente:

Si en la matriz $\begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & \frac{-1}{2} \end{pmatrix}$

seguimos el camino inverso al de obtener la matriz a partir de la regla de correspondencia, las columnas son las coordenadas en la base A del resultado de aplicar T^{-1} a los vectores de la base B . Por ello

$$T^{-1}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}x^2 + \frac{1}{2}$$
$$T^{-1}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{1}{3}x$$
$$T^{-1}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}x^2 - \frac{1}{2}$$

y

$$\begin{aligned} T^{1}\begin{pmatrix} a & b \\ b & c \end{pmatrix} &= a(\frac{1}{2}x^2 + \frac{1}{2}) + b\frac{1}{3}x + c(\frac{1}{2}x^2 - \frac{1}{2}) \\ &= (\frac{a+c}{2})x^2 + (\frac{b}{3})x + (\frac{a-c}{2})1 \end{aligned}$$

Ahora bien, no siempre es preciso recurrir a métodos matriciales, como ilustraremos en seguida.

Dado $w = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ lo que queremos es encontrar $T^{-1}(w)$ y sabemos que es un polinomio de grado no mayor a dos, digamos $\alpha x^2 + \beta x + \gamma$ de modo que

$$T^{-1}\begin{pmatrix} a & b \\ b & c \end{pmatrix} = \alpha x^2 + \beta x + \gamma$$

pero aplicando T a ambos miembros de la ecuación

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} = T(T^{-1}\begin{pmatrix} a & b \\ b & c \end{pmatrix}) = T(\alpha x^2 + \beta x + \gamma)$$

y usando la regla de correspondencia para T

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} = \begin{pmatrix} \alpha + \gamma & 3\beta \\ 3\beta & \alpha - \gamma \end{pmatrix}$$

de modo que

$$\begin{aligned} a &= \alpha + \gamma \\ b &= 3\beta \\ c &= \alpha - \gamma \end{aligned}$$

sistema que podemos resolver para α, β y γ obteniéndose

$$\alpha = \frac{a+c}{2}$$
$$\beta = \frac{b}{3}$$
$$\gamma = \frac{a-c}{2}$$

y

$$T^{-1}\begin{pmatrix} a & b \\ b & c \end{pmatrix} = \left(\frac{a+c}{2}\right)x^2 + \left(\frac{b}{3}\right)x + \left(\frac{a-c}{2}\right)\gamma$$

en concordancia con el resultado obtenido por métodos matriciales. En este caso los métodos matriciales resultaron más complicados, pero en general son preferibles pues proporcionan un camino seguro aunque, a veces, largo. El estudiante debe evaluar, ante un problema, si son preferibles métodos directos o matriciales. Casi todo problema matemático admite más de un camino hacia su solución.

El método matricial indica una extensión del principio del isomorfismo: los vectores se transforman en sus matrices de coordenadas, la transformación T se transforma en la matriz $M_B^A(T)$.

9.2. Cambio de base y matrices

Sean V y W dos espacios vectoriales y $T : V \to W$ una transformación lineal. Sean α y $\alpha\prime$ dos bases de V y β y $\beta\prime$ dos bases de W.

Sean $M = M_\beta^\alpha(T)$ y $N = M_{\beta\prime}^{\alpha\prime}(T)$ las correspondientes matrices.

Deseamos indagar cuál es la conexión entre M y N.

Para ello usaremos relaciones ya conocidas tales como

$$[T(v)]_\beta = M[v]_\alpha$$
$$[T(v)]_{\beta\prime} = N[v]_{\alpha\prime}$$

(aquí v es cualquier vector de v) y si P es la matriz de transición de α a $\alpha\prime$ en tanto que Q es la matriz de transición de β a $\beta\prime$ tendremos que, para cualesquiera $w \in V$ y $z \in W$

$$[w]_{\alpha\prime} = P[w]_\alpha$$
$$[z]_{\beta\prime} = Q[z]_\beta$$

Por ello

$$\begin{aligned} [T(v)]_{\beta\prime} &= Q[T(v)]_\beta \\ &= QM[v]_\alpha \\ &= QMP^{-1}[v]_{\alpha\prime} \\ &= N[v]_{\alpha\prime} \end{aligned}$$

y como esto vale para cualquier $[v]_{\alpha'}$ resulta que

$$QMP^{-1} = N$$

o

$$M = Q^{-1}NP$$

y estas últimas relaciones nos dicen cómo cambian las matrices cuando cambiamos de base tanto en el dominio como en el codominio.

Cuando $V = W$ (operador lineal) y $\alpha = \beta$ y $\alpha' = \beta'$ tendremos que $M = M_\alpha^\alpha(T)$ y $N = M_\beta^\beta(T)$ y también que $P = Q$ por lo que

$$M = P^{-1}NP$$

o

$$N = PMP^{-1}$$

Definición 181 *se dice que dos matrices M y N son similares si existe otra matriz P, no-singular y tal que $N = PMP^{-1}$. La relación de similaridad es una relación de equivalencia.*

9.3. Matrices “vacías”

En esta seccion consideraremos dos espacios V y W de dimensión finita con $\dim(V) > 0$ y $\dim(W) > 0$ (es decir, ninguno de los dos espacios es “trivial”). Considere el subespacio trivial $\Omega = \{0\}$ de V.

Ahora preguntémonos cómo serán las transformaciones lineales $A : \Omega \to W$. Como toda transformación lineal mapea el cero en el cero, concluimos que A es necesariamente la transformación cero. Note que, por lo anterior, sólo hay una trasnsformación lineal $A : \Omega \to W$.

¿y que pasa con las transformaciones lineales $B : W \to \Omega$? Resulta que necesariamente $B(w) = 0$ para todo $w \in W$. De nuevo la transformación cero es la única transformación lineal $B : W \to \Omega$.

Y, finalmente considere $C : \Omega \to \Omega$ lineal, para la cual también ocurre que es la transformación cero.

Si equipamos a W con una base, ¿cuál cree usted que serán las matrices de A, B y C? Si la dimensión de W es n, la matriz de A sería de $n \times 0$, la de B de $0 \times n$ y la de C ¡de 0×0!

De hecho algunos programas, como matlab, manejan estas “matrices vacías”. ¿podría usted elucidar cómo se multiplican las matrices vacías? ¿qué cree usted que resulta de multiplicar una matriz de $m \times 0$ con una de $0 \times n$? Respuesta: una matriz cero de $m \times n$.

Capítulo 10

Geometría de las transformaciones (I).

10.1. Haciendo geometría en $\mathbb{R}^2$

Es hora de recordar el curso de geometría analítica. Con cada vector de $\mathbb{R}^2$ asociamos tanto una "flecha" como un "punto" (la punta de la flecha). La flecha es el vector de posición del punto. La suma de vectores tiene una interpretación geométrica en términos de la regla del paralelogramo en tanto que la multiplicación por números (βa) cambia vectores en vectores paralelos pero con su tamaño alterado por un factor $\mid \beta \mid$ y su sentido cambiado si $\beta < 0$.

10.2. Efecto de transformaciones lineales sobre rectas y segmentos

Una recta puede pensarse como el conjunto

$$\mathcal{L} = \{P = P_0 + ta \mid t \in \mathbb{R}\}$$

donde $a \in \mathbb{R}^2$, $a \neq 0$, y P_0 es un vector dado de $\mathbb{R}^2$. Decimos que $\mathcal{L}$ es la recta paralela a a que pasa por P_0.

Un segmento de recta es un conjunto de la forma

$$\mathcal{L} = \{P = b + t(a - b) \mid t \in [0, 1]\}$$

es decir, pasa por b ($t = 0$), por a ($t = 1$) y es paralelo a $a - b$. Es el segmento que va de b a a y también podría escribirse como

$$\mathcal{L} = \{P = rb + sa \mid r, s \geq 0, r + s = 1\}$$

que se conoce como "combinación convexa de a y b". Las representaciones anteriores se llaman paramétricas pues todo depende de los parámetros reales r, s y t.

El punto clave es que las transformaciones lineales no singulares mapean rectas en rectas y segmento en segmentos. Veamos esto con mayor detalle.

Si $\mathcal{L}$ es una recta o un segmento de recta, y $T : \mathbb{R}^2 \to \mathbb{R}^2$ es lineal, definimos el efecto de T sobre $\mathcal{L}$ como

$$T(\mathcal{L}) = \{T(x) \mid x \in \mathcal{L}\}$$

y si $\mathcal{L}$ es una recta

$$T(\mathcal{L}) = \{T(P_0 + ta) \mid t \in \mathbb{R}\}$$

pero $T(P_0 + ta) = T(P_0) + tT(a)$ por lo que

$$T(\mathcal{L}) = \{T(P_0) + tT(a) \mid t \in \mathbb{R}\}$$

y vemos que $T(\mathcal{L})$ es una recta que pasa por $T(P_0)$ y es paralela a $T(a)$ (aquí hay que aclarar que si $T(a) = 0$ entonces la linea recta se mapearía al punto $T(P_0)$).

Algo completamente análogo ocurre con los segmentos.

También es fácil ver que si $\mathcal{L}$ y $\mathcal{L}'$ son dos rectas paralelas, entonces $T(\mathcal{L})$ y $T(\mathcal{L}')$ son paralelas. T mapea dos rectas que se cruzan en un punto en rectas que se cruzan en un punto. Por ello las transformaciones lineales mapean paralelogramos en paralelogramos.

A fin de asegurar que no vayamos a mapear una recta en un punto (como vimos más arriba) se suele restringir el ámbito de las transformaciones lineales a aquellas que son invertibles (el grupo lineal general $GL_2(\mathbb{R})$).

10.3. Areas de paralelogramos y de imágenes de paralelogramos bajo transformaciones lineales.

Sea $\mathbb{P}$ un paralelogramos generado por dos vectores a y b que son dos de sus lados adyacentes. El área de es el tamaño de su base$\mid a \mid$ multiplicado por su altura h de modo que

$$A =\mid a \mid h =\mid a \mid\mid b \mid sen(\theta) =\mid a \times b \mid$$

donde θ es el ángulo entre a y b (conceptos a ser definidos en el tema productos interiores). Pero de lo que sabemos del producto cruz

$$A =\mid a \times b \mid=\mid a_1 b_2 - a_2 b_1 \mid=\mid \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \mid$$

donde $a = (a_1, a_2)$ y $b = (b_1, b_2)$.

Si tenemos además la transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ cuya matriz, en la base canónica, es

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

De aquí vemos que si $\mathbb{P}$ es generado por los vectores a y b entonces $T(\mathbb{P})$ estará generado por los vectores $T(a)$ y $T(b)$ por lo que su área será

$$A' =\mid \det\left(\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}\right) = \det\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \det\left(\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}\right) \mid$$
$$=\mid \det\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \mid A$$

y tenemos el importante resultado de que la transformación cambia al paralelogramo en otro y multiplica su área por el determinante de la transformación.

10.4. Quiralidad.

Pero ¿y el signo del determinante? ¿significa algo?. Tiene que ver con quiralidad. Expliquemos.

Hay muchos objetos que vienen en dos variedades: derecha e izquierda. Por ejemplo, hay zapatos derechos y zapatos izquierdos, orejas derechas y orejas izquierdas, tornillos derechos (de rosca derecha) y tornillos izquierdos. Estos objetos se llaman quirales (del griego, keirós= mano, por ejemplo quiróptero quiere decir que tiene alas en las manos, la palabra cirujano quiere decir que cura con las manos, etc). Otros objetos no vienen en dos variedades derecha e izquierda, por ejemplo las pelotas de football; uno nunca pide una pelota derecha o izquierda. Tales objetos son aquirales. La propiedad de ser derecho o izquierdo se llama quiralidad (los físicos la llaman paridad, en ingles es handedness). Recomiendo al lector la obra de Lewis Carroll intitulada Alicia a través del espejo en donde la protagonista pasa al mundo del espejo, donde la quiralidad está cambiada (¿porqué un espejo cambia derecha e izquierda y no cambia arriba y abajo?).

El signo del determinante da la quiralidad. El área puede pensarse como con quiralidad, piense en el producto cruz, implica la regla de la mano derecha.

Si $\det(T) > 0$ no hay cambio en quiralidad, pero si $\det(T) <$ si lo hay y, como veremos en seguida, la transformación tiene un tanto el carácter de un espejo.

10.4.1. ejemplos

En los ejemplos que siguen, considere un cuadrado de área uno y su imagen bajo T(que es un paralelogramo).

- $M = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ es la identidad, no altera el cuadrado, ni le cambia la quiralidad pues $\det(M) = 1$
- $M = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ cambia cada vector por su negativo. Se llama inversión central y es, también, una rotación de 180 grados. No cambia la quiralidad pues $\det(M) = 1$ y no altera el área.

- $M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ refleja el cuadrado sobre el eje de las x. Como buen espejo cambia la quiralidad y $\det(M) = 1$, no altera el área
- .$M = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ refleja el cuadrado sobre el eje de las y. Como buen espejo cambia la quiralidad y $\det(M) = 1$, no altera el área
- $M = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ cambia al cuadrado en un paralelogramo que no es cuadrado. Sin embargo tiene la misma área pues $\det(M) = 1$ y no se altera la quiralidad. Es un ejemplo de corte ("shear") que también podría llamarse cizalla, cizalladura pero por favor evite la monstruosidad "trasquilado".
- $M = \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$ $k > 0$, alarga al cuadrado en la dirección x si k> 1(lo comprime si $k < 1$). $\det(M) = k$ refleja el cambio de área y muestra que no hay cambio de quiralidad.
- $M = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} = kI$ es una expansión o compresión uniforme en las dos direcciones. $\det(M) = k^2$.
- $M = \begin{pmatrix} k & 0 \\ 0 & l \end{pmatrix}$es una expansión o compresión diferente en los dos ejes. $\det(M) = kl$.
- $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, más dificil de visualizar. $\det(M) = -1$ por lo que cambia quiralidades sin alterar áreas. Se trata de una reflexión a través de un espejo $y = x$.
- $M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ es una transformación que todo lo proyecta sobre el eje x. (más adelante estudiaremos las proyecciones formalmente). $\det(M) = 0$ pues cambia el cuadrado en un segmento que no tiene área.
- $M = \begin{pmatrix} \cos(\theta) & -sen(\theta) \\ sen(\theta) & \cos(\theta) \end{pmatrix}$ representa una rotación por un ángulo θ. $\det(M) = 1$ por lo que no hay alteración ni en área ni en quiralidad. De hecho las rotaciones no cambian distancias ni ángulos: mapean un cuadrado en otro idéntico.

Lástima que este tema se ha colocado antes del tema de eigenvalores y eigenvectores, pues estos conceptos dicen mucho sobre el significado geométrico. Por ejemplo en una cizalladura siempre hay una (y sólo una) dirección en la que los vectores no cambian, es decir, hay una línea invariante.

Capítulo 11

eigenvalores y eigenvectores

11.1. algunas definiciones

Definición 182 *Sea V un espacio vectorial y $T : V \to V$ una transformación lineal. A estas transformaciones lineales de un espacio a él mismo se les llama operadores lineales.*

Definición 183 *Sea V un espacio vectorial sobre un campo F y $T : V \to V$ un operador lineal. Se dice que $v \in V$ es un eigenvector de T con eigenvalor $\lambda \in F$ si $v \neq 0$ y $Tv = \lambda v$.*

Algunos sinónimos de eigenvalor son: valor propio, autovalor, valor característico, valor latente.

Algunos sinónimos de eigenvector son: vector propio, vector característico, autovector, vector latente.

Comentario 184 *Algunos autores no exigen que $v \neq 0$, dicen ellos que 0 es eigenvector para cualquier eigenvalor.*

Definición 185 *Si V es un espacio vectorial, $T : V \to V$ es una transformación lineal y λ un eigenvalor de T, definimos el eigenespacio (o espacio propio, o espacio característico) E_λ mediante*

$$E_\lambda = \{v \in V \mid Tv = \lambda v\}$$

Note usted que el vector cero está en E_λ pero no es un eigenvector.

Teorema 186 *E_λ es un subespacio de V.*

Demostración. a) Claramente $E_\lambda \subset V$
b) Como $T(0) = 0 = \lambda 0$ tenemos que $0 \in E_\lambda$
c) Si $a, b \in E_\lambda$ entonces $Ta = \lambda a$ y $Tb = \lambda b$ por lo que $T(a+b) = T(a)+T(b) = \lambda a + \lambda b = \lambda(a + b)$ y $a + b \in E_\lambda$
d) Si $a \in E_\lambda$ y β es un escalar, como $T(a) = \lambda a$ tendremos que $T(\beta a) = \beta T(a) = \beta\lambda a = \lambda(\beta a)$ con lo que $\beta a \in E_\lambda$. ■

11.1.1. en términos de las matrices.

Suponga que $T : V \to V$ es un operador lineal, que v es un eigenvector con eigenvalor λ, que A es una base de V y que $M = M_A^A(T)$.

Entonces

$$\begin{aligned}[Tv]_A &= M[v]_A \\ [\lambda v]_A &= M[v]_A \\ \lambda[v]_A &= M[v]_A\end{aligned}$$

por lo que si llamamos $w = [v]_A$ tendremos que

$$Mw = \lambda w$$

y, en términos de coordenadas, la ecuación es de la misma forma $Tv = \lambda v$ o $Mw = \lambda w$. En lo sucesivo trabajaremos casi siempre con las matrices más que con las transformaciones.

Conversamente, si

$$M_A^A(T)[v]_A = \lambda[v]_A$$

tendremos que

$$\begin{aligned}[T(v)]_A &= M_A^A(T)[v]_A \\ &= \lambda[v]_A \\ &= [\lambda v]_A\end{aligned}$$

implicando que

$$T(v) = \lambda v$$

Es importante recalcar que el vector w contiene las coordenadas del eigenvector en la base A.

11.2. Hallando los eigenvalores

Como

$$\begin{aligned}Mw = \lambda w &\Leftrightarrow \\ Mw - \lambda w = 0 &\Leftrightarrow \\ Mw - \lambda I w = 0 &\Leftrightarrow \\ (M - \lambda I)w = 0 &\end{aligned}$$

vemos que los eigenvectores son los vectores no nulos en el núcleo de $(M - \lambda I)$. La ecuación $(M - \lambda I)w = 0$ representa un sistema lineal homogéneo cuya matriz de coeficientes es $M - \lambda I$. Sabemos (de lo visto en el tema de determinantes) que un sistema de este tipo tiene una solución no trivial si y sólo si el determinante de la matriz de coeficientes es cero, por ello

$$\det(M - \lambda I) = 0$$

ecuación que puede usarse para determinar los posibles valores de λ. Note que en esta ecuación no aparece w.

De hecho la función
$$P(\lambda) = \det(M - \lambda I)$$
es siempre un polinomio en λ. El grado de este polinomio es n si la matriz es de $n \times n$. El polinomio se llama polinomio característico de M(hay quien lo llama polinomio secular, o ecuación secular) y vemos, pues, que los eigenvalores son las raices de $P(\lambda)$.

Cuando $n = 2$ las raices del polinomio característico se hallan mediante la fórmula elemental para hallar las raices del polinomio $ax^2 + bx + c = 0$ que es $x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$. Cuando el polinomio es de grado 3 o 4 es mejor suponer que se trata de un polinomio con coeficientes enteros con una raiz entera: esta raiz debe de ser un divisor del término de grado cero. En los restantes casos se habrán de usar métodos numéricos ad hoc.

11.3. hallando los eigenvectores.

Sabiendo cuáles son los eigenvalores hay que resolver ahora
$$Mw = \lambda w$$
para cada eigenvalor. El conjunto solución proporcionará, además, bases para los eigenespacios.

11.3.1. primeros ejemplos

Sea $I_V : V \to V$ la transformación identidad en V ($\dim(V) = m$). Como
$$I_V x = x = 1x$$
para todo $x \in V$ resulta que todo vector $x \neq 0$ de V es un eigenvector de I_V con eigenvalor 1 . Sólo hay un eigenespacio
$$E_1 = V$$
pues todo vector x de V satisface $I_V x = 1x$.

El polinomio característico será
$$\begin{aligned} P_{I_V}(\lambda) &= \det(I_V - \lambda I_V) = \det((1-\lambda)I_V) \\ &= (1-\lambda)^m \det(I_V) = (1-\lambda)^m \end{aligned}$$
cuya única raíz es 1 con multiplicidad m .

A continuación considere $O_V : V \to V$, la transformación cero en V . Como para todo $x \in V$
$$O_V x = \overrightarrow{0} = 0\,\overrightarrow{0}$$

tenemos que todo vector $x \neq 0$ de V es un eigenvector de 0_V con eigenvalor 0. Sólo hay un eigenespacio

$$E_0 = V$$

pues todo vector x de V satisface $0_V x = 0x$.

El polinomio característico será

$$\begin{aligned} P_{0_V}(\lambda) &= \det(0_V - \lambda I_V) \\ &= \det(\lambda I_V) \\ &= \lambda^m \det(I_V) \\ &= \lambda^m \end{aligned}$$

y su única raíz es $\lambda = 0$ con multiplicidad m .

Por último considere el caso de un operador en un espacio REAL cuya matriz sea

$$M = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$$

con lo que

$$P_M(\lambda) = (1-\lambda)^2 + 4$$

que no tiene raices reales. Por lo tanto el operador no tiene ni eigenvalores ni eigenvectores.

Estos son ejemplos extremos. En general algunos vectores de V serán eigenvectores y otros no.

11.4. Un ejemplo sencillo.

Considere la matriz

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

un cálculo trivial indica que

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

por lo que $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ es un eigenvector de A con eigenvalor 2.

veamos ahora más sistemáticamente cuáles son los eigenvalores.

El polinomio característico es

$$\begin{aligned} p(\lambda) &= \det\left[A - \lambda I\right] \\ &= \det\left(\begin{bmatrix} 2 & 3 \\ 0 & 4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \\ &= \det\left(\begin{bmatrix} 2-\lambda & 3 \\ 0 & 4-\lambda \end{bmatrix} \right) \\ &= (2-\lambda)(4-\lambda) \end{aligned}$$

con lo que los eigenvalores son

$$\{2, 4\}$$

Para hallar los eigenvectores hay que usar la ecuación $Ax = \lambda x$ usando los eigenvalores encontrados. Así, para $\lambda = 2$ tendremos

$$\begin{bmatrix} 2 & 3 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2 \begin{bmatrix} x \\ y \end{bmatrix}$$

o

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

sistema que tiene como solución

$$y = 0$$

y cualquier x, es decir, el espacio solución tiene como una posible base

$$\{(1, 0)\}$$

que es también la base (eigenbase) del eigenespacio E_2 y que, por consiguiente tiene $\dim(E_2) = 1$.

Análogamente, para $\lambda = 4$ tenemos que

$$\begin{bmatrix} 2 & 3 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 4 \begin{bmatrix} x \\ y \end{bmatrix}$$

y

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

sistema que tiene como espacio solución los múltiplos de

$$\{(\frac{3}{2}, 1)\}$$

y también en este caso $\dim(E_4) = 1$.

11.5. Para obtener eigenvalores y eigenvectores.

- Obtenga el polinomio característico $P_T(\lambda) = \det(T - \lambda I_V)$
- Obtenga las raices λ_i $(i = 1...m)$ de $P_T(\lambda)$. Estos son los posibles eigenvalores.
- Para cada λ_i resuelva el sistema lineal homogéneo $Av_i - \lambda_i v_i = (A - \lambda_i)v_i = 0$. Toda solución no nula es un eigenvector. El eigenespacio correspondiente a λ_i , E_{λ_i} es simplemente el espacio solución de este sistema homogéneo. Una base del eigenespacio será, pués, una base del correspondiente sistema lineal homogéneo.

Con esto obtendremos todos los espacios característicos (o eigenespacios) y sus bases.

11.6. Matriz de un operador en una eigenbase.

Suponga ahora que la base α consiste de eigenvectores, con lo que $Tv_i = \lambda_i v_i$ $(i = 1, 2...n)$ y, por lo tanto, la matriz de T respecto a α es

$$M = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ \vdots & \vdots & & \\ 0 & 0 & ... & \lambda_n \end{pmatrix}$$

es decir, una matriz diagonal con los eigenvalores sobre la diagonal.

11.6.1. independencia de eigenvectores.

Tenemos que los eigenvectores correspondientes a diferentes eigenvalores son independientes.

Teorema 187 *Si $\{v_1, v_2..v_n\}$ son eigenvectores correspondientes a eigenvalores distintos, entonces son linealmente independientes.*

Demostración. : Por inducción en el número n de elementos. Para $n = 1$ sólo hay un eigenvector v_1, como $v_1 \neq 0$ (por ser eigenvector) el conjunto $\{v_1\}$ es independiente. Suponga ahora que el teorema es válido para $n = k$. Considere el conjunto $\{v_1, v_2..v_k, v_{k+1}\}$ y la ecuación de dependencia

$$c_1v_1 + c_2v_2 + ...c_kv_k + c_{k+1}v_{k+1} = 0$$

de donde

$$T(c_1v_1 + c_2v_2 + ...c_kv_k + c_{k+1}v_{k+1}) = 0$$
$$c_1\lambda_1v_1 + c_2\lambda_2v_2 + ...c_k\lambda_kv_k + c_{k+1}\lambda_{k+1}v_{k+1} = 0$$

y, multiplicando la ecuación $c_1v_1 + c_2v_2 + ...c_kv_k + c_{k+1}v_{k+1} = 0$ por $\lambda_{k+1}y$ restando ambas ecuaciones nos queda

$$(\lambda_1 - \lambda_{k+1})c_1v_1 + (\lambda_2 - \lambda_{k+1})c_2v_2 + ...(\lambda_k - \lambda_{k+1})c_kv_k = 0$$

pero, por la hipótesis de inducción $\{v_1, v_2..v_k\}$ es independiente de modo que

$$\begin{aligned}
(\lambda_1 - \lambda_{k+1})c_1 &= 0 \\
(\lambda_2 - \lambda_{k+1})c_2 &= 0 \\
&\vdots \\
(\lambda_k - \lambda_{k+1})c_k &= 0
\end{aligned}$$

y ninguno de los términos $(\lambda_1 - \lambda_{k+1})$ es cero pues los eigenvalores son todos distintos. Por ello

$$c_1 = c_2... = c_k = 0$$

y $c_{k+1}v_{k+1} = 0$ con lo que $c_{k+1} = 0$ y $\{v_1, v_2..v_k, v_{k+1}\}$ es independiente completando la demostración. ■

11.6.2. Matriz diagonalizadora

Vamos a considerar de nuevo los cambios de base. Lo haremos para operadores por lo que nuestro tratamiento será un caso particular de la discusión ya presentada anteriormente, pero es de tal importancia que bien vale la pena.

Sean V un espacio vectorial y $T : V \to V$ una transformación lineal. Sean α y $\alpha\prime$ dos bases de V y $M = M_\alpha^\alpha(T)$ y $N = M_{\alpha'}^{\alpha\prime}(T)$ las correspondientes matrices.

Deseamos indagar cuál es la conexión entre M y N.

Para ello usaremos relaciones ya conocidas tales como

$$\begin{aligned}
[T(v)]_\alpha &= M[v]_\alpha \\
[T(v)]_{\alpha'} &= N[v]_{\alpha\prime}
\end{aligned}$$

(aquí v es cualquier vector de v) y si P es la matriz de transición de α a $\alpha\prime$ tendremos que, para cualesquiera $w, z \in V$

$$[w]_{\alpha\prime} = P[w]_\alpha$$

Por ello

$$\begin{aligned}
[T(v)]_{\alpha\prime} &= P[T(v)]_\alpha \\
&= PM[v]_\alpha \\
&= PMP^{-1}[v]_{\alpha\prime} \\
&= N[v]_{\alpha\prime}
\end{aligned}$$

y como esto vale para cualquier $[v]_{\alpha\prime}$ resulta que

$$PMP^{-1} = N$$

o

$$M = P^{-1}NP$$

y estas últimas relaciones nos dicen cómo cambian las matrices cuando cambiamos de base.

Ahora tomaremos el caso en que α' es una cierta base y α es una eigenbase. Entonces M será una matriz diagonal en cuya diagonal se encuentran los eigenvalores, llamemos D a esta matriz para recordar que es diagonal. Entonces

$$D = P^{-1}NP$$

Recuerde que P es la matriz de transición de la base α a la base α'. Esta matriz tiene como columnas las coordenadas de la eigenbase en términos de la base original, es decir los eigenvectores tal y como se han calculado con el procedimiento indicado (polinomio característico y sus raices resolver el problema lineal homogéneo correspondiente $(A - \lambda I)v = 0$). En este caso P recibe el nombre de matriz diagonalizadora. Recuerde: P tiene como columnas los eigenvectores.

Comentario 188 *Si* $\dim(V) = n$ *entonces el polinomio característico será de grado n y tendrá n raices (sobre el campo de los números complejos). Si éstas son distintas, los eigenvectores formarán una eigenbase y el operador será diagonalizable. Pero si algunos eigenvalores se repiten, entonces el operador puede ser o no ser diagonalizable. Un teorema importante que no vamos a explicar mayormente establece que, si un eigenvalor λ aparece con multiplicidad m entonces* $\dim(E_\lambda) \leq m$.

11.7. Un ejemplo completo

Considere la matriz $A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$. ¿es A diagonalizable? En caso afirmativo calcular la matriz diagonalizadora.

El primer paso es el cálculo del polinomio característico. Por ello evaluemos

$$
\begin{aligned}
p(\lambda) &= \det \left[\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right] \\
&= \det \left[\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \right] \\
&= \det \begin{bmatrix} 2-\lambda & 1 & 1 \\ 2 & 3-\lambda & 2 \\ 3 & 3 & 4-\lambda \end{bmatrix} \\
&= (2-\lambda)\,[(3-\lambda)(4-\lambda)-(3)(2)] \\
&- (1)\,[(2)(4-\lambda)-(3)(2)] \\
&+ (1)\,[(2)(3)-(3)(3-\lambda)] \\
&= (2-\lambda)\left[(\lambda^2-7\lambda+12-6\right] \\
&- [(8-2\lambda-6]+[(6-9+3\lambda)] \\
&= (2-\lambda)\left[(\lambda^2-7\lambda+6\right]-(2-2\lambda)+(-3+3\lambda) \\
&= 2\lambda^2-14\lambda+12-\lambda^3+7\lambda^2-6\lambda-2+2\lambda-3+3\lambda \\
&= -\lambda^3+9\lambda^2-15\lambda+7
\end{aligned}
$$

o, en resumidas cuentas

$$p(\lambda) = -\lambda^3 + 9\lambda^2 - 15\lambda + 7$$

El segundo paso es hallar los eigenvalores, que son las raices del polinomio característico. En este caso el polinomio es de grado tres y tiene coeficientes enteros por lo que buscamos una raíz entre los divisores de 7 que son ± 1 y ± 7. Una evaluación explícita nos dice que

$$p(1) = -1 + 9 - 15 + 7 = 0$$

de modo que el primer eigenvalor encontrado es $\lambda_1 = 1$. Por ello el polinomio p puede factorizarse como

$$p(\lambda) = (\lambda - 1)q(\lambda)$$

donde q es un polinomio de grado dos. q puede ser hallado mediante división sintética $\left[\begin{array}{cccc|c} -1 & 9 & -15 & 7 & \\ & -1 & 8 & -7 & \\ \hline -1 & 8 & -7 & 0 & 1 \end{array}\right]$ de modo que

$$q(\lambda) = -\lambda^2 + 8\lambda - 7$$

cuyas raices son

$$
\begin{aligned}
\lambda_{2,3} &= \frac{-(8) \pm \sqrt{(8)^2 - (4)(-1)(-7)}}{(2)(-1)} \\
&= \frac{-8 \pm \sqrt{36}}{-2} = \frac{-8 \pm 6}{-2} \\
&= \begin{cases} 1 \\ 7 \end{cases}
\end{aligned}
$$

y

$$
\begin{aligned}
\lambda_1 &= 1 \\
\lambda_2 &= 1 \\
\lambda_3 &= 7
\end{aligned}
$$

observando que el eigenvalor 1 es de multiplicidad 2.

El siguiente paso es calcular los eigenvectores. Comencemos con $\lambda_3 = 7$. La ecuación de eigenvectores es

$$
\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 7 \begin{bmatrix} x \\ y \\ z \end{bmatrix}
$$

o

$$
\begin{aligned}
2x + y + z &= 7x \\
2x + 3y + 2z &= 7y \\
3x + 3y + 4z &= 7z
\end{aligned}
$$

que es equivalente a

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

cuya matriz de coeficientes es

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

y un sistema equivalente al original es

$$x - \frac{z}{3} = 0$$
$$y - \frac{2z}{3} = 0$$

y como x, y son variables ligadas en tanto que z es la única variable libre

$$(x, y, z) = (\frac{z}{3}, \frac{2z}{3}, z)$$

por lo que, si tomo $z = 3$ (para evitar fracciones), una base para E_7 será $\{(1, 2, 3)\}$ y $\dim(E_7) = 1$.

A continuación tomamos el eigenvalor 1 (de multiplicidad dos).

La ecuación de eigenvector es, en este caso

$$\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 1 \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

o

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

y pasando todo al lado izquierdo obtenemos el sistema homogéneo

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

cuya matriz de coeficientes es

$$\begin{aligned}
\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} &\approx \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \\
&\approx \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\end{aligned}$$

y un sistema equivalente es

$$x + y + z = 0$$

aquí x es la única variable ligada, las variables libres son y y z. Por ello

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

y una posible base para E_1 sería $\{(-1, 1, 0), (-1, 0, 1)\}$ y $\dim(E_1) = 2$.

Por todo esto, una base para $\mathbb{R}^3$ que sea eigenbase de A será

$$\beta = \{(1, 2, 3), (-1, 1, 0), (-1, 0, 1)\}$$

y si T es el operador cuya matriz es A (en alguna base no especificada α) entonces

$$\begin{aligned}
M_\alpha^\alpha(T) &= \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix} \\
M_\beta^\beta(T) &= \begin{bmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

note que la última matriz es diagonal y contiene a los eigenvalores en el orden elegido.

Por último, hay que calcular la matriz diagonalizadora. Para ello recuerde que los vectores obtenidos (es decir $(1,2,3), (-1,1,0), (-1,0,1)$) contienen las coordenadas de los eigenvectores en la base α (de la que no sabemos explícitamente quién es). Por ello la matriz que contiene, como columnas, estos números es la matriz de transición P de la base β a la base α, en nuestro caso

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

y podemos fácilmente corroborar que en efecto

$$\begin{aligned} & P^{-1}AP \\ &= \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Por cierto, si le piden que compruebe una ecuación del tipo $D = P^{-1}AP$ mejor compruebe la ecuación equivalente $AP = PD$, es más fácil (no hay que calcular ninguna inversa).

Comentario 189 *De la ecuación* $p(\lambda) = \det(A - \lambda I)$ *podemos ver que* $p(0) = \det(A)$ *por lo que el término de grado cero del polinomio característico (término independiente) es igual al determinante. También puede demostrarse que, si* A *es de* $n \times n$*, entonces el coeficiente del término de grado* $n-1$ *es igual a* $(-1)^{n+1} tr(A)$*.*

Comentario 190 *Dado que la traza y el determinante no cambian bajo transformaciones de semejanza; si* A *es diagonalizable entonces su traza es la suma de sus eigenvalores y su determinante el producto. En realidad esto es cierto aun sin la suposición de diagonalizabilidad para espacios complejos (toda matriz compleja es semejante a una matriz triangular superior donde los elementos de la diagonal son los eigenvalores).*

Capítulo 12

Interpretación geométrica de las transformaciones lineales (II).

Cuando la matriz de la transformación es $M = \begin{pmatrix} k & 0 \\ 0 & l \end{pmatrix}$ ya hemos visto que se trata de expansiones (o contracciones) por factores k y l en las direcciones x y y respectivamente. Notemos que esta matriz tiene eigenvalores precisamente k y l y que los eigenvectores son $(0,1)$ y $(1,0)$.

Si la matriz es real y del tipo $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, es decir, una matriz simétrica, ocurre (como veremos más adelante) que M es diagonalizable y existe una matriz ortogonal P tal que $P^{-1}MP$ es diagonal. Esto quiere decir que hay un sistema coordenado (dado por los eigenvectores de M) en el cual la transformación es del tipo $M = \begin{pmatrix} k & 0 \\ 0 & l \end{pmatrix}$. Luego, toda matriz simétrica representa expansiones o contracciones con respecto a dos ejes perpendiculares.

Los espejos, como $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, tienen dos eigenvectores perpendiculares, uno de los eigenvalores es 1 y el otro es -1, indicando que en la dirección del espejo nada cambia (linea invariante) y en la dirección perpendicular se cambia el signo ("se refleja").

Mucho más interesantes son los cortes del tipo $M = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$, estos operadores no son diagonalizables. Hay un sólo eigenvector ($(1,0)$ en este caso) y corresponde al eigenvalor 1. Decimos que hay tan sólo una línea invariante, es característica de los cortes el tener una línea (o plano en tres dimensiones) invariante. En cursos más avanzados una matriz como M se dice que está en forma de bloque de Jordan.

Parte IV

Espacios con producto interior

Capítulo 13

Espacios con producto interior

13.1. Motivación

Todo estudiante seguramente conoce el producto “punto” o “escalar” que le enseñaron desde preparatoria. Es el producto que usamos en los cursos elementales de física para calcular el trabajo hecho por una fuerza ($W = \vec{F} \cdot \vec{d}$).

Este producto se define mediante

$$\vec{a} \cdot \vec{b} = \mid \vec{a} \mid \mid \vec{b} \mid \cos(\theta)$$

donde $\mid \vec{a} \mid$y $\mid \vec{b} \mid$ representan los tamaños de $\vec{a}$ y de $\vec{b}$ respectivamente y θ es el ángulo entre los dos vectores

Aquí no podemos usar este enfoque porque:

- no se ha definido lo que significa el tamaño de un vector
- no se ha definido lo que significa el ángulo entre dos vectores.

13.2. Concepto de producto interior

Definición 191 *Decimos que V es un espacio con producto interior si es un espacio vectorial real (complejo) y hay una función*

$$\langle\,\rangle : V \times V \to \mathbb{R}$$

($\langle\,\rangle : V \times V \to \mathbb{C}$ según el caso) que asocia con cada par de vectores $a, b \in V$ un número real (complejo) $\langle a, b \rangle$ y que satisface los siguientes cuatro axiomas:

1)Si $a, b \in V$ entonces $\langle a, b \rangle = \overline{\langle b, a \rangle}$. La barra denota conjugación compleja.
2)Si $a, b, c \in V$ entonces $\langle a + b, c \rangle = \langle a, c \rangle + \langle b, c \rangle$
3) Si $a \in V$ y $\beta \in \mathbb{C}$ (o $\beta \in \mathbb{R}$) entonces $\langle \beta c, a \rangle = \beta \langle c, a \rangle$
4) Si $a \in V$ entonces $\langle a, a \rangle \geq 0$. $\langle a, a \rangle = 0$ si y sólo si $a = 0$.

Notación 192 *Los productos interiores pueden denotarse de diversas maneras amén de la ya explicada* $\langle a, b\rangle$, *entre otras podemos mencionar* $(a \mid b)$, $a \cdot b$, $\langle a \mid b\rangle$ *y* (a, b). *Entiéndase que no es terriblemente importante cuál signo de agrupación usemos (()⟨⟩) o cuál separador empleemos (,|). Los signos de agrupación* $<$ *y* $>$ *se conocen en inglés como "brackets" y (hasta donde sabemos) en español son "paréntesis angulares" (creemos que no se llaman ni llaves {} ni corchetes [] ni comillas «»). En estas notas usaremos preferentemente la notación* $\langle a, b\rangle$ *para los productos interiores. En los textos de mecánica cuántica se usa invariablemente la notación de Dirac* $\langle a \mid b\rangle$. *Como sinónimos podemos usar productos internos y productos escalares, aunque hay que aclarar que algunos autores reservan este último nombre para el producto habitual en espacios* $\mathbb{R}^n$ *(que se verá en seguida).*

Comentario 193 *En la mecánica cuántica el axioma tres tiene la forma: 3) Si* $a \in V$ *y* $k \in \mathbb{C}$ *(o* $k \in \mathbb{R}$ *) entonces* $\langle c, ka\rangle = k\langle c, a\rangle$. *¿no sería mejor usar en este curso las convenciones habituales en la mecánica cuántica, en donde el álgebra lineal tiene su más perfecta y acabada aplicación?. En lenguaje técnico; vamos a usar productos lineales con respecto a su primer factor y antilineales con respecto al segundo, pero en mecánica cuántica los productos son siempre antilineales en el primer factor y lineales en el segundo.*

Comentario 194 *Note que en esta sección es importante que los espacios sean específicamente reales o complejos. En las secciones anteriores casi todo era válido para espacios sobre campos arbitrarios; pero en la definición del producto interior se usan la conjugación de complejos y el orden (*$\leq$*) en los reales.*

13.3. Ejemplos de productos interiores

Ejemplo 195 *Considere* $\mathbb{R}^n$ *y defina*

$$\langle a, b\rangle = \sum_{i=1}^{n} a_i b_i$$

si

$$a = (a_1, a_2, ...a_n)$$
$$b = (b_1, b_2, ...b_n)$$

Entonces:
a) $\langle b, a\rangle = \sum_{i=1}^{n} b_i a_i = \sum_{i=1}^{n} a_i b_i = \langle a, b\rangle$ *(note que hemos usado la conmutatividad del producto de números reales)*

b) Si $c = (c_1, c_2, ...c_n)$ *tendremos que*

$$\begin{aligned}\langle a+b, c\rangle &= \sum_{i=1}^{n}(a_i+b_i)c_i \\ &= \sum_{i=1}^{n} a_i c_i + \sum_{i=1}^{n} b_i c_i \\ &= \langle a, c\rangle + \langle b, c\rangle\end{aligned}$$

c)si $\beta \in \mathbb{R}$

$$\begin{aligned}\langle \beta a, b\rangle &= \sum_{i=1}^{n}(\beta a_i)b_i \\ &= \beta\sum_{i=1}^{n} a_i b_i \\ &= \beta\langle a, b\rangle\end{aligned}$$

d) claramente si $a = 0$ *tendremos que*

$$\begin{aligned}\langle a, a\rangle &= \sum_{i=1}^{n} 00 \\ &= 0\end{aligned}$$

y, conversamente, si

$$\begin{aligned}\langle a, a\rangle &= 0 \\ &= \sum_{i=1}^{n} a_i^2\end{aligned}$$

pero la suma de números no negativos no puede ser cero a menos que todos los sumandos sean cero de modo que si $\langle a, a\rangle = 0$ *entonces* $a_i = 0$ $\forall i$ *y* $a = 0$. *Además, como* $\sum_{i=1}^{n} a_i^2$ *es la suma de términos no negativos, debe de satisfacer* $\sum_{i=1}^{n} a_i^2 \geq 0$.

Ejemplo 196 *Análogamente, en* $\mathbb{C}^n$ *definimos*

$$\langle a, b\rangle = \sum_{i=1}^{n} a_i \bar{b}_i$$

si

$$\begin{aligned}a &= (a_1, a_2, ...a_n) \\ b &= (b_1, b_2, ...b_n)\end{aligned}$$

y dejamos la demostración al lector. Note la barra de conjugación en la definición.

Comentario 197 *El producto interior en* $\mathbb{R}^n$ *puede escribirse como*

$$\langle a, b\rangle = ab^T$$

si pensamos que a *y* b *son un vectores renglón. Alternativamente,*

$$\langle a, b\rangle = b^T a$$

si pensamos que a *y* b *son un vectores columna.*

Comentario 198 *El producto en* $\mathbb{C}^n$ *puede escribirse como*

$$\langle a, b\rangle = ab^*$$

si pensamos que a *y* b *son un vectores renglón. Alternativamente,*

$$\langle a, b\rangle = b^* a$$

si pensamos que a *y* b *son un vectores columna.*

Ejemplo 199 *En* P_n *un producto frecuente es el definido por*

$$\langle P(x), Q(x)\rangle = \int_a^b P(x)Q(x)dx$$

donde a *y* b *son números reales fijos* $(a < b)$*. Es un producto interior pues:*
a)

$$\begin{aligned}
\langle P(x), Q(x)\rangle &= \int_a^b P(x)Q(x)dx \\
&= \int_a^b Q(x)P(x)dx \\
&= \langle Q(x), P(x)\rangle
\end{aligned}$$

b)

$$\begin{aligned}
\langle kP(x), Q(x)\rangle &= \int_a^b kP(x)Q(x)dx \\
&= k\int_a^b P(x)Q(x)dx \\
&= k\langle P(x), Q(x)\rangle
\end{aligned}$$

c)

$$\begin{aligned}
\langle P(x)+Q(x),R(x)\rangle &= \int_a^b (P(x)+Q(x))R(x)dx \\
&= \int_a^b (P(x)R(x)+Q(x)R(x))dx \\
&= \int_a^b P(x)R(x)dx + \int_a^b Q(x)R(x)dx \\
&= \langle P(x),R(x)\rangle + \langle Q(x),R(x)\rangle
\end{aligned}$$

donde $R(x) \in P_n$*.*
d) Si $P(x) = 0$ *claramente*

$$\langle P(x),P(x)\rangle = \int_a^b 0dx = 0$$

y, conversamente, si

$$\begin{aligned}
\langle P(x),P(x)\rangle &= 0 \\
&= \int_a^b P(x)^2 dx
\end{aligned}$$

pero siendo los polinomios funciones continuas, la única manera de tener una integral cero para un integrando no negativo es que el integrando sea cero. Siendo el integrando no negativo, necesariamente se cumple que $\int_a^b P(x)^2 dx \geq 0$*.*

Ejemplo 200 *Si sobre una partícula actúa una fuerza* f *y se desplaza* d *entonces el trabajo* W *estará dado por* $W = f \cdot d$ *donde* $\cdot$ *es el producto ordinario en espacios de dos o tres dimensiones. Cuando la fuerza es variable (en una dimensión)*

$$W = \int_a^b f(t)v(t)dt$$

que tiene la misma forma que el producto interior en P_n*. Similarmente, en un circuito con corriente* $i(t)$ *y diferencia de potencial* $v(t)$ *tendremos que*

$$W = \int_a^b i(t)v(t)dt$$

que de nuevo tiene la forma de producto interior entre funciones.

Ejemplo 201 *En espacios de matrices, digamos* $M^*(m,n)$ *un producto habitual es el dado por*

$$\langle M,N\rangle = tr(MN^*)$$

y dejo al lector la demostración. Algunos autores llaman a este producto "producto de Frobenius".

Ejemplo 202 *En espacios de polinomios P_n se define el producto*

$$\langle P(x), Q(x)\rangle = \sum_{i=1}^{n+1} P(x_i)Q(x_i)$$

donde los números $x_1 ... x_{n+1}$ son distintos. Demuestre usted que es un producto interior.

13.4. Algunas propiedades de los productos interiores

De los axiomas podemos obtener algunas propiedades básicas de los productos interiores.

Teorema 203 *Sea V un espacio vectorial con un producto interior. Entonces:*
a) $\langle a, b+c\rangle = \langle a, b\rangle + \langle a, c\rangle$ para cualesquiera a,b y $c \in V$
b) $\langle a, \beta b\rangle = \overline{\beta}\langle a, b\rangle$ para cualesquiera a,b $\in V$ y β escalar.
c) $\langle a, 0\rangle = \langle 0, a\rangle = 0$

Demostración. a)

$$\begin{aligned}\langle a, b+c\rangle &= \overline{\langle b+c, a\rangle} \\ &= \overline{\langle b, a\rangle + \langle c, a\rangle} \\ &= \overline{\langle b, a\rangle} + \overline{\langle c, a\rangle} \\ &= \langle a, b\rangle + \langle a, c\rangle\end{aligned}$$

b)

$$\begin{aligned}\langle a, \beta b\rangle &= \overline{\langle \beta b, a\rangle} \\ &= \overline{\beta\langle b, a\rangle} \\ &= \overline{\beta}\overline{\langle b, a\rangle} \\ &= \overline{\beta}\langle a, b\rangle\end{aligned}$$

c) $\langle a, \overrightarrow{0}\rangle = \langle a, 0b\rangle = 0\langle a, b\rangle = 0$ (b es cualquier vector, la demostración de $\langle 0, a\rangle = 0$ es similar) ■

13.5. La desigualdad de Cauchy-Schwarz

La llamada desigualdad de Cauchy-Schwarz juega un papel muy importante tanto en el álgebra lineal como en sus aplicaciones. En álgebra lineal está en el corazón de muchas construcciones geométricas tales como las normas y distancias. En mecánica cuántica da lugar a las famosas relaciones de incertidumbre de Heisenberg.

Teorema 204 *Sea V un espacio vectorial complejo con un producto interior. Entonces para cualesquiera vectores a y b tenemos que*

$$|\langle a,b\rangle|^2 \leq \langle a,a\rangle\langle b,b\rangle$$

y la igualdad se cumple si y sólo si los vectores son linealmente dependientes.

Demostración. Si $b=0$ la desigualdad se cumple trivialmente. Suponga, pues, que $b \neq 0$ y considere la expresión

$$\begin{aligned}
\rho &= \langle a - \frac{\langle a,b\rangle}{\langle b,b\rangle} b, a - \frac{\langle a,b\rangle}{\langle b,b\rangle} b\rangle \\
&= \langle a,a\rangle - \langle a, \frac{\langle a,b\rangle}{\langle b,b\rangle} b\rangle - \langle \frac{\langle a,b\rangle}{\langle b,b\rangle} b, a\rangle \\
&+ \langle \frac{\langle a,b\rangle}{\langle b,b\rangle} b, \frac{\langle a,b\rangle}{\langle b,b\rangle} b\rangle \\
&= \langle a,a\rangle - \overline{\frac{\langle a,b\rangle}{\langle b,b\rangle}} \langle a,b\rangle - \frac{\langle a,b\rangle}{\langle b,b\rangle}\langle b,a\rangle \\
&+ \frac{\langle a,b\rangle}{\langle b,b\rangle}\overline{\frac{\langle a,b\rangle}{\langle b,b\rangle}}\langle b,b\rangle \\
&= \langle a,a\rangle - 2\frac{\langle a,b\rangle}{\langle b,b\rangle}\langle b,a\rangle + \frac{\langle a,b\rangle}{\langle b,b\rangle}\langle b,a\rangle \\
&= \langle a,a\rangle - \frac{\langle a,b\rangle}{\langle b,b\rangle}\langle b,a\rangle
\end{aligned}$$

pero $\rho \geq 0$ pues es el producto de un vector consigo mismo, por ello

$$\begin{aligned}
\langle a,a\rangle - \frac{\langle a,b\rangle}{\langle b,b\rangle}\langle b,a\rangle &\geq 0 \\
\langle a,a\rangle &\geq \frac{\langle a,b\rangle}{\langle b,b\rangle}\langle b,a\rangle \\
\langle a,b\rangle\langle b,a\rangle &\leq \langle a,a\rangle\langle b,b\rangle \\
|\langle a,b\rangle|^2 &\leq \langle a,a\rangle\langle b,b\rangle
\end{aligned}$$

demostrando la primera parte del teorema.
Suponga ahora que

$$|\langle a,b\rangle|^2 = \langle a,a\rangle\langle b,b\rangle$$

de las ecuaciones anteriores se ve que esto implica que

$$\langle a - \frac{\langle a,b\rangle}{\langle b,b\rangle} b, a - \frac{\langle a,b\rangle}{\langle b,b\rangle} b\rangle = 0$$

y, por lo tanto, que

$$\begin{aligned}
a - \frac{\langle a,b\rangle}{\langle b,b\rangle} b &= 0 \\
a &= \frac{\langle a,b\rangle}{\langle b,b\rangle} b
\end{aligned}$$

probando que $\{a, b\}$ es dependiente.
Conversamente, si $\{a, b\}$ es dependiente, entonces $a = \beta b$ para algún β y

$$|\langle \beta b, b\rangle|^2 = |\beta|^2 |\langle b, b\rangle|^2$$

y

$$\begin{aligned}\langle a, a\rangle\langle b, b\rangle &= \langle \beta b, \beta b\rangle\langle b, b\rangle \\ &= \beta\overline{\beta}\langle b, b\rangle\langle b, b\rangle \\ &= |\beta|^2 |\langle b, b\rangle|^2\end{aligned}$$

■

13.6. Normas

¡El tamaño si importa!

13.6.1. motivación.

En el producto escalar elemental $a \cdot b = |a||b|\cos(\theta)$ obtenemos para el caso especial $a = b$ que $a \cdot a = |a|^2$ por lo que

$$|a| = \sqrt{a \cdot a}$$

y donde se entiende que tomamos la raíz positiva. Es decir, del producto podemos inferir el tamaño.

13.6.2. Caso general.

En un espacio vectorial con producto interior podemos definir la noción de "tamaño del vector" o norma.

Definición 205 *Sea V un espacio vectorial sobre el campo de los números reales o los números complejos. Por una norma en V entenderemos una función $\|\cdot\| : V \to \mathbb{R}$ que asocia con cada vector x un número real $\|x\|$ que cumple con las siguientes propiedades:*
a) $\|x\| \geq 0$ $\forall x \in V$ y $\|x\| = 0$ si y sólo si $x = 0$.
b) $\|\gamma x\| = |\gamma| \|x\|$ $\forall x \in V$ y $\forall\gamma \in F$ (algunos autores llaman a esto propiedad de homogeneidad).
c) $\|x + y\| \leq \|x\| + \|y\|$ (desigualdad del triángulo).

Definición 206 *Un espacio vectorial V con una norma $\|\cdot\|$ se llama espacio normado.*

Ejemplo 207 *En $\mathbb{C}^n$ y en $\mathbb{R}^n$ tenemos, como ejemplos de normas,*

$$\| v \|_p = \left[\sum_{i=1}^{n}(|v_i|)^p\right]^{\frac{1}{p}}$$

y, cuando $p \to \infty$,

$$\parallel v \parallel_\infty = \text{máx}\{\mid v_i \mid\}$$

Si $p = 1$ tendremos

$$\parallel v \parallel_1 = \sum_{i=1}^{n} (\mid v_i \mid)$$

En la norma $\parallel v \parallel_p$ cuando $p = 2$ tenemos la norma inducida por el producto interior habitual (ver el siguiente ejemplo).

Ejercicio 208 *Indague el significado de la homogeneidad de la norma en el caso de vectores geométricos en dos o tres dimensiones.*

En un espacio dado hay muchas normas posibles, sin embargo si el espacio está equipado con un producto interior hay una norma especial que proviene del producto, esta norma se llama "norma inducida por el producto interior".

Definición 209 *Sea V un espacio vectorial complejo con producto interior. La norma inducida por el producto interior se define mediante:*

$$\|x\| = \sqrt{\langle x, x \rangle}$$

(la raíz positiva).

Ejemplo 210 *En $\mathbb{R}^n$, con el producto interno habitual, tenemos la norma inducida*

$$\parallel v \parallel = \left[\sum_{i=1}^{n} (\mid v_i \mid)^2 \right]^{\frac{1}{2}}$$

que es una norma tipo $\parallel v \parallel_p$ con $p = 2$.

Teorema 211 *Sea V un espacio vectorial complejo con producto interior. La norma inducida por el producto interior es una norma.*

Demostración. a) De la definición se ve que $\|x\| \geq 0$. Si $x = 0$ $\|x\| = \sqrt{\langle 0, 0 \rangle} = \sqrt{0} = 0$ y. por último si $\|x\| = 0$ entonces $\langle x, x \rangle = 0$ y $x = 0$.

b)

$$\begin{aligned}
\|\gamma x\| &= \sqrt{\langle \gamma x, \gamma x \rangle} \\
&= \sqrt{\gamma \overline{\gamma} \langle x, x \rangle} \\
&= \sqrt{\mid \gamma \mid^2 \langle x, x \rangle} \\
&= \sqrt{\mid \gamma \mid^2} \sqrt{\|x\|^2} \\
&= \mid \gamma \mid \|x\|
\end{aligned}$$

c) Como

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \|x\|^2 + \|y\|^2 + \langle y,x\rangle + \langle x,y\rangle \\
&= \|x\|^2 + \|y\|^2 + \langle y,x\rangle + \overline{\langle y,x\rangle} \\
&= \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}(\langle y,x\rangle) \\
&\leq \|x\|^2 + \|y\|^2 + 2\mid \mathrm{Re}(\langle y,x\rangle)\mid \\
&\leq \|x\|^2 + \|y\|^2 + 2\mid \langle y,x\rangle \mid \\
&\leq \|x\|^2 + \|y\|^2 + 2\sqrt{\langle y,y\rangle\langle x,x\rangle} \\
&= \|x\|^2 + \|y\|^2 + 2\sqrt{\|x\|^2\|y\|^2} \\
&= \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \\
&= (\|x\| + \|y\|)^2
\end{aligned}$$

y por ello $\|x+y\| \leq \|x\| + \|y\|$ ■

No debe pensarse que toda norma en un espacio vectorial es inducida por un producto interior.

Si se sabe que en un espacio real la norma es inducida por un producto interior, entonces puede inferirse cuál es el producto interior.

Teorema 212 *Sea V un espacio vectorial real con un producto interior. Sea $\|\|$ la norma inducida por el producto interior. Entonces, para cualesquiera a, b en V tenemos que*

$$\langle a,b\rangle = \frac{1}{4}\left[\|a+b\|^2 - \|a-b\|^2\right]$$

Demostración.

$$\begin{aligned}
\|a+b\|^2 &= \langle a+b, a+b\rangle \\
&= \langle a,a\rangle + \langle a,b\rangle + \langle b,a\rangle + \langle b,b\rangle \\
&= \|a\|^2 + \|b\|^2 + 2\langle a,b\rangle
\end{aligned}$$

y

$$\begin{aligned}
\|a-b\|^2 &= \langle a-b, a-b\rangle \\
&= \langle a,a\rangle - \langle a,b\rangle - \langle b,a\rangle + \langle b,b\rangle \\
&= \|a\|^2 + \|b\|^2 - 2\langle a,b\rangle
\end{aligned}$$

por lo que

$$\|a+b\|^2 - \|a-b\|^2 = 4\langle a,b\rangle$$

y

$$\langle a,b\rangle = \frac{1}{4}\left[\|a+b\|^2 - \|a-b\|^2\right]$$

■

Teorema 213 *Sea V un espacio vectorial complejo con un producto interior. Sea $\|\|$ la norma inducida por el producto interior. Entonces, para cualesquiera a, b en V tenemos que*

$$\begin{aligned}\langle a,b\rangle &= \frac{1}{4}\left[\|a+b\|^2 - \|a-b\|^2\right] \\ &+ \frac{i}{4}\left[\|a+ib\|^2 - \|a-ib\|^2\right]\end{aligned}$$

Demostración.

$$\begin{aligned}\|a+b\|^2 &= \langle a+b, a+b\rangle \\ &= \langle a,a\rangle + \langle a,b\rangle + \langle b,a\rangle + \langle b,b\rangle \\ &= \|a\|^2 + \|b\|^2 + 2\operatorname{Re}\langle a,b\rangle\end{aligned}$$

$$\begin{aligned}\|a-b\|^2 &= \langle a-b, a-b\rangle \\ &= \langle a,a\rangle - \langle a,b\rangle - \langle b,a\rangle + \langle b,b\rangle \\ &= \|a\|^2 + \|b\|^2 - 2\operatorname{Re}\langle a,b\rangle\end{aligned}$$

$$\begin{aligned}\|a+ib\|^2 &= \langle a+ib, a+ib\rangle \\ &= \langle a,a\rangle + \langle a,ib\rangle + \langle ib,a\rangle + \langle b,b\rangle \\ &= \|a\|^2 + \|b\|^2 + i\left[\langle b,a\rangle - \langle a,b\rangle\right] \\ &= \|a\|^2 + \|b\|^2 + [2i]\, i\operatorname{Im}(\langle b,a\rangle) \\ &= \|a\|^2 + \|b\|^2 - 2\operatorname{Im}(\langle b,a\rangle) \\ &= \|a\|^2 + \|b\|^2 + 2\operatorname{Im}(\langle a,b\rangle)\end{aligned}$$

$$\begin{aligned}\|a-ib\|^2 &= \langle a-ib, a-ib\rangle \\ &= \langle a,a\rangle + \langle a,-ib\rangle + \langle -ib,a\rangle + \langle b,b\rangle \\ &= \|a\|^2 + \|b\|^2 + i\langle a,b\rangle - i\langle b,a\rangle \\ &= \|a\|^2 + \|b\|^2 + i\left[\langle a,b\rangle - \langle b,a\rangle\right] \\ &= \|a\|^2 + \|b\|^2 + i\left[2i\operatorname{Im}\langle a,b\rangle\right] \\ &= \|a\|^2 + \|b\|^2 - \left[2\operatorname{Im}\langle a,b\rangle\right]\end{aligned}$$

por lo que

$$\begin{aligned}\langle a,b\rangle &= \frac{1}{4}\left[\|a+b\|^2 - \|a-b\|^2\right] \\ &+ \frac{i}{4}\left[\|a+ib\|^2 - \|a-ib\|^2\right]\end{aligned}$$

■

Estas fórmulas que dan los productos interiores se conocen como identidades de polarización.

Ejemplo 214 *La norma* $\| \ v \ \|_\infty = \max\{| \ v_i \ |\}$ *no proviene de un producto interior. Por ejemplo en* $\mathbb{R}^3$ *, usando* $a = (1, 2, 3)$ *y* $b = (-2, -3, 0)$ *y con la identidad de polarización resulta que* $\langle 2a, b\rangle \neq 2\langle a, b\rangle$.

Comentario 215 *Como en el ejemplo anterior, a veces se nos pide que verifiquemos si una cierta norma proviene de un producto interior. Entre los caminos posibles señalaremos: 1) usar la identidad de polarización para obtener el presunto producto interior. Luego verificar que es en efecto un producto interior. Por último falta ver que ese producto interior dá la misma norma original; para ello basta con que* $\|\gamma a\|^2 = \gamma^2\|\gamma a\|^2$ *2) Si la norma proviene de un producto interior entonces debe satisfacerse para cualesquiera vectores* a *y* b *la llamada "identidad del paralelogramo" (que esperamos el lector demuestre)*

$$\|a + b\|^2 + \|a - b\|^2 = 2 \ \| \ a \ \|^2 + 2 \ \| \ b \ \|^2$$

si encontramos algún contraejemplo, entonces la norma no proviene de un producto interior. En el ejemplo anterior se puede ver que con $a = (1, 0)$ *y* $b = (0, 1)$ *se viola la identidad del paralelogramo, confirmando que esa norma no puede ser la norma inducida por un producto interior. En realidad la condición de que cumpla la identidad del paralelogramo es condición necesaria y suficiente para que la norma provenga de un producto interior, pero la demostración es bastante complicada.*

13.7. Distancias (métricas)

En un espacio V puede definirse la noción de distancia.

Sea V un espacio vectorial, una distancia (o métrica) en V es una función $d : V \times V \to \mathbb{R}$ que asocia con cada par x, y de vectores de V un número real $d(x, y)$ que además satisface las siguientes condiciones:
a) $d(x, y) \geq 0 \ \forall x, y \in V$ y $d(x, y) = 0$ si y sólo si $x = y$.
b) $d(x, y) = d(y, x) \ \ \forall x, y \in V$
c) $d(x, y) \leq d(x, z) + d(z, y) \ \forall x, y, z \in V$ (desigualdad del triángulo).

Definición 216 *Un espacio vectorial* V *con una métrica* d *se llama espacio métrico.*

En general hay muchas maneras de definir una distancia en un espacio dado. Sin embargo, si el espacio es normado, hay una métrica especial, llamada "métrica inducida por la norma".

Comentario 217 *De hecho las métricas pueden definirse en conjuntos que no sean espacios vectoriales, por ejemplo la métrica discreta mostrada más adelante se puede definir en cualquier conjunto.*

Definición 218 *Si* V *es un espacio vectorial normado, entonces la distancia inducida por la norma es la función definida mediante:*

$$d(x, y) = \|x - y\|$$

Teorema 219 *Si V es un espacio vectorial normado, entonces las distancia inducida por la norma es una métrica.*

Demostración. a) Como $d(x,y) = \|x-y\|$ tenemos (propiedades de las normas) que $d(x,y) \geq 0$. $d(x,x) = \|x-x\| = \|0\| = 0$. Además si $d(x,y) = 0$ entonces $\|x-y\| = 0$ y $x-y = 0$ y $x = y$.

b) $d(x,y) = \|x-y\| = \|y-x\| = d(y,x)$

c)

$$\begin{aligned} d(x,y) &= \|x-y\| \\ &= \|\,(x-z)+(z-y)\,\| \leq \|\,(x-z)\,\| + \|\,(z-y)\,\| \\ &= d(x,z)+d(z,y) \end{aligned}$$

■

Cuando la norma es inducida por un producto interior, la métrica inducida por esta norma se llama métrica inducida por el producto interior.

Si se sabe que la metrica proviene de una norma, es posible reconstruir la norma si nos dan la métrica.

Teorema 220 *Sea V un espacio normado y sea d la métrica inducida por la norma. Entonces, para todo $a \in V$*

$$\|a\| = d(a,0)$$

Comentario 221 *Hemos dado todas las demostraciones pensando en el caso de un espacio vectorial complejo. Dejamos al lector la tarea de cerciorarse de que estas demostraciones siguen siendo válidas en el caso real.*

No debe pensarse que todas las métricas provienen de una norma.

Ejemplo 222 *Sea V un espacio vectorial. Para $a,b \in V$ defina*

$$d(a,b) = \begin{cases} 0 & si\ a = b \\ 1 & si\ a \neq b \end{cases}$$

Entonces $d(a,b)$ es una métrica en V ("métrica discreta").

Ejemplo 223 *La métrica del ejemplo anterior no puede provenir de una norma pues, si este fuera el caso,*

$$\|a\| = d(a,0)$$

y, para $a \neq 0$ y $\gamma \neq 0$ $\mid \gamma \mid \neq 1$

$$\|\gamma a\| = d(\gamma a, 0) = 1 \neq \mid \gamma \mid \|a\| = \mid \gamma \mid d(a,0) = \mid \gamma \mid$$

13.8. Ángulos.

En espacios vectoriales reales puede definirse la noción de ángulo, en espacios complejos generalmente no se utilizan los ángulos. Consideremos de nuevo la ecuación $a \cdot b = \mid a \mid \mid b \mid \cos(\theta)$ de donde podemos despejar

$$\cos(\theta) = \frac{a \cdot b}{\mid a \mid \mid b \mid}$$

y definimos el ángulo entre los dos vectores a y b como el ángulo $\theta \in [0, \pi]$ que satisface la equación anterior.

Cuando $\theta = \frac{\pi}{2}$ $\cos(\theta) = 0$ y necesariamente $a \cdot b = 0$.

En términos de nuestra notación

$$\cos(\theta) = \frac{\langle a, b\rangle}{\|a\|\|b\|} = \frac{\langle a, b\rangle}{\sqrt{\langle a, a\rangle}\sqrt{\langle b, b\rangle}}$$

13.9. Conjuntos ortogonales, normales y ortonormales

Si un vector v en V tiene norma uno ($\|v\| = 1$) decimos que está normalizado (o que es unitario). Si $\|v\| \neq 1$ entonces podemos normalizarlo,esto es, obtener un vector w que es un múltiplo de v pero $\|w\| = 1$. Una manera simple de hacer esto es multiplicando v por el número $(\|v\|)^{-1}$ (suponiendo, por supuesto, que $\|v\| \neq 0$) pues

$$\|(\|v\|)^{-1}v\| = (\|v\|)^{-1}\|v\| = 1$$

Definición 224 *Un conjunto $\alpha = \{v_1, v_2..v_n\}$ se llama "normalizado"o "normal" si $\|v_i\| = 1$ para $i = 1, 2..n$.*

Definición 225 *Se dice que dos vectores v y w son ortogonales (perpendiculares) si $\langle v, w\rangle = 0$.*

Definición 226 *Se dice que un conjunto $\alpha = \{v_1, v_2..v_n\}$ es ortogonal si $\langle v_i, v_j\rangle = 0$ para $i \neq j$.*

Definición 227 *Se dice que un conjunto $\alpha = \{v_1, v_2..v_n\}$ es ortonormal si es normal y ortogonal, esto es, si*

$$\langle v_i, v_j\rangle = \delta_{ij}$$

para todo $i, j = 1, 2..n$.

Comentario 228 *Hemos usado en la definición precedente la llamada delta de Kronecker. Es una función definida mediante*

$$\delta_{ij} = \begin{cases} 1 \ si \ i = j \\ 0 \ si \ i \neq j \end{cases}$$

y donde i y j son enteros.

Definición 229 *Se dice que un conjunto* $\alpha = \{v_1, v_2..v_n\}$ *es una base ortonormal si es ortonormal y es una base. Si el conjunto* α *es ortogonal pero no normal, entonces se dice que es una base ortogonal.*

13.9.1. Los conjuntos ortogonales son independientes

Teorema 230 *Sea* V *un espacio vectorial con producto interior y* $S = \{v_1, v_2...v_n\}$ *un conjunto ortogonal de vectores no nulos. Entonces* S *es linealmente independiente.*

Demostración.

$$\begin{aligned}
\alpha_1 v_1 + \alpha_2 v_2 + ...\alpha_n v_n &= 0 \\
\langle \alpha_1 v_1 + \alpha_2 v_2 + ...\alpha_n v_n, v_j \rangle &= 0 \\
\langle \alpha_1 v_1, v_j \rangle + \langle \alpha_2 v_2, v_j \rangle + ...\langle \alpha_n v_n, v_j \rangle &= 0 \\
\alpha_1 \langle v_1, v_j \rangle + \alpha_2 \langle v_2, v_j \rangle + ...\alpha_n \langle v_n, v_j \rangle &= 0 \\
\alpha_j \langle v_j, v_j \rangle &= 0 \\
\alpha_j \|v_j\|^2 &= 0 \\
\alpha_j &= 0
\end{aligned}$$

donde hemos usado la ortogonalidad y también el hecho de que $v_j \neq 0$ ■

13.9.2. expansión con respecto a una base ortonormal .

Sea $\alpha = \{v_1, v_2..v_n\}$ una base ortonormal de un espacio vectorial V con un producto interior. Si $v \in V$ entonces existen números c_1, $c_2...c_n$, únicos, y con la propiedad de que

$$v = \sum_{i=1}^{n} c_i v_i$$

Tomando los productos interiores $\langle v, v_k \rangle$ obtenemos

$$\begin{aligned}
\langle v, v_k \rangle &= \langle \sum_{i=1}^{n} c_i v_i, v_k \rangle \\
&= \sum_{i=1}^{n} c_i \langle v_i, v_k \rangle \\
&= \sum_{i=1}^{n} c_i \delta_{ik} \\
&= c_k
\end{aligned}$$

de modo que

$$v = \sum_{i=1}^{n} v_i \langle v, v_i \rangle$$

Si la base es ortogonal pero no ortonormal entonces

$$v = \sum_{i=1}^{n} \frac{v_i\langle v, v_i\rangle}{\langle v_i, v_i\rangle} = \sum_{i=1}^{n} \frac{v_i\langle v, v_i\rangle}{\|v_i\|^2}$$

13.10. Matriz con respecto a una base ortonormal.

Sean V y W espacios vectoriales, suponga que W tiene un producto interior $\langle , \rangle$ y que $\alpha = \{v_1, v_2..v_n\}$ y $\beta = \{w_1, w_2..w_m\}$ son bases de V y W respectivamente. Suponga que β es una base ortonormal.

Si $T : V \to W$ es una transformación lineal, ya hemos visto que como $Tv_i \in W$ existen números M_{ij} tales que

$$Tv_j = \sum_{i=1}^{m} M_{ij} w_i$$

y tomando productos interiores

$$\begin{aligned} \langle Tv_j, w_h\rangle &= \langle \sum_{i=1}^{m} M_{ij} w_i, w_h\rangle \\ &= \sum_{i=1}^{m} M_{ij}\langle wi, w_h\rangle \\ &= \sum_{i=1}^{m} M_{ij}\delta_{ih} \\ &= M_{hj} \end{aligned}$$

y

$$M_{hj} = \langle Tv_j, w_h\rangle$$

En particular, cuando $V = W$ y $\alpha = \beta$

$$M_{hj} = \langle Tv_j, v_h\rangle$$

y las matrices se pueden calcular de una manera extraordinariamente simple.

13.11. Productos y cordenadas

Sea V un espacio con producto interior y $\alpha = \{v_1, v_2..v_n\}$ una base ortonormal. Entonces si

$$a = \sum_{i=1}^{n} a_i v_i$$

$$b = \sum_{j=1}^{n} b_j v_j$$

tendremos que

$$\begin{aligned}\langle a,b\rangle &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\bar{b}_j\langle v_i,v_j\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\bar{b}_j\delta_{ij} \\ &= \sum_{i=1}^{n} a_i\bar{b}_i \\ &= [b]_\alpha^*[a]_\alpha\end{aligned}$$

Pero note que $[b]_\alpha^*[a]_\alpha$ es meramente el producto canónico en $\mathbb{C}^n$.
En $\mathbb{R}^n$ tendremos que

$$\langle a,b\rangle = [b]_\alpha^T[a]_\alpha = [a]_\alpha^T[b]_\alpha$$

y $[b]_\alpha^T[a]_\alpha$ es el producto canónico en $\mathbb{R}^n$.

13.12. complementos ortogonales

Definición 231 *Sea V un espacio vectorial con un producto interior. Sea W un subespacio de V. El complemento ortogonal de W se define como el conjunto*

$$W^\perp = \{v \in V \mid \langle v,w\rangle = 0\ \forall w \in W\}$$

Teorema 232 *Sea V un espacio vectorial con un producto interior. Sean W un subespacio de V y $W^\perp$ el complemento ortogonal de W. Entonces $W^\perp$ es un subespacio de V.*

Demostración. a) Claramente (de la definición) $W^\perp \subset V$

b) Ciertamente $0 \in W^\perp$ pues $\langle 0,w\rangle = 0\ \forall w \in W$ y $W^\perp \neq \phi$

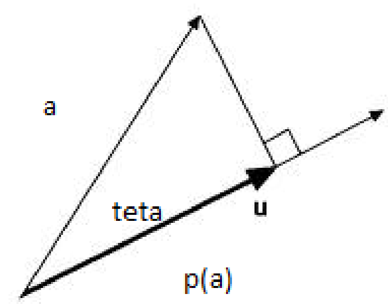
c) Si $a,b \in W^\perp$ tendremos que $\langle a,w\rangle = 0\ \forall w \in W$ y $\langle b,w\rangle = 0\ \forall w \in W$ pero entonces $\langle a+b,w\rangle = \langle a,w\rangle + \langle b,w\rangle = 0\ \forall w \in W$ y $a+b \in W^\perp$

d) sea β un escalar. Entonces si $a \in W^\perp$ tendremos que $\langle a,w\rangle = 0\ \forall w \in W$ pero entonces $\langle \beta a,w\rangle = \beta\langle a,w\rangle = 0\ \forall w \in W$ y $\beta a \in W^\perp$ ■

Teorema 233 *Sea V un espacio vectorial con un producto interior. Sean W un subespacio de V y $W^\perp$ el complemento ortogonal de W. Entonces* $\dim(W^\perp) = \dim(V) - \dim(W)$

Demostración. Sean $k = \dim(V)$, $n = \dim(W)$ y $A = \{v_1, v_2 ... v_n\}$ una base ortonormal de W. Extendemos esta base a una base ortonormal de V $B = \{v_1, v_2 ... v_n, v_{n+1}, .. v_k\}$. Aseveramos que los vectores $\{v_{n+1}, .. v_k\}$ son una base de $W^\perp$. Como son parte de una base (para V) son independientes. Ahora veremos que generan $W^\perp$. Sea $x \in W^\perp$, como $W^\perp \subset V$ tendremos que

$$x = \sum_{i=1}^{k} v_i\langle x,v_i\rangle$$



pero $\langle x, v_i \rangle = 0$ para $i = n+1...k$ y

$$x = \sum_{i=n+1}^{k} v_i \langle x, v_i \rangle$$

con lo que $\{v_{n+1}, ..v_k\}$ genera $W^{\perp}$ y es una base. Por ello $\dim(W^{\perp}) = \dim(V) - \dim(W)$ ■

Comentario 234 *En la práctica para ver que los vectores de un subespacio son ortogonales a TODOS los de otro basta con probar esto en una base, si* $\langle v, w_i \rangle = 0$ *para los vectores de una base* $\{w_i, w_2, ..w_n\}$ *entonces para* $w = \sum_{i=1}^{n} \gamma_i w_i$ *tendremos que*

$$\langle v, w \rangle = \langle v, \sum_{i=1}^{n} \gamma_1 w_i \rangle = \sum_{i=1}^{n} \overline{\gamma_i} \langle v, w_i \rangle = 0$$

13.13. Proyecciones

Como motivación primeramente vamos a revisar los procedimientos elementales usados en mecánica para proyectar vectores, luego abordamos el caso abstracto y general.

13.13.1. Motivación

Suponga que tenemos un vector $\vec{a}$ y que deseamos hallar su proyección en la dirección de otro vector $\vec{b}$. Sea θ el ángulo entre estos vectores, un vector unitario (de norma 1) en la dirección de $\vec{b}$ es $\vec{u} = \frac{\vec{b}}{|\vec{b}|}$. Del triángulo (ver figura) y por vil trigonometría tenemos que

$$\mid \pi(\vec{a}) \mid = \mid \vec{a} \mid \cos(\theta)$$

pero como

$$\begin{aligned}\vec{a}\cdot\vec{u} &=\mid \vec{a} \mid\mid \vec{u} \mid \cos(\theta)\\ &=\mid \vec{a} \mid \cos(\theta)\end{aligned}$$

tenemos que

$$\mid \pi(\vec{a}) \mid = \vec{a}\cdot\vec{u}$$

y un vector de tamaño $\mid \pi(\vec{a}) \mid$ en la dirección de $\vec{u}$ es

$$\begin{aligned}\pi(\vec{a}) &= \vec{a}\cdot\vec{u}\ \vec{u}\\ &= \frac{\vec{a}\cdot\vec{b}}{\mid \vec{b} \mid^2}\vec{b}\end{aligned}$$

En la notación propia del álgebra lineal

$$\pi(a) = \frac{\langle a, b\rangle}{\|b\|^2}b$$

y es esta expresión la que se generalizará a continuación.

13.13.2. Caso general

Sean V un espacio vectorial con producto interior y W un subespacio de V. Si $\alpha = \{w_1, w_2..w_n\}$ es una base ortogonal de W definimos la proyección ortogonal de $v \in V$ sobre el subespacio W como la transformación

$$\begin{aligned}\pi &: V \to V\\ \pi(v) &= \sum_{i=1}^{n} \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle}w_i\end{aligned}$$

Teorema 235 *π es lineal.*

Demostración. Sean $a, b \in V$ y β un escalar. Entonces:

a)

$$\begin{aligned}\pi(a+b) &= \sum_{i=1}^{n} \frac{\langle a+b, w_i\rangle}{\langle w_i, w_i\rangle}w_i\\ &= \sum_{i=1}^{n} \frac{\langle a, w_i\rangle}{\langle w_i, w_i\rangle}w_i + \frac{\langle b, w_i\rangle}{\langle w_i, w_i\rangle}w_i\\ &=\pi(a) + \pi(b)\end{aligned}$$

b)

$$\begin{aligned}
\pi(\beta a) &= \sum_{i=1}^{n} \frac{\langle \beta a, w_i\rangle}{\langle w_i, w_i\rangle} w_i \\
&= \sum_{i=1}^{n} \beta \frac{\langle a, w_i\rangle}{\langle w_i, w_i\rangle} w_i \\
&= \beta \sum_{i=1}^{n} \frac{\langle a, w_i\rangle}{\langle w_i, w_i\rangle} w_i \\
&= \beta \pi(a)
\end{aligned}$$

■

Ejemplo 236 *Veamos cuál es el núcleo de π.*

$$\begin{aligned}
x \in N(\pi) &\Leftrightarrow \pi(x) = 0 \\
&\Leftrightarrow \sum_{i=1}^{n} \frac{\langle x, w_i\rangle}{\langle w_i, w_i\rangle} w_i = 0 \\
&\Leftrightarrow \frac{\langle x, w_i\rangle}{\langle w_i, w_i\rangle} = 0 \ \forall i \\
&\Leftrightarrow \langle x, w_i\rangle = 0 \ \forall i
\end{aligned}$$

y por ello, el núcleo de π es, ni más ni menos, $W^{\perp}$.

Ejemplo 237 *Ahora veamos cuál es la imagen de π. De la definición de π vemos que $\mathrm{Im}(\pi) \subset W$. Pero del teorema de la dimensión*

$$\begin{aligned}
\dim(V) &= \dim(N(\pi)) + \dim(\mathrm{Im}(\pi)) \\
\dim(V) &= \dim(W^{\perp}) + \dim(\mathrm{Im}(\pi)) \\
\dim(\mathrm{Im}(\pi)) &= \dim(V) - \dim(W^{\perp}) = \dim(W)
\end{aligned}$$

de modo que $\mathrm{Im}(\pi) = W$.

Teorema 238 $\forall v \in V \ (v - \pi(v)) \in W^{\perp}$.

Demostración.

$$\begin{aligned}
&\langle v - \pi(v), w_j\rangle \\
&= \langle v, w_j\rangle - \langle \pi(v), w_j\rangle \\
&= \langle v, w_j\rangle - \langle \sum_{i=1}^{n} \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle} w_i, w_j\rangle \\
&= \langle v, w_j\rangle - \sum_{i=1}^{n} \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle} \langle w_i, w_j\rangle \\
&= \langle v, w_j\rangle - \langle v, w_j\rangle = 0
\end{aligned}$$

y $v - \pi(v)$ es ortogonal a todos los vectores de la base ortogonal de$W$. Por ello $v - \pi(v)$ es ortogonal a todos los vectores de W (pues son combinaciones lineales de los vectores de α). ■

Este teorema tiene una interpretación importante. Llamemos $w' = v - \pi(v)$ de modo que $v = \pi(v) + w'$ y tenemos a v expresado como la suma de un vector en W $(\pi(v))$ y uno en $W^{\perp}$.

Los proyectores (otro nombre para las proyecciones) son idempotentes:

Teorema 239 *Sea $\pi : V \to V$ una proyección ortogonal. Entonces $\pi^2 = \pi$. Dejamos la demostración al lector. Un operador con esta propiedad es llamado "idempotente."*

Comentario 240 *De la definición de proyector puede verse que cada uno de los vectores w_i es un eigenvector de π con eigenvalor 1. Análogamente, cada uno de los vectores de cualquier base de $W^{\perp}$ es eigenvector de π con eigenvalor 0. Los eigenvalores de π son 1 o 0. Dejamos al lector convencerse de que los proyectores son diagonalizables.*

13.14. Mejor aproximación.

Damos, sin demostración, el siguiente resultado:

Teorema 241 *(de "la mejor aproximación", también "teorema de proyección"). Sea V un espacio vectorial con producto interior y con la norma inducida $\|x\| = \sqrt{\langle x, x\rangle}$. Sea W un subespacio de V con una base ortogonal $\alpha = \{w_1, w_2..w_n\}$. Entonces para cualquier $v \in V$ $\pi(v)$ es el único punto en W más cercano a v y $\|v - \pi(v)\|$ es la distancia de v a W en el sentido de que $\pi(v)$ está en W y $\|v - \pi(v)\| \leq \|v - w\|$ para todo $w \in W$.*

Ejemplo 242 *Considere $V = \mathbb{R}^3$ (con el producto interior habitual) y W el espacio generado por $\{v_1, v_2\}$ donde $v_1 = (0, 1, 0)$ y $v_2 = (-\frac{4}{5}, 0, \frac{3}{5})$. ¿cuál es el vector de W más cercano a $u = (1, 1, 1)$?*
Calculemos

$$\begin{aligned}
\pi(u) &= \frac{\langle u, v_1\rangle}{\langle v_1, v_1\rangle} v_1 + \frac{\langle u, v_2\rangle}{\langle v_2, v_2\rangle} v_2 \\
&= \frac{(1)}{(1)} v_1 - \frac{(-\frac{1}{5})}{(1)} v_2 \\
&= (0, 1, 0) - \frac{1}{5}(-\frac{4}{5}, 0, \frac{3}{5}) \\
&= (\frac{4}{25}, 1, -\frac{3}{25})
\end{aligned}$$

Si llamamos w a

$$\begin{aligned} w &= u - \pi(u) \\ &= (1,1,1) - (\frac{4}{25}, 1, -\frac{3}{25}) \\ &= (\frac{21}{25}, 0, \frac{28}{25}) \end{aligned}$$

y

$$u = (\frac{4}{25}, 1, -\frac{3}{25}) + (\frac{21}{25}, 0, \frac{28}{25})$$

que es la suma de un vector en W y otro en $W^\perp$.

13.15. Método de Gram-Schmidt

Es un método que nos dice cómo obtener bases ortogonales de un espacio tomando como punto de partida alguna base.

13.15.1. Formulación del problema.

Sea V un espacio vectorial con producto interior y sea $\alpha = \{v_1, v_2 ... v_n\}$ una base no-ortogonal. Lo que se desea es construir explícitamente una base $\beta = \{w_1, w_2 ... w_n\}$ que sea ortogonal.

13.15.2. Idea básica

La idea tras el método (o algoritmo) de Gram-Schmidt es comenzar tomando $w_1 = v_1$. No podemos tomar $w_2 = v_2$ pues, en general, v_2 no va a ser ortogonal a v_1. Por ello tomamos mejor w_2 como v_2 menos su proyección sobre el espacio generado por w_1. De lo dicho en secciones anteriores, este vector sí será ortogonal a w_1. Por ello ya tendremos un conjunto ortogonal $\{w_1, w_2\}$. El siguiente paso es tomar w_3 como v_3 menos su proyección sobre el espacio generado por $\{w_1, w_2\}$ y así hasta acabar con una base ortogonal $\{w_1, w_2 ... w_n\}$.

13.15.3. Formulación explícita

En términos matemáticos,

$$\begin{aligned}
w_1 &= v_1 \\
w_2 &= v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 \\
w_3 &= v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 \\
&\vdots \\
w_k &= v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, w_i \rangle}{\langle w_i, w_i \rangle} w_i \\
&\vdots \\
w_n &= v_n - \sum_{i=1}^{n-1} \frac{\langle v_n, w_i \rangle}{\langle w_i, w_i \rangle} w_i
\end{aligned}$$

Una base ortonormal puede obtenerse normalizando β.

Ejemplo 243 *Considere el subespacio de $\mathbb{R}^3$ generado por $\{(1,0,2),(0,1,3)\}$, hallar una base ortogonal de dicho subespacio, usando el producto interior habitual en $\mathbb{R}^3$.*
Usando el método de Gram-Schmidt

$$\begin{aligned}
w_1 &= (1,0,2) \\
w_2 &= (0,1,3) - \frac{\langle (0,1,3),(1,0,2) \rangle}{\langle (1,0,2),(1,0,2) \rangle}(1,0,2) \\
&= (0,1,3) - \frac{6}{5}(1,0,2) \\
&= \frac{1}{5}\left[(0,5,15) - (6,0,12)\right] \\
&= \frac{1}{5}(-6,5,3)
\end{aligned}$$

Capítulo 14

Cuadrados mínimos.

14.0.4. motivación

Una ecuación de la forma

$$Ax = y$$

donde A es una matriz de $m \times n$, x es de $n \times 1$ y y es de $m \times 1$, puede:

- no tener soluciones
- tener una solución única
- tener un número infinito de soluciones

Sin embargo a veces lo que nos interesa es saber cuál es la x que más se acerca a ser una solución de la ecuación. Esto es relevante en contextos donde hay errores experimentales en las medidas, quizá en sentido estricto no haya una solución.

Por ello nos planteamos el problema de hallar x_0 tal que

$$\|Ax_0 - y\|^2 \leq \|Ax - y\|^2$$

para cualquier x. Esto es lo que se llama problema de cuadrados mínimos.

14.0.5. Teoría

Podemos razonarlo de la siguiente manera:

- Ax_0 está en el espacio columna de A, $\mathrm{col}(A)$
- por ello, la pregunta es equivalente a inquirir cuál es el vector de $\mathrm{col}(A)$ más cercano a y
- pero por el teorema de la proyección sabemos que ese vector es la proyección de y sobre $\mathrm{col}(A)$ de modo que $Ax_0 = \pi_{\mathrm{col}(A)}(y)$

- además sabemos que $y - Ax_0$ es perpendicular a todo $\mathrm{col}(A)$ y por ello perpendicular a Ax para cualquier x
- entonces $\langle y - Ax_0, Ax\rangle = 0$ para todo x
- y se sigue que $\langle y, Ax\rangle = \langle Ax_0, Ax\rangle$
- note que $\langle x, y\rangle = x^T y$ por lo que $\langle Ax, y\rangle = (Ax)^T y = x^T A^T y = x^T (A^T y) = \langle x, A^T y\rangle$
- y $\langle A^T y, x\rangle = \langle A^T Ax_0, x\rangle$
- como esto vale para toda x tenemos que

  $$A^T y = A^T Ax_0$$

  es la condición a satisfacerse. Este resultado suele llamarse "ecuaciones normales" para el problema de cuadrados mínimos.

Cuando $A^T A$ es invertible (lo cual es el caso si las columnas de A son independientes),

$$x_0 = \left(A^T A\right)^{-1} A^T y$$

da la solución al problema.

Comentario 244 *Una mnemotecnia: Si el problema es* $Ax \approx y$ *entonces* $A^T Ax_0 = A^T y$ *, es decir, multiplique ambos miembros de* $Ax \approx y$ *por la izquierda por* A^T *, cambie el signo de "aproximadamente igual"* $\approx$ *por el de igualdad* $=$ *y cambie* x *por* x_0*.*

Comentario 245 *De* $Ax_0 = \pi_{\mathrm{col}(A)}(y)$ *y* $x_0 = \left(A^T A\right)^{-1} A^T y$ *vemos que* $\pi_{\mathrm{col}(A)}(y) = A\left(A^T A\right)^{-1} A^T y$ *, expresión que a veces es útil.*

Ejemplo 246 *Considere el problema lineal inhomogéneo*

$$Ax = b$$

donde

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

$$b = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix}$$

y primero veamos si podemos resolverlo por los métodos habituales.
La matriz aumentada es

$$A' = \begin{pmatrix} 4 & 0 & 2 \\ 0 & 2 & 0 \\ 1 & 1 & 11 \end{pmatrix}$$

cuya forma escalonada reducida por renglones es

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

indicando que el sistema es inconsistente (¿recuerda Usted que la huella digital de un sistema incompatible es el tener un uno delantero en la columna del término inhomogéneo?).
Sin embargo tiene solución en el sentido de mínimos cuadrados. Usando las ecuaciones normales

$$A^T A = \begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix}$$
$$A^T b = \begin{pmatrix} 19 \\ 11 \end{pmatrix}$$

y por ello, la solución óptima x_0 será

$$\begin{aligned} x_0 &= \left(A^T A\right)^{-1} A^T b \\ &= \begin{pmatrix} \frac{5}{84} & -\frac{1}{84} \\ -\frac{1}{84} & \frac{17}{84} \end{pmatrix} \begin{pmatrix} 19 \\ 11 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{aligned}$$

14.1. Aplicaciones

Una aplicación importante surge cuando deseamos "ajustar" alguna curva a un conjunto de datos.

Suponga que tenemos una serie de datos (x_i, y_i) $(i = 1, 2...n)$ y que sospechamos que representan (con cierto error experimental) una línea recta $y = mx + b$.

Para cada x_i definimos el residual como $y_i - (mx_i + b)$ y el problema aquí es encontrar la m y la b que minimizan la suma de los cuadrados de los residuales.

Si los residuales fueran cero

$$y_i = (mx_i + b)$$

que se puede reescribir en forma matricial como (consideramos aquí a m y b como las incógnitas)

$$AX = Y$$

donde

$$X = \begin{pmatrix} b \\ m \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

ecuación que podemos resolver en el sentido de cuadrados mínimos con las ecuaciones normales.

Con tres datos tendríamos

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$\begin{aligned}
A &= \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{pmatrix} \\
A^T &= \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \end{pmatrix} \\
A^T A &= \begin{pmatrix} 3 & x_1 + x_2 + x_3 \\ x_1 + x_2 + x_3 & x_1^2 + x_2^2 + x_3^2 \end{pmatrix} \\
A^T A X &= \begin{pmatrix} 3b + m\,(x_1 + x_2 + x_3) \\ b\,(x_1 + x_2 + x_3) + m\left(x_1^2 + x_2^2 + x_3^2\right) \end{pmatrix}
\end{aligned}$$

y

$$A^T Y = \begin{pmatrix} y_1 + y_2 + y_3 \\ x_1 y_1 + x_2 y_2 + x_3 y_3 \end{pmatrix}$$

de modo que

$$\begin{aligned}
3b + m\,(x_1 + x_2 + x_3) &= y_1 + y_2 + y_3 \\
b\,(x_1 + x_2 + x_3) + m\left(x_1^2 + x_2^2 + x_3^2\right) &= x_1 y_1 + x_2 y_2 + x_3 y_3
\end{aligned}$$

y, generalizando,

$$\begin{aligned}
nb + m \sum_{i=1}^{n} x_i &= \sum_{i=1}^{n} y_i \\
b \sum_{i=1}^{n} x_i + m \sum_{i=1}^{n} x_i^2 &= \sum_{i=1}^{n} x_i y_i
\end{aligned}$$

tenemos las ecuaciones normales para este caso.

14.1.1. usando un poco de cálculo

$$AX = \begin{pmatrix} b + mx_1 \\ b + mx_2 \\ \vdots \\ b + mx_n \end{pmatrix}$$

$$AX - Y = \begin{pmatrix} b + mx_1 - y_1 \\ b + mx_2 - y_2 \\ \vdots \\ b + mx_n - y_n \end{pmatrix}$$

$$\delta = \|AX - Y\|^2 = \sum_{i=1}^{n} (b + mx_i - y_i)^2$$

(hemos llamado δ al "residual" $\|AX - Y\|^2$) y por lo tanto minimizaremos δ cuando

$$\frac{\partial \delta}{\partial b} = 0$$
$$\frac{\partial \delta}{\partial m} = 0$$

o sea

$$\sum_{i=1}^{n} 2\,(b + mx_i - y_i) = 0$$
$$\sum_{i=1}^{n} 2\,(b + mx_i - y_i)\, x_i = 0$$

y

$$nb + m \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$
$$b \sum_{i=1}^{n} x_i + m \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$

exactamente igual que en el cálculo puramente algebraico.

14.1.2. otros comentarios

No entiendo porqué el concepto de pseudoinversa no es parte del curso.

Una solución al problema

$$Ax \approx y$$

de cuadrados mínimos es

$$x_0 = A^{\dagger} y$$

donde $A^{\dagger}$ es la pseudoinversa (de Moore-Penrose) de A. Además como

$$Ax_0 = AA^{\dagger} y = \pi_{\mathrm{col}(A)}(y)$$

tenemos que

$$\pi_{\mathrm{col}(A)} = AA^{\dagger}$$

Otra manera de enfocar el problema de los cuadrados mínimos es mediante la descomposición QR, que puede verse como una pequeña modificación del método de Gram-Schmidt.

Una medida de qué tan bueno fue el ajuste está dada por

$$\delta = \|AX - Y\|^2$$

si δ fuera cero tendríamos, entonces una solución exacta. Una δ pequeña indica un buén ajuste en tanto que grandes valores de δ señalan un pobre ajuste.

14.1.3. un caso general y típico

Suponga usted que hemos medido una serie de parejas (x_i, y_i) con $i = 1...n$ y que sospechamos que son puntos que, salvo los inevitables errores experimentales, son descritos por un polinomio de grado m, a saber

$$y = a_0 + a_1 x + a_2 x^2 ... a_m x^m$$

Si no hubiera errores, tendríamos que

$$a_0 + a_1 x_i + a_2 x_i^2 ... a_m x_i^m = y_i$$

para toda i. Pero este sistema de ecuaciones se puede re-escribir en forma vectorial como

$$a_0 \begin{bmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_n^0 \end{bmatrix} + a_1 \begin{bmatrix} x_1^1 \\ x_2^1 \\ \vdots \\ x_n^1 \end{bmatrix} + a_2 \begin{bmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{bmatrix} + ... a_m \begin{bmatrix} x_1^m \\ x_2^m \\ \vdots \\ x_n^m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

y en forma matricial como

$$\begin{bmatrix} 1 & x_1^1 & \cdots & x_1^m \\ 1 & x_2^1 & \cdots & x_2^m \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n^1 & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

y si definimos

$$A = \begin{bmatrix} 1 & x_1^1 & \cdots & x_1^m \\ 1 & x_2^1 & \cdots & x_2^m \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n^1 & \cdots & x_n^m \end{bmatrix}$$

$$a = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix}$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

la ecuación exacta sería

$$Aa = Y$$

Pero como hemos señalado, en presencias de errores esta ecuación no tiene (en general) solución exacta y buscamos la mejor aproximación del tipo

$$Aa \sim Y$$

y sabemos que esta mejor aproximación está dada por las ecuaciones normales

$$A^T Aa = A^T Y$$

14.1.4. otro enfoque

Consideremos el problema de encontrar la mejor solución (en el sentido de cuadrados mínimos) de la ecuación $Ax = y$.

El problema consiste en minimizar la función δ donde

$$\delta(x) = \parallel Ax - y \parallel^2$$

o sea

$$\begin{aligned} \delta(x) &= \langle Ax - y, Ax - y \rangle \\ &= \langle Ax, Ax \rangle - \langle y, Ax \rangle - \langle Ax, y \rangle + \langle y, y \rangle \end{aligned}$$

y por ello si pasamos de x a $x + \Delta x$ tendremos que

$$\begin{aligned} \delta(x + \Delta x) = &\langle A(x + \Delta x), A(x + \Delta x) \rangle \\ &- \langle y, A(x + \Delta x) \rangle \\ &- \langle A(x + \Delta x), y \rangle \\ &+ \langle y, y \rangle \end{aligned}$$

o sea,

$$\begin{aligned}\delta(x+\Delta x) = &\langle Ax, Ax\rangle + \langle Ax, A\Delta x\rangle + \langle A\Delta x, Ax\rangle + \langle A\Delta x, A\Delta x\rangle\\ &- \langle y, Ax\rangle - \langle y, A\Delta x\rangle\\ &- \langle Ax, y\rangle - \langle A\Delta x, y\rangle\\ &+ \langle y, y\rangle\end{aligned}$$

por lo que

$$\begin{aligned}&\delta(x+\Delta x) - \delta(x) =\\ &\qquad \langle Ax, A\Delta x\rangle + \langle A\Delta x, Ax\rangle + \langle A\Delta x, A\Delta x\rangle\\ &\qquad - \langle y, A\Delta x\rangle - \langle A\Delta x, y\rangle\end{aligned}$$

y la parte lineal (es decir la diferencial de la función), es

$$\begin{aligned}D\delta(x, \Delta x) &=\\ \langle Ax, A\Delta x\rangle + \langle A\Delta x, Ax\rangle - \langle y, A\Delta x\rangle - \langle A\Delta x, y\rangle &=\\ \langle \Delta x, A^T Ax\rangle + \langle \Delta x, A^T Ax\rangle - \langle \Delta x, A^T y\rangle - \langle \Delta x, A^T y\rangle &=\\ &2\langle \Delta x, A^T Ax\rangle - 2\langle \Delta x, A^T y\rangle\end{aligned}$$

y en el punto extremo x_0 necesariamente para toda Δx tendremos

$$D\delta(x, \Delta x) = 2\langle \Delta x, A^T Ax\rangle - 2\langle \Delta x, A^T y\rangle = 0$$

o

$$\langle \Delta x, A^T Ax_0\rangle - \langle \Delta x, A^T y\rangle = 0$$

y

$$A^T Ax_0 = A^T y$$

que son las llamadas ecuaciones normales.

En este enfoque, de hecho, estamos usando la definición de derivada como mejor aproximación lineal. Lo interesante es que no tuvimos que escribir una sola derivada parcial.

Parte V

Operadores lineales en espacios con producto interior.

Capítulo 15

productos interiores y funciones lineales

Hay una conexión entre los productos interiores y las transformaciones lineales del espacio al campo F

$$F : V \to F$$

En primer lugar, dado un vector $v \in V$ podemos definir una funcion lineal mediante

$$\phi(x) = \langle x, v \rangle$$

y dejamos al lector la comprobación de que ϕ es lineal.

Conversamente, vamos a ver que si $\psi : V \to F$ es lineal entonces existirá algún $v \in V$ tal que $\psi(x) = \langle x, v \rangle$.

Para ello recordemos que una transformación lineal queda fija una vez que damos su efecto en una base. Sea pues $\alpha = \{v_1, v_2..v_n\}$ una base ortonormal de V. Claramente si

$$x = \sum_{i=1}^{n} c_i v_i$$

entonces

$$\psi(x) = \psi\left(\sum_{i=1}^{n} x_i v_i\right) = \sum_{i=1}^{n} x_i \psi(v_i)$$

Defina ahora

$$v = \sum_{i=1}^{n} [\psi(v_i)]^* \, v_i$$

Un cálculo sencillo revela que

$$\begin{aligned}
\langle x, v \rangle &= \langle x, \sum_{i=1}^{n} [\psi(v_i)]^* v_i \rangle \\
&= \sum_{i=1}^{n} [\psi(v_i)] \langle x, v_i \rangle \\
&= \sum_{i=1}^{n} [\psi(v_i)] x_i
\end{aligned}$$

y

$$\langle x, v \rangle = \psi(x)$$

En palabras, a cada vector v le corresponde una transformación lineal $\phi(x) = \langle x, v \rangle$ y a cada transformación lineal ψ le corresponde un vector $v = \sum_{i=1}^{n} [\psi(v_i)]^* v_i$. El vector v que le corresponde a una transformación lineal ψ es único pues si v' fuera otro vector con las mismas propiedades que v tendríamos que

$$\langle x, v \rangle = \langle x, v' \rangle$$

y

$$\langle x, v - v' \rangle = 0$$

para todo $x \in V$. Por ello $v = v'$.

Capítulo 16

Operadores adjuntos

Sean V un espacio vectorial sobre F de dimensión finita con producto interior y $T : V \to V$ un operador lineal.

Considere la expresión

$$\begin{gathered}\phi_a : V \to F \\ \phi_a(x) = \langle T(x), a\rangle\end{gathered}$$

Claramente ϕ_a es lineal para cualquier $a \in V$. Por ello existe un (único) vector $\psi_a \in V$ tal que

$$\phi_a(x) = \langle T(x), a\rangle = \langle x, \psi_a\rangle$$

Ahora prestemos atención al mapeo

$$\begin{gathered}T^* : V \to V \\ a \mapsto \psi_a\end{gathered}$$

aseveramos que es un mapeo lineal, es decir, un operador lineal. Este mapeo recibe el nombre de adjunto de T, también a veces se le llama mapeo dual. Note que la palabra adjunto o adjunta es ambigua. En teoría de matrices se suele llamar adjunta de A a la matríz $A^\dagger$ que tiene la propiedad de que $AA^\dagger = A^\dagger A = \det(A)I$. Algunos autores, para resolver la confusión, llaman a $A^\dagger$ "adjunta clásica" y a veces "adyugada".

Por la definición de T^* vemos que

$$\begin{aligned}\langle T(x), a\rangle &= \langle x, \psi_a\rangle \\ &= \langle x, T^*(a)\rangle\end{aligned}$$

ecuación básica para deducir las propiedades de los adjuntos.

Para ver que el adjunto es lineal considere

$$\begin{aligned}\langle T(x), a+b\rangle &= \langle T(x), a\rangle + \langle T(x), b\rangle \\ &= \langle x, T^*(a)\rangle + \langle x, T^*(b)\rangle \\ &= \langle x, T^*(a) + T^*(b)\rangle \\ &= \langle x, T^*(a+b)\rangle\end{aligned}$$

y por ello para todo $x \in V$

$$\langle x, T^*(a+b) - (T^*(a) + T^*(b))\rangle = 0$$

implicando que

$$T^*(a+b) = (T^*(a) + T^*(b))$$

Análogamente como

$$\begin{aligned}\langle \beta b, T(x)\rangle &= \beta\langle b, T(x)\rangle \\ &= \beta\langle T^*(b), x\rangle \\ &= \langle \beta T^*(b), x\rangle \\ &= \langle T^*(\beta b), x\rangle\end{aligned}$$

tenemos que para todo $x \in V$

$$\langle T^*(\beta b) - \beta T^*(b), x\rangle = 0$$

y

$$T^*(\beta b) = \beta T^*(b)$$

Por ello llamamos a T^* operador (lineal) adjunto a T.

Comentario 247 *¿y si tuvieramos $\langle x, T(y)\rangle$?*

$$\begin{aligned}\langle x, T(y)\rangle &= \overline{\langle T(y), x\rangle} \\ &= \overline{\langle y, T^* x\rangle} \\ &= \langle T^* y, x\rangle\end{aligned}$$

16.1. Matriz del operador adjunto con respecto a una base ortonormal.

Sea V un espacio vectorial (real o complejo) con producto interior, $T : V \to V$ un operador lineal y $\alpha = \{v_1, v_2..v_n\}$ una base ortonormal de V. Sean $M = M_\alpha^\alpha(T)$ y $N = M_\alpha^\alpha(T^*)$. Tal y como ya hemos visto

$$T(v_j) = \sum_{i=1}^{n} M_{ij} v_i$$

por lo que

$$\begin{aligned}\langle T(v_j), v_k\rangle &= \langle\sum_{i=1}^{n} M_{ij} v_i, v_k\rangle \\ &= \sum_{i=1}^{n} M_{ij}\langle v_i, v_k\rangle \\ &= \sum_{i=1}^{n} M_{ij}\delta_{ik} \\ &= M_{kj}\end{aligned}$$

Exactamente lo mismo ocurre para T^* y

$$\begin{aligned}\langle T^*(v_j), v_k\rangle &= \langle\sum_{i=1}^{n} N_{ij} v_i, v_k\rangle \\ &= \sum_{i=1}^{n} N_{ij}\langle v_i, v_k\rangle \\ &= \sum_{i=1}^{n} N_{ij}\delta_{ik} \\ &= N_{kj}\end{aligned}$$

pero, por la definición de adjunto,

$$\begin{aligned}N_{kj} &= \langle T^*(v_j), v_k\rangle \\ &= \langle v_j, T(v_k)\rangle \\ &= \overline{\langle T(v_k), v_j\rangle} \\ &= \overline{M}_{jk}\end{aligned}$$

y las matrices M y N son conjugadas transpuestas (adjuntas) una de la otra.

En la notación habitual

$$M_\alpha^\alpha(T^*) = [M_\alpha^\alpha(T)]^*$$

donde $*$ en la izquierda quiere decir "adjunta del operador" y la $*$ en la derecha quiere decir "transpuesta conjugada".

16.2. Propiedades del adjunto

De la definición de adjunto se sigue que:

- $(T^*)^* = T$ (involución).
  Demostración: Como $\langle a, T(b)\rangle = \langle T^*(a), b\rangle$ tenemos que el adjunto de T^*ha de ser T.

- $(T+R)^* = T^* + R^*$ (para cualesquiera operadores $T, R : V \to V$)
  Demostración: $\langle a, (T+R)(b)\rangle = \langle a, T(b) + R(b)\rangle = \langle a, T(b)\rangle + \langle a, R(b)\rangle = \langle T^*(A), b\rangle + \langle R^*(b), a\rangle = \langle (T^* + R^*)(a), b\rangle$ de donde $(T+R)^* = T^* + R^*$.

- $(S \circ T)^* = T^* \circ S^*$ (si $T : V \to W$ y $S : W \to \Omega$)
  Demostración: $\langle a, (S \circ T)(a)\rangle = \langle a, S(T(b))\rangle = \langle S^*(a), T(b)\rangle = \langle T^*(S^*(a)), (b)\rangle = \langle (T^* \circ S^*)(a), b\rangle$ de donde se sigue que $(S \circ T)^* = T^* \circ S^*$.

- $(\alpha T)^* = \overline{\alpha} T^*$ para cualquier operador lineal T y cualquier número α
  Demostración: $\langle a, (\alpha T)(b)\rangle = \langle a, \alpha T(b)\rangle = \bar{\alpha}\langle a, T(b)\rangle = \bar{\alpha}\langle T^*(a), b\rangle = \langle \bar{\alpha} T^*(a), b\rangle$ de donde
  se sigue que $(\alpha T)^* = \overline{\alpha} T^*$.

- $I_V^* = I_V$
  Demostración: $\langle a, I_V b\rangle = \langle a, b\rangle = \langle I_V a, b\rangle$ con lo que $I_V^* = I_V$.

A continuación suponga que $x \in N(T)$ de modo que $T(x) = 0$. Para cualquier $y \in V$ tendremos que

$$\begin{aligned} 0 &= \langle y, 0\rangle \\ &= \langle y, T(x)\rangle \\ &= \langle T^* y, x\rangle \end{aligned}$$

de modo que $T^* y$ es ortogonal a cualquier elemento de $N(T)$ y por ello

$$\operatorname{Im}(T^*) = [N(T)]^{\perp}$$

y, análogamente

$$[\operatorname{Im}(T^*)]^{\perp} = N(T)$$

e intercambiando los roles de T y T^* (recuerde que $(T^*)^* = T$). En resumidas cuentas:

$$\begin{aligned} \operatorname{Im}(T) &= [N(T^*)]^{\perp} \\ [\operatorname{Im}(T)]^{\perp} &= N(T^*) \end{aligned}$$

16.3. regla de correspondencia del adjunto

A veces tenemos un operador lineal T y se nos pide la regla de correspondencia de su adjunto T^*.

Ejemplo 248 *Considere el espacio $\mathbb{R}^2$ equipado con el producto interior ordinario. Si $T : \mathbb{R}^2 \to \mathbb{R}^2$ está dado por*

$$T(x, y) = (2x + y, x - y)$$

podemos preguntarnos cuál es la regla de correspondencia de

$$T^* : \mathbb{R}^2 \to \mathbb{R}^2$$

en cuyo caso tendremos que

$$\begin{aligned}
\langle (u,v), T(x,y)\rangle &= \langle (u,v), (2x+y, x-y)\rangle \\
&= u(2x+y) + v(x-y) \\
&= (2u+v)x + (u-v)y \\
&= \langle (2u+v, u-v), (x,y)\rangle \\
&= \langle T^*(u,v), (x,y)\rangle
\end{aligned}$$

con lo que

$$T^*(u,v) = (2u+v, u-v)$$

y

$$T^* = T$$

Una manera más sistemática de resolver estos problemas comprende los siguientes pasos:

- equipe su espacio con una base ortonormal α.
- calcule la matriz $M_\alpha^\alpha(T)$
- por lo señalado arriba, la matriz $M_\alpha^\alpha(T^*)$ está dada por la transpuesta conjugada (adjunta) de $M_\alpha^\alpha(T)$
- teniendo $M_\alpha^\alpha(T^*)$ es fácil deducir la regla de correspondencia de T^* usando el hecho de que $[T^*(v)]_\alpha = M_\alpha^\alpha(T^*)[v]_\alpha$.

Es posible también seguir una variante: Sea $\alpha = \{v_1, v_2..v_n\}$ una base ortonormal de V, entonces

$$\begin{aligned}
T^*(x) &= \sum_{i=1}^{n} \langle T^*(x), v_i\rangle v_i \\
&= \sum_{i=1}^{n} \langle x, T(v_i)\rangle v_i
\end{aligned}$$

fórmula que nos permite calcular $T^*(x)$. Por supuesto, ambos métodos son enteramente equivalentes.

Ejemplo 249 *Para el ejemplo anterior, en la base canónica, la matriz de T es*

$$M(T) = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$$

con lo que la matriz de T^* es la misma que la de T (pues $M(T)$ es simétrica). Por ello $T^* = T$.

16.3.1. ¿y si la base no era ortonormal?

Si la base dada no es ortonormal podríamos ortogonalizarla mediante Gram-Schmidt, lo cual es bastante engorroso.

16.4. Un ejemplo

Considere el espacio P_2 de los polinomios de grado menor o igual a dos con coeficientes reales y en donde se ha definido el producto interior

$$\langle P, Q\rangle = a_0b_0 + a_1b_1 + a_2b_2$$

para

$$\begin{aligned} P(x) &= a_0 + a_1x + a_2x^2 \\ Q(x) &= b_0 + b_1x + b_2x^2 \end{aligned}$$

Considere también el operador $D : P_2 \to P_2$ (derivación) dado por

$$D(P(x)) = \frac{dP(x)}{dx}$$

Encuentre la matriz asociada a D y a su adjunta D^* si la base es $\alpha = \{1, x, x^2\}$ y dé la regla de correspondencia de D^*.

Para la primera parte calculamos

$$\begin{aligned} D(1) &= 0 = (0)1 + (0)x + (0)x^2 \\ D(x) &= 1 = (1)1 + (0)x + (0)x^2 \\ D(x^2) &= 2x = (0)1 + (2)x + (0)x^2 \end{aligned}$$

de modo que

$$M_\alpha^\alpha(D) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Ahora bien, la base α es ortonormal pues

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

por lo que la matriz de D^* será simplemente la transpuesta de la de D y

$$M_\alpha^\alpha(D^*) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}$$

La regla de correspondencia de D^* se infiere facilmente del hecho de que

$$M_\alpha^\alpha(D^*) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}$$

con lo que

$$[D^*(v)]_\alpha = M_\alpha^\alpha(D^*)[v]_\alpha$$

por lo que si

$$\begin{aligned} v &= a + bx + cx^2 \\ [v]_\alpha &= \begin{bmatrix} a \\ b \\ c \end{bmatrix} \end{aligned}$$

tendremos que

$$\begin{aligned} [D^*(a + bx + cx^2)]_\alpha &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ a \\ 2b \end{bmatrix} \end{aligned}$$

y

$$D^*(a + bx + cx^2) = ax + 2bx^2$$

Capítulo 17

Operadores y matrices Hermiteanos, antihermiteanos, simétricos, antisimétricos, unitarios, ortogonales y normales.

En las aplicaciones tenemos casi siempre operadores lineales en espacios que también tienen productos interiores. De aquí surgen diversas propiedades; las principales tienen que ver con los operadores Hermiteanos. En lo que sigue trabajaremos casi exclusivamente con espacios complejos aunque casi todos los resultados se aplican al caso real.

Cuando el espacio vectorial es real, el adjunto de un operador T también puede ser llamado "operador transpuesto" y se denota como T^T.

17.1. Algunas definiciones básicas

Definición 250 *Se dice que un operador lineal $T : V \to V$ es Hermiteano (o autoadjunto) si $T = T^*$*

Definición 251 *Se dice que un operador lineal $T : V \to V$ es antihermitiano si $T = -T^*$*

Definición 252 *Se dice que un operador lineal en un espacio real $T : V \to V$ es simétrico si $T = T^T$*

Definición 253 *Se dice que un operador lineal en un espacio real $T : V \to V$ es antisimétrico si $T = -T^T$*

Definición 254 *Se dice que un operador lineal* $T : V \to V$ *es unitario si* $T^{-1} = T^*$

Definición 255 *Se dice que un operador lineal en un espacio real* $T : V \to V$ *es ortogonal si* $T^{-1} = T^T$

Definición 256 *Se dice que un operador lineal* $T : V \to V$ *es normal si* $TT^* = T^*T$.

Similarmente:

Definición 257 *Se dice que una matriz* M *es Hermiteana (o autoadjunta) si* $M = M^*$

Definición 258 *Se dice que una matriz* M *es antihermitiana si* $M = -M^*$

Definición 259 *Se dice que una matriz* M *es simétrica si* $M = M^T$ *(aquí la matriz puede ser real o compleja).*

Definición 260 *Se dice que una matriz* M *es antisimétrica si* $M = -M^T$

Definición 261 *Se dice que una matriz* M *es unitaria si* $M^{-1} = M^*$

Definición 262 *Se dice que una matriz* M *es ortogonal si* $M^{-1} = M^T$

Definición 263 *Se dice que una matriz* M *es normal si* $MM^* = M^*M$.

17.2. Algunos teoremas

Al tomar en cuenta que las transformaciones están asociadas con matrices, al introducir bases ortonormales, tendremos que estas definiciones son lógicas, adjuntos de operadores corresponden a adjuntos (conjugadas transpuestas) de matrices, operadores normales corresponden a matrices normales, operadores hermitianos corresponden a matrices hermitianas y operadores unitarios corresponden a matrices unitarias.

Teorema 264 *Sea* V *un espacio vectorial con producto interior,* α *una base ortonormal y* $T : V \to V$ *un operador lineal. Sea* $M = M_\alpha^\alpha(T)$*, entonces* T *es normal si y sólo si* $MM^* = M^*M$.

Demostración. Suponga que T es normal. Entonces

$$\begin{aligned}
MM^* &= M_\alpha^\alpha(T)M_\alpha^\alpha(T)^* \\
&= M_\alpha^\alpha(T)M_\alpha^\alpha(T^*) \\
&= M_\alpha^\alpha(TT^*) \\
&= M_\alpha^\alpha(T^*T) \\
&= M_\alpha^\alpha(T^*)M_\alpha^\alpha(T) \\
&= M_\alpha^\alpha(T)^*M_\alpha^\alpha(T) \\
&= M^*M
\end{aligned}$$

Conversamente, si $MM^* = M^*M$ entonces

$$\begin{aligned} MM^* - M^*M &= 0 \\ M_\alpha^\alpha(TT^* - T^*T) &= 0 \\ TT^* - T^*T &= 0 \end{aligned}$$

■

Teorema 265 *Sea V un espacio vectorial con producto interior, α una base ortonormal , $T : V \to V$ un operador lineal y $M = M_\alpha^\alpha(T)$. Entonces T es hermitiano si y sólo si $M = M^*$*

Demostración. Sea T hermiteano. Entonces

$$\begin{aligned} M^* &= M_\alpha^\alpha(T^*) \\ &= M_\alpha^\alpha(T) \\ &= M \end{aligned}$$

Conversamente, si $M^* = M$ entonces

$$\begin{aligned} M^* - M &= 0 \\ M_\alpha^\alpha(T^* - T) &= 0 \\ T^* - T &= 0 \end{aligned}$$

■

Teorema 266 *Sea V un espacio vectorial con producto interior, α una base ortonormal $T : V \to V$ un operador lineal y $M = M_\alpha^\alpha(T)$. Entonces T es unitario si y sólo si $M^{-1} = M^*$*

Demostración. Sea T unitario. Entonces M^*

$$\begin{aligned} M^* &= M_\alpha^\alpha(T)^* \\ &= M_\alpha^\alpha(T^*) \\ &= M_\alpha^\alpha(T^{-1}) \\ &= M_\alpha^\alpha(T)^{-1} \\ &= M^{-1} \end{aligned}$$

Conversamente, si $M = M^{-1}$ entonces:

$$\begin{aligned} M^* - M^{-1} &= 0 \\ M_\alpha^\alpha(T^* - T^{-1}) &= 0 \\ T^* - T^{-1} &= 0 \end{aligned}$$

■

Teorema 267 *Si H es un operador Hermiteano en un espacio vectorial complejo con producto interior, entonces es normal.*

Demostración. Como

$$H^* H = HH$$
$$HH^* = HH$$

tenemos que H es trivialmente normal. ■

Teorema 268 *Si U es un operador unitario en un espacio vectorial complejo con producto interior, entonces es normal.*

Demostración. Como

$$U^* U = U^{-1} U = U U^{-1} = U U^*$$

tenemos que U es normal. ■

Teorema 269 *Si U es un operador unitario en un espacio vectorial complejo con producto interior, entonces para cualesquiera vectores u y v*

$$\langle Uu, Uv \rangle = \langle u, v \rangle$$

Demostración. Tenemos que

$$\begin{aligned} \langle Uu, Uv \rangle &= \langle u, U^* U v \rangle \\ &= \langle u, U^{-1} U v \rangle \\ &= \langle u, I v \rangle \\ &= \langle u, v \rangle \end{aligned}$$

■

Teorema 270 *Si λ es un eigenvalor de un operador Hermiteano H en un espacio vectorial complejo con producto interior; entonces λ es real.*

Demostración. Sea $v \neq 0$ un eigenvector de H con eigenvalor λ. Entonces

$$\begin{aligned} \langle Hv, v \rangle &= \langle \lambda v, v \rangle = \lambda \langle v, v \rangle \\ &= \langle v, Hv \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle \end{aligned}$$

con lo que $\lambda = \bar{\lambda}$ (pues $\langle v, v \rangle \neq 0$). ■

Teorema 271 *Si v_1 y v_2 son dos eigenvectores de un operador Hermiteano H en un espacio vectorial complejo con producto interior correspondientes a eigenvalores diferentes λ_1 y λ_2 entonces $\langle v_1, v_2 \rangle = 0$.*

Demostración. tenemos que

$$\begin{aligned}\langle Hv_1, v_2\rangle &= \langle \lambda_1 v_1, v_2\rangle = \lambda_1\langle v_1, v_2\rangle \\ &= \langle v_1, Hv_2\rangle = \langle v_1, \lambda_2 v_2\rangle = \bar{\lambda}_2\langle v_1, v_2\rangle = \lambda_2\langle v_1, v_2\rangle\end{aligned}$$

con lo que

$$\begin{aligned}\lambda_1\langle v_1, v_2\rangle &= \lambda_2\langle v_1, v_2\rangle \\ (\lambda_1 - \lambda_2)\langle v_1, v_2\rangle &= 0 \\ \langle v_1, v_2\rangle &= 0\end{aligned}$$

pues $(\lambda_1 - \lambda_2) \neq 0$. ■

Teorema 272 *Si λ es un eigenvalor de un operador unitario U en un espacio vectorial complejo con producto interior; entonces $\mid \lambda \mid = 1$.*

Demostración. tenemos que si $Uv = \lambda v$

$$\begin{aligned}\langle Uv, Uv\rangle &= \langle \lambda v, \lambda v\rangle = \lambda\bar{\lambda}\langle v, v\rangle = \mid \lambda \mid^2 \langle v, v\rangle \\ &= \langle v, U^* Uv\rangle = \langle v, U^{-1}Uv\rangle = \langle v, Iv\rangle = \langle v, v\rangle\end{aligned}$$

con lo que

$$\begin{aligned}\left(\mid \lambda \mid^2 - 1\right)\langle v, v\rangle &= 0 \\ \mid \lambda \mid^2 &= 1\end{aligned}$$

pues $\langle v, v\rangle \neq 0$ ■

Tenemos resultados más generales

Teorema 273 *Si v es un eigenvector de un operador normal N en un espacio vectorial complejo con producto interior correspondiente al eigenvalor λ, entonces v es eigenvector de N^* con eigenvalor $\bar{\lambda}$*

Demostración. Sea $Nv = \lambda v$ de modo que

$$\begin{aligned}0 &= \parallel Nv - \lambda v \parallel^2 = \langle Nv - \lambda v, Nv - \lambda v\rangle \\ &= \langle Nv, Nv\rangle - \langle Nv, \lambda v\rangle - \langle \lambda v, Nv\rangle + \langle \lambda v, \lambda v\rangle \\ &= \langle v, N^*Nv\rangle - \langle \bar{\lambda}v, N^*v\rangle - \langle N^*v, \bar{\lambda}v\rangle + \langle \bar{\lambda}v, \bar{\lambda}v\rangle \\ &= \langle v, NN^*v\rangle - \langle \bar{\lambda}v, N^*v\rangle - \langle N^*v, \bar{\lambda}v\rangle + \langle \bar{\lambda}v, \bar{\lambda}v\rangle \\ &= \langle N^*v, N^*v\rangle - \langle \bar{\lambda}v, N^*v\rangle - \langle N^*v, \bar{\lambda}v\rangle + \langle \bar{\lambda}v, \bar{\lambda}v\rangle \\ &= \parallel N^*v - \bar{\lambda}v \parallel^2\end{aligned}$$

de modo que $N^*v = \bar{\lambda}v$ ■

Teorema 274 *Si v_1 y v_2 son dos eigenvectores de un operador normal N en un espacio vectorial complejo con producto interior correspondientes a eigenvalores diferentes λ_1 y λ_2 entonces $\langle v_1, v_2\rangle = 0$.*

Demostración. Tenemos que si

$$Nv_1 = \lambda_1 v_1$$
$$Nv_2 = \lambda_2 v_2$$

entonces

$$\begin{aligned}\bar{\lambda}_2 \langle v_1, v_2 \rangle &= \langle v_1, \lambda_2 v_2 \rangle = \langle v_1, N v_2 \rangle \\ &= \langle N^* v_1, v_2 \rangle = \langle \bar{\lambda}_1 v_1, v_2 \rangle = \bar{\lambda}_1 \langle v_1, v_2 \rangle\end{aligned}$$

y

$$\begin{aligned}\left(\bar{\lambda}_2 - \bar{\lambda}_1\right) \langle v_1, v_2 \rangle &= 0 \\ \langle v_1, v_2 \rangle &= 0\end{aligned}$$

pues $\lambda_1 \neq \lambda_2 \square$ ■

Comentario 275 *Las matrices unitarias y ortogonales tienen las siguientes propiedades:*
a) los renglones y las columnas forman conjuntos ortonormales (con respecto a los productos habituales en $\mathbb{R}^n$ *y en* $\mathbb{C}^n$ *)*
b) sus determinantes tienen módulo uno.

17.3. Teorema espectral.

Muy frecuentemente se llama espectro de un operador al conjunto de sus eigenvalores. Por ello las cuestiones relativas a los eigenvalores suelen recibir el calificativo de "espectrales" (nada que ver con los fantasmas).

Más formalmente:

Definición 276 *Sean V un espacio vectorial de dimensión finita y $T : V \to V$ un operador lineal. Entonces el espectro de T se define como*

$$spec(T) = \{\lambda \mid \lambda \text{ es un eigenvalor de } T\}$$

si el espacio no es de dimensión finita hay que usar una definición ligeramente diferente (y más general):

$$spec(T) = \{\lambda \mid (T - \lambda I) \text{ es singular}\}$$

ambas definiciones son equivalentes en el caso de dimensión finita.

Teorema 277 *Si $\alpha = \{v_1, v_2..v_n\}$ y $\beta = \{w_1, w_2, ...w_n\}$ son bases ortonormales en un espacio V con producto interior, entonces la matriz de transición de α a β es una matriz unitaria (ortogonal en el caso real).*

Demostración. Sean M la matriz de transición de β a α y M^{-1} la matriz de transición de α a β. Entonces

$$\begin{aligned} w_j &= \sum_{i=1}^{n} M_{ij} v_i \\ \langle w_j, v_k \rangle &= \langle \sum_{i=1}^{n} M_{ij} v_i, v_k \rangle \\ &= \sum_{i=1}^{n} M_{ij} \langle v_i, v_k \rangle \\ &= \sum_{i=1}^{n} M_{ij} \delta_{ik} \\ &= M_{kj} \end{aligned}$$

Similarmente

$$\begin{aligned} v_k &= \sum_{i=1}^{n} M_{ik}^{-1} w_i \\ \langle v_k, w_j \rangle &= \langle \sum_{i=1}^{n} M_{ik}^{-1} w_i, w_j \rangle \\ &= \sum_{i=1}^{n} M_{ik}^{-1} \langle w_i, w_j \rangle \\ &= \sum_{i=1}^{n} M_{ik}^{-1} \langle w_i, w_j \rangle \\ &= \sum_{i=1}^{n} M_{ik}^{-1} \delta_{ij} \\ &= M_{jk}^{-1} \end{aligned}$$

Por ello

$$\begin{aligned} M_{jk}^{-1} &= \langle v_k, w_j \rangle \\ &= \overline{\langle w_j, v_k \rangle} \\ &= \overline{M}_{kj} \end{aligned}$$

y $M^{-1} = M^* \square$ ■

Daremos ahora, sin demostración, varios teoremas relativos a la diagonalización de diversos operadores.

Teorema 278 *Sea V un espacio vectorial complejo de dimensión finita con producto interior y sea $T : V \to V$ un operador lineal normal. Entonces T es diagonalizable y tiene una base ortonormal de eigenvectores de T.*

Teorema 279 *(converso) Sea V un espacio vectorial complejo de dimensión finita con producto interior y sea $T : V \to V$ un operador lineal tal que hay una base ortonormal de V formada por eigenvectores de T. Entonces T es normal.*

Teorema 280 *Si M es una matriz normal entonces existe una matriz unitaria U tal que $U^{-1}MU$ es diagonal (la matriz diagonalizadora es unitaria).*

Teorema 281 *Sea V un espacio vectorial real de dimensión finita con producto interior y sea $T : V \to V$ un operador lineal simétrico. Entonces T es diagonalizable y tiene una base ortonormal de eigenvectores de T.*

Teorema 282 *(converso) Sea V un espacio vectorial real de dimensión finita con producto interior y sea $T : V \to V$ un operador lineal diagonalizable y tal que hay una base ortonormal de eigenvectores de T. Entonces T es simétrico.*

Teorema 283 *Si M es una matriz simétrica real entonces existe una matriz ortogonal O tal que $O^{-1}MO$ es diagonal (la matriz diagonalizadora es ortogonal).*

17.4. resolución espectral

17.4.1. productos exteriores

Suponga que a y b son matrices de $n \times 1$ (vectores columna) de modo que

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{bmatrix}$$

y

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{bmatrix}$$

con lo que si definimos $M = ab^T$

$$\begin{aligned} M &= ab^T \\ &= \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} \\ &= \begin{bmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \vdots & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{bmatrix} \end{aligned}$$

y tenemos que

$$M_{ij} = a_i b_j$$

Conversamente, si para una matriz se tiene que $M_{ij} = a_i b_j$ entonces $M = ab^T$ .Esto resultará importante en la siguiente sección. Los productos del tipo $M = ab^T$ (con a y b vectores columna) se llaman productos exteriores.

17.4.2. resolución espectral del operador y de la identidad

Sean V un espacio vectorial complejo con producto interior y $T : V \to V$ un operador lineal normal. Sean $\alpha = \{v_1, v_2 ... v_n\}$ una eigenbase ortonormal de T y $S = \{\lambda_1, \lambda_2 ... \lambda_n\}$ el espectro de T.

Considere el operador lineal definido por

$$T'(v) = \sum_{i=1}^{n} \lambda_i \langle v, v_i \rangle v_i$$

Este operador es lineal pues para $\forall v, w \in V$ y $\alpha \in \mathbb{C}$

$$\begin{aligned} T'(v+w) &= \sum_{i=1}^{n} \lambda_i \langle v+w, v_i \rangle v_i \\ &= \sum_{i=1}^{n} [\lambda_i \langle v, v_i \rangle v_i + \lambda_i \langle w, v_i \rangle v_i] \\ &= \sum_{i=1}^{n} \lambda_i \langle v, v_i \rangle v_i + \sum_{i=1}^{n} \lambda_i \langle w, v_i \rangle v_i \\ &= T'(v) + T'(w) \end{aligned}$$

y

$$\begin{aligned} T'(\alpha v) &= \sum_{i=1}^{n} \lambda_i \langle \alpha v, v_i \rangle v_i \\ &= \alpha \sum_{i=1}^{n} \lambda_i \langle v, v_i \rangle v_i \\ &= \alpha T'(v) \end{aligned}$$

Evaluando T' en la base, vemos que

$$\begin{aligned} T'(v_j) &= \sum_{i=1}^{n} \lambda_i \langle v_j, v_i \rangle v_i \\ &= \sum_{i=1}^{n} \lambda_i \delta_{ij} v_i \\ &= \lambda_j v_j \end{aligned}$$

por lo que $T = T'$ (ambos operadores lineales coinciden en una base).

En conclusión
$$T(v) = \sum_{i=1}^{n} \lambda_i \langle v, v_i \rangle v_i$$

Además cada operador
$$\pi_i(v) = \langle v, v_i \rangle v_i$$
es un proyector ortogonal sobre el espacio generado por $\{v_i\}$ y

$$T(v) = \sum_{i=1}^{n} \lambda_i \pi_i(v)$$

o
$$T = \sum_{i=1}^{n} \lambda_i \pi_i$$

Ésta es la llamada descomposición espectral de T.

Además como para cualquier vector $v \in V$ tenemos que

$$v = \sum_{i=1}^{n} \langle v, v_i \rangle v_1$$

obtenemos
$$I_V = \sum_{i=1}^{n} \pi_i$$

que es una (¡hay muchas!) descomposición espectral de la identidad.

En términos de alguna otra base ortonormal $\beta = \{w_1, w_2 ... w_n\}$ la matriz del proyector π_i será $M^i = M_\beta^\beta(\pi_i)$ con

$$\begin{aligned} M_{kh}^i &= \langle \pi_i w_h, w_k \rangle \\ &= \langle \langle w_h, v_i \rangle v_i, w_k \rangle \\ &= \langle w_h, v_i \rangle \langle v_i, w_k \rangle \\ &= \langle v_i, w_k \rangle \overline{\langle v_i, w_h \rangle} \end{aligned}$$

que tiene la forma de un producto exterior,

$$M^i = [v_i]_\beta [v_i]_\beta^*$$

Es por esto que, en términos de matrices, nuestro teorema dice que

$$M = \sum_{i=1}^{n} \lambda_i [v_i]_\beta [v_i]_\beta^*$$

El caso más común es cuando β es una base canónica, entonces los $[v_i]_\beta$ son los eigenvectores expresados en dicha base.

Ejemplo 284 *Consideremos el operador simétrico $k : \mathbb{R}^2 \to \mathbb{R}^2$ definido mediante $k(x, y) = (2x + 2y, 2x + 5y)$.*
La matriz de k referida a la base canónica de $\mathbb{R}^2$ es

$$M = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$$

El polinomio característico es

$$\begin{aligned} P(\lambda) &= \det \begin{bmatrix} 2-\lambda & 2 \\ 2 & 5-\lambda \end{bmatrix} \\ &= \lambda^2 - 7\lambda + 6 \end{aligned}$$

por lo que los eigenvalores son $\lambda_1 = 1$ y $\lambda_2 = 6$.
Para λ_1 tendremos

$$\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

por lo que

$$\begin{aligned} 2x + 2y &= x \\ 2x + 5y &= y \end{aligned}$$

o

$$\begin{aligned} x + 2y &= 0 \\ 2x + 4y &= 0 \end{aligned}$$

y una base normalizada para este eigenespacio E_1 es

$$V^1 = \frac{1}{\sqrt{5}}(-2, 1)$$

Para λ_2 tendremos

$$\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 6 \begin{bmatrix} x \\ y \end{bmatrix}$$

de modo que

$$\begin{aligned} 2x + 2y &= 6x \\ 2x + 5y &= 6y \end{aligned}$$

y

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

con lo que una base normalizada para este eigenespacio E_6 es

$$V^2 = \frac{1}{\sqrt{5}}(1, 2)$$

y la descomposición espectral de k es simplemente

$$\begin{aligned} k &= \lambda_1 \pi_{V^1} + \lambda_2 \pi_{V^2} \\ &= \pi_{V^1} + 6\pi_{V^2} \end{aligned}$$

o, explícitamente

$$\begin{aligned} \pi_{V^1}(x,y) &= \frac{1}{5}\langle (x,y),(-2,1)\rangle(-2,1) \\ &= \frac{1}{5}(-2x+y)(-2,1) \\ &= \frac{1}{5}(4x-2y,-2x+y) \end{aligned}$$

y

$$\begin{aligned} \pi_{V^2}(x,y) &= \frac{1}{5}\langle (x,y),(1,2)\rangle(1,2) \\ &= \frac{1}{5}(x+2y)(1,2) \\ &= \frac{1}{5}(x+2y,2x+4y) \end{aligned}$$

y facilmente podemos ver que

$$\begin{aligned} \pi_{V^1}(x,y) + \pi_{V^2}(x,y) &= \frac{1}{5}(4x-2y,-2x+y) + \frac{1}{5}(x+2y,2x+4y) \\ &= (x,y) = I(x,y) \end{aligned}$$

nos da la resolución de la identidad y que

$$\begin{aligned} \pi_{V^1}(x,y) + 6\pi_{V^2}(x,y) &= \frac{1}{5}(4x-2y,-2x+y) + \frac{6}{5}(x+2y,2x+4y) \\ &= (2x+2y,2x+5y) = k(x,y) \end{aligned}$$

En términos puramente matriciales

$$\frac{1}{\sqrt{5}}\begin{bmatrix} -2 \\ 1 \end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix} -2 & 1 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}$$

$$\frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

y tenemos que

$$\frac{1}{5}\begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} + \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

dando la resolución espectral de la identidad y

$$(1)\frac{1}{5}\begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} + (6)\frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$$

dándonos la resolución espectral de M .

17.4.3. Proyectores de nuevo

Cuando vimos los proyectores dimos una expresión en términos de una base ortonormal del subespacio sobre el cual se proyecta. Eso lo hicimos así porque no teníamos las herramientas adecuadas para dar la definición "correcta". La definición dada es defectuosa pues requiere dar una base ortonormal; una buena definición no debe depender de tal elección, debe de ser válida para cualquier selección de base.

Definición 285 *Sea V un espacio vectorial con producto interior. Un proyector ortogonal es un operador lineal $\pi : V \to V$ que es Hermiteano e idempotente ($\pi^2 = \pi$). Un proyector (no necesariamente ortogonal) es un operador idempotente.*

Capítulo 18

Cuádricas.

18.1. Introducción a las ideas principales

En geometría analítica frecuentemente nos preguntamos cuál es el lugar geométrico de los puntos en C donde

$$C = \{(x, y) \mid ax^2 + bxy + cy^2 + dx + ey + f = 0\}$$

En primer lugar note que si definimos la matriz simétrica A mediante

$$A = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}$$

entonces

$$\begin{aligned} &[x, y] \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= ax^2 + bxy + cy^2 \end{aligned}$$

por lo que si

$$v = \begin{bmatrix} x \\ y \end{bmatrix}$$

el término $ax^2 + bxy + cy^2$ es simplemente

$$v^T A v$$

Podemos completar la descripción puramente matricial del conjunto A si definimos,

$$T = [d, e]$$

de modo que el término $dx + ey$ es

$$Tv$$

y el conjunto C es simplemente

$$C = \{v \in \mathbb{R}^2 \mid v^T Av + Tv + f = 0\}$$

De la geometría analítica elemental sabemos que A representa una cónica (círculo, parábola, elipse, hipérbola o un caso "degenerado").

El propósito de este capítulo es mostrar cómo la teoría espectral de los operadores (y las matrices) simétricos nos permite caracterizar la cónica dada por A. Resultados similares se aplican en $\mathbb{R}^3$.

18.1.1. propiedades de A

- A es simétrica
- A es diagonalizable
- A tiene una eigenbase ortonormal
- La matriz diagonalizadora, siendo la matriz de cambio de base entre bases ortonormales, es ortogonal.

18.2. diagonalización de formas cuadráticas.

Consideremos la ecuación

$$v^T Av + Tv + f = 0$$

donde $T = [t_1, t_2...t_n]$.

Pensamos que la base es la canónica (α) e introducimos otra base (β), eigenbase ortonormal de A; de esta manera podemos suponer que hay un vector x de $\mathbb{R}^3$ tal que

$$\begin{aligned} v &= [x]_\alpha \\ w &= [x]_\beta \end{aligned}$$

y sea P la matriz de transición de β hacia α (recuerde, ésta es también la matriz diagonalizadora de A), de modo que $v = Pw$, entonces

$$\begin{aligned} v^T Av + Tv + f &= (Pw)^T A(Pw) + T(Pw) + f \\ &= w^T(P^T AP)w + (TP)w + f \\ &= w^T(P^{-1}AP)w + (TP)w + f \\ &= 0 \end{aligned}$$

Como P es la matriz diagonalizadora de A resulta que

$$D = P^{-1}AP$$

donde D es diagonal y tiene en su dioagonal los eigenvalores de A. Entonces

$$w^T(D)w + (T')w + f = 0$$

en donde hemos definido

$$T' = TP$$

y se dice que la forma cuadrática ha sido diagonalizada.

Si los elementos diagonales de D son λ_i y las componentes de w son w_i tendremos que

$$w^T(D)w = \lambda_1 w_1^2 + \lambda_2 w_2^2 + ...\lambda_n w_n^2$$

y no aparecen términos cruzados $w_i w_j$.

En resumen, una forma en la que hay términos cruzados puede convertirse, mediante un cambio de coordenadas, en otra sin términos cruzados.

18.3. cónicas

En $\mathbb{R}^2$ la forma cuadrática más general es del tipo

$$\begin{aligned} Q(x,y) &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= ax^2 + 2bxy + cy^2 \end{aligned}$$

que, mediante el cambio de coordenadas P descrito anteriormente se reduce a

$$w^T(D)w = \lambda_1 w_1^2 + \lambda_2 w_2^2$$

Ahora nos preguntamos por el lugar geométrico de todos los puntos que satisfacen

$$w^T(D)w = \lambda_1 w_1^2 + \lambda_2 w_2^2 = 1$$

- Cuando $\lambda_1 = \lambda_2$ y $\lambda_1 > 0$ tenemos un círculo de radio $\sqrt{1/\lambda_1}$
- Cuando $\lambda_1 = \lambda_2$ y $\lambda_1 \leq 0$ tenemos el conjunto vacío.
- Cuando $\lambda_1 \neq \lambda_2$ pero ambas positivas, se trata de una elipse
- Cuando $\lambda_1 \neq \lambda_2$ pero una positiva y la otra negativa se trata de una hipérbola.
- Cuando $\lambda_1 \neq \lambda_2$ pero ambas negativas, se trata del conjunto vacío
- Cuando λ_1 es cero y $\lambda_2 > 0$, se trata de dos rectas.
- Cuando λ_2 es cero y $\lambda_1 > 0$, se trata de dos rectas.
- Cuando λ_1 es cero y $\lambda_2 < 0$, es el conjunto vacío
- Cuando λ_2 es cero y $\lambda_1 < 0$, es el conjunto vacío

Los términos del tipo $T'w$ también pueden ser incluidos. El punto importante es que la diagonalización nos permite librarnos de los términos cruzados, el resto del análisis es como en geometría analítica.

18.4. superficies cuádricas.

En $\mathbb{R}^3$ tenemos una situación similar que resumimos en la siguiente tabla:

Eigenvalores λ_1, λ_2 y λ_3	Superficie(s) posibles
Todos con el mismo signo	Elipsoide
Dos de un signo	Cono elíptico, hiperboloide de dos hojas o hiperboloide de
Uno es cero, dos del mismop signo	Paraboloide eliptico o cilindro elíptico (caso degenerado)
Uno es cero, dos de signos opuestos	Paraboloide hiperbólico o cilindro hiperbólico (caso degen
Dos son cero y el tercero diferente de cero	Cilindro parabólico o dos planos paralelos (caso degenera

Apéndice A

Dependencia e independencia lineales.

Cuando usted considera un fenómeno que requiere ciertas variables $x_1, x_2...x_n$ frecuentemente nos preguntamos ¿cuántas y cuáles de esas varibles puedo elegir libremente?. Por ejemplo, al considerar una esfera en $\mathbb{R}^3$

$$x^2 + y^2 + z^2 = 25$$

uno podría elegir dos de las variables con cierta libertad (por ejemplo cualesquiera x o y tales que $x^2 \leq 25$) pero entonces el valor de z quedaría muy constreñido y $z = \pm\sqrt{25 - x^2 - y^2}$.

En general diremos que las variables están ligadas, o que hay constricciones entre ellas o que son dependientes entre sí si existe alguna función $F : \mathbb{R}^n \to \mathbb{R}$ tal que

$$F(x_1, x_2, ...x_n) = 0$$

Pero debe de tenerse el cuidado de exigir que la función F no sea trivial, es decir, la función cero $\widehat{0} : \mathbb{R}^n \to \mathbb{R}$ pues, trivialmente,

$$\widehat{0}(x_1, x_2, ...x_n) = 0$$

para cualquier selección de los valores $x_1, x_2...x_n$.

Cuando haya un enlace F no trivial entre las variables podemos, frecuentemente, "despejarünas vasriables en términos de otras.

Usted ya conoce un ejemplo importante de esto en álgebra. Cuando usted quiere resolver un sistema homogéneo de ecuaciones mediante los métodos de Gauss o de Gauss-Jordan lo que hace es aplicar operaciones elementales a la matriz de coeficientes. Cuando llega a una forma escalonada (o escalonada reducida por renglones) usted tiene que hallar las variables "ligadas"que son aquellas cuyas columnas corresponden a los pivotes (unos delanteros, unos principales) de la matriz escalonada. Las restantes variables son "libres". El algoritmo Gaussiano nos dice precisamente cuáles variables pueden asumir los valores que deseemos (son "libres) y cómo las restantes variables tienen valores completamente fijados por los de las variables libres.

En un espacio vectorial V diremos que los vectores $v_1, v_2...v_n$ son dependientes si hay una función no trivial F tal que

$$F(v_1, v_2, ...v_n) = 0$$

pero hay que preguntarse ¿qué clase de funciones F son admisibles en V donde únicamente tenemos la estructura de espacio vectorial?. En V

lo único que podemos hacer con los vectores es multiplicarlos por números y sumarlos, es decir, las funciones F han de ser combinaciones lineales. No podemos, por ejemplo, decir $\cos(v_1) - \exp(v_2) = 0$.

Así pues, los vectores $v_1, v_2...v_n$ son dependientes cuando existe una combinación lineal no-trivial

$$F(v_1, v_2, ...v_n) = \sum_{i=1}^{n} \alpha_i v_i$$

(con escalares $\alpha_1, \alpha_2, ...\alpha_n$) tal que

$$F(v_1, v_2, ...v_n) = \sum_{i=1}^{n} \alpha_i v_i = 0$$

La combinación trivial corrresponde a la función cero: todos los escalares son cero.

Apéndice B

Isomorfismos

¿cuándo dos cosas son "la misma cosa?' o al menos idénticas. Según Leibnitz dos objetos O_1 y O_2 son el mismo si y sólo si cualquier cosa cierta sobre O_1 es también cierta de O_2 y viceversa. Este criterio es tan restrictivo que en realidad se convierte en una trivialidad çada objeto es idéntico a si mismo y a nada más".

Pero en lenguaje ordinario decimos cosas del tipo .ᵉˢtas dos llaves de mi casa son iguales". Pero dos llaves nunca podrán ser idénticas, una podría ser de latón y la otra de cobre, hay siempre pequeñasa diferencias de forma etc. Pero cuando decimos que son iguales lo que queremos decir es que abren las mismas cerraduras. Es decir, queremos un concepto de igualdad que no sea tan estricto como el de Leibnitz sino que más bien indique que son iguales para ciertos fines.

Si defino la relación ˜ para indicar $ll_1 \tilde{} ll_2$ si y solo si ll_1 y ll_2 abren las mismas cerraduras, es fácil ver que esta relación tiene las siguientes propiedades:

1. Para cualquier llave ll tenemos que $ll \tilde{} ll$ (reflexividad)

2. Si $ll_1 \tilde{} ll_2$ entonces $ll_2 \tilde{} ll_1$ (simetría)

3. Si $ll_1 \tilde{} ll_2$ y $ll_2 \tilde{} ll_3$ entonces $ll_1 \tilde{} ll_3$ (transitividad)

Una relación que es reflexiva, simétrica y transitiva, recibe el nombre de relación de equivalencia. Si dos cosas están reelacionadas por una relación de equivalencia que no sea la relación de igualdad estricta, podemos decir que son lo mismo desde cierto punto de vista, para ciertos propósitos (como el de abrir puertas) .

Algo similar ocurre con los juegos, digamos el ajedrez. Si tenemos dos jugadores en paises distintos, uno tiene un tablero de marfil, el otro de cartón; el primero tiene piezas de platino, el segundo de plático. ¿están jugando el mismo juego? Lo primero a verificar es que su tablero y piezas son similares en el sentido de que cada uno tiene un tablero y tienen el mismo número de piezas. Es decir necesitamos una función ϕ que asocie a cada pieza del primer jugador una y sólo una pieza de otro. También asocia con el tablero de uno el tablero del otro. Esta función debe de der.pues, biyectiva. Debe de ser inyectiva pues dos piezas del primer jugador no han de corresponder a una misma pieza del segundo jugador. Debe de ser suprayectiva pues no debe haber piezas del segundo que no correspondan a una del primero. Por lo tanto ϕ es invertible.

Pero esto no basta, aunque se parezcan las piezas para que jueguen el mismo juego deben de seguir las mismas reglas. No se vale que uno juegue ajedrez y el otro damas.

Vamos ahora a los espacios vectoriales. Aquí las piezas son vectores, diremos entonces que para que en dos espacios vectoriales V y W estemos teniendo el mismo juego necesitamos un mapeo biyectivo $\phi : V \to W$ que además "preserve estructura". Por ejemplo su en V tomo dos vectores a y b tendré en W sus imágenes $\phi(a)$ y $\phi(b)$. Si los sumo para obtener $a + b$, su imagen será $\phi(a + b)$. Pero el hecho crucial es que en V a , b y $a + b$ están relacionados, el tercero es la suma de los dos primeros. Diremos que se "preserva"la suma si $\phi(a)$, $\phi(b)$ y $\phi(a + b)$ están relacionados de la misma manera y

$$\phi(a + b) = \phi(a) + \phi(b)$$

Similarmente, si tomo $a \in V$ y $\alpha \in F$ (F es el campo), puedo formar αa cuya imagen en W es $\phi(\alpha a)$. Pero a y αa están relacionados, uno es α veces el otro. Decimos que se preserva la multiplicación por escalares si la misma relación ocurre en W es decir

$$\phi(\alpha a) = \alpha\phi(a)$$

Cuando las dos relaciones anteriores se cumplen diremos que ϕ preserva la estructura aditiva y multiplicativa de los espacios. Un isomorfismo es un mapeo biyectivo que cumple las dos condiciones. Un mapeo que nada más cumpla las condiciones se denomina transformación lineal. Un isomorfismo es una transformación lineal invertrible.

Dos espacios con un isomorfismo entre ellos se denominan isomorfos (o isomórficos).

La relación de isomorfismo es una relación de equivalencia.

Apéndice C

Comentarios sobre las referencias

El lector podrá constatar facilmente que existen centenas de libros de álgebra lineal, algunos en español. La lista que sigue comprende los que conozco y uso. Tienen en común el que tratan tanto de espacios reales como de espacios complejos; muchos de los libros más usados tratan tan sólo de los espacios reales, pero nuestro temario exige incluir el caso complejo. Existe traducción al español del libro de Friedberg et al. Para la parte histórica recomiendo el libro de Kleiner.

Bibliografía

[1] Serge Lang. Linear Algebra. Addisson Wesley

[2] Jesús Rojo Álgebra Lineal. Mc Graw Hill

[3] Seymour Lipschutz. Álgebra Lineal. Serie Schaum, Mc Graw Hill

[4] Stephen Friedberg, Lawrence Spence, Arnold Insel. Linear Algebra. Prentice Hall

[5] Kenneth Hoffman, Ray Kunze. Álgebra Lineal. Prentice Hall.

[6] Israel Kleiner. History of Abstract Algebra.

Índice alfabético

ángulo entre vectores, 148

ajuste de rectas, 161
anillo, 15
asociatividad, 4

base, 46
  cambio de, 56
  canónica, 47

campo, 17
cancelación, 6
Cauchy-Schwarz, 141
cerradura, 3
combinaciones lineales, 36
complemento ortogonal, 151
conjunto
  normal, 148
  ortogonal, 148
  ortonormal, 148
conmutatividad, 4
coordenadas , 53
cuadrados mínimos, 159

dependencia lineal, 41
dimensión, 47
  finita, 47
  infinita, 47
distributividad, 5
dominio entero, 16

ecuaciones lineales homogéneas, 35
eigenespacio, 117
eigenvalor, 117
eigenvector, 117
espacio
  normado, 142
espacio columna, 63
espacio renglón, 63
espacio vectorial, 21
  complejo, 22
espacio vectorial
  real, 22
espacios $\mathbb{R}^n$, 23
espacios $\mathbb{C}^n$, 26
espacios $M(m,n)$, 28
espacios $M^*(m,n)$, 29
espacios P_n, 27
espacios vectoriales
  de funciones, 69
  intersección de subespacios, 34
  propiedades básicas, 30
  subespacio, 33
  unión de subespacios, 34
estructura algebráica, 9

función delta de Kronecker, 148

generadores, 39
Gram-Schmidt, 156
grupo, 11
  abeliano, 11
grupoide, 9

identidad, 4
  derecha, 4
  izquierda, 4
imagen, 84
independencia lineal, 41
inversa, 6
  derecha, 6
  izquierda, 6
invertibilidad, 105
isomorfismo, 60

de coordenadas, 61
principio del, 62

métrica, 146
magma, 9
matriz
de transición, 57
rango de, 64
matriz de coordenadas, 54
mejor aproximación, 155
monoide, 9
multiplicación
tabla de , 5

norma, 142
inducida por un producto interior, 143
nucleo, 84

operación
binaria, 3
operación
unaria, 3
operador
identidad, 81
operadores lineales, 117

polinomio característico, 119
producto interior, 135
proyecciones ortogonales, 153

quiralidad, 115

semigrupo, 9
sistema algebraico, 9
subgrupo, 13

teorema
de la dimensión, 88
transformación
inversa, 103
transformación cero, 81
transformación lineal
matriz de, 89
transformaciones
álgebra de , 93
composición de , 98
interpretación geométrica , 113
producto por un número, 94
suma de , 93
transformaciones lineales, 79

variable libre, 40
variable ligada, 40
vector
geométrico (flecha), 22
vector de coordenadas, 54
vectores
ortogonales, 148

wronskiano, 72