

Libros de Cátedra

ÁLGEBRA LINEAL APLICADA

Victoria Vampa

FACULTAD DE
CIENCIAS ASTRONÓMICAS Y GEOFÍSICAS

e
exactas


EDITORIAL DE LA UNLP



UNIVERSIDAD
NACIONAL
DE LA PLATA

ÁLGEBRA LINEAL APLICADA

Victoria Vampa

Facultad de Ciencias Astronómicas y Geofísicas



UNIVERSIDAD
NACIONAL
DE LA PLATA





Agradecimientos

Quiero expresar mi profundo agradecimiento a la Facultad de Ciencias Astronómicas y Geofísicas, y en particular al Dr. Adrián Brunini, por haberme invitado en el año 2012 a elaborar una propuesta para incorporar el dictado de la asignatura Álgebra Lineal en la facultad, y por haber respaldado la redacción de este texto.

Mi reconocimiento especial es para lxs estudiantes, cuyas preguntas y comentarios han enriquecido profundamente mis cursos y han contribuido de manera significativa al desarrollo de este material.



Prólogo

Este libro es el resultado de enseñar durante muchos años los temas de Álgebra Lineal tanto en la Facultad de Ingeniería como en la Facultad de Ciencias Astronómicas y Geofísicas y de las contribuciones importantes que han realizado lxs estudiantes y que han servido notablemente al mejoramiento de mis clases.

En el texto, los desarrollos teóricos y las técnicas se presentan de manera equilibrada. Se describen con detalle y se complementan con numerosos ejemplos, enfatizando la importancia de la interpretación geométrica. La Licenciada en Astronomía Lucía Rizzo Buschiazzo, contribuyó con la elaboración de los problemas de aplicación y de las guías de trabajos prácticos.

Consideramos que el material redactado es adecuado como apoyo y guía de estudio para el desarrollo de la asignatura Álgebra Lineal, cursada por lxs estudiantes de segundo año de la carrera Licenciatura en Astronomía.



Índice

	Introducción	11
1	Espacios Vectoriales	15
1.1	Definición de espacio vectorial. Ejemplos	15
1.2	Subespacio vectorial	18
1.3	Base y dimensión de un espacio vectorial	24
1.4	Intersección y suma de subespacios vectoriales	34
1.5	Cambio de base	39
1.6	Actividades propuestas	47
1.6.1	Ejercicios	48
1.6.2	Autoevaluación	54
2	Transformaciones Lineales	55
2.1	Definición de transformación lineal. Ejemplos.	55
2.2	Matriz de una transformación lineal	61
2.3	Transformaciones lineales inyectivas y suryectivas	69
2.4	Núcleo e imagen de una transformación lineal	70
2.5	Geometría de las transformaciones lineales de \mathbb{R}^2 en \mathbb{R}^2	77

2.6	Cambio de base para transformaciones lineales	79
2.7	Espacio dual de un espacio Vectorial	84
2.8	Actividades propuestas	97
2.8.1	Ejercicios	97
2.8.2	Autoevaluación	103
3	Autovalores y autovectores	105
3.1	Introducción	105
3.2	Subespacios invariantes. Valores y vectores propios	110
3.2.1	Localización de autovalores	118
3.3	Polinomio característico	119
3.4	Diagonalización	123
3.4.1	Espacios propios	128
3.4.2	Polinomios minimales	131
3.5	Teorema de la descomposición prima	138
3.6	Forma canónica de Jordan	144
3.7	Actividades propuestas	152
3.7.1	Ejercicios	153
3.7.2	Autoevaluación	159
4	Espacios vectoriales con producto interno	161
4.1	Producto interno. Ejemplos	161
4.2	Matriz de un producto Interno	163
4.3	Longitud, ángulos, distancia y ortogonalidad	164
4.4	Distancia entre vectores	168
4.4.1	Complemento Ortogonal	177
4.5	Proyección ortogonal	181
4.5.1	Problema de cuadrados mínimos	185
4.6	Endomorfismos de espacios vectoriales con producto interno	190
4.7	Actividades propuestas	201
4.7.1	Ejercicios	202

4.7.2	Autoevaluación	207
5	Formas bilineales y cuadráticas	209
5.1	Formas bilineales y cuadráticas	209
5.1.1	Formas cuadráticas. Aplicación a las secciones cónicas	214
5.1.2	Formas cuadráticas: aplicación a las superficies cuádricas	217
5.2	Actividades propuestas	222
5.2.1	Ejercicios	222
5.2.2	Autoevaluación	224
6	Cálculo tensorial	225
6.1	Invariancia y representación	225
6.2	Convenio de suma de Einstein	226
6.3	Notación indicial	227
6.3.1	Delta de Kronecker	227
6.3.2	Símbolo de Levi-Civita	229
6.4	Operaciones con vectores	231
6.5	Transformaciones lineales	234
6.6	Definición de tensor	237
6.6.1	Cambio de coordenadas	237
6.6.2	Tensores de orden 0, 1 y 2	238
6.6.3	Suma. Contracción de índices	240
6.6.4	Transformaciones ortogonales	243
6.7	Tensores cartesianos	244
6.8	Sistema de coordenadas curvilíneas	245
6.8.1	Tensor fundamental (o Tensor Métrico)	247
6.8.2	Bases en coordenadas curvilíneas	251
6.9	Diagonalización de tensores de segundo orden. Invariantes	259
6.10	Tensores de mayor orden	261
6.11	Actividades propuestas	263
6.11.1	Ejercicios	263
6.11.2	Autoevaluación	266

7	Aplicaciones	267
7.1	Ecuaciones diferenciales	267
7.2	Problemas de aproximación de funciones	272
7.2.1	Series de Fourier	273
7.2.2	Series de Haar. Bases de wavelets ortogonales	276
	Apéndice: Ejercicios preliminares	283
	Glosario	291
	Índice alfabético	293
	La autora	297
	Referencias	301
	Bibliografía recomendada	305
	Respuestas a las autoevaluaciones	307



Introducción

Introducción

El Álgebra Lineal es una rama de la Matemática en la que se introducen numerosos conceptos abstractos. Es una disciplina de gran utilidad en la actualidad, en la resolución de problemas complejos y de grandes dimensiones.

Este libro abarca los temas básicos de Álgebra Lineal como son: espacios vectoriales, transformaciones lineales, diagonalización de una matriz y espacios vectoriales con producto interno. Si bien los temas tratados son los mismos que aparecen en la mayoría de los textos introductorios al Álgebra Lineal, el punto de vista con que se enfoca la teoría y la ejercitación se aparta del enfoque tradicional, y se enfatizan las aplicaciones. En todos los temas se establece la conexión fundamental con la interpretación geométrica. Se presenta una gran variedad de ejemplos y se proponen, además de ejercicios, actividades de investigación especialmente diseñadas para estudiantes de Astronomía. El texto tiene además un capítulo de cálculo tensorial y otro capítulo que describe aplicaciones en la resolución de sistemas ecuaciones diferenciales y en la aproximación de funciones.

En cuanto al origen, la palabra *Álgebra* procede del título de un tratado de un matemático, geógrafo y astrónomo persa conocido como Al-Juarismi. Vivió aproximadamente entre los años 780 y 850, en un tiempo de esplendor del mundo islámico. Su tratado, el *Hisab al-yabr wa'l muqabala* es un Compendio de cálculo por restauración y reducción [1]:


Al-yabr

Al-yabr, restauración, la palabra del título que ha dado origen al término álgebra, es una de las operaciones básicas que ofrece para resolver ecuaciones y que consiste en pasar los términos negativos de un lado de la ecuación como positivos al otro. Mientras que la otra operación, la muqabala, consiste en simplificar la ecuación agrupando los términos similares.

La historia del Álgebra Lineal moderna se remonta a mediados del siglo XIX con los trabajos de William Hamilton, quien introdujo el uso del término vector. Sin embargo, fue recién en la segunda mitad del siglo XX, cuando se incorporó al Álgebra Lineal como una materia básica e introductoria en las matemáticas universitarias.

Por sus múltiples aplicaciones, el estudio del Álgebra Lineal cobra cada día más importancia. Su teoría es extensamente usada en el análisis funcional, en el análisis vectorial y en las ecuaciones diferenciales, entre otras áreas. Cabe señalar que sus numerosas aplicaciones no se restringen al campo de las ciencias exactas, sino que se extienden también al campo de las ciencias naturales y de las ciencias sociales.

Con la escritura de este libro he intentado hacer interesantes y accesibles los temas de Álgebra Lineal, equilibrando los desarrollos teóricos con las técnicas que se utilizan en las aplicaciones, pretendiendo proporcionar a lxs estudiantes las habilidades algebraicas necesarias para resolver problemas. He resaltado las interpretaciones geométricas de conceptos importantes, como las transformaciones lineales y el producto interno.

El texto tiene siete capítulos, con una breve introducción al comienzo de cada uno de ellos. Para facilitar la lectura, en todos los capítulos se ha indicado con  a las observaciones importantes. Además, para una mejor comprensión de los temas, se han incluido numerosos ejemplos.

A continuación de la teoría, en cada capítulo, y fruto de la colaboración de Lucía Rizzo Buschiazzo, se proponen actividades a desarrollar por lxs estudiantes. Su tarea consistió en el diseño de un problema de aplicación, la selección de ejercicios y la elaboración de una autoevaluación en cada capítulo. Propuso, además, una serie de ejercicios para realizar como precalentamiento antes de abordar el libro que se incluye en el apéndice. Estos temas, como también los relacionados con las plantillas de edición e imágenes de este proyecto, se definieron a partir del intercambio de ideas y de un trabajo conjunto.

Por último, y para estimular al lector el interés sobre el desarrollo histórico de los temas, se incluyen varias notas históricas dispersas a lo largo del libro, y semblanzas breves sobre científicos

que han realizado aportes muy valiosos al desarrollo del Álgebra Lineal.

Template: the Legrand Orange Book Template

<http://www.latextemplates.com/template/the-legrand-orange-book>

Imagen en la portada de los capítulos: Nebulosa Roseta, IC 1396B,

obtenida por el relevamiento fotométrico IPHAS/N, preparada por Nick Wright.

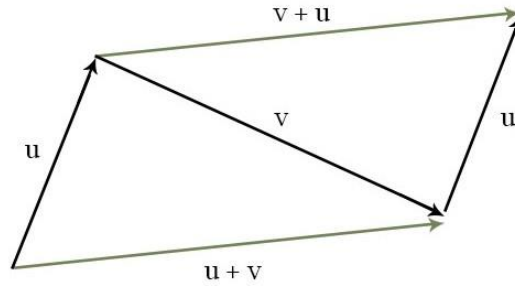


1. Espacios Vectoriales

El concepto de espacio vectorial generaliza las propiedades que tienen las operaciones de suma y producto por escalares para los vectores de \mathbb{R}^2 y \mathbb{R}^3 . Abordaremos en este capítulo la estructura de espacio vectorial, objeto básico de estudio del Álgebra Lineal. A sus elementos se los denomina vectores, independientemente de su naturaleza.

1.1 Definición de espacio vectorial. Ejemplos

El conjunto de los números reales y el conjunto de los números complejos, con los cuales ya se trabajó, tienen propiedades similares. En ambos conjuntos pueden definirse dos operaciones $+$ y \cdot que satisfacen ciertas propiedades y reciben el nombre de *cuerpo*. Trabajaremos tanto con el cuerpo de los reales, \mathbb{R} como con el cuerpo de los complejos \mathbb{C} , denotándolos por K . Al estudiar vectores en el plano y en el espacio, se ha definido la *suma de vectores* y la *multiplicación por un número real*, y se vieron las propiedades que satisfacían. También para el conjunto de polinomios. Cuando en varios conjuntos distintos aparecen estructuras similares es conveniente axiomatizar éstas y darles un nombre al ente resultante, con la ventaja de que estudiando esta estructura, quedan estudiadas todas las estructuras que en ella se encuadran. Cuando en un conjunto se da una estructura similar a la de los ejemplos anteriores, se dice que se tiene un *espacio vectorial*.

Figura 1.1: $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ 

Definición 1.1.1 Un conjunto V , cuyos elementos se denotan mediante \vec{u} , \vec{v} , \vec{w} , se dice que es un *espacio vectorial* sobre el cuerpo K , si en él se han definido dos operaciones: la *suma*, de manera que a cada par de elementos \vec{u} y \vec{v} de V se le hace corresponder el elemento $\vec{u} + \vec{v}$ de V , denominado suma de \vec{u} y \vec{v} , y la *multiplicación por escalares*, de manera que a todo elemento \vec{u} de V y a todo elemento a de K se le hace corresponder el elemento $a\vec{u}$ de V , y se satisfacen las siguientes propiedades:

1. Conmutativa $\vec{u} + \vec{v} = \vec{v} + \vec{u} \quad \forall \vec{u} \text{ y } \vec{v} \in V$.
2. Asociativa $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}, \quad \forall \vec{u}, \vec{v} \text{ y } \vec{w} \in V$.
3. Existe un elemento de V , designado por $\vec{0}$ y denominado *elemento neutro*, tal que $\vec{u} + \vec{0} = \vec{u} \quad \forall \vec{u} \in V$.
4. Para todo elemento $\vec{u} \in V$, existe un elemento designado por $-\vec{u}$ y denominado *elemento opuesto* de \vec{u} , tal que $\vec{u} + (-\vec{u}) = \vec{0}$.
5. $1\vec{u} = \vec{u} \quad \forall \vec{u} \in V$, donde 1 denota el elemento unidad del cuerpo K .
6. $a(b\vec{u}) = (ab)\vec{u} \quad \forall \vec{u} \in V, \text{ y } \forall a \text{ y } b \in K$.
7. $(a+b)\vec{u} = a\vec{u} + b\vec{u} \quad \forall \vec{u} \in V, \text{ y todo } a \text{ y } b \in K$.
8. $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v} \quad \forall \vec{u}, \vec{v} \in V, \text{ y } \forall a \in K$.

En la Figura 1.1 se muestra la propiedad conmutativa 1.



- Los elementos del espacio vectorial reciben el nombre genérico de *vectores*.

- Las primeras cuatro propiedades se refieren a la suma en V , las dos que siguen a la multiplicación de elementos de V por escalares y las dos últimas son las propiedades distributivas de una operación con respecto a la otra.
- Si K es \mathbb{R} se dice que V es un *espacio vectorial real*, y si K es \mathbb{C} se dice que es un *espacio vectorial complejo*.

■ **Ejemplo 1.1** \mathbb{R} es un espacio vectorial sobre \mathbb{Q} , \mathbb{C} es un espacio vectorial sobre \mathbb{R} y sobre \mathbb{Q} . \mathbb{R}^2 o \mathbb{R}^3 (vectores en el plano, o en el espacio), con las operaciones usuales son espacios vectoriales sobre \mathbb{R} . ■

■ **Ejemplo 1.2**

$$K^n = \{(x_1, x_2, \dots, x_n), x_j \in K, j = 1, 2, \dots, n\}$$

con las operaciones usuales es un espacio vectorial sobre K . En particular, \mathbb{R}^n es un espacio vectorial real y \mathbb{C}^n es un espacio vectorial complejo. ■

■ **Ejemplo 1.3** Sea $P_K[x]$ el conjunto de todos los polinomios en la variable x sobre el cuerpo K , es decir, todos los elementos de la forma

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n + \dots$$

donde los coeficientes $a_j \in K$ con las operaciones suma de polinomios y multiplicación por escalares. $P_K[x]$ es un espacio vectorial sobre K . ■

■ **Ejemplo 1.4** Sea $C([a, b])$ el conjunto de todas las funciones continuas definidas en el intervalo real $[a, b]$, con valores en \mathbb{R} , $\{f : [a, b] \rightarrow \mathbb{R}\}$ con las operaciones suma de funciones,

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

y multiplicación de una función por un escalar,

$$(af)(x) = a(f(x)).$$

Puede comprobarse fácilmente que $C([a, b])$ es un espacio vectorial. El elemento neutro es la función nula. ■

■ **Ejemplo 1.5** El conjunto $S(A)$ de soluciones del sistema homogéneo $A\vec{X} = \vec{0}$, donde $A \in \mathbb{R}^{m \times n}$ y $\vec{X} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ es un espacio vectorial sobre \mathbb{R} . Es un subespacio de \mathbb{R}^n . ■

i Los siguientes son algunos resultados que se deducen de las propiedades que definen un espacio vectorial y se dejan como ejercicio para el lector.

- El elemento neutro de un espacio vectorial es único.
- El opuesto de cada elemento en un espacio vectorial es único.
- Para todo \vec{u} de un espacio vectorial V , $0 \cdot \vec{u} = \vec{0}$.
- Para todo elemento \vec{u} de un espacio vectorial V , $(-1)\vec{u}$ es su opuesto.
- En todo espacio vectorial V , $a \cdot \vec{0} = \vec{0}$, donde $a \in K$ y $\vec{0}$ es el elemento neutro de V .

1.2 Subespacio vectorial

Algunos subconjuntos de un espacio vectorial V son a su vez espacios vectoriales con las operaciones definidas en V ; estos subconjuntos especiales reciben el nombre de *subespacios vectoriales* de V .

Definición 1.2.1 Un *subespacio vectorial* de un espacio vectorial V es un subconjunto S no vacío de V , que a su vez es un espacio vectorial con las operaciones definidas en V .

i Para demostrar que un subconjunto S es un subespacio vectorial no es necesario comprobar de nuevo que satisface todas las propiedades del espacio vectorial. Es suficiente demostrar que contiene al vector nulo, que la suma de dos elementos de S es otro elemento de S , y que la multiplicación de un elemento de S por un elemento del cuerpo K , es otro elemento de S :

1. $\vec{0} \in S$
2. Si \vec{u} y $\vec{v} \in S$, $\vec{u} + \vec{v} \in S$
3. Si $a \in K$ y $\vec{u} \in S$, $a \vec{u} \in S$

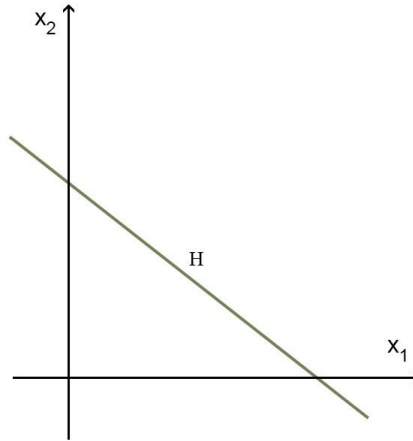
■ **Ejemplo 1.6** Sea V un espacio vectorial sobre K . $S = \{\vec{0}\}$ es un subespacio de V . ■

■ **Ejemplo 1.7** V es un subespacio de V . ■

■ **Ejemplo 1.8** Veamos cómo caracterizar los subespacios de \mathbb{R}^2 .

1. $S = \{(0,0)\}$ es un subespacio.
2. Supongamos S un subespacio que contiene algún elemento \vec{u} no nulo. Entonces para todo $a \in R$, $a\vec{u} \in S$. Si esos son todos los elementos de S , S es un subespacio y gráficamente es una recta por el origen.
3. Si S contiene a un \vec{v} que no es $a\vec{u}$, contiene a sus múltiplos $a\vec{v}$. Luego contiene a dos rectas $L_{\vec{u}}$ y $L_{\vec{v}}$ por el origen. Por la regla del paralelogramo cualquier $\vec{w} \in \mathbb{R}^2$ es suma de un elemento de $L_{\vec{u}}$ y uno de $L_{\vec{v}}$. En consecuencia $S = \mathbb{R}^2$.

Figura 1.2: Recta que no pasa por el origen



Los subespacios de \mathbb{R}^2 son entonces, el vector nulo, las rectas por el origen y todo \mathbb{R}^2 . ■

■ **Ejemplo 1.9** Sea $H = \{(x, y) \text{ tales que } y = mx + b \mid m, b \in \mathbb{R}, b \neq 0\}$ (Ver Figura 1.2). H no es un subespacio de \mathbb{R}^2 . Ya que si (x_1, y_1) y (x_2, y_2) son 2 puntos sobre la recta $y = mx + b$, $y_1 = mx_1 + b$ e $y_2 = mx_2 + b$, se tiene que $y_1 + y_2 = m(x_1 + x_2) + 2b$, y entonces, $y_1 + y_2 \notin H$. O bien, directamente, no es subespacio de \mathbb{R}^2 porque $(0, 0) \notin H$. ■

■ **Ejemplo 1.10** Si $\vec{v} \in V$, $S = \{a\vec{v}, a \in K\}$ es un subespacio de V . Este subespacio se denomina el subespacio generado por \vec{v} , y se nota $S = \langle \vec{v} \rangle$. ■

Proposición 1.2.1 Sean $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n \in V$. Entonces $S = \{a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_n\vec{v}_n, a_i \in K\}$ es un subespacio de V .

Demostración:

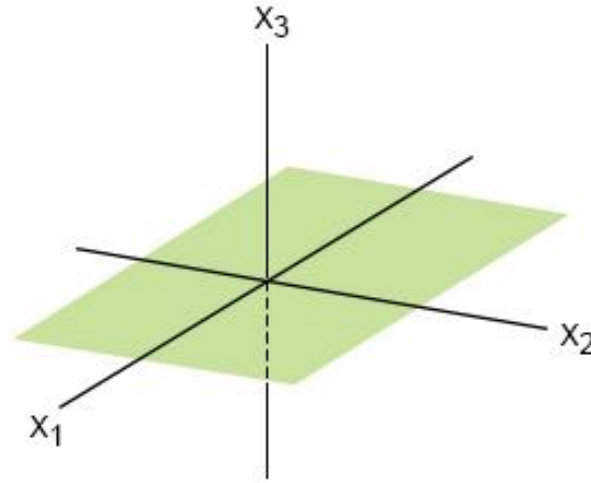
- $\vec{0} \in S$ ya que $\vec{0} = 0\vec{v}_1 + 0\vec{v}_2 + \dots + 0\vec{v}_n$, $0 \in K$.
- Si $\vec{u} = a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_n\vec{v}_n$, $a_i \in K$ y $\vec{w} = b_1\vec{v}_1 + b_2\vec{v}_2 + \dots + b_n\vec{v}_n$, $b_i \in K$, entonces $\vec{u} + \vec{w} = (a_1 + b_1)\vec{v}_1 + (a_2 + b_2)\vec{v}_2 + \dots + (a_n + b_n)\vec{v}_n$, $(a_i + b_i) \in K$, por lo tanto, $\vec{u} + \vec{w} \in S$.
- Si $\alpha \in K$ $\alpha\vec{u} = (\alpha a_1)\vec{v}_1 + (\alpha a_2)\vec{v}_2 + \dots + (\alpha a_n)\vec{v}_n$, $(\alpha a_i) \in K$, por lo tanto $\alpha\vec{u} \in S$

Se tiene, entonces, que S es un subespacio de V . □



El espacio vectorial \mathbb{R}^2 no es un subespacio de \mathbb{R}^3 . \mathbb{R}^2 ni siquiera es un subconjunto de \mathbb{R}^3 . En cambio el conjunto

Figura 1.3: El plano x_1x_2 , ($x_3 = 0$) es un subespacio de \mathbb{R}^3



$$S = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R} \right\}$$

sí es un subconjunto y un subespacio de \mathbb{R}^3 (Ver Figura 1.3).

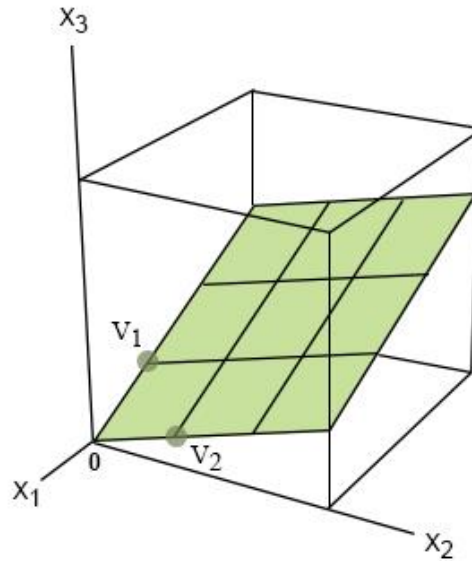
■ **Ejemplo 1.11** $P_{\mathbb{R}}^{(n)}[t]$ (polinomios en t de grado $\leq n$, con coeficientes reales) es un subespacio vectorial de $P_{\mathbb{R}}[t]$; a su vez, $P_{\mathbb{R}}[t]$ es un subespacio vectorial del espacio vectorial de las funciones continuas en \mathbb{R} . ■

■ **Ejemplo 1.12** Sea $S = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ un conjunto de n vectores de un espacio vectorial V . Consideremos como en el Ejemplo 1.10 pero con más vectores. Se define el conjunto de todas las combinaciones lineales de los vectores de S ,

$$\mathbf{L}(S) = \mathbf{L}(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n) = \left\{ \sum_{j=1}^n a_j \vec{u}_j, \quad a_j \in K, j = 1, 2, \dots, n \right\}$$

El conjunto $\mathbf{L}(S)$ es un subespacio vectorial de V (ver Proposición 1.2.1), que recibe el nombre de *subespacio vectorial generado por S* .

En \mathbb{R}^3 , si \vec{v}_1 y \vec{v}_2 son dos vectores tales que uno no es múltiplo del otro, entonces, $\mathbf{L}(\vec{v}_1, \vec{v}_2)$ es un plano que pasa por el origen. Es un subespacio de \mathbb{R}^3 ; se muestra en la Figura 1.4. ■

Figura 1.4: Subespacio generado por \vec{v}_1 y \vec{v}_2 

■ **Ejemplo 1.13** Sean $a_1, a_2, \dots, a_n \in K$ fijos. $S = \{(x_1, x_2, \dots, x_n) \in K^n, a_1x_1 + a_2x_2 + \dots + a_nx_n = 0\}$ es un subespacio de K^n . ■

■ **Ejemplo 1.14** Dada una matriz $A \in \mathbb{R}^{m \times n}$, y de rango r , todas las soluciones del sistema de ecuaciones homogéneo

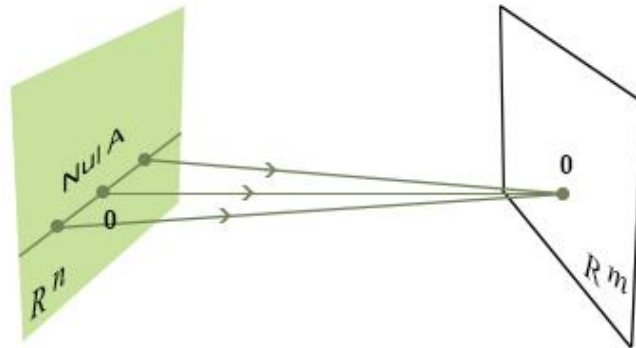
$$A\vec{X} = \vec{0}, \quad \vec{X} \in \mathbb{R}^n$$

constituyen un subespacio vectorial de \mathbb{R}^n , conocido como *espacio nulo* de la matriz A . Se anota $Nul(A)$ y se muestra en la Figura 1.5.

Para el sistema homogéneo:

$$\begin{cases} 2y - z + w &= 0 \\ 3x + y + 10z + 5w &= 0 \\ x + 3z + w &= 0 \end{cases}$$

luego de realizar operaciones elementales sobre las filas de la matriz de coeficientes del sistema (método de eliminación gaussiana), se llega a la matriz escalonada:

Figura 1.5: Espacio nulo de una matriz A 

$$\begin{pmatrix} 3 & 1 & 10 & 5 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & -1/2 & 1/2 \end{pmatrix}$$

de donde la solución es $z = -w$, $y = -w$ y $x = 2w$. El subespacio de soluciones del sistema homogéneo es, entonces,

$$S = \text{Nul}(A) = \langle (2, -1, -1, 1) \rangle$$

■

i Así como vimos que las rectas que no pasan por el origen no son un subespacio de \mathbb{R}^2 (Ejemplo 1.9), las soluciones de un sistema no homogéneo

$$A\vec{x} = \vec{b}, \quad \vec{b} \neq \vec{0}$$

son un subconjunto pero no un subespacio de \mathbb{R}^n .

William Rowan Hamilton (1805 - 1865)

Fue un matemático británico. Fue uno de los fundadores de la escuela británica moderna de matemáticas puras e hizo importantes contribuciones al desarrollo de la óptica, la dinámica, y el álgebra. Su descubrimiento del cuaternión, junto con su sistematización de la dinámica, son sus trabajos más conocidos. Este último trabajo sería decisivo en el desarrollo de la mecánica cuántica, donde un concepto fundamental llamado hamiltoniano lleva su nombre. Hamilton fue el cuarto de los nueve hijos. Vivían en Dublín. Se dice que Hamilton demostró un inmenso talento a una edad muy temprana. Su tío observó que Hamilton, había mostrado una asombrosa habilidad para aprender idiomas. A la edad de siete años, ya había hecho un progreso considerable con el hebreo, y antes de los trece años, bajo la supervisión de su tío (un lingüista), había adquirido conocimientos casi en tantos idiomas como años de edad tenía (idiomas europeos clásicos y modernos, y persa, árabe, hindustaní, sánscrito e incluso maratí y malayo). Hamilton es reconocido como uno de los científicos más destacados de Irlanda, y a medida que la nación se vuelve más consciente de su herencia científica, cada vez se lo celebra más. Se dice que se le permitía pisar el césped de la Universidad, algo totalmente prohibido. Este hecho camina entre la realidad y la ficción. Posiblemente ocurriera que, absorto en sus meditaciones, descuidara esta prohibición y accidentalmente caminase por los jardines. Esta anécdota seguramente sirve para dar idea de la categoría de Hamilton como uno de los grandes matemáticos de su tiempo y de la historia. El Instituto Hamilton está dedicado a la investigación sobre matemáticas aplicadas en la Universidad Maynooth. Irlanda emitió dos sellos conmemorativos en 1943 para celebrar el centenario del anuncio de los cuaterniones. El Banco Central de Irlanda acuñó en 2005 una moneda de plata conmemorativa de 10 euros para conmemorar los 200 años desde su nacimiento. Los talleres de mantenimiento más nuevos del sistema de tranvías de Dublín (LUAS), llevan su nombre.

En su juventud, Hamilton tuvo un telescopio y se convirtió en un experto en el cálculo de fenómenos celestes, como por ejemplo, la determinación de la visibilidad de los eclipses de luna. Fue elegido Astrónomo Real de Irlanda y se instaló en el Observatorio de Dunsink, donde permaneció hasta su muerte en 1865. Hoy en día, Hamilton no es reconocido como un gran astrónomo, aunque durante su vida sí gozó de esta consideración. Sus conferencias de introducción a la astronomía fueron famosas; además de sus alumnos, atrajeron a muchos eruditos y poetas, e incluso a damas; en aquellos días una hazaña notable. La poetisa Felicia Hemans escribió su poema "La oración del estudiante solitario" después de escuchar una de sus conferencias. [28]

1.3 Base y dimensión de un espacio vectorial

Sea V un espacio vectorial sobre un cuerpo K ; un número finito de vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ se dice que son *linealmente dependientes* si existen n elementos de K , a_1, a_2, \dots, a_n no todos nulos, tal que

$$a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_n \vec{v}_n = \vec{0}$$

Si los vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ no son linealmente dependientes, se dice que son *linealmente independientes*; por lo tanto, los vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ son linealmente independientes si cualquier igualdad como la anterior implica que todos los elementos de K , a_1, a_2, \dots, a_n son nulos.

Si en la igualdad anterior a_n es no nulo, podemos escribir

$$\vec{v}_n = -\frac{a_1}{a_n} \vec{v}_1 - \frac{a_2}{a_n} \vec{v}_2 + \dots - \frac{a_{n-1}}{a_n} \vec{v}_{n-1}$$

y decimos que \vec{v}_n es una *combinación lineal* de los vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{n-1}$. En general, se dice que \vec{v} es combinación lineal de los vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$, si existen $a_1, a_2, \dots, a_k \in K$ tal que

$$\vec{v} = a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_k \vec{v}_k$$

Un conjunto finito de vectores $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ de un espacio vectorial V se dice que es *un sistema de generadores* de V si todo elemento de V se puede escribir como una combinación lineal de los vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$.

Proposición 1.3.1 Un conjunto finito de vectores linealmente independientes de un espacio vectorial V no puede contener un subconjunto de vectores que sean linealmente dependientes.

Demostración:

Si $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ son linealmente independientes y suponemos que $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$, $k \leq n$ son linealmente dependientes se tendría

$$\vec{v} = a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_k \vec{v}_k = \vec{0}$$

con no todos los a_j nulos; basta observar que, entonces,

$$a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_k \vec{v}_k + 0 \vec{v}_{k+1} + \dots + 0 \vec{v}_n = \vec{0}$$

con lo cual los originales serían linealmente dependientes. □

Antes de exponer algunos ejemplos es conveniente realizar algunas observaciones.



- Todo conjunto finito de vectores que contiene al elemento neutro (o nulo) es linealmente dependiente; basta observar que

$$a\vec{0} + 0\vec{v}_2 + \cdots + 0\vec{v}_n = \vec{0}$$

para cualquier $a \in K$.

- Tres vectores no nulos de \mathbb{R}^2 son siempre linealmente dependientes.
- En general, $n + 1$ vectores de K^n son siempre linealmente dependientes.

■ **Ejemplo 1.15** Si $\vec{u}_1 = (1, 0, 1)$, $\vec{u}_2 = (-1, 1, 0)$ y $\vec{u}_3 = (1, 1, 2)$, $L(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ es un subespacio vectorial de \mathbb{R}^3 . No es todo \mathbb{R}^3 porque estos vectores no son linealmente independientes, ya que, se anula el determinante de la matriz que tiene esos vectores como filas:

$$\begin{vmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 1 & 2 \end{vmatrix} = 0$$

Para hallar el subespacio que generan esos vectores se realizan operaciones elementales sobre las filas, y se llega a la matriz escalonada:

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

La última fila de ceros indica que el vector \vec{u}_3 es combinación lineal de \vec{u}_1 y \vec{u}_2 . Entonces los vectores generados por \vec{u}_1 , \vec{u}_2 y \vec{u}_3 son de la forma $\alpha(1, 0, 1) + \beta(0, 1, 1) = (\alpha, \beta, \alpha + \beta)$ por lo que $L(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ es el plano por el origen $z = x + y$. Considerando la matriz que tiene los vectores \vec{u}_i como filas, $L(S)$ es el espacio generado por las filas de la matriz, conocido como *espacio fila*. ■

■ **Ejemplo 1.16** Las funciones $p_0(t) = 1$, $p_1(t) = t$, $p_2(t) = t^2$, \dots , $p_n(t) = t^n$, son linealmente independientes, ya que si tenemos la igualdad

$$a_0 1 + a_1 t + a_2 t^2 + \cdots + a_n t^n = 0$$

para todo $t \in \mathbb{R}$, resultan $a_0 = a_1 = a_2 = \cdots a_n = 0$.

Para demostrarlo basta tomar n puntos t_i distintos y resolver el sistema. Tiene como única solución la trivial. El determinante del sistema es conocido como determinante de Vandermonde. ■

■ **Ejemplo 1.17** Las funciones $f(t) = \cos^2(t)$, $g(t) = \sin^2(t)$ y $h(t) = 1$ son linealmente dependientes en $C([0, 2\pi])$ ya que $\cos^2(t) + \sin^2(t) = 1$, y entonces es posible escribir al vector nulo con coeficientes no todos nulos

$$1\cos^2(t) + 1\sin^2(t) + (-1)1 = 0.$$

Por otro lado, ejemplos de funciones linealmente independientes son $f_1(t) = e^{k_1 t}$ y $f_2(t) = e^{k_2 t}$ con $k_1 \neq k_2$. ■

i La independencia lineal de funciones es de importancia para describir el conjunto solución de ecuaciones diferenciales y se determina a partir del cálculo de un determinante conocido como Wronskiano (ver [20]).

■ **Ejemplo 1.18** $S = \{\cos(nx), \sin(mx)\}_{n,m \in \mathbb{N}}$ es un conjunto de funciones linealmente independiente en $C([0, 2\pi])$. ■

i Al desarrollo en serie de una función en términos de las funciones $\cos(nx)$ y $\sin(mx)$ con $n, m \in \mathbb{N}$ se lo conoce como Serie de Fourier.

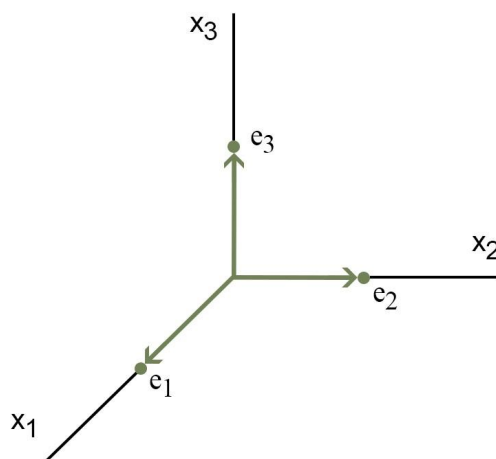
Definición 1.3.1 Un conjunto finito de vectores $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ se dice que es una *base* de un espacio vectorial V si se cumplen las dos condiciones siguientes:

1. Los vectores $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ son linealmente independientes.
2. Todo elemento de V es una combinación lineal de los vectores $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$.

i La segunda condición de esta definición es equivalente al hecho de que el conjunto de vectores $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ sea un sistema de generadores de V . Sin embargo, no todo sistema de generadores de un espacio vectorial V es una base. Se deja al lector pensar ejemplos.

■ **Ejemplo 1.19** Si $\vec{e}_j = (0, 0, \dots, 1, \dots, 0) \in K^n$, donde 1 ocupa el lugar j , se tiene que $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ son linealmente independientes y además si $\vec{x} = (x_1, x_2, \dots, x_n) \in K^n$, se tiene que

$$\vec{x} = \sum_{j=1}^n x_j \vec{e}_j$$

Figura 1.6: Base canónica de \mathbb{R}^3 

Por lo tanto $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ es una base de K^n , que recibe el nombre de *base canónica* de este espacio. En la Figura 1.6 se muestra para el caso $n = 3$.

■

■ **Ejemplo 1.20** Dada una matriz A de m filas y n columnas, y de rango r , todas las soluciones del sistema de ecuaciones homogéneo

$$A\vec{X} = \vec{0}, \quad \vec{X} \in \mathbb{R}^n$$

constituyen un subespacio vectorial de \mathbb{R}^n generado por $n - r$ vectores. Recordar que r es la cantidad de pivotes al realizar operaciones elementales sobre las filas de la matriz en eliminación gaussiana.

En el Ejemplo 1.14 se tiene que $m = 3$, $n = 4$ y el rango $r = 3$. $S = \text{Nul}(A) = \langle (2, -1, -1, 1) \rangle$, es un subespacio de dimensión $n - r = 4 - 3 = 1$.

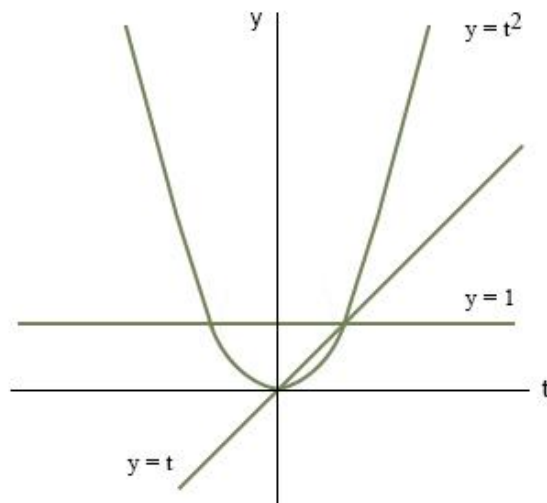
■

■ **Ejemplo 1.21** El conjunto $\{1, t, \dots, t^n\}$ es una base de $P_K^{(n)}[t]$, ya que son polinomios linealmente independientes de acuerdo con el resultado del Ejemplo 1.16 (para $K = \mathbb{R}$), y todo polinomio p de grado inferior o igual a n puede escribirse de la forma

$$p(t) = a_0 1 + a_1 t + a_2 t^2 + \dots + a_n t^n$$

Para para el caso $n = 2$ se tiene la base $\{1, t, t^2\}$ que se muestra en la Figura 1.7.

■

Figura 1.7: Base canónica de $P_R^{(2)}[t]$ 

- i** El conjunto $\{1, t, \dots, t^n\}$ no es una base de $P_K[t]$, ya que el polinomio $p(t) = t^{n+1}$ no es combinación lineal de estos. Se puede ver que ningún conjunto finito de polinomios genera a $P_K[t]$ (ver [20]).

Coordenadas de un vector

Si $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ es una base de un espacio vectorial V y \vec{v} es cualquier elemento de V podemos escribir a \vec{v} como combinación lineal de $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$, de la forma

$$\vec{v} = a_1 \vec{e}_1 + a_2 \vec{e}_2 + \dots + a_n \vec{e}_n$$

con $a_j \in K$. Los números a_1, a_2, \dots, a_n se denominan *coordenadas* de \vec{v} con respecto a la base $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$.

Proposición 1.3.2 Las coordenadas de un vector \vec{v} con respecto a una base son únicas.

Demostración:

Si suponemos se tienen coordenadas a_i y b_i , $i = 1, \dots, n$ para un mismo vector \vec{v} ,

$$\vec{v} = a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_n\vec{e}_n$$

y

$$\vec{v} = b_1\vec{e}_1 + b_2\vec{e}_2 + \cdots + b_n\vec{e}_n$$

se tiene que

$$\vec{0} = (b_1 - a_1)\vec{e}_1 + (b_2 - a_2)\vec{e}_2 + \cdots + (b_n - a_n)\vec{e}_n$$

Como $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ son linealmente independientes, $b_1 = a_1, b_2 = a_2, \dots, b_n = a_n$. □

■ **Ejemplo 1.22** Sea $V = \mathbb{R}^3$ y sea E la base canónica. Las coordenadas de un vector \vec{v} se anotan $(x, y, z)_E = (x, y, z)$.

Si en lugar de la base canónica la base es $B = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$, las coordenadas de un vector (x, y, z) son $(z, y - z, x - y)_B$ y se escribe la igualdad $(x, y, z) = (z, y - z, x - y)_B$.

Esto se obtiene escribiendo (x, y, z) como combinación lineal de los vectores de B ,

$$(x, y, z) = a(1, 1, 1) + b(1, 1, 0) + c(1, 0, 0),$$

y resolviendo el sistema lineal:

$$\begin{cases} a + b + c &= x \\ a + b + 0c &= y \\ a + 0b + 0c &= z \end{cases}$$

■

Un mismo espacio vectorial puede poseer varias bases; nuestro próximo objetivo es demostrar que todas ellas han de poseer el mismo número de elementos.

Proposición 1.3.3 Si V es un espacio vectorial que posee una base con n elementos, cualesquiera $n + 1$ vectores de V son linealmente dependientes.

Demostración:

Sea $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ una base de V y sean $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n, \vec{x}_{n+1}$, $n + 1$ vectores de V ,

que pueden escribirse como combinación lineal de la base dada:

$$\vec{x}_1 = \sum_{j=1}^n x_{j1} \vec{e}_j, \quad \vec{x}_2 = \sum_{j=1}^n x_{j2} \vec{e}_j, \quad \vec{x}_n = \sum_{j=1}^n x_{jn} \vec{e}_j, \quad \text{y } \vec{x}_{n+1} = \sum_{j=1}^n x_{jn+1} \vec{e}_j$$

Se quiere ver si son linealmente independientes. Nos preguntamos si existen a_i no todos nulos tales que

$$a_1 \vec{x}_1 + a_2 \vec{x}_2 + \cdots + a_n \vec{x}_n + a_{n+1} \vec{x}_{n+1} = \vec{0}$$

Reemplazando, se tiene

$$a_1 (\sum_{j=1}^n x_{j1} \vec{e}_j) + a_2 (\sum_{j=1}^n x_{j2} \vec{e}_j) + a_n (\sum_{j=1}^n x_{jn} \vec{e}_j) + a_{n+1} (\sum_{j=1}^n x_{jn+1} \vec{e}_j) = \vec{0}$$

Al desarrollar las sumas anteriores y reordenar sacando factor común los vectores \vec{e}_j , se obtiene

$$\vec{e}_1 (x_{11} a_1 + x_{12} a_2 + \cdots + x_{1n} a_n + x_{1n+1} a_{n+1}) = 0$$

$$\vec{e}_2 (x_{21} a_1 + x_{22} a_2 + \cdots + x_{2n} a_n + x_{2n+1} a_{n+1}) = 0$$

...

$$\vec{e}_n (x_{n1} a_1 + x_{n2} a_2 + \cdots + x_{nn} a_n + x_{nn+1} a_{n+1}) = 0$$

Los términos entre paréntesis constituyen un sistema homogéneo de n ecuaciones con $n+1$ incógnitas, a_1, a_2, \dots, a_{n+1} por lo que existe una solución no trivial (a_i no todos nulos).

Se concluye, entonces, que los vectores $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n, \vec{x}_{n+1}$ son linealmente dependientes.

□



De la proposición anterior se deduce un resultado un poco más general: en un espacio vectorial V que posee una base con n elementos, cualesquiera m vectores de V , con $m > n$ son linealmente dependientes. Basta observar que $n+1$ de los m vectores dados han de ser linealmente dependientes, debido a la proposición anterior, y por lo tanto, todos ellos han de formar un conjunto de vectores linealmente dependiente. Este resultado se aplica en la demostración del teorema que sigue.

Jean-Baptiste Joseph Fourier (1768 - 1830)

Fue un matemático y físico francés conocido por sus trabajos sobre la descomposición de funciones periódicas en series trigonométricas convergentes llamadas Series de Fourier, método con el cual consiguió resolver la ecuación del calor. La transformada de Fourier recibe su nombre en su honor. Fue el primero en dar una explicación científica al efecto invernadero en un tratado. Inició sus estudios en la Escuela Superior Benedictina de Auxerre, orientándose inicialmente a la carrera religiosa, hasta que el monarca Luis XV la convirtió en academia militar. Jean-Baptiste fue seleccionado como estudiante en la institución ya reformada, donde permanecería hasta los 14 años de edad, y empezó a ser instruido en idiomas, música, álgebra y matemáticas, materia en la que destacó, lo que le encaminó a dedicarse al estudio de las ciencias. Posteriormente, participó en la Revolución francesa y, gracias a la caída del poder de Robespierre, se salvó de ser guillotinado. Se incorporó a la Escuela Normal Superior de París en donde tuvo entre sus profesores a los matemáticos Joseph Louis Lagrange y Pierre Simon Laplace. Posteriormente, ocupó una cátedra como docente en la prestigiosa École polytechnique. Fourier participó en la expedición de Napoleón Bonaparte a Egipto en 1798. Entró a la Academia de Ciencias Francesa en 1817 y al cabo de cinco años se convirtió en el secretario perpetuo de las secciones de matemáticas y física. Fue en Grenoble donde condujo sus experimentos sobre la propagación del calor que le permitieron modelar la evolución de la temperatura a través de series trigonométricas. Estos trabajos mejoraron el modelado matemático de fenómenos físicos y contribuyeron a los fundamentos de la termodinámica. Sin embargo, la simplificación excesiva que proponen estas herramientas fue muy debatida, principalmente por sus maestros Laplace y Lagrange. Publicó en 1822 su *Théorie analytique de la chaleur* (Teoría analítica del calor), tratado en el cual estableció la ecuación diferencial parcial que gobierna la difusión del calor solucionándola mediante el uso de series infinitas de funciones trigonométricas, lo que establece la representación de cualquier función como series de senos y cosenos, ahora conocidas como las series de Fourier. El trabajo de Fourier provee el impulso para trabajar más tarde en las series trigonométricas y la teoría de las funciones de variables reales. Fourier en esta obra dedujo la ecuación en derivadas parciales que rige tal fenómeno, la cual es conocida como la ecuación del calor. [12]

Proposición 1.3.4 Todas las bases de un mismo espacio vectorial V poseen el mismo número de elementos.

Demostración:


Sean $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ y $\{\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_m\}$ dos bases de un espacio vectorial V ; por lo anterior $m \leq n$, ya que en caso contrario los vectores de la segunda base serían linealmente dependientes. Similarmente $n \leq m$ ya que en caso contrario los vectores de la primera base serían linealmente dependientes. Se tiene, por lo tanto, que $n = m$.

□

El número de elementos que posee una base cualquiera de un espacio vectorial V recibe el nombre de *dimensión* de V ; este número será designado mediante $\dim(V)$. Si el espacio vectorial sólo contiene un elemento, es decir $V = \{\vec{0}\}$ tiene *dimensión cero*.


De los ejemplos anteriores podemos deducir los siguientes resultados:

1. La dimensión de K^n es n .
2. La dimensión de $P_K^{(n)}[t]$ es $n + 1$.
3. En el Ejemplo 1.15 se puede ver que $\mathbf{L}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$, donde $\vec{u}_1 = (1, 0, 1)$, $\vec{u}_2 = (-1, 1, 0)$ y $\vec{u}_3 = (1, 1, 2)$, es un subespacio vectorial de \mathbb{R}^3 de dimensión 2 (un plano por el origen).

 En \mathbb{R}^n un *hiperplano* que contiene al vector nulo es un subespacio H de dimensión $n - 1$. O sea

$$H = \{(x_1, x_2, \dots, x_n) : a_1x_1 + a_2x_2 + \dots + a_nx_n = 0\}$$

donde a_1, a_2, \dots, a_n son números reales fijos, no todos nulos. Es decir, un hiperplano generaliza la noción de plano en \mathbb{R}^3 .

 Se llama *nulidad* de una matriz a la dimensión del espacio nulo.

Proposición 1.3.5 Sea V un espacio vectorial de dimensión n . Todo conjunto de n vectores de V que sean linealmente independientes son una base de V .

Demostración:

Sean $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$, n vectores linealmente independientes. Si \vec{v} es otro vector de V , por la Proposición 1.3.3, $\vec{v}, \vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$ son linealmente dependientes.

Entonces, $\exists a_0, a_1, \dots, a_n$ tales que


$$a_0 \vec{v} + a_1 \vec{x}_1 + a_2 \vec{x}_2 + \dots + a_n \vec{x}_n = \vec{0}, \quad \text{con algún } a_j \text{ no nulo.}$$

En realidad a_0 debe ser no nulo, ya que si fuera 0, los vectores $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$ serían linealmente dependientes. Se tiene, entonces,

$$\vec{v} = \frac{-a_1}{a_0} \vec{x}_1 + \frac{-a_2}{a_0} \vec{x}_2 + \dots + \frac{-a_n}{a_0} \vec{x}_n$$

y por lo tanto $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$ generan V .

□

-  Una forma sencilla de encontrar una base de un espacio vectorial V es agregar vectores a un conjunto de vectores linealmente independientes de V . En la demostración de la proposición que sigue se explica la forma de agregarlos.

Proposición 1.3.6 Sea V un espacio de dimensión finita n ; todo conjunto de vectores linealmente independientes de V puede completarse para obtener una base, es decir, dados k vectores $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k$, con $k < n$, de V , linealmente independientes, existen $n - k$ vectores $\vec{e}_{k+1}, \vec{e}_{k+2}, \dots, \vec{e}_n$ de V tal que el conjunto $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k, \vec{e}_{k+1}, \vec{e}_{k+2}, \dots, \vec{e}_n\}$ es una base de V .

Demostración:

Como $k < n$, puedo encontrar un elemento de V linealmente independiente con $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k$ (sino, $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k\}$ serían base de V). Lo llamo \vec{e}_{k+1} . Se repite con $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k, \vec{e}_{k+1}\}$ hasta encontrar n vectores linealmente independientes, que necesariamente serán base de V . □

Si V es un espacio vectorial de dimensión finita n en la Proposición 1.3.6 probamos que k vectores linealmente independientes de V pueden *completarse* para obtener una base. Puede demostrarse

también que si S es un sistema de generadores de V , de él puede extraerse un subconjunto S_1 que sea una base de V .

Ahora unos comentarios acerca de la dependencia o independencia lineal de subconjuntos infinitos de un espacio vectorial.



- Un *conjunto infinito* S de elementos de un espacio vectorial V se dice linealmente independiente si cualquier subconjunto finito de S es linealmente independiente. En caso contrario, S se dice *linealmente dependiente*; es decir S es linealmente dependiente si existe un subconjunto finito de él que es linealmente dependiente.
- Un espacio vectorial V en el que se puede encontrar un subconjunto S linealmente independiente y con infinitos elementos, se dice que tiene *dimensión infinita*.
- Los espacios vectoriales $P_K[t]$, y $C([0, 2\pi])$, introducidos en las secciones anteriores, son espacios vectoriales de dimensión infinita.
- El conjunto $S = \{t^n, n \in \mathbb{N}\}$ es un conjunto linealmente independiente de $P_K[t]$ mientras que el conjunto $S = \{\cos(nx), \sin(mx)\}_{n,m \in \mathbb{N}}$ es un conjunto linealmente independiente en el espacio vectorial de las funciones continuas $C([0, 2\pi])$.

1.4 Intersección y suma de subespacios vectoriales

Una pregunta que surge es si al considerar las operaciones de unión e intersección entre subespacios de un espacio vectorial V (que son subconjuntos de V) se preserva la estructura de subespacio. Veremos que se preserva en la intersección pero no en la unión.

Dados dos subespacios V_1 y V_2 de un espacio vectorial V podemos definir su *intersección*

$$V_1 \cap V_2 = \{\vec{u}, \vec{u} \in V_1 \wedge \vec{u} \in V_2\}$$

y se demuestra fácilmente que $V_1 \cap V_2$ es un subespacio.

Por otro lado, con un ejemplo se puede ver que con la unión de dos subespacios V_1 y V_2 , no ocurre lo mismo. Si V_1 y V_2 son los subespacios generados por los vectores $(1, 0)$ y $(0, 1)$ (los ejes x e y respectivamente) la unión de V_1 y V_2 son los vectores que están sobre un eje o el otro. $V_1 \cup V_2$ no es un subespacio, ya que la suma no es cerrada: la suma de los vectores $(1, 0) + (0, 1) = (1, 1) \notin V_1 \cup V_2$ pues $(1, 1) \notin V_1$ y $(1, 1) \notin V_2$ siendo que $(1, 0) \in V_1$ y $(0, 1) \in V_2$.

Se define, entonces, para sí obtener un subespacio, la *suma* de dos subespacios V_1 y V_2 de la forma siguiente

$$V_1 + V_2 = \{\vec{u}_1 + \vec{u}_2, \vec{u}_1 \in V_1 \wedge \vec{u}_2 \in V_2\}$$

y se puede demostrar que $V_1 + V_2$ es un subespacio vectorial de V (y contiene a $V_1 \cup V_2$).

La relación que existe entre las dimensiones de estos subespacios vectoriales y las dimensiones de los subespacios V_1 y V_2 queda plasmada en el siguiente resultado:

Proposición 1.4.1

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2)$$

para cualesquiera subespacios vectoriales V_1 y V_2 de un espacio vectorial V de dimensión finita.

Demostración:

Sea $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_l\}$ una base de $V_1 \cap V_2$. Es posible completarla:

por un lado hasta obtener una base de V_1 , $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_l, \vec{f}_{l+1}, \vec{f}_{l+2}, \dots, \vec{f}_k\}$ (de $l + (k - l)$ vectores)

y por otro, hasta obtener una base de V_2 , $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_l, \vec{g}_{l+1}, \vec{g}_{l+2}, \dots, \vec{g}_m\}$ (de $l + (m - l)$ vectores).

Veremos que

$$\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_l, \vec{f}_{l+1}, \vec{f}_{l+2}, \dots, \vec{f}_k, \vec{g}_{l+1}, \vec{g}_{l+2}, \dots, \vec{g}_m\} \text{ es base de } V_1 + V_2.$$

Está claro que es un sistema de generadores de $V_1 + V_2$. Veamos que es un conjunto linealmente independiente. Se considera una combinación lineal igual al vector nulo:

$$a_1\vec{e}_1 + a_2\vec{e}_2 + \dots + a_l\vec{e}_l + b_{l+1}\vec{f}_{l+1} + b_{l+2}\vec{f}_{l+2} + \dots + b_k\vec{f}_k + c_{l+1}\vec{g}_{l+1} + c_{l+2}\vec{g}_{l+2} + \dots + c_m\vec{g}_m = \vec{0}$$

que puede escribirse en forma equivalente

$$a_1\vec{e}_1 + a_2\vec{e}_2 + \dots + a_l\vec{e}_l + b_{l+1}\vec{f}_{l+1} + b_{l+2}\vec{f}_{l+2} + \dots + b_k\vec{f}_k = -c_{l+1}\vec{g}_{l+1} - c_{l+2}\vec{g}_{l+2} - \dots - c_m\vec{g}_m$$

El término del lado izquierdo $\in V_1$, mientras que el lado derecho $\in V_2$. Es decir que está en $V_1 \cap V_2$ cuya base son los vectores $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_l\}$. Entonces es posible escribir el término de la derecha como combinación lineal de la base $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_l\}$ con coordenadas δ_i . Es decir, se tiene la igualdad

$$-c_{l+1}\vec{g}_{l+1} - c_{l+2}\vec{g}_{l+2} - \cdots - c_m\vec{g}_m = \delta_1\vec{e}_1 + \delta_2\vec{e}_2 + \cdots + \delta_l\vec{e}_l$$

que puede reescribirse,

$$\delta_1\vec{e}_1 + \delta_2\vec{e}_2 + \cdots + \delta_l\vec{e}_l + c_{l+1}\vec{g}_{l+1} + c_{l+2}\vec{g}_{l+2} + \cdots + c_m\vec{g}_m = \vec{0}$$

Por ser $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_l, \vec{g}_{l+1}, \vec{g}_{l+2}, \dots, \vec{g}_m\}$ una base de V_2 , se tiene que

$$c_{l+1} = c_{l+2} = \cdots = c_m = \delta_1 = \delta_2 = \cdots = \delta_l = 0$$

y, entonces, el término de la izquierda,

$$a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_l\vec{e}_l + b_{l+1}\vec{f}_{l+1} + b_{l+2}\vec{f}_{l+2} + \cdots + b_k\vec{f}_k = \vec{0}$$

y como $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_l, \vec{f}_{l+1}, \vec{f}_{l+2}, \dots, \vec{f}_k\}$ son base de V_1 , son linealmente independientes,

$$a_1 = a_2 = \cdots = a_l = b_{l+1} = b_{l+2} = \cdots = b_k = 0$$

Luego,

$$\dim(V_1 + V_2) = l + (k - l) + (m - l) = k + m - l = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2) \quad \square$$



- Si V_1 y V_2 son los subespacios generados por los vectores $(1, 0)$ y $(0, 1)$ respectivamente, la suma de V_1 y V_2 da todo \mathbb{R}^2 , ya que $(x, y) = (x, 0) + (0, y)$, $(x, 0) \in V_1$ y $(0, y) \in V_2$.
- Es importante notar que $V_1 + V_2$ es el menor subespacio (con respecto a la inclusión) que contiene a $V_1 \cup V_2$.

■ **Ejemplo 1.23** El subespacio S de las soluciones del sistema homogéneo,

$$S = \{(x_1, x_2, \dots, x_n) \in K^n \mid \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}\}$$

es un subespacio de K^n . Es intersección de m subespacios, $S = \bigcap_{i=1}^m S_i$, donde $S_i = \{(x_1, x_2, \dots, x_n) \in K^n, a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0\}$, $1 \leq i \leq m$. Cada S_i es un subespacio de K^n (correspondiente a las soluciones de cada una de las m ecuaciones, como se vió en el Ejemplo 1.13, Sección 1.2). ■

■ **Ejemplo 1.24** Dos subespacios vectoriales distintos de \mathbb{R}^2 , V_1 y V_2 , ambos de dimensión 1, tienen una suma que coincide con todo \mathbb{R}^2 , ya que

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2) = 1 + 1 - 0 = 2$$

■

Definición 1.4.1 Un espacio vectorial V es *suma directa* de dos subespacios V_1 y V_2 si

1. $V_1 + V_2 = V$
2. $V_1 \cap V_2 = \vec{0}$

Utilizaremos la notación $V = V_1 \oplus V_2$ para indicar que V es suma directa de los subespacios V_1 y V_2 .



- El plano \mathbb{R}^2 puede escribirse como suma directa de dos rectas no coincidentes que pasan por el origen.
- El espacio \mathbb{R}^3 puede escribirse como suma directa de un plano que pasa por el origen y una recta que le corta en ese punto.
- De acuerdo a la Proposición 1.4.1 anterior si $V = V_1 \oplus V_2$, se tiene que

$$\dim(V) = \dim(V_1 \oplus V_2) = \dim(V_1) + \dim(V_2)$$

ya que el subespacio $V_1 \cap V_2 = \vec{0}$ tiene dimensión 0. Además, si B_1 es base de V_1 y B_2 es base de V_2 , $B = B_1 \cup B_2$ es una base de V .

■ **Ejemplo 1.25** Sean los subespacios de \mathbb{R}^3 , $S = \{\vec{x} \in \mathbb{R}^3, x_1 + x_2 + x_3 = 0, \}$ y $T = \langle (1, 1, 1) \rangle$. Se tiene que $\dim(S) = 2$, $\dim(T) = 1$ y $S \cap T = \vec{0}$. Entonces, $\dim(S + T) = 3$, de donde, $S + T = \mathbb{R}^3$. ■



Si $V = V_1 + V_2$ todo elemento $\vec{v} \in V$ puede escribirse de la forma $\vec{v} = \vec{v}_1 + \vec{v}_2$ con $\vec{v}_1 \in V_1$ y $\vec{v}_2 \in V_2$. Si la suma es directa, esta descomposición es única.

Proposición 1.4.2 Sean V_1 y V_2 subespacios vectoriales de un espacio vectorial V . Las siguientes afirmaciones son equivalentes:

1. $V = V_1 \oplus V_2$
2. Para todo $\vec{v} \in V$ existe una descomposición única de la forma $\vec{v} = \vec{v}_1 + \vec{v}_2$ con $\vec{v}_1 \in V_1$ y $\vec{v}_2 \in V_2$.

Demostración:

Veamos que $1 \rightarrow 2$

Supongamos $\vec{v} = \vec{v}_1 + \vec{v}_2$ con $\vec{v}_1 \in V_1$ y $\vec{v}_2 \in V_2$ y $\vec{v} = \vec{u}_1 + \vec{u}_2$ con $\vec{u}_1 \in V_1$ y $\vec{u}_2 \in V_2$.

Se tiene que,

$$\vec{v}_1 + \vec{v}_2 = \vec{u}_1 + \vec{u}_2,$$

igualdad que puede reescribirse

$$\vec{v}_1 - \vec{u}_1 = \vec{u}_2 - \vec{v}_2$$

De donde $\vec{v}_1 - \vec{u}_1 \in V_1 \cap V_2$, como por hipótesis, $V_1 \cap V_2 = \vec{0}$, $\vec{v}_1 = \vec{u}_1$ y de la misma forma, $\vec{v}_2 = \vec{u}_2$.

Para ver que $2 \rightarrow 1$ alcanza con demostrar que $V_1 \cap V_2 = \vec{0}$.

Si $\vec{v} \in V_1 \cap V_2 = \vec{0}$, \vec{v} se puede escribir $\vec{v} = \vec{v} + \vec{0}$ y $\vec{v} = \vec{0} + \vec{v}$, como la descomposición es única, $\vec{v} = \vec{0}$.

□

En general, dados n subespacios vectoriales V_1, V_2, \dots, V_n de un espacio vectorial V , definimos

$$\bigcap_{j=1}^n V_j = \{ \vec{u} \in V, \vec{u} \in V_j, j = 1, \dots, n \}$$

y

$$\sum_{j=1}^n V_j = \left\{ \sum_{j=1}^n \vec{u}_j, \vec{u}_j \in V_j, j = 1, \dots, n \right\}$$

que reciben el nombre de *intersección* y *suma*, respectivamente, de los subespacios vectoriales V_j dados. Estos dos nuevos subespacios son también subespacios vectoriales de V .

La definición de *suma directa* de varios subespacios vectoriales es un poco más complicada en general, que si solamente hay dos. Se dice que V es *suma directa* de los subespacios vectoriales V_1, V_2, \dots, V_n y se escribe

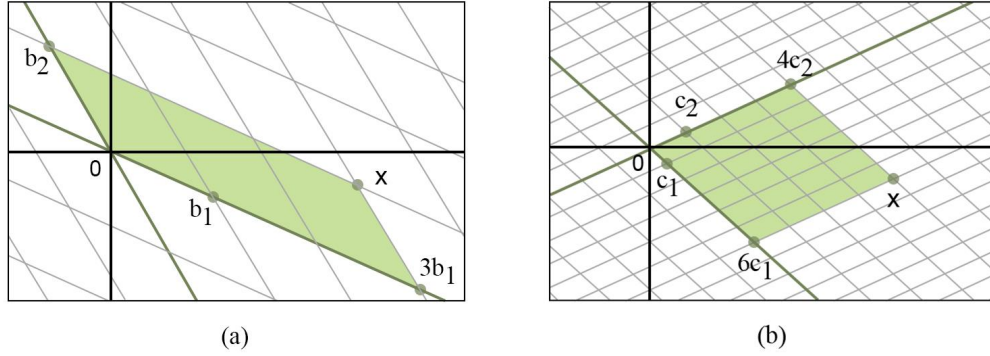
$$V = V_1 \oplus V_2 \oplus V_3 \oplus \dots \oplus V_n$$

si todo vector $\vec{v} \in V$ tiene una descomposición *única* de la forma $\vec{v} = \sum_{i=1}^n \vec{v}_i$ con $\vec{v}_i \in V_i, i = 1, \dots, n$.



- Si $n = 2$, esta última definición y la dada anteriormente son equivalentes.
- Para $n \geq 3$ se puede demostrar que las siguientes afirmaciones son equivalentes:
 - $V = V_1 \oplus V_2 \oplus V_3 \oplus \dots \oplus V_n$
 - $V = \sum_{i=1}^n V_i$ y $V_i \cap (\sum_{k \neq i} V_k) = \vec{0}$

Figura 1.8: Coordenadas de un vector \vec{x} , en las bases $\{\vec{b}_1, \vec{b}_2\}$ (a) y $\{\vec{c}_1, \vec{c}_2\}$ (b)



1.5 Cambio de base

Para visualizar el problema de cambio de base, considere los dos sistemas de coordenadas que se muestran en la Figura 1.8. En (a), $\vec{x} = 3\vec{b}_1 + 1\vec{b}_2$, mientras que en (b), el mismo vector \vec{x} se expresa como $\vec{x} = 6\vec{c}_1 + 4\vec{c}_2$. El problema consiste en encontrar la relación que hay entre las coordenadas de un mismo vector \vec{x} en las dos bases $\{\vec{b}_1, \vec{b}_2\}$ y $\{\vec{c}_1, \vec{c}_2\}$.


Para el caso general, supongamos que se tienen dos bases B y B' de un espacio vectorial V de dimensión finita. Se verá que con la ayuda de una matriz se pueden obtener las coordenadas de un vector con respecto a una base de V a partir de las coordenadas del vector en la otra base. Llamaremos base antigua a la base B y base nueva a la base B' .

Si $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ y $B' = \{\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_n\}$ son dos bases de un espacio vectorial V de dimensión n , todo elemento de la base B' puede escribirse como combinación lineal de los elementos de la base B :

$$\begin{cases} \vec{e}'_1 &= a_{11}\vec{e}_1 + a_{21}\vec{e}_2 + \dots + a_{n1}\vec{e}_n \\ \vec{e}'_2 &= a_{12}\vec{e}_1 + a_{22}\vec{e}_2 + \dots + a_{n2}\vec{e}_n \\ \dots & \\ \vec{e}'_n &= a_{1n}\vec{e}_1 + a_{2n}\vec{e}_2 + \dots + a_{nn}\vec{e}_n \end{cases} \quad (1.1)$$

que en forma abreviada puede escribirse

$$\vec{e}'_j = \sum_{i=1}^n a_{ij}\vec{e}_i, \quad j = 1, \dots, n$$

-  Puede escribirse en forma más concisa usando convenio de Einstein (índice repetido indica suma)


$$\vec{e}'_j = a_{ij}\vec{e}_i, \quad j = 1, \dots, n$$

La nueva base B' se obtiene de la base B mediante la siguiente matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

donde la j -ésima columna de A son las coordenadas del vector \vec{e}'_j con respecto a la base antigua \vec{e}_j , $j = 1, 2, \dots, n$.

La matriz A se denomina *matriz del cambio de base* de B' a B y se denota $P_{B,B'}$.

- 
- A la matriz del cambio de base de B' a B también se la denomina *matriz de transición* de la base B' a la base B .
 - Cuando sea necesario hacer constar las bases B y B' escribiremos $P_{B',B}$ para denotar la matriz de cambio de base de B a B' .
 - Si B y B' coinciden se tiene que $P_{B',B} = I_{n \times n}$.

Proposición 1.5.1 La matriz A del cambio de base de B' a B es invertible y su inversa es la matriz de cambio de base de B a B' . Podemos, por lo tanto, escribir

$$A^{-1} = P_{B,B'}^{-1} = P_{B',B}$$

Demostración:

El determinante de la matriz es no nulo, ya que sus columnas son las coordenadas de los vectores \vec{e}_j que por ser una base son linealmente independientes.

Si A' es la matriz de cambio de base de B a B' , entonces se tiene

$$\vec{e}_j = \sum_{i=1}^n a'_{ij}\vec{e}'_i, \quad j = 1, \dots, n$$

y por la definición de la matriz A ,

$$\vec{e}'_i = \sum_{k=1}^n a_{ki} \vec{e}_k, \quad i = 1, \dots, n.$$

por lo tanto,

$$\vec{e}_j = \sum_{i=1}^n a'_{ij} \left(\sum_{k=1}^n a_{ki} \vec{e}_k \right) = \sum_{k=1}^n \left(\sum_{i=1}^n a_{ki} a'_{ij} \right) \vec{e}_k, \quad j = 1, \dots, n.$$

Como el término dentro de la sumatoria en k , $\sum_{i=1}^n a_{ki} a'_{ij}$, es el elemento que ocupa el lugar (k, j) del producto de las matrices A y A' , de acuerdo a la igualdad, ese término vale 1 si $k = j$, y 0 si no. Es decir que se tiene que el producto de las matrices A y A' , es $AA' = I_{n \times n}$. □



- Las expresiones anteriores pueden reescribirse en forma sintética considerando que el índice repetido se suma de 1 a n : $\vec{e}_j = a'_{ij} \vec{e}_i$
- Usando la delta de Kronecker, δ_{kj} (definida $\delta_{kj} = 1$ si $k = j$ y $\delta_{kj} = 0$ si $k \neq j$), la expresión se escribe $\vec{e}_j = a_{ki} a'_{ij} \vec{e}_k = \delta_{kj} \vec{e}_k$

■ **Ejemplo 1.26** Dadas las bases de \mathbb{R}^3 , $B = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ y $B' = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ (base canónica), la matriz de cambio de base de B a B' , de acuerdo a las Ecs.(1.1), es

$$P_{B',B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

■



Si la base nueva, B' es la base la canónica, la matriz de cambio de base de B a B' se obtiene directamente poniendo las coordenadas de los vectores de la base B en cada columna (ver Ejemplo 1.26).

Relación entre las coordenadas en la base B y en la B'

Tratemos ahora de relacionar entre sí las coordenadas de un mismo vector en las bases nueva B' y antigua B . Supongamos que $\vec{x} \in V$,

$$\vec{x} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + \cdots + x_n \vec{e}_n \quad (1.2)$$

y además

$$\vec{x} = x'_1 \vec{e}'_1 + x'_2 \vec{e}'_2 + \cdots + x'_n \vec{e}'_n \quad (1.3)$$

Sustituyendo Ec.(1.1) en la segunda expresión, Ec.(1.3), obtenemos

$$\begin{aligned} \vec{x} &= x'_1 \left(\sum_{i=1}^n a_{i1} \vec{e}_i \right) + x'_2 \left(\sum_{i=1}^n a_{i2} \vec{e}_i \right) + \cdots + x'_n \left(\sum_{i=1}^n a_{in} \vec{e}_i \right) \\ &= (a_{11}x'_1 + a_{12}x'_2 + \cdots + a_{1n}x'_n) \vec{e}_1 + (a_{21}x'_1 + a_{22}x'_2 + \cdots \\ &\quad + a_{2n}x'_n) \vec{e}_2 + \cdots + (a_{n1}x'_1 + a_{n2}x'_2 + \cdots + a_{nn}x'_n) \vec{e}_n \end{aligned}$$

Comparando esta última igualdad con la primera expresión Ec.(1.2) podemos escribir

$$\begin{cases} x_1 &= a_{11}x'_1 + a_{12}x'_2 + \cdots + a_{1n}x'_n \\ x_2 &= a_{21}x'_1 + a_{22}x'_2 + \cdots + a_{2n}x'_n \\ \cdots & \\ x_n &= a_{n1}x'_1 + a_{n2}x'_2 + \cdots + a_{nn}x'_n \end{cases} \quad (1.4)$$

$$\text{Si convenimos en escribir } X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_n \end{pmatrix} \text{ y } X' = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ \cdots \\ x'_n \end{pmatrix}$$

a las coordenadas del vector \vec{x} en la antigua base B y en la nueva base B' , respectivamente, las Ecs.(1.4) se escribe de la forma

$$X = AX' \quad (1.5)$$

Esto permite obtener las coordenadas del vector \vec{x} en la base antigua conociendo las coordenadas del mismo vector en la base nueva.

■ **Ejemplo 1.27** Dadas las mismas bases del Ejemplo 1.26, $B = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ y $B' = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, se quieren encontrar las coordenadas del vector $\vec{x} = (3, 2, 3)_{B'}$ en la base B , para lo cual se necesita la matriz de cambio de base, $P_{B,B'}$.

De acuerdo a las Ecs. (1.1), la matriz de cambio de base, $P_{B,B'}$ se obtiene a partir de encontrar las coordenadas de los vectores de la base B' en la base B y colocarlos como columnas. Después de resolver el sistema de ecuaciones se obtuvo

$$\begin{cases} (1, 0, 0) &= 0(1, 1, 1) + 0(1, 1, 0) + 1(1, 0, 0) \\ (0, 1, 0) &= 0(1, 1, 1) + 1(1, 1, 0) + (-1)(1, 0, 0) \\ (0, 0, 1) &= 1(1, 1, 1) + (-1)(1, 1, 0) + 0(1, 0, 0) \end{cases} \quad (1.6)$$

y entonces,

$$P_{B,B'} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}$$

Otra alternativa para hallar la matriz $P_{B,B'}$ es usar la Proposición 1.5.1 y entonces $P_{B,B'} = P_{B',B}^{-1}$ dado

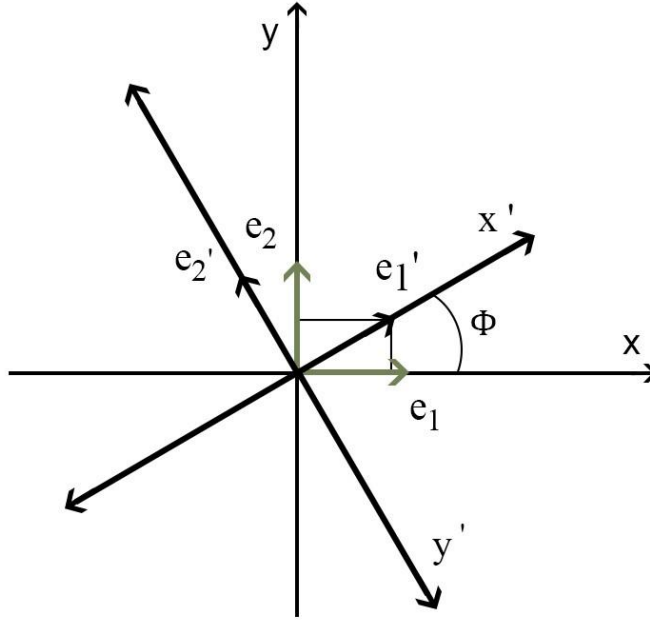
que $P_{B',B}$ se encontró en el Ejemplo 1.26.

Luego, usando (1.5), las coordenadas son

$$X = P_{B,B'} X' = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}_{B'} = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}_B$$

Así se obtuvieron las coordenadas del vector \vec{x} en la base B conociendo las coordenadas del mismo vector en la base B' . ■

■ **Ejemplo 1.28** Sean \vec{e}_1 y \vec{e}_2 dos vectores perpendiculares unitarios en \mathbb{R}^2 en la dirección de los ejes de coordenadas cartesianas. Girando los ejes de coordenadas un ángulo ϕ en sentido positivo, contrario a las agujas del reloj, se obtiene una nueva base $B' = \{\vec{e}'_1, \vec{e}'_2\}$.

Figura 1.9: Cambio de base. Rotación de ejes en un ángulo ϕ 

De acuerdo a la Figura 1.9, se observa que

$$\begin{cases} \vec{e}'_1 &= \cos(\phi)\vec{e}_1 + \sin(\phi)\vec{e}_2 \\ \vec{e}'_2 &= -\sin(\phi)\vec{e}_1 + \cos(\phi)\vec{e}_2 \end{cases} \quad (1.7)$$

por lo cual, teniendo en cuenta el sistema (1.1), la matriz del cambio de base A es

$$A = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \quad (1.8)$$

Así, si $\phi = \pi/4$, las coordenadas respecto a la base $B = \{\vec{e}'_1, \vec{e}'_2\}$ del vector $(2, 3)_B$

$$\begin{pmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 3 \end{pmatrix}_B \approx \begin{pmatrix} 3.53 \\ 0.71 \end{pmatrix}_{B'}$$

■

■ **Ejemplo 1.29** Dadas las bases de $P_{\mathbb{R}}^{(2)}[x]$, $B = \{3, 1+x, x^2\}$ y $B' = \{1, x+3, x^2+x\}$, se quiere hallar la matriz de cambio de base de B' a B .

Teniendo en cuenta la Observación i, luego del Ejemplo 1.26 resulta más simple hallar $P_{E,B}$ y $P_{E,B'}$, donde $E = \{1, x, x^2\}$ es la base canónica. Como

$$\begin{aligned} 1 &= 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ x+3 &= 3 \cdot 1 + 1 \cdot x + 0 \cdot x^2 \\ x^2+x &= 0 \cdot 1 + 1 \cdot x + 1 \cdot x^2 \end{aligned}$$

Se tiene que

$$P_{E,B'} = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

y de la misma forma,

$$P_{E,B} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Luego, la matriz de cambio de base de B' a B sale de multiplicar las matrices de cambio de base de E a B y de B' a E , es decir,

$$P_{B,B'} = P_{B,E}P_{E,B'} = (P_{E,B})^{-1}P_{E,B'} = \begin{pmatrix} 1/3 & 2/3 & -1/3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1/3 & 2/3 & -1/3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}_{B'} = \begin{pmatrix} -1/3 \\ 1 \\ 1 \end{pmatrix}_B = x^2 + x$$

■

Ariel Goobar

Nacido en Argentina en 1962. Es un astrofísico sueco que trabaja en la Universidad de Estocolmo en astrofísica de altas energías y teoría de la relatividad. Es integrante del equipo que ganó el Nobel de Física en 2011 por descubrir mediante las mediciones con supernovas, la expansión acelerada del universo, actualmente es Director del Oskar Klein Centre for Cosmoparticle Physics. Emigró de Argentina a los 13 años. Se dedica al contenido de lo que se llama materia oscura. La gran esperanza –aunque a lo mejor me equivoco dijo– desde el punto de vista experimental es que si entendemos qué es la energía oscura, muy probablemente eso nos vaya a dar un indicio importante de cómo se puede entender la fuerza de gravedad, cómo casar la teoría de la relatividad con la física cuántica. Muy probablemente estén relacionadas, pero hoy no tenemos idea de cómo. Hoy, si yo pudiera soñar algo, creo que es entender eso, sería alucinante. [4]

Figura 1.10



1.6 Actividades propuestas

Problema de Aplicación 1 Para realizar un cambio de coordenadas celestes a coordenadas horizontales, es necesario hacer dos rotaciones:

$$\vec{r}_{el} = R_z(TSL) * \vec{r}_{ec}$$

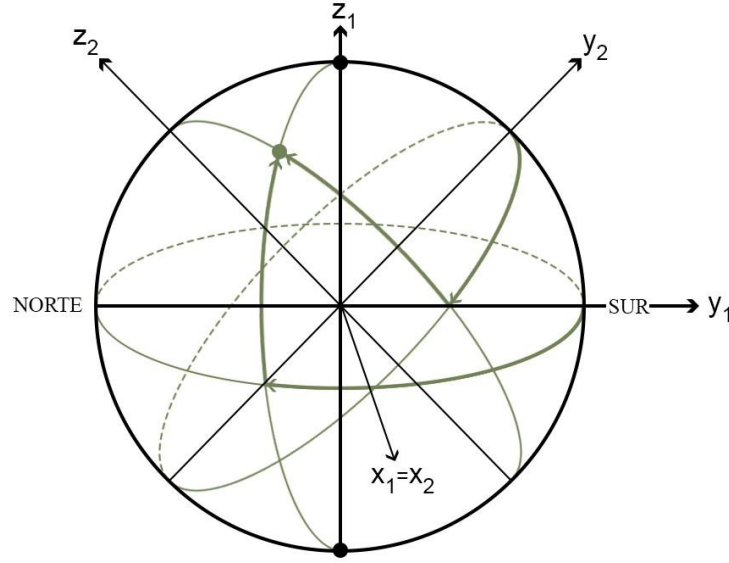
$$\vec{r}_h = R_y(90 - \varphi) * \vec{r}_{el}$$

Esto se debe a que entre el sistema ecuatorial celeste y el ecuatorial local, el polo celeste (eje z) permanece fijo para ambos, pero el origen desde donde medimos uno y otro sistema en el ecuador celeste, cambia en una cantidad TSL . Usaremos TSL como tiempo siderio local, $TSL = 18:31:31$, Recuerde pasar de horas a grados para poder operar. Luego para realizar el cambio de coordenadas de celeste locales a horizontales, se mantiene en común el eje y , la línea este-oeste, y el ángulo en que se rotan el plano x, z es $90 - \varphi$. Para el problema actual usaremos la $\varphi = -34^\circ 50'$ que corresponde a la ciudad de La Plata.

El vector de las coordenadas ecuatoriales celestes se escribe como:

$$\vec{r}_{ec} = \begin{pmatrix} \cos(\delta)\cos(\alpha) \\ \cos(\delta)\sin(\alpha) \\ \sin(\delta) \end{pmatrix}$$

Figura 1.11: Coordenadas (x_1, y_1, z_1) y (x_2, y_2, z_2) en el sistema cartesiano horizontal y en el cartesiano ecuatorial



y el vector de las coordenadas horizontales:

$$\vec{r}_h = \begin{pmatrix} \cos(h)\cos(A) \\ -\cos(h)\sin(A) \\ \sin(h) \end{pmatrix}$$

donde $h = \sin^{-1}(z)$ y $A = \tan^{-1}(-y/x)$.

Amplíe la Tabla 1.1 con las coordenadas horizontales para cada cúmulo. Recuerde pasar de horas a grados para poder operar. Nota: No tenga en cuenta la precesión. **Recomendación:** Realice un programa computacional para hacer los cálculos.

1.6.1 Ejercicios

Ejercicio 1.1

Analice si los siguientes conjuntos son espacios vectoriales sobre \mathbb{R} .

- a) El conjunto $S = \{(x, y) : y = 2x + 1\} \subseteq \mathbb{R}^2$

Tabla 1.1: Coordenadas celestes de 13 cúmulos abiertos

Cúmulo	RA(J2000)	Dec(J2000)
NGC6192	16:40:16.40	−43:30:31.0
NGC6242	16:55:32.38	−39:28:02.0
NGC6322	17:18:25.13	−42:56:03.3
NGC6704	18:50:42.00	−05:12:42.5
NGC6737	19:02:16.30	−18:32:56.5
Rup 102	12:13:32.95	−62:43:18.7
Rup 166	13:25:38.14	−63:27:54.6
SLS4565	18:01:59.55	−23:41:06.3
Lynga 14	16:55:03.40	−45:14:09.1
Trumpler 22	14:31:03.33	−61:09:57.0
Trumpler 24	16:56:11.14	−40:40:01.1
Dominici 11	18:57:36.31	−10:23:39.9
Dominici 12	18:51:24.93	−13:18:50.2

- b) El semiplano en \mathbb{R}^2 , $S = \{(x, y) : y \geq 0\}$,
c) Los polinomios de grado menor o igual que 2, $P_{\mathbb{R}}^{(2)}[x]$.

Ejercicio 1.2

Dé al menos 5 ejemplos de espacios vectoriales y escriba, según su opinión, qué utilidad tendría saber que su estructura es de espacio vectorial.

Ejercicio 1.3

De acuerdo con la definición, S es un subespacio de un espacio vectorial V sí y sólo sí, se cumplen las siguientes condiciones:

- S contiene al vector $\vec{0}$ de V .
- Si \vec{u} y \vec{v} están en S , entonces $\vec{u} + \vec{v}$ está en S .
- Si \vec{u} está en S y α es un escalar, entonces $\alpha\vec{u}$ está en S .

Compruebe si valen las siguientes afirmaciones:

- $S = \{(x, y, z) : z = 0\}$, es un subespacio de \mathbb{R}^3 .
- El conjunto de polinomios $P_{\mathbb{R}}^{(2)}[x]$, de grado menor o igual que 2, es subespacio vectorial del espacio vectorial $P_{\mathbb{R}}^{(n)}[x]$ de todos los polinomios con coeficientes reales.

Ejercicio 1.4

Dados los subespacios de \mathbb{R}^3 , $S = \{(x, y, z) \in \mathbb{R}^3 : x = y = 0\}$ y $T = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$ calcule: a) Una base y la dimensión de ambos subespacios. b) $S + T$ y $S \cap T$, dando las bases de dichos subespacios. c) ¿La suma $S + T$ es directa?

Ejercicio 1.5

Encuentre en cada uno de los ejemplos siguientes la suma y la intersección del par de subespacios dados, y compruebe que se verifica la ecuación:

$$\dim(V_1) + \dim(V_2) = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)$$

a) Los subespacios que corresponden al conmutador para cada una de las matrices siguientes.

$$A_1 = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$$

(El conmutador de una matriz A , se define como $C(A) = \{B \in \mathbb{R}^{2 \times 2} \text{ tal que } BA = AB\}$).

b) Los subespacios formados por las bases $\{\sin(t), \cos(t)\}$ y $\{e^{it}, e^{-it}\}$ considerados en el espacio de las funciones complejas continuas en el intervalo $[0, 1]$ con escalares en \mathbb{R} .

Ejercicio 1.6

- Demuestre que el conjunto de soluciones de la ecuación diferencial de primer orden homogénea, con coeficientes constantes: $y' + ky = 0$ es un espacio vectorial de dimensión uno, siendo $\{e^{-kx}\}$ una base. A su vez el conjunto de soluciones de esta ecuación es un subespacio vectorial del espacio de las funciones derivables cuya dimensión es infinita.
- Luego resuelva la ecuación diferencial homogénea de segundo orden:
 $y'' - y' - 6y = 0$, con las condiciones iniciales $y(0) = 3$ y $y'(0) = -1$.
 Este tipo de ecuaciones se resuelve proponiendo la solución $y = e^{\lambda x}$ y resolviendo la ecuación característica, que para éste caso sería: $\lambda^2 - \lambda - 6 = 0$.
 - Encuentre las raíces λ_1 y λ_2 y reemplace en la solución general: $y(x) = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x}$.
 - Escriba la base del conjunto solución, y especifique que dimensión tiene.
 - Halle la solución particular definiendo los coeficientes C_1 y C_2 , con ayuda de las condiciones iniciales.
- Investigue que pasaría si las soluciones de la ecuación característica de una ecuación diferencial homogénea fueran raíces dobles. Escriba su base.

4. Cite al menos un ejemplo de física donde se necesita usar ecuaciones diferenciales.

Ejercicio 1.7

¿Son los vectores $\vec{u} = (4, -2, 5)$ y $\vec{v} = (1, -1, -1)$ de \mathbb{R}^3 combinación lineal de $\vec{x}_1 = (1, -1, 2)$ y $\vec{x}_2 = (2, 0, 1)$? Interprete geoméricamente y conecte con los subespacios de \mathbb{R}^3 .

Ejercicio 1.8

Los primeros cuatro polinomios de Laguerre son $\{1, 1-x, 2-4x+x^2, 6-18x+9x^2-x^3\}$. Demuestre que estos polinomios forman una base de $P_{\mathbb{R}}^{(3)}[x]$.

Ejercicio 1.9

Compruebe que $B = \{1, x, x^2\}$ es una base del espacio vectorial $P_{\mathbb{R}}^{(2)}[x]$. En consecuencia, $\dim(P_{\mathbb{R}}^{(2)}[x]) = 3$. ¿Es correcta la afirmación $\dim(P_{\mathbb{R}}^{(n)}[x]) = n+1$?

Ejercicio 1.10

Sea la matriz,

$$A = \begin{pmatrix} x & a & b \\ a & x & b \\ a & b & x \end{pmatrix}$$

Encuentre los valores de x para los que el $\text{Det}(A) = 0$. Lo cual es equivalente a decir que columnas o filas son linealmente dependientes. ¿Cuáles son las dimensiones posibles del espacio generado por las filas?

Ejercicio 1.11

Encuentre las coordenadas del vector $\vec{x} = (1, 3, -2)$ con respecto a la base $B = \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ donde $\vec{b}_1 = (1, 0, 0)$, $\vec{b}_2 = (1, 1, 0)$, $\vec{b}_3 = (1, 1, 1)$.

Ejercicio 1.12

Calcule las coordenadas del vector \vec{w} relativas a la base $B = \{\vec{u}_1, \vec{u}_2\}$.

- a) $\vec{u}_1 = (1, 0)$, $\vec{u}_2 = (0, 1)$; $\vec{w} = (3, 7)$
- b) $\vec{u}_1 = (2, -4)$, $\vec{u}_2 = (3, 8)$; $\vec{w} = (1, 1)$
- c) $\vec{u}_1 = (1, 1)$, $\vec{u}_2 = (0, 2)$; $\vec{w} = (a, b)$

Ejercicio 1.13

Encuentre las coordenadas del vector \vec{v} relativas a la base $B = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

$$\vec{v}_1 = (1, 0, 0), \vec{v}_2 = (2, 2, 0), \vec{v}_3 = (3, 3, 3); \vec{v} = (2, -1, 3).$$

Ejercicio 1.14

Calcule las coordenadas del vector $p \in P_{\mathbb{R}}^{(2)}[x]$ relativas a la base $B = \{p_1(x), p_2(x), p_3(x)\}$.

$$p_1(x) = 1 + x, p_2(x) = 1 + x^2, p_3(x) = x + x^2; p(x) = 4 - 3x + x^2$$

Ejercicio 1.15

En $\mathbb{R}^{2 \times 2}$, encuentre las coordenadas de la matriz $A = \begin{pmatrix} 2 & 0 \\ -1 & 3 \end{pmatrix}$ relativas a la base

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

Ejercicio 1.16

Considere las bases $B = \{\vec{u}_1, \vec{u}_2\}$ y $B' = \{\vec{v}_1, \vec{v}_2\}$ para \mathbb{R}^2 , donde $\vec{u}_1 = (1, 0)$, $\vec{u}_2 = (0, 1)$, $\vec{v}_1 = (2, 1)$, $\vec{v}_2 = (-3, 4)$.

- Halle la matriz de cambio de base de B a B' , $P_{B',B}$.
- Utilice la matriz anterior para obtener las coordenadas en la base B' de $\vec{w} = (3, -5)_B$.
- Verifique lo obtenido en b) haciéndolo directamente.
- Calcule la matriz de transición $P_{B,B'}$ y verifique que $P_{B,B'} = P_{B',B}^{-1}$.

Ejercicio 1.17

Considere las bases $B = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ y $B' = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ para \mathbb{R}^3 , donde $\vec{u}_1 = (-3, 0, -3)$, $\vec{u}_2 = (-3, 2, 1)$, $\vec{u}_3 = (1, 6, -1)$, $\vec{v}_1 = (-6, -6, 0)$, $\vec{v}_2 = (-2, -6, 4)$ y $\vec{v}_3 = (-2, -3, 7)$.

- Halle la matriz de cambio de base de B a B' .
- Utilice la matriz anterior para obtener las coordenadas en la base B' de $\vec{w} = (-5, 8, -5)_B$.

Ejercicio 1.18

Considere las bases $B = \{p_1(x), p_2(x)\}$ y $B' = \{q_1(x), q_2(x)\}$ para $P_{\mathbb{R}}^{(1)}[x]$, donde $p_1(x) = 6 + 3x$, $p_2(x) = 10 + 2x$, $q_1(x) = 2$, $q_2(x) = 3 + 2x$.

- a) Halle la matriz de transición de B a B' .
- b) Utilice la matriz anterior para obtener las coordenadas de $p(x) = -4 + x$ en la base B' .

Ejercicio 1.19

Si se quiere obtener un sistema de coordenadas rectangulares $x'y'$ haciendo girar un sistema de coordenadas rectangulares xy hasta describir un ángulo de $\theta = \frac{3}{4}\pi$.

- a) Halle las coordenadas $x'y'$ del punto cuyas coordenadas xy son $(-2, 6)$.
- b) Calcule las coordenadas xy del punto cuyas coordenadas $x'y'$ son $(5, 2)$.

Ejercicio 1.20

Si se quiere obtener un sistema de coordenadas rectangulares $x'y'z'$ haciendo girar un sistema de coordenadas rectangulares xyz en sentido contrario a las agujas del reloj alrededor del eje z , cuando se observa hacia abajo a lo largo del eje z hasta describir un ángulo de $\theta = \frac{\pi}{4}$.

- a) Halle las coordenadas $x'y'z'$ del punto cuyas coordenadas xyz son $(-1, 2, 5)$.
- b) Calcule las coordenadas xyz del punto cuyas coordenadas $x'y'z'$ son $(1, 6, -3)$.

Ejercicio 1.21

Sea V un espacio vectorial de dimensión n . Demuestre que todo conjunto linealmente independiente de n elementos es una base de V .

Ejercicio 1.22

Sean $p_0(x), p_1(x), \dots, p_n(x)$ polinomios cualesquiera de $P_{\mathbb{R}}^{(n)}[x]$ de grado $0, 1, \dots, n$ respectivamente: demuestre que $\{p_0(x), p_1(x), \dots, p_n(x)\}$ es una base de $P_{\mathbb{R}}^{(n)}[x]$. ¿Podría encontrar alguna relación entre este teorema y el teorema del resto? ¿Y con la fórmula de Taylor?

Ejercicio 1.23

En el supuesto que $V = V_1 \oplus V_2$ donde V_1 y V_2 son dos subespacios de V de dimensiones n y m respectivamente y sean $B_1 = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ y $B_2 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ sus bases, compruebe que $B = B_1 \cup B_2$ es una base de $V_1 \oplus V_2$.

Ejercicio 1.24

Sea $A = LU$, donde L es una matriz triangular inferior invertible y U es triangular superior. Explique por qué la primera columna de A es un múltiplo de la primera columna de L . ¿Cómo se relaciona la segunda columna de A con las columnas de L ?

Ejercicio 1.25

Sea $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ la base canónica de \mathbb{R}^n , y sean $\vec{u}_1 = \vec{e}_2 - \vec{e}_1$, $\vec{u}_2 = \vec{e}_3 - \vec{e}_2, \dots, \vec{u}_{n-1} = \vec{e}_n - \vec{e}_{n-1}, \vec{u}_n = \vec{e}_n$. Demuestre que $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ es una base de \mathbb{R}^n . Exprese el vector $\vec{v} = \vec{e}_1 + \vec{e}_2 + \dots + \vec{e}_n$ como una combinación lineal de los vectores $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$.

1.6.2 Autoevaluación**Verdadero o Falso**

1. Si una matriz tiene dos filas iguales, su determinante vale 0.
2. Si $F = F_1 \oplus F_2 \oplus \dots \oplus F_p$ entonces $\dim F \neq \dim F_1 + \dim F_2 + \dots + \dim F_p$.
3. Si $P_{B,A}$ es invertible, entonces $P_{B,A}^{-1} = P_{B,A}$.
4. Siendo A, B y C bases de un espacio vectorial, se cumple que $P_{C,B} \cdot P_{B,A} = P_{C,A}$.
5. Sea S un conjunto de un espacio vectorial V de dimensión n y además S contiene menos de n vectores, entonces S no puede generar V .
6. Un plano en \mathbb{R}^3 es un subespacio de dimensión 2 de \mathbb{R}^3 .
7. Si un conjunto $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$ genera un espacio vectorial V de dimensión finita y si U es un conjunto de más de p vectores de V , entonces U es linealmente dependiente.
8. La suma del subespacio de las matrices simétricas de $\mathbb{R}^{n \times n}$ con las matrices antisimétricas de $\mathbb{R}^{n \times n}$ es directa generando el espacio vectorial de las matrices de $\mathbb{R}^{n \times n}$.
9. La suma del subespacio de las funciones pares con el subespacio de las funciones impares no es directa.



2. Transformaciones Lineales

En este capítulo nos interesamos por aquellas aplicaciones entre espacios vectoriales que preservan las operaciones de suma y producto por escalares. Son las aplicaciones o transformaciones lineales. Son las funciones con las que se trabaja en Álgebra Lineal y tienen una amplia variedad de aplicaciones. Se verá que, en el caso de dimensión finita, es posible asociarles una matriz. Se estudian los espacios asociados; el espacio nulo y el espacio que generan las columnas de la matriz. Un resultado útil e importante es que las sumas de las dimensiones de esos subespacios dan la dimensión del espacio de partida. Mostramos la interpretación geométrica tanto de las aplicaciones lineales en el plano como en el espacio. Se presentan una gran variedad de ejemplos. Se estudia, además, el espacio dual de un espacio vectorial. Es el conjunto de todas las transformaciones lineales entre un espacio vectorial y el cuerpo de los escalares, conocidas como funcionales lineales.

2.1 Definición de transformación lineal. Ejemplos

Definición 2.1.1 Sean V y W dos espacios vectoriales, una *transformación lineal* T de V en W es una aplicación $T : V \rightarrow W$ tal que:

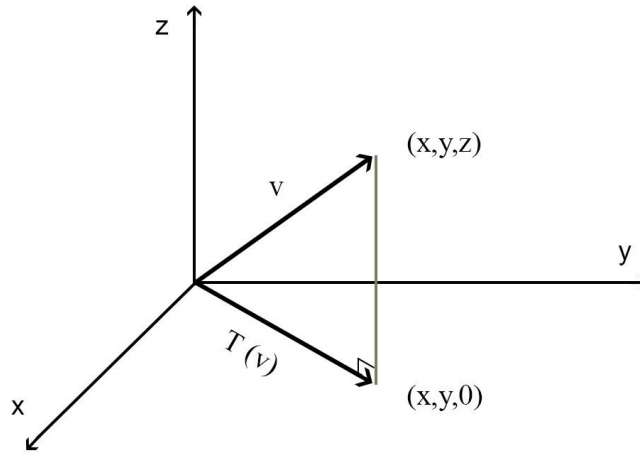
1. $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$ para todo $\vec{v}, \vec{w} \in V$.
2. $T(a \vec{v}) = a T(\vec{v})$ para todo $a \in K$ y todo $\vec{v} \in V$.

Entre las transformaciones lineales más utilizadas están las proyecciones. En la Figura 2.1 se muestra la proyección ortogonal de un vector $\vec{v} = (x, y, z)$ sobre el plano xy .

Se tiene $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, donde $T((x, y, z)) = (x, y, 0)$. Es una transformación lineal ya que se cumplen para $\forall \vec{v} = (x, y, z)$, $\vec{w} = (x', y', z')$ y $\forall a \in K$:

1. $T(\vec{v} + \vec{w}) = T((x + x', y + y', z + z')) = (x + x', y + y', 0) = (x, y, 0) + (x', y', 0) = T(\vec{v}) + T(\vec{w})$.
2. $T(a \vec{v}) = T(a(x, y, z)) = T((ax, ay, az)) = (ax, ay, 0) = a(x, y, 0) = a T(\vec{v})$.

Figura 2.1: Proyección ortogonal de un vector \vec{v} sobre el plano xy



- En una transformación lineal V y W deben ser espacios vectoriales sobre el mismo cuerpo K .
- Las aplicaciones $O : V \rightarrow W$, $O(\vec{u}) = \vec{0}_W$ para todo $\vec{u} \in V$ y $I_d : V \rightarrow V$, $I_d(\vec{u}) = \vec{u}$ son transformaciones lineales.
- La traza de una matriz,
 $Tr : K^{n \times n} \rightarrow K$ dada por $Tr(A) = \sum_{i=1}^n a_{ii}$, es una transformación lineal.
- $T : \mathbb{R} \rightarrow \mathbb{R}$ dada por $T(x) = x^2$ es un ejemplo de transformación *no lineal*. Se tiene que $T(x + y) \neq T(x) + T(y)$, ya que $T(x + y) = (x + y)^2 = x^2 + y^2 + 2xy$, mientras que $T(x) + T(y) = x^2 + y^2$.
- Otro ejemplo de transformación no lineal es el determinante de una matriz, ya que, en general $Det(A + B) \neq Det(A) + Det(B)$.

■ **Ejemplo 2.1** Dado un número real a , la aplicación que asocia a cada polinomio p del conjunto $P_{\mathbb{R}}[t]$ su valor en $x = a$, $p(a)$, es una transformación lineal. Está definida mediante las siguientes expresiones:

$$T : P_{\mathbb{R}}[t] \rightarrow \mathbb{R}$$

$$T(p(t)) = p(a)$$

El hecho que T es una transformación lineal se deduce de las igualdades $T(p+q) = T(p) + T(q)$ y $T(cp) = cT(p)$ que se prueban a continuación:

$$T(p+q)(t) = (p+q)(a) = p(a) + q(a) = T(p)(t) + T(q)(t)$$

y

$$T(cp)(t) = (cp)(a) = c(p(a)) = cT(p)(t)$$

para todo número real c . ■

Aplicando repetidas veces las propiedades 1 y 2 de la definición de transformación lineal entre espacios vectoriales V y W se puede ver que la imagen de una combinación lineal de vectores del espacio vectorial inicial V es una combinación lineal de vectores del espacio vectorial final W , es decir

$$T\left(\sum_{j=1}^n c_j \vec{v}_j\right) = \sum_{j=1}^n c_j T(\vec{v}_j)$$

donde $c_j \in K$ y $\vec{v}_j \in V$ para todo $j = 1, 2, \dots, n$.

Otras propiedades de las transformaciones lineales que se deducen de la definición se enuncian en las proposiciones a continuación.

Proposición 2.1.1 Sea T una aplicación lineal entre dos espacios vectoriales V y W . Se tienen los siguientes resultados:

1. La imagen del elemento neutro de V mediante T es el neutro de W , es decir, $T(\vec{0}_V) = \vec{0}_W$
2. La imagen mediante T del opuesto de un elemento $\vec{v} \in V$ es el opuesto de $T(\vec{v})$, es decir, $T(-\vec{v}) = -T(\vec{v})$

Demostración:

1. $T(\vec{0}_V) = T(\vec{0}_V + \vec{0}_V) = T(\vec{0}_V) + T(\vec{0}_V)$

Restando $T(\vec{0}_V)$ en ambos miembros de la igualdad, se tiene

$$T(\vec{0}_V) - T(\vec{0}_V) = T(\vec{0}_V) \text{ de donde, } \vec{0}_W = T(\vec{0}_V).$$

$$2. T(-\vec{v}) = T((-1)\vec{v}) = (-1)T(\vec{v}) = -T(\vec{v}), \forall \vec{v} \in V.$$

□

Proposición 2.1.2 Sea $T : V \rightarrow W$ una transformación lineal entre espacios vectoriales. La imagen mediante T de cualquier subespacio vectorial V_1 de V , $W_1 = T(V_1)$ es un subespacio vectorial de W .

Demostración:

- $0_W \in T(V_1)$, ya que $T(0_V) = 0_W$.
- Sean \vec{w}_1 y $\vec{w}_2 \in T(V_1)$. Existen $\vec{v}_1, \vec{v}_2 \in V_1$ tales que $T(\vec{v}_1) = \vec{w}_1$ y $T(\vec{v}_2) = \vec{w}_2$. Para ver que $\vec{w}_1 + \vec{w}_2 \in T(V_1)$ basta ver que, por ser V_1 subespacio, $\vec{v}_1 + \vec{v}_2 \in V_1$ y $T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) = \vec{w}_1 + \vec{w}_2$.
- Y también $\alpha\vec{w}_1 \in T(V_1)$, ya que $\alpha\vec{v}_1 \in V_1$, por ser V_1 subespacio, y $T(\alpha\vec{v}_1) = \alpha T(\vec{v}_1) = \alpha\vec{w}_1$.

□

Proposición 2.1.3 Sea $T : V \rightarrow W$ una transformación lineal entre espacios vectoriales. Si U es un subespacio de W , entonces $T^{-1}(U) = \{\vec{v}/\vec{v} \in V, T(\vec{v}) \in U\}$ es un subespacio de V .

Demostración:

- $0_V \in T^{-1}(U)$, ya que $T^{-1}(0_W) = 0_V$.
- Sean \vec{v}_1 y $\vec{v}_2 \in T^{-1}(U)$. Existen \vec{u}_1 y $\vec{u}_2 \in U$ tales que $T(\vec{v}_1) = \vec{u}_1$ y $T(\vec{v}_2) = \vec{u}_2$. Como $T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) \in U$, $\vec{v}_1 + \vec{v}_2 \in T^{-1}(U)$.
- De la misma forma, si $\vec{v}_1 \in T^{-1}(U)$, $\alpha\vec{v}_1 \in T^{-1}(U)$ pues $T(\alpha\vec{v}_1) = \alpha T(\vec{v}_1) = \alpha\vec{u}_1$.

□

Proposición 2.1.4 La imagen mediante una transformación lineal de un subespacio vectorial de dimensión k es un subespacio vectorial de dimensión no superior a k .

Demostración:

Por la Proposición 2.1.2, si el subespacio V_1 de V tiene $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ como base, todo elemento w de la imagen de $W_1 = T(V_1)$ puede escribirse como combinación lineal de los vectores $T(\vec{v}_1), T(\vec{v}_2), \dots, T(\vec{v}_k)$. Esto es cierto ya que tomando $\vec{v} \in V_1$ tal que $T(\vec{v}) = \vec{w}$ se tiene que

$$\vec{w} = T(\vec{v}) = T\left(\sum_{j=1}^k c_j \vec{v}_j\right) = \sum_{j=1}^k c_j T(\vec{v}_j).$$

Por lo tanto, W_1 coincide con el subespacio generado por los vectores $T(\vec{v}_1), T(\vec{v}_2), \dots, T(\vec{v}_k)$,

$$\langle T(\vec{v}_1), T(\vec{v}_2), \dots, T(\vec{v}_k) \rangle$$

Es decir $W_1 = \mathbf{L}(T(\vec{v}_1), T(\vec{v}_2), \dots, T(\vec{v}_k))$. T preserva las combinaciones lineales.

En consecuencia, la dimensión de W_1 no puede superar k .

□

Demostraremos con el teorema que sigue que una transformación queda determinada cuando se conocen las imágenes de los elementos de una base del espacio vectorial inicial.

Proposición 2.1.5 Sea $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ una base de un espacio vectorial V y sean $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_n$ n vectores cualesquiera de otro espacio vectorial W . En estas condiciones, existe una única transformación lineal T de V en W tal que

$$T(\vec{e}_j) = \vec{w}_j, \quad j = 1, 2, \dots, n$$

Demostración:

- Existencia. Dado $\vec{v} \in V$, $\vec{v} = \sum_{j=1}^n \alpha_j \vec{e}_j$ con $\alpha_j \in K$. Se define $T(\vec{v}) = \sum_{j=1}^n \alpha_j \vec{w}_j$.
- T es lineal

Sean \vec{v} y \vec{v}' tales que $\vec{v} = \sum_{j=1}^n \alpha_j \vec{e}_j$ y $\vec{v}' = \sum_{j=1}^n \alpha'_j \vec{e}_j$. Entonces,

$$\vec{v} + \vec{v}' = \sum_{j=1}^n \alpha_j \vec{e}_j + \sum_{j=1}^n \alpha'_j \vec{e}_j = \sum_{j=1}^n (\alpha_j + \alpha'_j) \vec{e}_j$$

y

$$T(\vec{v} + \vec{v}') = \sum_{j=1}^n (\alpha_j + \alpha'_j) \vec{w}_j = \sum_{j=1}^n \alpha_j \vec{w}_j + \sum_{j=1}^n \alpha'_j \vec{w}_j = T(\vec{v}) + T(\vec{v}')$$

De la misma forma,

$$T(c\vec{v}) = \sum_{j=1}^n (c\alpha_j) \vec{w}_j = c \sum_{j=1}^n \alpha_j \vec{w}_j = cT(\vec{v})$$

- T es única

Si T' cumple $T'(\vec{e}_j) = \vec{w}_j$ y $\vec{v} = \sum_{j=1}^n \alpha_j \vec{e}_j$, se tiene que

$$T'(\vec{v}) = \sum_{j=1}^n \alpha_j T'(\vec{e}_j) = \sum_{j=1}^n \alpha_j T(\vec{e}_j) = T(\vec{v}).$$

Luego $T(\vec{v}) = T'(\vec{v})$, $\forall \vec{v} \in V$, de donde $T = T'$.

□

Se presentan a continuación ejemplos de transformaciones lineales conocidas como la derivada, la integral definida (entre espacios vectoriales de funciones) y la multiplicación de una matriz por un vector. Se deja al lector la verificación de que son transformaciones lineales.

- **Ejemplo 2.2** $D : P_{\mathbb{R}}^{(n)}[x] \rightarrow P_{\mathbb{R}}^{(n-1)}[x]$ (derivada)

$$D(a_0 1 + a_1 x + a_2 x^2 + \cdots + a_n x^n) = a_1 + 2a_2 x + \cdots + na_n x^{n-1}$$

■

- **Ejemplo 2.3** $J : C([0, 1]) \rightarrow \mathbb{R}$ (integral definida)

$$J(f) = \int_0^1 f(x) dx$$

■

- **Ejemplo 2.4** Dada la matriz,

$$A = \begin{pmatrix} i & 1 & 0 \\ 1 & i & 0 \end{pmatrix}$$

es posible definir la transformación que multiplica la matriz por un vector $\vec{v} = (z_1, z_2, z_3) \in \mathbb{C}^3$, es decir, $A : \mathbb{C}^3 \rightarrow \mathbb{C}^2$, está dada por $A((z_1, z_2, z_3)) = A(z_1, z_2, z_3)^t$.

Es una transformación lineal, ya que se cumple

$$A(\vec{v} + \vec{v}') = A\vec{v} + A\vec{v}'$$

y

$$A(c\vec{v}) = cA\vec{v}$$

■



- Dada una matriz $A \in K^{m \times n}$, la transformación $A : K^n \rightarrow K^m$, dada por

$$A((z_1, z_2, \dots, z_n)) = A(z_1, z_2, \dots, z_n)^T,$$

es la transformación lineal asociada con la matriz A .

- Se verá que, recíprocamente, dada una transformación lineal es posible hallar la matriz que la representa.

2.2 Matriz de una transformación lineal

Sean V y W dos espacios vectoriales sobre el mismo cuerpo K . Sea $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ una base de V y $\vec{B} = \{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ una base de W . El elemento $T(\vec{e}_1)$ es un vector de W , y por lo tanto puede escribirse como combinación lineal de los vectores de la base \vec{B} :

$$T(\vec{e}_1) = a_{11}\vec{f}_1 + a_{21}\vec{f}_2 + \dots + a_{m1}\vec{f}_m$$

Análogamente,

$$T(\vec{e}_2) = a_{12}\vec{f}_1 + a_{22}\vec{f}_2 + \dots + a_{m2}\vec{f}_m$$

.....

$$T(\vec{e}_n) = a_{1n}\vec{f}_1 + a_{2n}\vec{f}_2 + \dots + a_{mn}\vec{f}_m.$$

Estas igualdades se escriben de la forma

$$T(\vec{e}_j) = \sum_{i=1}^m a_{ij}\vec{f}_i, \quad j = 1, 2, \dots, n \quad (2.1)$$

O en forma más abreviada, usando notación indicial, es decir, sumando sobre el índice repetido i de 1 a m ,

$$T(\vec{e}_j) = a_{ij}\vec{f}_i, \quad j = 1, 2, \dots, n$$

En estas condiciones diremos que

$$T = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \dots\dots\dots & & & & \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

es la matriz de la aplicación T con respecto a las bases B y \bar{B} .



- En la j -ésima columna de la matriz de la aplicación lineal T están las coordenadas de $T(\vec{e}_j)$ con respecto a la base \bar{B} de W . Ver Sección 1.5
- El cambio de base para obtener las coordenadas de un vector visto en la Sección 1.5 es una transformación lineal. Las nuevas coordenadas del vector se obtienen al multiplicar por la matriz de cambio de base. La matriz de una transformación lineal, se construye, entonces, de la misma forma que lo hicimos con la matriz de cambio de base.
- A veces se agregan en la notación las bases B y \bar{B} , para indicar las bases consideradas en los espacios V y W .
- Para el caso dimensión finita, y si están especificadas las bases B y \bar{B} , es posible denominar a la matriz de la transformación lineal con la misma letra que la transformación lineal.

Dado $\vec{x} \in V$, se puede escribir

$$\vec{x} = \sum_{j=1}^n x_j \vec{e}_j \quad e$$

$$\vec{y} = T(\vec{x}) = \sum_{i=1}^m y_i \vec{f}_i,$$

la relación entre las coordenadas y_i y x_j de \vec{y} y \vec{x} viene dada por la matriz T .

En efecto, teniendo en cuenta la expresión para $T(\vec{e}_j)$, Ec.(2.1),

$$\sum_{i=1}^m y_i \vec{f}_i = T(\vec{x}) = T\left(\sum_{j=1}^n x_j \vec{e}_j\right) = \sum_{j=1}^n x_j T(\vec{e}_j) = \sum_{j=1}^n x_j \left(\sum_{i=1}^m a_{ij} \vec{f}_i\right) = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j\right) \vec{f}_i$$

del primer y del último término de la igualdad anterior, se tiene que

$$y_i = \sum_{j=1}^n a_{ij} x_j, \quad i = 1, 2, \dots, m$$

es decir que la relación entre las coordenadas y_i y x_j viene dada por los elementos a_{ij} (ver Sección 1.5).

■ **Ejemplo 2.5** Sea P la proyección ortogonal sobre el plano xy , (representada en la Figura 2.1). $P : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. P es una transformación lineal que verifica

$$P(\vec{e}_1) = \vec{e}_1, \quad P(\vec{e}_2) = \vec{e}_2, \quad P(\vec{e}_3) = \vec{0}$$

por lo tanto, su matriz con respecto a la base canónica es:

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

■

■ **Ejemplo 2.6** Sea R_ϕ la transformación que corresponde a una rotación en un ángulo ϕ en sentido positivo (antihorario) alrededor del origen (Ver Figura (2.2)). $R_\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Su matriz en la base canónica es:

$$R_\phi = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}. \quad (2.2)$$

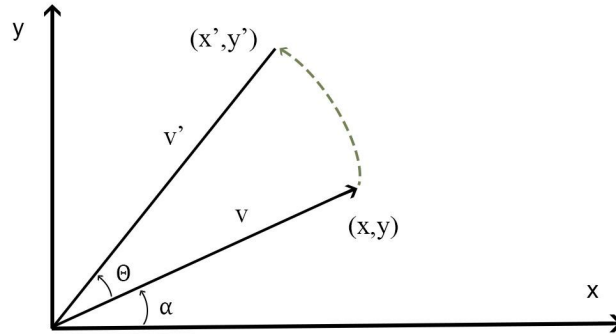
De acuerdo a la Figura 2.2, se tiene que

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = R_\phi \begin{pmatrix} x \\ y \end{pmatrix}$$

Esta matriz es la matriz de cambio de base del Ejemplo 1.28.

■

Figura 2.2: Transformación de rotación



■ **Ejemplo 2.7** En \mathbb{R}^3 , consideramos el subespacio V_1 correspondiente al plano xy . Si $B = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ es la base canónica, y S la transformación lineal que para cada vector \vec{v} da el vector simétrico con respecto al plano xy , como se muestra en la Figura 2.3, se tiene que

$$S(\vec{e}_1) = \vec{e}_1, \quad S(\vec{e}_2) = \vec{e}_2, \quad S(\vec{e}_3) = -\vec{e}_3.$$

Por lo tanto, su matriz con respecto a la base canónica es

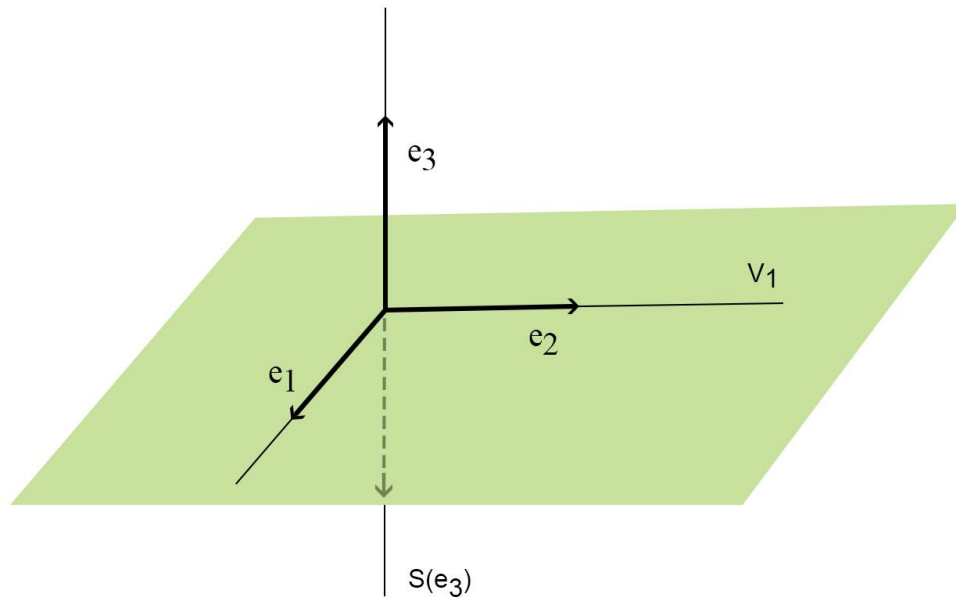
$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

■



Si se quiere hallar la matriz que corresponde a la transformación lineal que a cada vector le hace corresponder el vector simétrico con respecto a un plano cualquiera, es conveniente hallar una base del plano \vec{u}_1, \vec{u}_2 y un vector \vec{u}_3 perpendicular. Así, en la base $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, la matriz de la simetría con respecto al plano es la misma que la matriz del Ejemplo 2.7. Una vez obtenida la matriz, se realiza el cambio de base a la base deseada.

■ **Ejemplo 2.8** De acuerdo al corolario anterior, para hallar la matriz correspondiente a la simetría con respecto al plano $x + y + z = 0$ (Figura 2.4). se busca una base del plano (como $x = -y - z$, los vectores en el plano son de la forma $(-y - z, y, z) = y(-1, 1, 0) + z(-1, 0, 1)$, es decir que los vectores $\vec{u}_1 = (-1, 1, 0)$ y $\vec{u}_2 = (-1, 0, 1)$ son una base del mismo). Y un vector perpendicular es

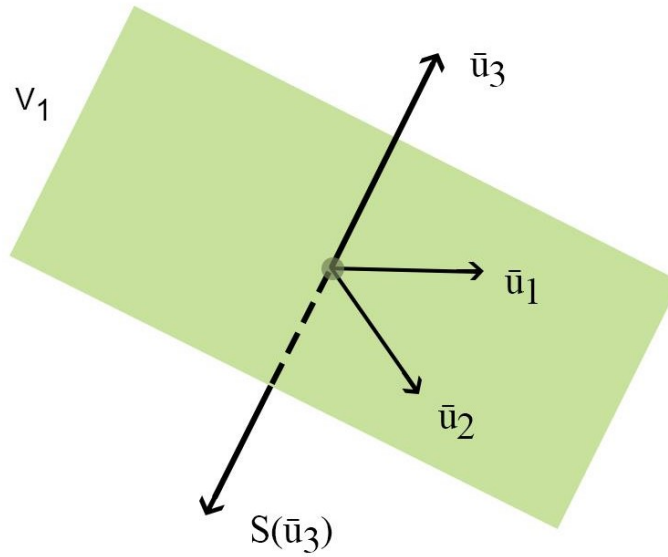
Figura 2.3: Simetría con respecto al plano xy 

$\vec{u}_3 = (1, 1, 1)$. En esa base $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ la matriz de la simetría con respecto al plano $x + y + z = 0$ es, entonces,

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

■

Si en los espacios vectoriales V y W , de dimensiones finitas n y m , respectivamente, se fijan bases, existe una correspondencia biunívoca entre las transformaciones lineales de V en W y el conjunto de las matrices $K^{m \times n}$ (de orden $m \times n$) sobre el cuerpo K . Puesto que el conjunto $K^{m \times n}$ posee una estructura de espacio vectorial, también tiene esa estructura el conjunto de todas las transformaciones lineales entre dos espacios vectoriales sobre el mismo cuerpo K . A ese espacio vectorial se lo denomina $L(V, W)$.

Figura 2.4: Simetría respecto al plano $x + y + z = 0$ 

Teorema 2.2.1 Sean V y W dos espacios vectoriales sobre un mismo cuerpo K ; el conjunto $L(V, W)$ de las aplicaciones lineales entre V y W es un espacio vectorial sobre el cuerpo K .

Demostración:

■ Suma.

Dadas T_1 y $T_2 \in L(V, W)$, se define $T_1 + T_2, T_1 + T_2 : V \rightarrow W$ como

$$(T_1 + T_2)(\vec{v}) = T_1(\vec{v}) + T_2(\vec{v}) \quad \forall \vec{v} \in V$$

Veamos que $T_1 + T_2$ es una transformación lineal

- $(T_1 + T_2)(\vec{v} + \vec{w}) = T_1(\vec{v} + \vec{w}) + T_2(\vec{v} + \vec{w})$
 $= T_1(\vec{v}) + T_1(\vec{w}) + T_2(\vec{v}) + T_2(\vec{w})$
 $= (T_1 + T_2)(\vec{v}) + (T_1 + T_2)(\vec{w})$
- $(T_1 + T_2)(\alpha \vec{v}) = T_1(\alpha \vec{v}) + T_2(\alpha \vec{v})$
 $= \alpha T_1(\vec{v}) + \alpha T_2(\vec{v})$
 $= \alpha(T_1(\vec{v}) + T_2(\vec{v})) = \alpha(T_1 + T_2)(\vec{v})$

■ Producto por escalares.

Dada $T \in L(V, W)$ y $\alpha \in K$, se define $(\alpha T), (\alpha T) : V \rightarrow W$ como

$$(\alpha T)(\vec{v}) = \alpha T(\vec{v})$$

Veamos que (αT) es una transformación lineal

- $(\alpha T)(\vec{v} + \vec{w}) = \alpha(T(\vec{v} + \vec{w}))$
 $= \alpha(T(\vec{v}) + T(\vec{w})) = \alpha T(\vec{v}) + \alpha T(\vec{w})$
 $= (\alpha T)(\vec{v}) + (\alpha T)(\vec{w})$
- $(\alpha T)(\beta \vec{v}) = \alpha(T(\beta \vec{v}))$
 $= \alpha(\beta T(\vec{v})) = (\alpha \beta)T(\vec{v})$
 $= \beta(\alpha T(\vec{v})) = \beta(\alpha T)(\vec{v})$

□

Como toda transformación lineal puede representarse mediante una matriz y recíprocamente, se tiene el siguiente resultado:

Teorema 2.2.2 Sean V y W espacios vectoriales de dimensiones n y m , respectivamente, entonces, el espacio vectorial de las transformaciones lineales del espacio vectorial V al espacio vectorial W , $L(V, W)$, tiene dimensión $m \times n$.

Demostración:

Se puede ver en el libro de E. Hernández [21]. En él se construye una base de $L(V, W)$. También puede demostrarse a partir de la correspondencia biyectiva entre el espacio vectorial de las matrices $K^{m \times n}$ (de dimensión $m \times n$) y $L(V, W)$.

□



- Si V y W coinciden escribimos $L(V)$ en lugar de $L(V, V)$.
- La matriz de la aplicación lineal *suma* coincide con la suma de las matrices de cada una de las aplicaciones y la matriz de la aplicación lineal cT coincide con el producto de la matriz T por el escalar c . Si llamamos $M(T)$ a la matriz de la aplicación lineal T , esto se escribe

$$M(T + T') = M(T) + M(T') \quad \text{y} \quad M(cT) = cM(T), \quad c \in K$$

Se verá a continuación que la composición de funciones usual puede realizarse entre dos transformaciones lineales y el resultado es otra transformación lineal.

Proposición 2.2.3 Sean V, W y X espacios vectoriales sobre el cuerpo K . Sean $T \in L(V, W)$ y $T' \in L(W, X)$. Entonces

$$T' \circ T \in L(V, X)$$

Demostración:

Sean $\vec{v}_1, \vec{v}_2 \in V$, entonces

$$(T' \circ T)(\vec{v}_1 + \vec{v}_2) = T'(T(\vec{v}_1 + \vec{v}_2)) = T'(T(\vec{v}_1)) + T'(T(\vec{v}_2))$$

$$= (T' \circ T)(\vec{v}_1) + (T' \circ T)(\vec{v}_2)$$

Análogamente,

$$(T' \circ T)(\alpha \vec{v}) = T'(T(\alpha \vec{v})) = T'(\alpha T(\vec{v})) = \alpha(T' \circ T)(\alpha \vec{v})$$

□

Proposición 2.2.4 Si los espacios V, W y X tienen dimensión finita y si denotamos por $M(T)$, $M(T')$ y $M(T' \circ T)$ las matrices de T , T' y $T' \circ T$, respectivamente, con respecto a las bases de antemano fijadas, se tiene el siguiente resultado:

$$M(T' \circ T) = M(T')M(T)$$

Demostración:

$$T(\vec{e}_j) = \sum_{i=1}^m a_{ij} \vec{f}_i, \quad j = 1, 2, \dots, n \tag{2.3}$$

donde a_{ij} son los elementos de la matriz $M(T)$.

Sean $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$, $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ y $\{\vec{g}_1, \vec{g}_2, \dots, \vec{g}_p\}$ bases de V , W y X , respectivamente. En la i -ésima columna de la matriz $M(T' \circ T)$ están las coordenadas del vector $(T' \circ T)(\vec{e}_i)$ con respecto a la base g_k .

$$\begin{aligned} T'(T(\vec{e}_i)) &= T'\left(\sum_{j=1}^m a_{ji} \vec{f}_j\right) \\ &= \sum_{j=1}^m a_{ji} T'(\vec{f}_j) \\ &= \sum_{j=1}^m a_{ji} \sum_{k=1}^p b_{kj} \vec{g}_k \\ &= \sum_{k=1}^p \sum_{j=1}^m b_{kj} a_{ji} \vec{g}_k \end{aligned}$$

donde b_{ij} son los elementos de la matriz $M(T')$.

Esto prueba que $\sum_{j=1}^m b_{kj} a_{ji}$ es el elemento que ocupa el lugar (k, i) de la matriz $M(T' \circ T)$ y este valor coincide con el elemento (k, i) del producto de las matrices $M(T')$ y $M(T)$.

□

- i** El resultado anterior se generaliza para el caso de una sucesión de transformaciones lineales, T_i , $i = 1, \dots, k$ aplicadas a un vector \vec{v} . Se tendrá entonces que resulta equivalente a aplicar a \vec{v} una única matriz T tal que

$$M(T) = M(T_k)M(T_{k-1}) \cdots M(T_2)M(T_1).$$

2.3 Transformaciones lineales inyectivas y suryectivas

Sean V y W dos espacios vectoriales sobre el mismo cuerpo K y T una aplicación lineal de V en W . Recordamos que T es inyectiva si $T(\vec{x}) = T(\vec{y})$ implica $\vec{x} = \vec{y}$ y T es suryectiva si para todo $\vec{y} \in W$ existe $\vec{x} \in V$ tal que $T(\vec{x}) = \vec{y}$ (o equivalentemente $T(V) = W$, donde $T(V)$ denota la imagen de V mediante T). Finalmente recordamos que T es *biyectiva* si es a la vez inyectiva y suryectiva.

- i** // En el caso de transformaciones lineales cada uno de los tipos anteriores recibe un nombre especial: una aplicación lineal *inyectiva* recibe el nombre de *monomorfismo*; si es *suryectiva* se le da el nombre de *epimorfismo*; finalmente si la aplicación es *biyectiva* se dice que es un *isomorfismo*.

Encontraremos ahora condiciones sencillas que sirvan para determinar si una aplicación lineal es de cualquiera de los tipos anteriores. Comenzaremos con las aplicaciones inyectivas, y para ello necesitamos definir el concepto de *núcleo* de una aplicación lineal.

2.4 Núcleo e imagen de una transformación lineal

Definición 2.4.1 Dada una aplicación lineal $T : V \rightarrow W$, definimos el *núcleo* de T , que se denota por $N(T)$ (o $\text{Ker}(T)$, del inglés *kernel* significa núcleo), como el conjunto de todos los $\vec{v} \in V$ tales que $T(\vec{v}) = \vec{0}$, es decir

$$N(T) = \left\{ \vec{v} \in V, / T(\vec{v}) = \vec{0} \right\}$$

El subconjunto $N(T)$ nunca es vacío, ya que $\vec{0} \in N(T)$ y esto se deduce de que $T(\vec{0}) = \vec{0}$ como ya fue demostrado. Se tiene además, el siguiente resultado:

Proposición 2.4.1 Si $T : V \rightarrow W$ es una aplicación lineal entre espacios vectoriales, $N(T)$ es un subespacio vectorial de V .

Demostración:

Esta propiedad es consecuencia de la Proposición 2.1.3. Por definición $N(T)$ es la preimagen de $\vec{0}_W$ que es un subespacio de W . □

Proposición 2.4.2 Una aplicación lineal $T : V \rightarrow W$ es inyectiva si y solo si $N(T) = \{\vec{0}\}$.

Demostración:

- Si T es inyectiva se tiene que $T(\vec{v}) = T(\vec{v}')$ implica que $\vec{v} = \vec{v}'$. Si $\exists \vec{v} \in N(T)$ tal que $T(\vec{v}) = \vec{0}$ como $T(\vec{0}) = \vec{0}$ resulta $\vec{v} = \vec{0}$.
- Para ver que T es inyectiva suponemos $\exists \vec{v}$ y \vec{v}' , tales que $T(\vec{v}) = T(\vec{v}')$. Por ser T una transformación lineal $T(\vec{v}) = T(\vec{v}') = T(\vec{v} - \vec{v}')$, y si $T(\vec{v} - \vec{v}') = \vec{0}$, entonces, $\vec{v} - \vec{v}' \in N(T)$. Como $N(T) = \{\vec{0}\}$, se tiene que $\vec{v} = \vec{v}'$ y por lo tanto T es inyectiva. □

■ **Ejemplo 2.9** Para la proyección ortogonal P de la Figura 2.1 (Ejemplo 2.5),

$$N(P) = \{(x, y, z) \in \mathbb{R}^3 \text{ tales que } P(x, y, z) = (x, y, 0) = (0, 0, 0)\}.$$

O sea $N(P) = \{(0, 0, z), z \in \mathbb{R}\}$, todo el eje z . Por la Proposición 2.4.2, se tiene que P no es inyectiva (intuitivamente se ve que muchos vectores de \mathbb{R}^3 dan el mismo vector al proyectar los sobre el plano xy). Además, se tiene que las dimensiones de $N(P)$ y de la $Im(P)$ son 1 y 2, respectivamente, suman 3, que es la dimensión de \mathbb{R}^3 . ■

i Si bien las soluciones de un sistema $A\vec{X} = \vec{b}$, $\vec{b} \neq \vec{0}$ son un subconjunto pero no un subespacio de \mathbb{R}^n (ver Observación i en 1.2), toda solución puede expresarse de la forma $\vec{X} = \vec{X}_{NH} + \vec{X}_H$, donde \vec{X}_{NH} es solución de $A\vec{X} = \vec{b}$ mientras que $\vec{X}_H \in Nul(A)$. Esto sale porque si \vec{X}_{NH} es solución de $A\vec{X} = \vec{b}$, $A(\vec{X}_{NH} + \vec{X}_H) = A\vec{X}_{NH} + A\vec{X}_H = \vec{b} + \vec{0} = \vec{b}$ y si \vec{X} es otra solución de $A\vec{X} = \vec{b}$, entonces $A(\vec{X} - \vec{X}_{NH}) = A\vec{X} - A\vec{X}_{NH} = \vec{b} - \vec{b} = \vec{0}$, de donde $\vec{X} - \vec{X}_{NH} \in Nul(A)$ y por lo tanto $\vec{X} = \vec{X}_{NH} + \vec{X}_H$.

Como ejemplo se deja al lector verificar que la solución del sistema no homogéneo

$$\begin{cases} x + y + w &= 1 \\ 2x + 3y + z + 2w &= 1 \end{cases}$$

es $\vec{X} = \vec{X}_{NH} + \vec{X}_H = (2, -1, 0, 0) + \alpha(1, -1, 1, 0) + \beta(-1, 0, 0, 1)$ con α y $\beta \in \mathbb{R}$. ▲

Como se demostró en la Proposición 2.1.2, la imagen de una transformación lineal T es un subespacio. En el teorema que sigue se enuncia la relación que existe entre las dimensiones del núcleo de T , de la imagen de T y la dimensión de V :

Teorema 2.4.3 Sean V y W dos espacios vectoriales de los cuales V es de dimensión finita y $T : V \rightarrow W$ es una aplicación lineal. Entonces

$$\dim(N(T)) + \dim(Im(T)) = \dim(V).$$

Demostración:

Sea $n = \dim(V)$ y $k = \dim(N(T))$,

si $k = n$, entonces T es la aplicación nula y la $\dim(Im(T)) = 0$. Por lo tanto el teorema vale.

Si $k = 0$, entonces T es un monomorfismo. Si B es una base de V , $T(V)$ es base de la $Im(T)$. Luego la $dim(Im(T)) = dim(V)$ y el teorema vale.

Supongamos que $0 < k < n$ y sea $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ una base de $N(T)$. Sean $\vec{v}_{k+1}, \vec{v}_{k+2}, \dots, \vec{v}_n$ tales que $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k, \vec{v}_{k+1}, \vec{v}_{k+2}, \dots, \vec{v}_n\}$ es una base de V .

Veamos que $\{T(\vec{v}_{k+1}), T(\vec{v}_{k+2}), \dots, T(\vec{v}_n)\}$ es una base de $Im(T)$.

En ese caso se tendrá que $dim(N(T)) + dim(Im(T)) = k + (n - k) = n = dim(V)$.

- Si $\vec{w} \in Im(T)$, $\exists \vec{v} \in V$ tal que $T(\vec{v}) = \vec{w}$.

Como $\vec{v} = \sum_{j=1}^n c_j \vec{v}_j$, $\vec{w} = T(\vec{v}) = \sum_{j=1}^n c_j T(\vec{v}_j) = \sum_{j=k+1}^n c_j T(\vec{v}_j)$, ya que $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ es una base de $N(T)$.

Entonces, $\{T(\vec{v}_{k+1}), T(\vec{v}_{k+2}), \dots, T(\vec{v}_n)\}$ es un sistema de generadores de $Im(T)$.

- Para ser si en un conjunto linealmente independiente, supongamos

$\sum_{j=k+1}^n c_j T(\vec{v}_j) = \vec{0} = T(\sum_{j=k+1}^n c_j \vec{v}_j)$, entonces $\sum_{j=k+1}^n c_j \vec{v}_j \in N(T)$. Como $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ es una base de $N(T)$, existen escalares c_1, c_2, \dots, c_k tales que

$$\sum_{j=k+1}^n c_j \vec{v}_j = \sum_{j=1}^k c_j \vec{v}_j$$

que puede reescribirse

$$\sum_{j=1}^k (-c_j) \vec{v}_j + \sum_{j=k+1}^n c_j \vec{v}_j = \vec{0}$$

Se tiene que $c_i = 0$, $\forall 1 \leq i \leq k$, por ser $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k, \vec{v}_{k+1}, \vec{v}_{k+2}, \dots, \vec{v}_n\}$ una base de V .

En particular $c_i = 0$, $\forall k+1 \leq i \leq n$. Luego $\{T(\vec{v}_{k+1}), T(\vec{v}_{k+2}), \dots, T(\vec{v}_n)\}$ es un conjunto linealmente independiente.

□

- **Ejemplo 2.10** Se verificará el teorema anterior para la transformación lineal $T : \mathbb{R}^5 \rightarrow \mathbb{R}^3$, dada por $T((z_1, z_2, \dots, z_5)) = A(z_1, z_2, \dots, z_5)^T$, donde A es la matriz (ver Observación 1 al final de la Sección 2.1).

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

En primer lugar, se resuelve el sistema homogéneo utilizando eliminación gaussiana (con la matriz ampliada):

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

Al quedar solo dos pivotes, hay 2 variables dependientes y $n - 2 = 3$ variables independientes. Se tiene que $\dim(N(T))=3$.

Para estudiar cuál es el subespacio que corresponde a la imagen de T , se debe hallar el subespacio de \mathbb{R}^3 que generan las columnas. Puede repetirse la eliminación anterior con término independiente (x, y, z) .

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & x \\ 0 & 1 & 0 & -1 & 1 & y \\ 1 & 0 & 1 & 2 & 0 & z \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & x \\ 0 & 1 & 0 & -1 & 1 & y \\ 0 & -1 & 0 & 1 & -1 & z-x \end{pmatrix} \rightarrow \begin{pmatrix} \underline{1} & 1 & 1 & 1 & 1 & x \\ 0 & \underline{1} & 0 & -1 & 1 & y \\ 0 & 0 & 0 & 0 & 0 & z-x+y \end{pmatrix}$$

Se tiene, entonces, que $Im(T) = \{(x, y, z) = (x, y, x - y)\}$, es el plano por el origen $z = x - y$, y

$$\dim(N(T)) + \dim(Im(T)) = 3 + 2 = 5 = \dim(V),$$

ya que $V = \mathbb{R}^5$. ■

Para deducir algunas consecuencias del Teorema 2.4.3 es necesario hacer uso del concepto de *rango*.

Sea T una aplicación entre los espacios vectoriales V y W , ambos de dimensión finita, m y n respectivamente. Sea A la matriz de la aplicación lineal en dos bases cualesquiera de V y W . Para encontrar el núcleo de T , $N(T)$, es necesario resolver el sistema homogéneo $A\vec{x} = \vec{0}$ (como se hizo en el ejemplo anterior).

Si $r(A)$ es el rango de la matriz A , se obtienen, $r(A)$ soluciones dependientes y $n - r(A)$ soluciones linealmente independientes. Es decir que

$$\dim(N(T)) = \dim(V) - r(A)$$

y comparando con el Teorema 2.4.3, se tiene que

$$\dim(\text{Im}(T)) = r(A)$$

Puesto que la dimensión de la $\text{Im}(T)$ no depende de las bases que se elijan en V y W , de la igualdad anterior se deduce que las matrices de la aplicación T en cualquier base tienen el mismo rango.

Como consecuencia de lo anterior, es posible definir el *rango de una transformación lineal* T , que escribiremos $r(T)$ como el rango de cualquiera de sus representaciones matriciales, y resumir los resultados anteriores en el Corolario que sigue:

Corolario Sean V y W dos espacios vectoriales de los cuales V es de dimensión finita y T una transformación lineal, $T : V \rightarrow W$. Entonces

1. T es inyectiva sí y sólo sí $r(T) = \dim(V)$.
2. T es suryectiva sí y sólo sí $r(T) = \dim(W)$.

Por último, se estudian las aplicaciones lineales entre espacios vectoriales de igual dimensión que son biyectivas. Se conocen como isomorfismos. Supongamos que V y W son espacios vectoriales de dimensión finita, y que T es un isomorfismo entre ellos. Del Corolario anterior deducimos que

$$\dim(V) = r(T) = \dim(W)$$

Otras consecuencias de resultados anteriores se resumen en el teorema que sigue.

Teorema 2.4.4 Sean V y W espacios vectoriales de dimensión finita n y sea $T : V \rightarrow W$ es una aplicación lineal. Las siguientes condiciones son equivalentes:

1. T es biyectiva
2. T es inyectiva
3. $N(T) = \{\vec{0}\}$
4. T es suryectiva
5. El rango de T es n

Demostración:

Entre 2, 3, 4 y 5 se tienen las equivalencias siguientes:

$$\begin{array}{ccc}
 2 & \xLeftrightarrow{\text{Prop. 2.4.2}} & 3 \\
 \Downarrow \text{Corolario} & & \\
 5 & \xLeftrightarrow{\text{Corolario}} & 4
 \end{array}$$

Además, $1 \rightarrow 2$ porque toda transformación biyectiva es inyectiva. Como 2 y 4 son equivalentes en este contexto y ambas implican 1, se tiene que también $2 \rightarrow 1$, y queda demostrado. \square

Decimos que dos espacios vectoriales cualesquiera son *isomorfos* si podemos encontrar un isomorfismo entre ellos. Para que esto ocurra entre espacios vectoriales de dimensión finita ya vimos que ambos han de tener la misma dimensión. El recíproco también es cierto.

Teorema 2.4.5 Dado cualquier número natural n , todos los espacios vectoriales de dimensión n sobre un mismo cuerpo son isomorfos.

Demostración:

Sean V y W dos espacios vectoriales de dimensión n con bases $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ y $\{\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_n\}$ respectivamente. Existe una transformación lineal T_1 , $T_1 : V \rightarrow K^n$ definida de la forma siguiente:

$$\text{Si } \vec{v} = \sum_{i=1}^n \alpha_i \vec{e}_i,$$

$$T_1(\vec{v}) = (\alpha_1, \alpha_2, \dots, \alpha_n)$$

Es decir que la transformación da el vector con las coordenadas de \vec{v} . Se demuestra fácilmente (y se deja al lector) que esta transformación es lineal, inyectiva y suryectiva. Al ser biyectiva, existe también su transformación inversa (Proposición 2.4.4). Utilizando este isomorfismo de un espacio vectorial V con K^n , se tiene que dos espacios cualesquiera de la misma dimensión son isomorfos. Para encontrar la transformación entre V y W hay que componer la transformación T_1 entre V y K^n con la transformación T_2 entre K^n y el espacio vectorial W , $T_2 : K^n \rightarrow W$. Esta última está dada por

$$T_2(\alpha_1, \alpha_2, \dots, \alpha_n) = \sum_{i=1}^n \alpha_i \vec{e}'_i.$$

El isomorfismo entre V y W está dado por la transformación

$$(T_2 \circ T_1)(\vec{v}) = T_2(T_1(\vec{v})) = \vec{w} = \sum_{i=1}^n \alpha_i \vec{e}'_i.$$

□

■ **Ejemplo 2.11** Aplicando el Teorema anterior, el isomorfismo entre $P_{\mathbb{R}}^{(2)}[t]$ y \mathbb{R}^3 está dado por

$$T(a_0 1 + a_1 t + a_2 t^2) = (a_0, a_1, a_2)^t,$$

mientras que el isomorfismo entre \mathbb{R}^3 y las matrices simétricas de $\mathbb{R}^{2 \times 2}$ está dado por:

$$T(x, y, z) = \begin{pmatrix} x & y \\ y & z \end{pmatrix}.$$

Se consideraron las bases canónicas de \mathbb{R}^3 , de $P_{\mathbb{R}}^{(2)}[t]$ y de las matrices simétricas de $\mathbb{R}^{2 \times 2}$, espacios vectoriales de dimensión 3. Se deja al lector la verificación de estos resultados. ■



Si T es un isomorfismo entre dos espacios vectoriales V y W de dimensión n , por el Teorema 2.4.4, su rango es n , y por lo tanto la matriz $M(T)$ de T en cualesquiera bases de V y W es *invertible*. Además, la inversa de $M(T)$ es la matriz de la aplicación inversa de T .

■ **Ejemplo 2.12** Sean V y W dos espacios vectoriales de funciones, de dimensión infinita:

$$V = \{f \in C^1[0, 1] / f(0) = 0\} \text{ y } W = C[0, 1].$$

Sea la transformación

$D : V \rightarrow W$ dada por $D(f) = f'$. D es una transformación lineal (En el Ejemplo 2.2 se vió para polinomios en $P_R^{(n)}[x]$).

■ D es monomorfismo

Supongamos $D(f) = D(g)$, entonces $f' = g'$ o, en forma equivalente $(f - g)' = 0$. Entonces $f(x) - g(x) = cte$. Como $f(0) = g(0) = 0$, se tiene que la $cte = 0$, por lo tanto $f = g$.

- D es epimorfismo

Sea $g \in W$ y sea

$$f(x) = \int_0^x g(t)dt.$$

Entonces, por el Teorema Fundamental del cálculo, $f \in C^1[0, 1]$ y $f' = g(x)$, $\forall x \in [0, 1]$. Más aún, como

$$\int_0^0 g(t)dt = 0,$$

se tiene que $f(0) = 0$. Por lo tanto, $\forall g \in W$, $\exists f \in V$ tal que $Df = g$. O sea D es epimorfismo. Resulta, entonces, que V y W son espacios isomorfos. ■

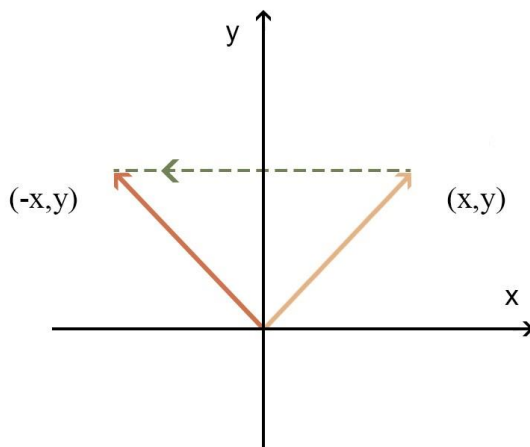
Gabriela González

Gabriela González es una física, investigadora y profesora argentina. Nació en 1965. Fue portavoz y coordinó durante seis años un equipo de mil especialistas, que trabajó en las detecciones de ondas gravitacionales efectuadas desde el proyecto LIGO (Ondas Gravitacionales con Interferómetro Láser, por sus siglas en inglés). En febrero de 2016 fue uno de los cuatro científicos de LIGO que anunciaron la primera observación ondulatoria gravitacional, detectada en septiembre de 2015. Egresada de la Universidad Nacional de Córdoba y actual profesora en el departamento de física y astronomía de la Universidad de Louisiana, fue reconocida en 2016 como una de los diez científicos más destacados del mundo por la revista académica Nature. Además, a partir de 2018 forma parte de la Academia de Ciencias de Estados Unidos, institución de máximo prestigio internacional. [9]

2.5 Geometría de las transformaciones lineales de \mathbb{R}^2 en \mathbb{R}^2

Se verán en esta sección algunas propiedades geométricas de las transformaciones lineales en el plano. Dada la matriz

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

Figura 2.5: Transformación de simetría (o reflexión) respecto del eje y 

la transformación $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $L((x, y)) = A(x, y)^t$ es

$$L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

■ **Ejemplo 2.13** En la Figura 2.5 se muestra la transformación que a cada vector le hace corresponder el simétrico respecto del eje y .

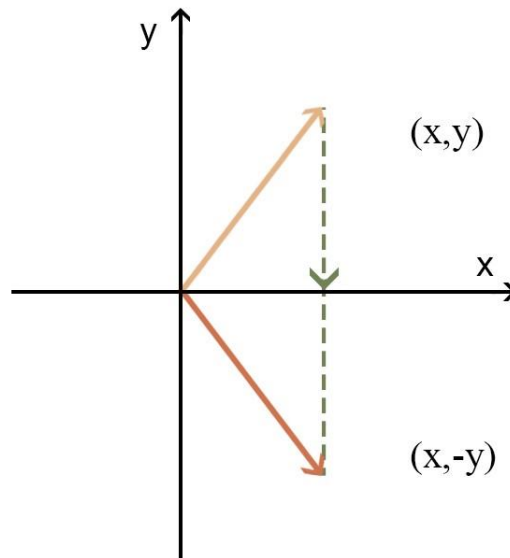
$$L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}$$

■ **Ejemplo 2.14** En la Figura 2.6 se muestra la transformación que a cada vector le hace corresponder el simétrico respecto del eje x

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

■ **Ejemplo 2.15** En la Figura 2.7 se muestra la transformación que a cada vector le hace corresponder su proyección ortogonal sobre el eje x :

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

Figura 2.6: Transformación de simetría respecto del eje x 

■ **Ejemplo 2.16** En la Figura 2.8 se muestra la transformación que a cada vector le hace su reflexión respecto de la recta $y = x$:

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

i Otras transformaciones se obtienen al multiplicar una de las coordenadas por una constante k . Así el efecto es comprimir o dilatar en esa dirección, dependiendo si $k < 1$ o $k > 1$. También están las transformaciones llamadas de *trasquillado*, dadas por matrices de la forma:

$$L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & k \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + ky \\ y \end{pmatrix}.$$

Estos casos se analizarán en los ejercicios propuestos.

2.6 Cambio de base para transformaciones lineales

Sean V y W dos espacios vectoriales sobre el mismo cuerpo K de dimensiones n y m , respectivamente. Sea T una transformación lineal de V en W con matriz

Figura 2.7: Proyección ortogonal sobre el eje x

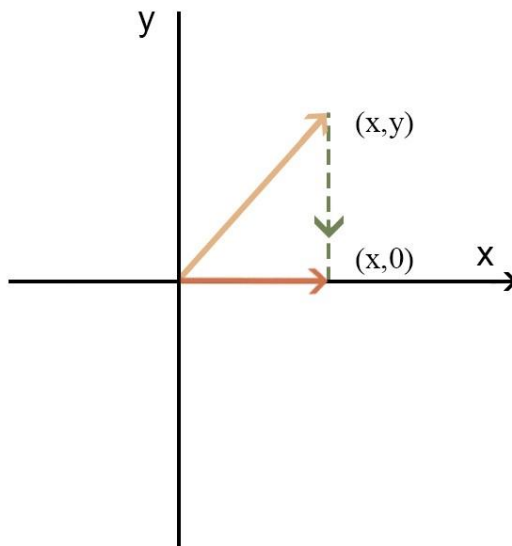
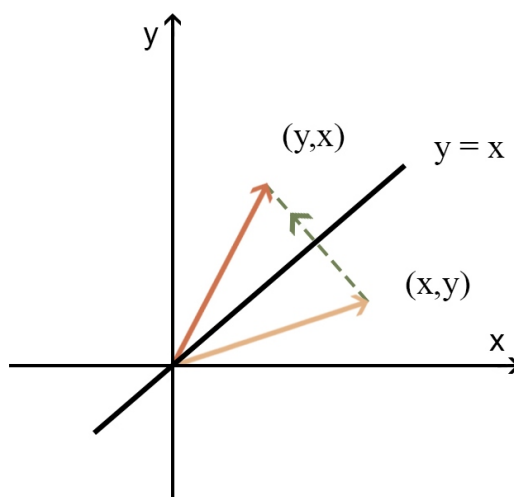


Figura 2.8: Transformación de reflexión respecto de la recta $y = x$



$$T = (a_{i,j})_{i=1,m,j=1,n}$$

con respecto a las bases $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ y $\bar{B} = \{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ de V y W , respectivamente.

Si queremos conocer la matriz $T' = (a'_{i,j})_{i,j=1,n}$ de la misma aplicación T respecto a dos nuevas bases $B' = \{\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_n\}$ y $\bar{B}' = \{\vec{f}'_1, \vec{f}'_2, \dots, \vec{f}'_m\}$ de V y W , respectivamente, es necesario realizar los cambios de base adecuados en los espacios vectoriales inicial y final, V y W .

Para seguir el razonamiento, veamos el diagrama siguiente:

$$\begin{array}{ccc}
 V & & W \\
 & y = T(x) & \\
 B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\} & \xrightarrow{a_{i,j}} & \bar{B} = \{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\} \\
 C \uparrow & & D \uparrow \\
 B' = \{\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_n\} & \xrightarrow{a'_{i,j}} & \bar{B}' = \{\vec{f}'_1, \vec{f}'_2, \dots, \vec{f}'_m\}
 \end{array}$$

En el diagrama

- a_{ij} son los elementos de la matriz de la transformación T tomando la base B en V y la base \bar{B} en W
- a'_{ij} son los elementos de la matriz de la transformación T tomando la base B' en V y la base \bar{B}' en W
- C y D son las matrices del cambio de base de B' a B y de \bar{B}' a \bar{B} , respectivamente.

Se tiene que $\vec{x} \in V$ puede escribirse de dos formas

$$x_1 \vec{e}_1 + x_2 \vec{e}_2 + \dots + x_n \vec{e}_n = x'_1 \vec{e}'_1 + x'_2 \vec{e}'_2 + \dots + x'_n \vec{e}'_n$$

e $\vec{y} = T(\vec{x})$ también

$$y_1 \vec{f}_1 + y_2 \vec{f}_2 + \dots + y_m \vec{f}_m = y'_1 \vec{f}'_1 + y'_2 \vec{f}'_2 + \dots + y'_m \vec{f}'_m$$

En primer lugar, se tiene que,

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_m \end{pmatrix} = T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{pmatrix} \quad (2.4)$$

y de acuerdo a los cambios de base en un mismo espacio vectorial que se estudiaron antes,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{pmatrix} = C \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ \dots \\ x'_n \end{pmatrix} \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_m \end{pmatrix} = D \begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \\ \dots \\ y'_m \end{pmatrix}$$

mientras que, considerando las bases B' y \bar{B}' , se tiene que,

$$\begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \\ \dots \\ y'_m \end{pmatrix} = T' \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ \dots \\ x'_n \end{pmatrix} \quad (2.5)$$

Sustituyendo en (2.4), se obtiene,

$$D \begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \\ \dots \\ y'_m \end{pmatrix} = TC \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ \dots \\ x'_n \end{pmatrix}$$

y como D es una matriz de cambio de base, tiene inversa, por lo tanto

$$\begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \\ \dots \\ y'_m \end{pmatrix} = D^{-1}TC \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ \dots \\ x'_n \end{pmatrix}$$

Comparando esta expresión con (2.5) se obtiene, finalmente que

$$T' = D^{-1}TC,$$

y es posible calcular la matriz T' de la aplicación T con respecto a las bases B' y \bar{B}' , conocidas la matriz T de la misma aplicación con respecto a las bases \bar{B} y \bar{B}' y las matrices C y D del cambio de base B' a B y de \bar{B}' a \bar{B} , respectivamente.



- Cuando entre dos matrices T y T' se tiene la relación $T' = D^{-1}TC$, se dice que las matrices T y T' son *equivalentes* (D y C son matrices invertibles). Y si las matrices D y C coinciden, se dice que las matrices T y T' son *semejantes*.
- En muchos casos los espacios inicial V y final W de una transformación lineal coinciden, y se anota $T \in L(V)$.
- Cuando B y \bar{B} coinciden y B' y \bar{B}' coinciden, la fórmula del cambio de base es más sencilla. Si T es la matriz de la aplicación $T \in L(V)$ con respecto a la base B de V , la matriz T' de la misma aplicación con respecto a una base B' de V está dada por

$$T' = C^{-1}TC.$$

donde C es la matriz del cambio de base de B' a B . Al tener esta relación entre las matrices, por lo anterior, T y T' son *semejantes*.

- Si $T' = C^{-1}TC$, $\text{Det}(T') = \text{Det}(T)$, ya que

$$\begin{aligned} \text{Det}(T') &= \text{Det}(C^{-1}TC) = \text{Det}(C)^{-1}\text{Det}(T).\text{Det}(C) \\ &= \text{Det}(C^{-1}C)\text{Det}(T) = \text{Det}(I)\text{Det}(T) = \text{Det}(T). \end{aligned}$$

■ **Ejemplo 2.17** Si B es la base canónica, $B' = \{(1, 2), (2, 3)\}$ y

$$(T)_B = \begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix}$$

se tiene que

$$(T)_{B'} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

■

2.7 Espacio dual de un espacio Vectorial

Dado un espacio vectorial V sobre un cuerpo K , podemos considerar el conjunto $L(V, K)$ de todas las transformaciones lineales de V en el espacio vectorial K (de dimensión 1 sobre K).

Este espacio vectorial es un caso particular del estudiado anteriormente ($L(V, W)$), y el Teorema 2.2.1 de la sección nos permite concluir que $L(V, K)$ es un espacio vectorial sobre K . Este espacio vectorial recibe el nombre de *espacio dual* del espacio vectorial V y para indicarlo se utiliza comúnmente el símbolo V^* , en lugar de $L(V, K)$. En otras palabras, V^* es el espacio vectorial de todas las aplicaciones lineales de V en K , también llamados funcionales lineales.

Los elementos de V^* son transformaciones lineales. Si V es un espacio vectorial de dimensión finita n , del Teorema 2.2.2 se deduce que el espacio dual V^* tiene dimensión n . El teorema que sigue exhibe una base B^* asociada de manera única y natural a una base B de V . La demostración del Teorema 2.2.2, más general, fue citada. Se presenta a continuación la demostración para este caso particular.

Teorema 2.7.1 Sea V un espacio vectorial de dimensión finita n y $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ una base de V . Existe una única base

$$B^* = \{\varphi_1, \varphi_2, \dots, \varphi_n\}$$

de V^* tal que $\varphi_i(\vec{e}_i) = 1$ para todo $i = 1, \dots, n$, y $\varphi_j(\vec{e}_i) = 0$ si $i \neq j$. Es decir, los elementos de la base dual de B satisfacen

$$\varphi_j(\vec{e}_i) = \delta_{ij}, \quad j, i = 1, 2, \dots, n$$

La base B^* se denomina *base dual* de B .

Demostración:

Para $\vec{v} = \sum_{j=1}^n v_j \vec{e}_j$, definimos los funcionales φ_j , $j = 1, 2, \dots, n$ de la forma siguiente,

$$\varphi_j(\vec{v}) = v_j$$

es decir, da la coordenada j -ésima.

- $\varphi_j \in V^*$ y satisface $\varphi_j(\vec{e}_i) = \delta_{ij}$
- ¿ $\{\varphi_1, \varphi_2, \dots, \varphi_n\}$ son linealmente independientes ?

Si $\sum_{j=1}^n c_j \varphi_j = \vec{0}$, ¿ Se cumple que $c_j = 0$, $\forall j$?

Notar que el término $\vec{0}$ del miembro derecho de la igualdad es la aplicación nula (la imagen de $\vec{0}$ es $0 \in K \forall \vec{v} \in V$). Se deberá cumplir esa igualdad al evaluar las transformaciones lineales de ambos lados en cualquier vector \vec{v} . En particular, si se evalúan en los vectores de la base,

$$\sum_{j=1}^n c_j \varphi_j(\vec{e}_i) = \vec{0}(\vec{e}_i) = 0, \quad \forall i = 1, 2, \dots, n.$$

Por lo tanto, $c_i = 0, i = 1, 2, \dots, n$, ya que $\varphi_j(\vec{e}_i)$ es no nulo solo cuando $j = i$. De ahí que B^* es un conjunto de aplicaciones linealmente independientes.

- Finalmente, para ver que generan, si $A \in V^*$,

$$\begin{aligned} A(\vec{v}) &= A\left(\sum_{j=1}^n v_j \vec{e}_j\right) = \sum_{j=1}^n A(v_j \vec{e}_j) \\ &= \sum_{j=1}^n v_j A(\vec{e}_j) = \sum_{j=1}^n \varphi_j(\vec{v}) A(\vec{e}_j) = \sum_{j=1}^n A(\vec{e}_j) \varphi_j(\vec{v}). \end{aligned}$$

Entonces, se tiene la igualdad

$$A = \sum_{j=1}^n A(\vec{e}_j) \varphi_j$$

y por lo tanto, $B^* = \{\varphi_1, \varphi_2, \dots, \varphi_n\}$ es una base y queda demostrado el teorema.

□

■ **Ejemplo 2.18** Se quiere hallar la base dual $B^* = \{\varphi_1, \varphi_2, \varphi_3\}$ de la base canónica $B = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ de \mathbb{R}^3 .

La transformación lineal φ_1 debe satisfacer

$$\varphi_1(\vec{e}_1) = 1, \varphi_1(\vec{e}_2) = 0, \text{ y } \varphi_1(\vec{e}_3) = 0, \text{ de donde se obtiene,}$$

$$\varphi_1(x_1, x_2, x_3) = \varphi_1(x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3) = x_1.$$

$$\text{De manera similar, } \varphi_2(x_1, x_2, x_3) = x_2 \text{ y } \varphi_3(x_1, x_2, x_3) = x_3.$$

■

■ **Ejemplo 2.19** Sean $L_i, i = 1, 2, 3$, funcionales sobre $P_{\mathbb{R}}^{(2)}[t]$, definidos como $L_i(p(t)) = p(t_i)$ donde los t_i son distintos.

Son aplicaciones lineales y son linealmente independientes, ya que si $c_1L_1 + c_2L_2 + c_3L_3 = \vec{0}$, para todo $p \in P_{\mathbb{R}}^{(2)}[t]$, entonces $c_1 = c_2 = c_3 = 0$.

$V = P_{\mathbb{R}}^{(2)}[t]$ tiene dimensión 3. L_1, L_2 y $L_3 \in V^*$ y entonces, $\{L_1, L_2, L_3\}$ es una base de V^* . (Recordar que V y V^* tienen la misma dimensión).

¿Existe una base B de V para la cual $\{L_1, L_2, L_3\}$ es su base dual B^* ?

Es decir, se quieren hallar $\{p_1, p_2, p_3\} \in V$ tales que

$$L_j(p_i) = \delta_{ji}$$

$$p_1(t_1) = 1, p_1(t_2) = 0, p_1(t_3) = 0$$

$$p_2(t_1) = 0, p_2(t_2) = 1, p_2(t_3) = 0$$

$$p_3(t_1) = 0, p_3(t_2) = 0, p_3(t_3) = 1$$

De donde,

$$p_1(t) = \frac{(t-t_2)(t-t_3)}{(t_1-t_2)(t_1-t_3)}$$

$$p_2(t) = \frac{(t-t_1)(t-t_3)}{(t_2-t_1)(t_2-t_3)}$$

$$p_3(t) = \frac{(t-t_1)(t-t_2)}{(t_3-t_1)(t_3-t_2)}$$

Entonces, para cada $p \in V$, sus coordenadas son:

$$p(t) = L_1(p(t))p_1(t) + L_2(p(t))p_2(t) + L_3(p(t))p_3(t) = p(t_1)p_1(t) + p(t_2)p_2(t) + p(t_3)p_3(t)$$

$p_1(t), p_2(t), p_3(t)$ son los polinomios de interpolación de Lagrange. Es importante señalar que estos polinomios tienen muchas aplicaciones en aproximación de funciones y en integración numérica.

Es posible, por ejemplo, expresar el polinomio $p(t) = t^2 + 1$ como combinación lineal de los funcionales L_i si $t_1 = 0, t_2 = 1$ y $t_3 = -1$.

$$p_1(t) = \frac{(t-t_2)(t-t_3)}{(t_1-t_2)(t_1-t_3)} = \frac{(t-1)(t+1)}{(-1)(1)} = 1-t^2$$

$$p_2(t) = \frac{(t-t_1)(t-t_3)}{(t_2-t_1)(t_2-t_3)} = \frac{(t)(t+1)}{(1)(2)} = \frac{(t^2+t)}{2}$$

$$p_3(t) = \frac{(t-t_1)(t-t_2)}{(t_3-t_1)(t_3-t_2)} = \frac{(t)(t-1)}{(-1)(-2)} = \frac{(t^2-t)}{2}$$

Como $L_i(p) = p(t_i)$, se tiene que

$$L_1(p) = p(t_1) = p(0) = 1, \quad L_2(p) = p(t_2) = p(1) = 2, \quad \text{y} \quad L_3(p) = p(t_3) = p(-1) = 2.$$

Finalmente,

$$p(t) = (1)(1-t^2) + 2\frac{(t^2+t)}{2} + 2\frac{(t^2-t)}{2} = t^2 + 1$$

Esta última es la expresión de $p(t)$ en la base $\{p_1, p_2, p_3\}$. Notar que sus coordenadas están dadas por $L_i(p) = p(t_i)$, $i = 1, 2, 3$. ■

Siempre es posible hallar la base de B como se hizo en el ejemplo anterior. Así como toda base de V de dimensión finita tiene una base dual asociada, toda base de V^* es la base dual de una base de V . Esta propiedad importante -de la cual no incluimos la demostración- se enuncia en el teorema a continuación.

Teorema 2.7.2 Sea V un espacio vectorial de dimensión finita n y sea V^* su espacio dual. Sea $B' = \{\phi_1, \phi_2, \dots, \phi_n\}$ una base de V^* . Existe una única base $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ de V que satisface $B^* = B'$.

Relación entre las coordenadas en las bases B y B^* Si B es una base de un espacio vectorial V de dimensión finita y

$$B^* = \{\phi_1, \phi_2, \dots, \phi_n\}$$

es su base dual, es posible calcular fácilmente las coordenadas de un elemento de V usando la base B^* como se realizó al final del Ejemplo 2.19. Y recíprocamente es posible hallar las coordenadas de un elemento de V^* utilizando la base B . Esto se muestra en el ejemplo que sigue:

■ **Ejemplo 2.20** Si $B = \{\vec{e}_1, \vec{e}_2\} = \{(1, 1), (1, -1)\}$ y

$$B^* = \{\phi_1, \phi_2\} = \left\{ \frac{x+y}{2}, \frac{x-y}{2} \right\}$$

La relación entre las coordenadas en las bases B y B^* es:

$$(5, 5) = \alpha(1, 1) + \beta(1, -1)$$

$$\phi_1(5, 5) = \alpha\phi_1(1, 1) + \beta\phi_1(1, -1) = \alpha$$

$$\phi_2(5, 5) = \alpha\phi_2(1, 1) + \beta\phi_2(1, -1) = \beta$$

ya que $\phi_i(\vec{e}_j) = \delta_{ij}$.

Por otro lado, dado un funcional $\varphi(x, y) \in V^*$, $\varphi(x, y) = \alpha^*\phi_1(x, y) + \beta^*\phi_2(x, y)$

Así, para

$$\varphi(x, y) = 3x + 5y = \alpha^*\left(\frac{x+y}{2}\right) + \beta^*\left(\frac{x-y}{2}\right)$$

sus coordenadas son, $\alpha^* = \varphi(1, 1)$ y $\beta^* = \varphi(1, -1)$

Entonces, la relación entre las coordenadas es la siguiente:

$$\alpha = \varphi_1(5, 5) = 5 \quad \beta = \varphi_2(5, 5) = 0$$

$$\alpha^* = \varphi(1, 1) = 8 \quad \beta^* = \varphi(1, -1) = -2$$

■

Se puede generalizar lo que vimos en el ejemplo anterior.

Sean $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ una base de V y $B^* = \{\varphi_1, \varphi_2, \dots, \varphi_n\}$ una base de V^* .

■ Dado $\vec{v} \in V$, $\vec{v} = \sum_{i=1}^n \alpha_i \vec{e}_i$, $\alpha_i \in K$

$$\varphi_j(\vec{v}) = \varphi_j\left(\sum_{i=1}^n \alpha_i \vec{e}_i\right) = \left(\sum_{i=1}^n \alpha_i \varphi_j(\vec{e}_i)\right) = \alpha_j. \quad (2.6)$$

Luego,

$$(\vec{v})_B = (\varphi_1(\vec{v}), \varphi_2(\vec{v}), \dots, \varphi_n(\vec{v}))$$


■ Dada $\varphi \in V^*$, $\exists \beta_i \in K$ tal que $\varphi = \sum_{i=1}^n \beta_i \varphi_i$

Para cada j , $1 \leq j \leq n$,

$$\varphi(\vec{e}_j) = \left(\sum_{i=1}^n \beta_i \varphi_i\right)(\vec{e}_j) = \sum_{i=1}^n \beta_i \varphi_i(\vec{e}_j) = \beta_j \quad (2.7)$$

Luego,

$$(\varphi)_{B^*} = (\varphi(\vec{e}_1), \varphi(\vec{e}_2), \dots, \varphi(\vec{e}_n))$$

 En el Ejemplo 2.19 las coordenadas de $p(t)$ son, de acuerdo a (2.6), $L_i(p(t)) = p(t_i)$, $i = 1, 2, 3$.

Anulador de un subespacio

Existe también relación entre los subespacios de V con ciertos subespacios de V^* . En particular, dado un subespacio S de V si consideramos el conjunto de todas los funcionales lineales que se anulan en S se prueba que tiene una estructura de subespacio.

Definición 2.7.1 Sea V un espacio vectorial y sea S un subespacio de V . Se llama *anulador* al conjunto

$$\begin{aligned} S^0 &= \{\varphi \in V^*, \varphi(\vec{s}) = 0 \forall \vec{s} \in S\} \\ &= \{\varphi \in V^*, S \subseteq N(\varphi)\} \end{aligned}$$

Proposición 2.7.3 S^0 es un subespacio de V^* .

Demostración:

- $\vec{0} \in S^0$ (se anula en todo S y $\vec{0} \in S$).
- Si φ_1 y $\varphi_2 \in S^0$, entonces $\varphi_1(\vec{s}) = 0$ y $\varphi_2(\vec{s}) = 0 \forall \vec{s} \in S$, de donde $(\varphi_1 + \varphi_2)(\vec{s}) = \varphi_1(\vec{s}) + \varphi_2(\vec{s}) = 0 \forall \vec{s} \in S$. Se tiene, que $\varphi_1 + \varphi_2 \in S^0$.
- Si $c \in K$ y $\varphi_1 \in S^0$, $(c\varphi_1)(\vec{s}) = c\varphi_1(\vec{s}) = c \cdot 0 = 0 \forall \vec{s} \in S$. Luego, $c\varphi_1 \in S^0$.

□

Proposición 2.7.4 Sea V un espacio vectorial de dimensión n y sea S un subespacio de V . Entonces

$$\dim(S) + \dim(S^0) = n$$

Demostración:

Sea $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ una base de S , y sean $\vec{v}_{k+1}, \vec{v}_{k+2}, \dots, \vec{v}_n \in V$ tales que $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k, \vec{v}_{k+1}, \dots, \vec{v}_n\}$ sea una base de V . Sea $B^* = \{\varphi_1, \varphi_2, \dots, \varphi_k, \varphi_{k+1}, \dots, \varphi_n\} \subset V^*$ la base dual de B . Entonces, para cada $k+1 \leq i \leq n$, se tiene que $\varphi_i(\vec{v}_1) = \varphi_i(\vec{v}_2) = \dots = \varphi_i(\vec{v}_k) = 0$, y por lo tanto φ_i se anula en todo S . Se tiene que $\{\varphi_{k+1}, \dots, \varphi_n\} \subseteq S^0$.

Como $\{\varphi_{k+1}, \dots, \varphi_n\}$ es parte de una base, es un conjunto linealmente independiente. Veamos que es un sistema de generadores de S^0 y entonces es base de S^0 .

Sea $\psi \in S^0$. Como B^* es una base de V^* , existen $c_1, c_2, \dots, c_n \in K$ tales que $\psi = \sum_{i=1}^n c_i \varphi_i$. Por la relación entre las coordenadas (2.7) se tiene que $c_i = \psi(\vec{v}_i)$. Además, como $\psi \in S^0$ y $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ una base de S , $\psi(\vec{v}_i) = 0$ para cada $1 \leq i \leq k$. En consecuencia, $c_i = 0$ para $1 \leq i \leq k$, y por lo tanto $\psi \in \langle \varphi_{k+1}, \dots, \varphi_n \rangle$.

Luego, $\{\varphi_{k+1}, \dots, \varphi_n\}$ es una base de S^0 y entonces,

$$\dim(S^0) = n - k = n - \dim(S).$$

□

■ **Ejemplo 2.21** Si se desea hallar el subespacio anulador S^0 para $S = \langle (1, 1, 1), (1, 2, 1) \rangle \subset \mathbb{R}^3$. De acuerdo a la demostración de la Proposición 2.7.4, $S^0 = \langle \varphi_3 \rangle$, se completa S para tener una base de \mathbb{R}^3 , $B = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \{(1, 1, 1), (1, 2, 1), (1, 0, 0)\}$ y luego, a partir de escribir (x, y, z) como combinación lineal de esa base y teniendo en cuenta que $\varphi_i(\vec{v}_j) = \delta_{ij}$, se obtiene que $\varphi_3(x, y, z) = x - z$.

■

■ **Ejemplo 2.22** De la misma forma, para hallar una base de S^0 , donde $S = \langle (-1, -1, 1) \rangle$:

- Se completa la base de S para tener una base de \mathbb{R}^3 , $B = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$, por ejemplo,
 $B = \{(-1, -1, 1), (1, 0, 0), (0, 0, 1)\}$.
- Se halla la base $B^* = \{\varphi_1, \varphi_2, \varphi_3\}$ tal que $\varphi_j(\vec{v}_i) = \delta_{ij}$, $i, j = 1, 2, 3$.
- La base de S^0 está dada por $\{\varphi_2, \varphi_3\} = \{x_1 - x_2, x_2 + x_3\}$.

■

El ejemplo que sigue muestra que los sistemas de ecuaciones lineales pueden estudiarse desde el punto de vista de los funcionales lineales.

■ **Ejemplo 2.23** Sea el sistema lineal homogéneo:

$$\begin{cases} x_1 + x_3 &= 0 \\ 2x_1 - x_2 + x_3 &= 0 \end{cases}$$

Sea S el subespacio de \mathbb{R}^3 generado por $\alpha_1 = (1, 0, 1)$, $\alpha_2 = (2, -1, 1)$. Entonces el espacio solución es el espacio anulador, S^0 . Es decir, $\varphi \in S^0 \Leftrightarrow \varphi(\alpha_i) = 0$, para $i = 1, 2$

■

Se verá a continuación cómo se comporta el anulador con la suma y la intersección de subespacios:

Proposición 2.7.5 Sea V un espacio vectorial de dimensión n y sean S y T subespacios de V . Entonces,

1. $(S + T)^0 = S^0 \cap T^0$ y

$$2. (S \cap T)^0 = S^0 + T^0$$

Demostración:

1. Sea $\varphi \in V^*$. Se tiene que $\varphi \in (S + T)^0$ sí y solo sí $\varphi(\vec{s} + \vec{t}) = 0$ para todo $\vec{s} \in S$ y para todo $\vec{t} \in T$. Y esto es equivalente a que $\varphi(\vec{s}) = 0$ para todo $\vec{s} \in S$ y $\varphi(\vec{t}) = 0$ para todo $\vec{t} \in T$, es decir que $\varphi \in S^0 \cap T^0$.
2. Sea $\varphi \in S^0 + T^0$. Entonces $\varphi = \varphi_S + \varphi_T$, $\varphi_S \in S^0$ y $\varphi_T \in T^0$. Para cada $\vec{x} \in S \cap T$ se tiene que $\varphi(\vec{x}) = \varphi_S(\vec{x}) + \varphi_T(\vec{x}) = 0 + 0 = 0$. Luego $\varphi \in (S \cap T)^0$. Por lo tanto $S^0 + T^0 \subseteq (S \cap T)^0$. Por el Teorema de la dimensión para la suma de subespacios (Proposición 1.4.1), teniendo en cuenta que $(S + T)^0 = S^0 \cap T^0$ y la Proposición 2.7.4,

$$\begin{aligned}
 \dim(S^0 + T^0) &= \dim(S^0) + \dim(T^0) - \dim(S^0 \cap T^0) \\
 &= \dim(S^0) + \dim(T^0) - \dim((S + T)^0) \\
 &= (n - \dim(S)) + (n - \dim(T)) - (n - \dim(S + T)) \\
 &= n - (\dim(S) + \dim(T) - \dim(S + T)) \\
 &= n - \dim(S \cap T) \\
 &= \dim((S \cap T)^0)
 \end{aligned}$$

En consecuencia, $(S \cap T)^0 = S^0 + T^0$.

□

El doble dual

Como $V^* = L(V, K)$ es un espacio vectorial, es posible definir su espacio dual. Entonces $(V^*)^* = L(V^*, K)$. Por las propiedades vistas antes, se tiene que $\dim(V) = \dim(V^*) = \dim(V^{**})$. Si V es de dimensión finita, sabemos que los espacios V y V^* , al tener la misma dimensión, son isomorfos. Para hallar el isomorfismo se debe hallar una base de V y su base dual. En forma similar, como la dimensión de V^* es igual a la dimensión de V^{**} , existe un isomorfismo entre los espacios V^* y V^{**} . En este segundo caso, el isomorfismo no requiere la elección de bases, y se dice que es natural.

Proposición 2.7.6 Sea V un espacio vectorial, $\dim(V) < \infty$ sobre K . Para cada vector $v \in V$, se define $L_v : V^* \rightarrow K$, $L_v(\phi) = \phi(\vec{v})$, $\phi \in V^*$.

La aplicación $\phi : V \rightarrow V^{**}$ que a \vec{v} hace corresponder $L_v(\phi)$ es un isomorfismo.

Notar que L_v se aplica a elementos de V^* y debe dar un escalar. La definición indica que, aplicado a ϕ ese escalar es $\phi(\vec{v})$, de ahí que el nombre de este isomorfismo es la *evaluación*, ya que para cada ϕ da su evaluación en \vec{v} , $\phi(\vec{v})$.

Demostración:

Se deja al lector probar que L_v es una transformación lineal. Como V y V^{**} tienen la misma dimensión, alcanza con probar que la transformación es inyectiva.

Sea $\vec{v} \in N(\phi)$. Entonces $L_v(\phi) = 0$ para todo ϕ en V^* , o sea $\phi(\vec{v}) = 0$ para todo ϕ en V^* . El único elemento de V con esta propiedad (da 0 en cualquier funcional V^*), es $\vec{0}$. Es decir que $N(\phi) = \vec{0}$. Así que, ϕ es inyectiva, y por lo tanto, es un isomorfismo.

□



Si V tiene dimensión infinita la aplicación $\phi : V \rightarrow V^{**}$ sigue siendo inyectiva, pero no suryectiva. Así que no todo espacio de dimensión infinita es isomorfo a su doble dual.

Aplicación transpuesta

Sean V y W dos espacios vectoriales de $\dim < \infty$ sobre K . Sea $T : V \rightarrow W$ una aplicación lineal. T induce una aplicación lineal de W^* en V^* que llamaremos *transpuesta* de T ,

Si $\phi : W \rightarrow K$,

$$T^t : W^* \rightarrow V^*$$

$$T^t(\phi) = \phi \circ T$$

$$T^t(\phi)(\vec{v}) = \phi(T(\vec{v}))$$

Se deja al lector probar que $T^t(\phi)$ es una transformación lineal.

Proposición 2.7.7 Sean V y W dos espacios vectoriales sobre el mismo cuerpo K . Sea T una aplicación lineal de V en W , entonces

$$N(T^t) = \text{Im}(T)^0$$

Se deja la prueba al lector.

Relación entre las matrices que representan a una aplicación lineal y su traspuesta

Sean $T : V \rightarrow W$ y $T^t : W^* \rightarrow V^*$. Y sean B y B' bases de V y W respectivamente, y B^* y B'^* bases de V^* y W^* , respectivamente.

Si $A = [T]_{B',B}$ y $C = [T^t]_{B^*,B'^*}$, entonces $C = A^t$

■ **Ejemplo 2.24** En este ejemplo se muestra la relación entre las matrices de la transformación T y la de su aplicación traspuesta T^t , en el caso $V = \mathbb{R}^2$, $W = \mathbb{R}^3$, y $T : V \rightarrow W$ definida por $T(x, y) = (2x - y, 3x, x - 2y)$.

Sean las bases de V y W ,

$B = \{(1, 2), (1, 3)\}$ y $B' = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$, respectivamente.

■ Como

$$T((1, 2)) = (0, 3, -3) = -3(1, 1, 1) + 6(1, 1, 0) - 3(1, 0, 0)$$

$$T((1, 3)) = (-1, 3, -5) = -5(1, 1, 1) + 8(1, 1, 0) - 4(1, 0, 0)$$

la matriz de T en las bases B y B' es $\begin{pmatrix} -3 & -5 \\ 6 & 8 \\ -3 & -4 \end{pmatrix}$

Veamos ahora la matriz correspondiente a la transformación traspuesta T^t , $T^t : W^* \rightarrow V^*$, definida como $T^t(\varphi)(\vec{v}) = \varphi(T(\vec{v}))$ y considerando las bases B'^* de W^* y B^* de V^* .

■ $B^* = \{\varphi_1, \varphi_2\}$

$(x_1, x_2) = \alpha(1, 2) + \beta(1, 3)$ de donde $\alpha + \beta = x_1$ y $2\alpha + 3\beta = x_2$.

Se resuelve el sistema y resulta

$$\varphi_1(x_1, x_2) = \alpha = 3x_1 - x_2$$

$$\varphi_2(x_1, x_2) = \beta = x_2 - 2x_1$$

$$\blacksquare B'^* = \{\varphi'_1, \varphi'_2, \varphi'_3\}$$

Para hallar esta base se hace lo mismo que para B^* , partiendo de

$$(x_1, x_2, x_3) = \alpha(1, 1, 1) + \beta(1, 1, 0) + \gamma(1, 0, 0)$$

se aplican φ'_i , $i = 1, 2, 3$, y se obtiene

$$\varphi'_1(x_1, x_2, x_3) = \alpha = x_3$$

$$\varphi'_2(x_1, x_2, x_3) = \beta = x_2 - x_3$$

$$\varphi'_3(x_1, x_2, x_3) = \gamma = x_1 - x_2$$

■ Para obtener la matriz de T^t se debe calcular

$$T^t(\varphi'_1)(x_1, x_2, x_3) = \varphi'_1(T(x_1, x_2, x_3)) = x_1 - 2x_2 = -3\varphi_1 - 5\varphi_2$$

$$T^t(\varphi'_2)(x_1, x_2, x_3) = \varphi'_2(T(x_1, x_2, x_3)) = 6\varphi_1 + 8\varphi_2$$

$$T^t(\varphi'_3)(x_1, x_2, x_3) = \varphi'_3(T(x_1, x_2, x_3)) = -3\varphi_1 - 4\varphi_2$$

La matriz de T^t en las bases B'^* de W^* y B^* de V^* es $\begin{pmatrix} -3 & 6 & -3 \\ -5 & 8 & -4 \end{pmatrix}$

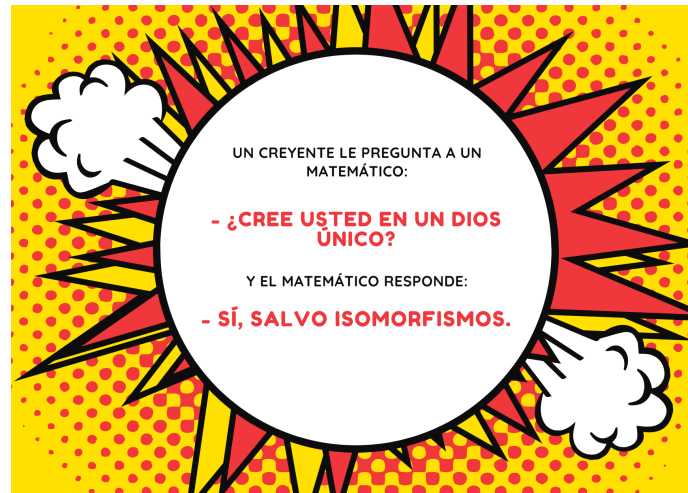
Se verifica que la matriz correspondiente a la transformación transpuesta, T^t es la matriz transpuesta de la matriz correspondiente a la transformación T .

■

Alicia Dickestein

Nació en Argentina en 1955. Es una matemática, investigadora y profesora argentina conocida por su trabajo en geometría algebraica, particularmente geometría teórica. En 1982, obtuvo el título de Doctora en Ciencias Matemáticas en la Facultad de Ciencias Exactas y Naturales de la Universidad de Buenos Aires, con una tesis sobre geometría analítica compleja. Fue la primera directora del Departamento de Matemática de la Facultad de Ciencias Exactas y Naturales de la UBA (en el período 1996-1998), donde desde 2009 pasó a ser profesora regular titular plenaria. Allí se desempeña como investigadora superior del CONICET, organismo en el que ingresó en 1985. En la actualidad Dickenstein es Investigadora Superior del Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina (CONICET). En los últimos años concentró su trabajo en las aplicaciones de la geometría algebraica en el ámbito de la biología molecular. En 2019 se incorporó como Académica Titular de la Academia Nacional de Ciencias Exactas, Físicas y Naturales (ANCEFN). A lo largo de su vida participó de una gran cantidad de congresos y reuniones científicas. Dirigió numerosas tesis de licenciatura, maestría y doctorado. Se desempeñó también como jurado de diversas tesis doctorales y de maestría tanto en su país como en el extranjero. Publicó varios libros, entre los cuales se encuentra *Mate max: la matemática en todas partes*, que presenta problemas matemáticos destinados a niñas y niños de los últimos años de educación básica. Fue merecedora del Premio TWAS (de la Academia Mundial de Ciencias) en el área de Matemática, en el año 2015, por su destacada contribución a la comprensión de discriminantes. En 2017 le fue entregado el Premio Consagración en Matemática de la Academia Nacional de Ciencias Exactas, Físicas y Naturales. En 2021 recibió el premio L'Oréal-UNESCO a Mujeres en Ciencia por su trayectoria en geometría algebraica. También en 2021 le fue otorgado el reconocimiento como "Personalidad destacada de la Universidad de Buenos Aires", durante los festejos por el Bicentenario de dicha universidad, recibiendo también una medalla personalizada, una moneda acuñada por la Casa de la Moneda y un sello postal del Correo Argentino. [3]

Figura 2.9



2.8 Actividades propuestas

Problema de Aplicación 2 Dada una variable aleatoria discreta X con distribución de probabilidad $P[X = x_i]$ con $i = 1, 2, \dots, n$, la esperanza (o valor medio) de X se define como:

$$E[X] = \sum_{i=1}^n x_i P[X = x_i]$$

Indique los espacios vectoriales involucrados en esta aplicación. Demuestre que la Esperanza $E[\cdot]$ es una transformación lineal, es decir que para cualquier par de variables aleatorias, X e Y y cualquier $c \in \mathbb{R}$ se cumple que:

$$E[X + Y] = E[X] + E[Y]$$

$$E[cX] = cE[X]$$

(Para una analogía física, si los x_i son masas puntuales en la recta, cada una con peso $P[X = x_i]$, entonces $E[X]$ es el centro de gravedad de estas masas.)

Busque otros ejemplos e investigue el caso de una variable aleatoria continua, cómo se define su esperanza en ese caso y si también se tiene una transformación lineal.

2.8.1 Ejercicios

Ejercicio 2.1

Utilice la definición de transformación lineal para justificar por qué: $T : \mathbb{R} \rightarrow \mathbb{R}$ tal que $T(\varepsilon) = \frac{1}{3}E\varepsilon^2$ no la cumple (E es una constante). Grafique.

Ejercicio 2.2

Sea A la matriz de la transformación lineal $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$. Encuentre una \vec{x} en \mathbb{R}^2 cuya imagen bajo T sea \vec{b} y responda si existe más de una \vec{x} cuya imagen bajo T sea \vec{b} .

$$A = \begin{pmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 3 \\ 2 \\ -5 \end{pmatrix}$$

Ejercicio 2.3

La transformación de trasquilado deforma un cuadrado como si este se empujara hacia la derecha manteniendo fija la base. Grafique el producto de multiplicar por A los vértices del cuadrado: $(0,0)$, $(0,2)$, $(2,0)$ y $(2,2)$.

$$A = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$$

Ejercicio 2.4

Encuentre la expresión de la transformación proyección $T(\vec{x})$ donde $\vec{x}=(x_1, x_2, x_3)$ y describa la aplicación. Su matriz es:

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

Ejercicio 2.5

Dado un escalar r , si $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es tal que $T(\vec{x}) = r\vec{x}$, identifique qué valores debe tomar r para que T sea una contracción y cuáles para que T sea una dilatación.

Dadas:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

$$E = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad F = \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}, \quad H = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

$$I = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}, \quad J = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad K = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Indique cuál de ellas es una contracción o expansión vertical y cuál una horizontal, cuál una proyección sobre el eje x_1 y cuál sobre el eje x_2 , cuál es un trasquillado vertical y cuál horizontal, cuál es una reflexión con respecto al origen, cuál a través del eje x_1 y cuál a través del eje x_2 , por último cuál representa una reflexión a través de la recta $x_1=x_2$ y cuál a través de la recta $x_1=-x_2$.

Ejercicio 2.6

Para las matrices del ejercicio anterior calcule su determinante. Con el resultado exprese de qué modo el valor del determinante determina el área que se crea entre dos vectores antes y después de la transformación. Comente cómo el signo del determinante está conectado con la *quiralidad* para estas transformaciones.

Hay objetos que vienen en dos variedades: derecha e izquierda. Por ejemplo, hay zapatos derechos y zapatos izquierdos, orejas derechas y orejas izquierdas. Estos objetos se llaman quirales (del griego, keirós= mano). Otros objetos no vienen en dos variedades derecha e izquierda, por ejemplo las pelotas de fútbol; 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). Es así que, los espejos cambian la quiralidad. El área puede pensarse con quiralidad, piense en el producto cruz, implica la regla de la mano derecha. Si el $\text{Det}(T) \geq 0$ no hay cambio de quiralidad, pero si $\text{Det}(T) \leq 0$ sí lo hay y la transformación tiene el carácter de un espejo.

Ejercicio 2.7

Sea $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definida por $T((x,y)) = (x+2y, -x, 0)$

a) Encuentre la matriz de la transformación lineal T respecto a las bases $B = \{\vec{u}_1, \vec{u}_2\}$ y $B' = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$, donde

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \vec{u}_2 = \begin{pmatrix} -2 \\ 4 \end{pmatrix}, \vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \vec{v}_3 = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}.$$

b) Verifique utilizando la matriz anterior que $T\left(\begin{pmatrix} 8 \\ 3 \end{pmatrix}\right) = \begin{pmatrix} 14 \\ -8 \\ 0 \end{pmatrix}$

Ejercicio 2.8

Una rotación de Givens es una transformación lineal de $\mathbb{R}^n \rightarrow \mathbb{R}^n$ que se utiliza para crear una entrada cero en un vector. Sería como generar la rotación en vez de cambiar el sistema de referencia.

Para $n = 2$ la rotación de Givens tiene la forma general:

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad a^2 + b^2 = 1$$

Encuentre a y b tales que $\begin{pmatrix} 4 \\ 3 \end{pmatrix}$ gire a $\begin{pmatrix} 5 \\ 0 \end{pmatrix}$

Ejercicio 2.9

Sea $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, dada por $T((x, y, z)) = (x + y, x + z)$. Determinar el núcleo y la imagen de T , y sus dimensiones. Caracterice el conjunto $T^{-1}(C)$, siendo $C = \{(x, y) : x = 1\}$.

Ejercicio 2.10

Sea $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, la multiplicación por la matriz $\begin{pmatrix} 2 & -1 \\ -8 & -4 \end{pmatrix}$

Indique cuáles vectores están en el núcleo de T y cuáles en la imagen de T :

$$\begin{pmatrix} 1 \\ -4 \end{pmatrix}, \begin{pmatrix} 5 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 12 \end{pmatrix}, \begin{pmatrix} 5 \\ 10 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Ejercicio 2.11

Sea $T : P_{\mathbb{R}}^{(2)}[x] \rightarrow P_{\mathbb{R}}^{(3)}[x]$ la transformación lineal definida por $T(p(x)) = xp(x)$.

- Indique cuáles polinomios de los siguientes están en el núcleo de T : x^2 , 0 , $1 + x$.
- Y cuáles polinomios están en la imagen de T : $x + x^2$, $1 + x$, $3 - x^2$

Ejercicio 2.12

Sea $T : P_{\mathbb{R}}^{(2)}[x] \rightarrow P_{\mathbb{R}}^{(1)}[x]$ la transformación lineal definida por $T(a_0 + a_1x + a_2x^2) = (a_0 + a_1) - (2a_1 + 3a_2)x$.

Halle la matriz de T con respecto a las bases canónicas de $P_{\mathbb{R}}^{(2)}[x]$ y $P_{\mathbb{R}}^{(1)}[x]$.

Ejercicio 2.13

En cada caso utilice la información que se da para hallar la nulidad de T .

- a) $T : \mathbb{R}^5 \rightarrow \mathbb{R}^7$ tiene rango 3.
- b) La imagen de $T : \mathbb{R}^6 \rightarrow \mathbb{R}^3$ es \mathbb{R}^3 .
- c) $T : \mathbb{R}^{2 \times 2} \rightarrow \mathbb{R}^{2 \times 2}$ tiene rango 3.

Ejercicio 2.14

Sea $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ la multiplicación de un vector de \mathbb{R}^3 por la matriz $\begin{pmatrix} 1 & 3 & 4 \\ 3 & 4 & 7 \\ -2 & 2 & 0 \end{pmatrix}$

- a) Demuestre que el núcleo de T es una recta por el origen y encuentre sus ecuaciones paramétricas.
- b) Demuestre que la imagen de T es un plano por el origen y encuentre su ecuación.

Ejercicio 2.15

Sea $D : P_{\mathbb{R}}^{(3)}[t] \rightarrow P_{\mathbb{R}}^{(2)}[t]$ la transformación derivación. Describa el núcleo de D , $N(D)$.

Ejercicio 2.16

Sea $I : P_{\mathbb{R}}^{(1)}[t] \rightarrow \mathbb{R}$ la transformación integración de p , $\int_{-1}^1 p(x)dx$. Describa el núcleo de I , $N(I)$.

Ejercicio 2.17

Explique de qué modo se obtiene la base dual de un espacio vectorial dado; utilice un ejemplo. Luego, de un ejemplo a la inversa, esto es, teniendo la base dual, encuentre la base del espacio vectorial.

Ejercicio 2.18

Responda:

- a) ¿Qué relación hay entre las dimensiones del espacio dual con su espacio vectorial V ?
- b) ¿A qué se llama espacio anulador?
- c) ¿Los elementos del espacio anulador pertenecen al dual?
- d) ¿A qué se llama aplicación transpuesta? De un ejemplo.
- e) ¿A qué se denomina doble dual?

Ejercicio 2.19

Dado un cuerpo K , sea $T : V \rightarrow V'$ una transformación lineal entre dos espacios vectoriales sobre el cuerpo K . Sean S y S' subespacios de V y V' respectivamente. Pruebe que $T(S)$, $T^{-1}(S')$ y $N(T)$ son subespacios vectoriales.

Ejercicio 2.20

Sea V un espacio vectorial de dimensión n sobre el cuerpo K y sea $T : V \rightarrow V$ una transformación lineal. Pruebe que T es inyectiva si y solo si $T^{-1}(\vec{0}) = \{\vec{0}\}$.

Ejercicio 2.21

Sea V un espacio vectorial sobre K de dimensión n y $X = \{\vec{v}_1, \dots, \vec{v}_m\}$ un conjunto finito de vectores de V . Considere la aplicación lineal $E : K^m \rightarrow V$ definida por $E(k_1, \dots, k_m) = k_1\vec{v}_1 + k_2\vec{v}_2 + \dots + k_m\vec{v}_m$.

- Pruebe que E es una transformación lineal.
- Demuestre que E es inyectiva si y sólo si los vectores del conjunto X son linealmente independientes.
- Pruebe que E es suryectiva si y sólo si X es un conjunto de generadores de V .

Ejercicio 2.22

Dado V un espacio vectorial sobre el cuerpo K de dimensión n , sea $\varphi \in V^*$. Pruebe que $\text{Im}(\varphi) = K$ y que $\dim(N(\varphi)) = n - 1$.

Ejercicio 2.23

Sea V un \mathbb{C} -espacio vectorial. Dados φ y ϕ funcionales lineales sobre V , suponga que la función ψ definida por $\psi(v) = \varphi(v) \cdot \phi(v)$ también es un funcional lineal sobre V . Demuestre que $\varphi = \mathbf{0}$ o $\phi = \mathbf{0}$.

Ejercicio 2.24

Sea V un K -espacio vectorial de dimensión finita (K cuerpo). Demuestre:

- Si A y B son subconjuntos de V tales que $A \subseteq B$, entonces $B^0 \subseteq A^0$.
- Dados S y T subespacios de V ,
 $(S + T)^0 = S^0 \cap T^0$ y $(S \cap T)^0 = S^0 + T^0$

Ejercicio 2.25

Sean V y W dos espacios vectoriales sobre K y sea $T : V \rightarrow W$ una transformación lineal. La *traspuesta* de T es la función $T^t : W^* \rightarrow V^*$ que aplica a un funcional $\varphi \in W^*$ en el funcional $T^t(\varphi) \in V^*$ definido por

$$(T^t(\varphi))(v) = \varphi(T(v)), \text{ para todo } v \in V$$

- Pruebe que T^t está bien definida y que es una aplicación lineal.
- Pruebe que $Nu(T^t) = (\text{Im}(T))^0$.

Ejercicio 2.26

Sea V un espacio vectorial sobre K y sea $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ una base ordenada de V , se define la transformación lineal $T : V \rightarrow V$ de la siguiente forma:

$$T(\vec{v}_i) = \begin{cases} \vec{v}_{i+1} & 1 \leq i \leq n-1 \\ 0 & i = n \end{cases}$$

- Encuentre la matriz correspondiente a T en la base B .
- Pruebe que $T^n = \mathbf{0}$, pero $T^{n-1} \neq \mathbf{0}$.
- Sea S cualquier transformación lineal sobre V tal que $S^n = 0$, pero $S^{n-1} \neq \mathbf{0}$. Demuestre que existe una base ordenada B' de V tal que la matriz de S en la base B' coincide con la matriz hallada en el primer inciso.
- Demuestre que si M y N son matrices de $n \times n$ sobre K tales que $M^n = N^n = \mathbf{0}$, con $M^{n-1} \neq \mathbf{0}$ y $N^{n-1} \neq \mathbf{0}$, entonces M y N son semejantes.

2.8.2 Autoevaluación**Verdadero o Falso**

- El rango de T es el conjunto de todas las combinaciones lineales de las columnas de A .
- La generalización $T(c_1\vec{v}_1 + \dots + c_p\vec{v}_p) = c_1T(\vec{v}_1) + \dots + c_pT(\vec{v}_p)$ es lo que se conoce como principio de superposición en física.
- El conocimiento de $T(\vec{e}_1)$ y $T(\vec{e}_2)$ siendo \vec{e}_1 y \vec{e}_2 los vectores canónicos, no basta para encontrar T sabiendo que T es lineal.
- $A = [T(\vec{e}_1) \dots T(\vec{e}_n)]$ se llama matriz estándar de T .
- T es suryectiva si y sólo si las filas de A generan la imagen de T .
- T es inyectiva si y sólo si las columnas de A son linealmente independientes y la ecuación $T(\vec{x}) = \vec{0}$ tiene únicamente la solución trivial.
- Para que una transformación sea isomorfa el dominio y la imagen deben coincidir.
- A es invertible si la dimensión del núcleo de A es cero.
- Si los vectores en el dominio generan un área y después que se les aplica una transformación lineal siguen generando un área, cuánto aumente o disminuya el área dependerá del determinante de la matriz A de la transformación lineal.
- Una transformación lineal siempre lleva el vector nulo del dominio al vector nulo de la imagen.
- Existe un isomorfismo $T : P_{\mathbb{R}}^{(3)}[x] \rightarrow \mathbb{R}^4$.



3. Autovalores y autovectores

En este capítulo se tratará sobre autovectores y autovalores que es una herramienta matemática muy útil a la hora de resolver diversos problemas. Nos ocuparemos de resolver el problema $A\vec{v} = \lambda\vec{v}$, donde A es una matriz cuadrada. G. Strang [27] lo llama el *segundo problema* de Álgebra lineal, considerando que el *primer problema* es resolver $Ax = b$. Notar que $A\vec{v} = \lambda\vec{v}$ es una ecuación no lineal, ya que λ multiplica a \vec{v} y ambos λ y \vec{v} son desconocidos. El método de eliminación gaussiana, adecuado para el problema $Ax = b$, no es una herramienta útil, ya que las operaciones elementales sobre las filas de una matriz pueden modificar a los autovalores λ . El problema se resuelve a partir de simplificar la matriz, y eso es haciéndola lo más diagonal posible. A partir del cálculo de un determinante se obtiene un polinomio cuyas raíces son los autovalores. La obtención de una forma casi diagonal de la matriz A tiene muchas aplicaciones, entre ellas el cálculo de las potencias de una matriz y la resolución de sistemas de ecuaciones diferenciales.

3.1 Introducción

Presentamos a modo de introducción un modelo lineal que representa la dinámica de la infección y de la propagación de una epidemia. En este modelo, la enfermedad se introduce en una población, y en cada día se cuenta la fracción de la población que se encuentra dividida en cuatro estados o compartimentos:

- Susceptibles: son los individuos que pueden adquirir la enfermedad al día siguiente.
- Infectados: son los individuos con la enfermedad.

- Recuperados: son los individuos que tuvieron la enfermedad y se recuperaron. Ahora tienen inmunidad.
- Fallecidos: son los individuos que tuvieron la enfermedad y fallecieron a causa de ella.

Son llamados *modelos compartimentales*. A este, en particular, se lo conoce como modelo SIRD (Susceptible, Infectado, Recuperado, Fallecido) y las variables que indican la cantidad de individuos en cada compartimento al día st son X_t^1 , X_t^2 , X_t^3 y X_t^4 . En este caso, conocidos los valores al día t , se supone que al día siguiente $t + 1$:

- El 6 % de la población de individuos Susceptibles adquirirá la enfermedad (el 94 % restante sigue siendo Susceptible)
- El 1 % de la población infectada morirá a causa de la enfermedad, el 16 % se recuperará y adquirirá inmunidad, y el 3 % se recuperará y no adquirirá inmunidad y por lo tanto pasará a ser Susceptible. El 80 % restante seguirá Infectado.
- Los individuos Recuperados con inmunidad y los Fallecidos permanecen esos estados

Si X_t^1 es la proporción de individuos Susceptibles al día t , al día siguiente, X_{t+1}^1 está dada por los Susceptibles de hoy que no se infectaron, $0.94 * X_t^1$, más los infectados que se recuperaron sin inmunidad $0.03 * X_t^2$. La proporción de Infectados, X_{t+1}^2 estará dada por los Susceptibles que adquieren la enfermedad $0.06 * X_t^1$, más los infectados que siguen infectados $0.80 * X_t^2$. De igual forma, la cantidad de Recuperados, X_{t+1}^3 estará dada por los infectados que se recuperen con inmunidad $0.16 * X_t^2$ más los que permanecen Recuperados, X_t^3 . Y los Fallecidos, X_{t+1}^4 comprenden el 1 % de los Infectados X_t^2 más los Fallecidos, X_t^4 .

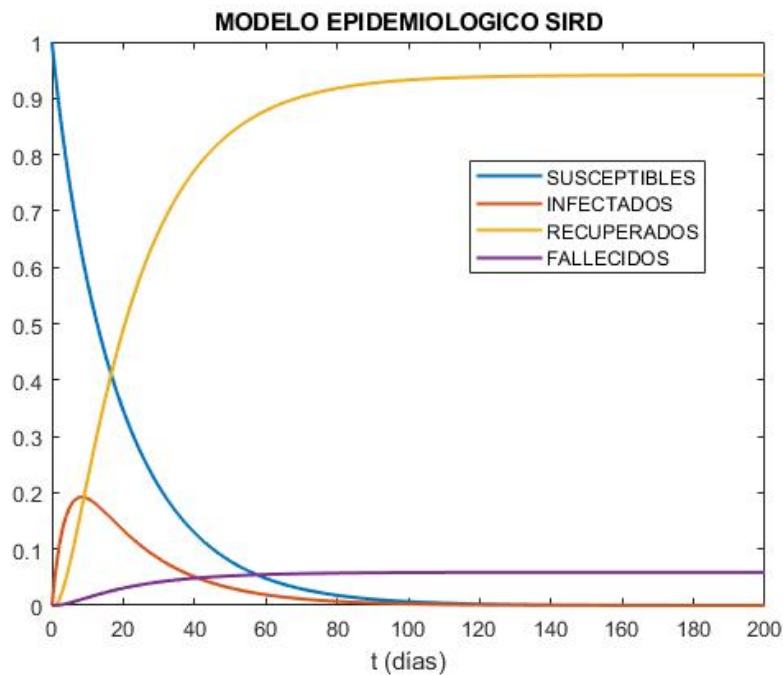
Es decir que la modelización tiene la expresión

$$X_{t+1} = AX_t$$

donde

$$A = \begin{pmatrix} 0.94 & 0.03 & 0 & 0 \\ 0.06 & 0.80 & 0 & 0 \\ 0 & 0.16 & 1 & 0 \\ 0 & 0.01 & 0 & 1 \end{pmatrix}$$

Si se desea saber la evolución de la cantidad de individuos en cada estado después de 6 meses, se debe calcular ¡ A^{180} !




Si A fuera una matriz diagonal esto llevaría un bajo costo computacional, ya que quedarían elevados al exponente 180 los elementos de la diagonal.

El programa en lenguaje de programación GNU Octave que sigue diagonaliza la matriz A del modelo.

```

1  % MODELO SIRD (SUSCEPTIBLES, INFECTADOS, RECUPERADOS, FALLECIDOS)
2  clear all
3  close all
4  A=[.94 .03 0 0 ; 0.06 .8 0 0 ; 0 0.16 1 0 ; 0 0.01 0 1];
5  X=[1 0 0 0 ]';
6  t=0:200;
7  for k= 1:200
8      X(:,k+1)=A*X(:,k);
9  end
10 for k=1:4
11 %plot(t,X(k,:), 'linewidth ',1.5)
12 hold on
13 axis tight
14 end
15 XF=A^180
16 [U,D]=eig(A);
17 XFD=U*D^180*inv(U)

```

 Este ejemplo estuvo inspirado en la pandemia del COVID-19 del año 2020 y en estudios realizados en el tema, como el del trabajo [19]. Los modelos matemáticos de epidemia que se utilizaron para su análisis y para hacer predicciones son modelos de última generación, representados mediante un sistema de ecuaciones diferenciales con muchas variables y poblaciones y que, entre otros parámetros, contemplan los que miden el comportamiento social.

En la Sección 2.6 vimos, en el Ejemplo 2.17, que dada una aplicación lineal T cuya matriz en la base canónica es

$$T = \begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix}$$

es posible hallar una base tal que la matriz sea diagonal.

$$T' = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

La matriz T' en esa nueva base es mucho más sencilla que la matriz T y se tiene que $T' = C^{-1}TC$ donde C es la matriz de cambio de base. Nos surge la pregunta si esto siempre es posible.

Un ejemplo sencillo que nos permite responder que no siempre es posible es la matriz

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

Supongamos que existe una matriz C de cambio de base tal que

$$C^{-1}TC = T' = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$$

o en forma equivalente,

$$T = C \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} C^{-1}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \frac{1}{\text{Det}(C)}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a\alpha & b\beta \\ c\alpha & d\beta \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \frac{1}{\text{Det}(C)}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a\alpha d - b\beta c & -a\alpha b + b\beta a \\ c\alpha d - cd\beta & -c\alpha b + d\beta a \end{pmatrix} \frac{1}{\text{Det}(C)}$$

Igualando los elementos de ambas matrices, se tiene un sistema de ecuaciones:

$$\text{Det}(C) = a\alpha d - b\beta c \quad (3.1)$$

$$\text{Det}(C) = ba(-\alpha + \beta) \quad (3.2)$$

$$0 = cd(-\alpha + \beta) \quad (3.3)$$

$$\text{Det}(C) = -c\alpha b + d\beta a \quad (3.4)$$

De la igualdad (3.3) se tiene que $c = 0$ o $d = 0$ o $\alpha = \beta$. Si $c = 0$, de (3.1) y (3.4), queda $\text{Det}(C) = a\alpha d = d\beta a$, de donde $\alpha = \beta$ y en (3.2) se tiene $1 = 0$. Si $d = 0$, de (3.1) y 3.4, $\text{Det}(C) = -b\beta c = -c\alpha b$, de donde también resulta $\alpha = \beta$ y en (3.2) se tiene $1 = 0$. Lo mismo si $\alpha = \beta$.

Se llega a una contradicción. Concluimos que no siempre es posible diagonalizar una matriz.



- No todas las matrices son diagonalizables.
- Si es digonalizable, A es semejante a una matriz diagonal y tendremos ventaja al calcular su potencia. En otros casos será semejante a una matriz casi diagonal.
- Siempre es posible encontrar una forma *más sencilla* de una matriz dada mediante un cambio de base. Se denomina *matriz de Jordan de la matriz dada* y el nombre se debe al matemático Camille Jordan (1838-1922).

Si se desea calcular

$$T^6 = \begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix}^6$$

como $T = CT'C^{-1}$, $T^2 = CT'C^{-1}CT'C^{-1} = CT'IT'C^{-1} = C(T')^2C^{-1}$, y en general,

$$T^n = C(T')^nC^{-1}, \quad n \in \mathbb{Z}$$

de donde se tiene que

$$T^6 = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 2^6 & 0 \\ 0 & 3^6 \end{pmatrix} \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$$

3.2 Subespacios invariantes. Valores y vectores propios

Dado un espacio vectorial V y una aplicación lineal $T : V \rightarrow V$, es decir $T \in L(V)$, un subespacio vectorial W de V se dice *invariante* respecto a T si $T(W) \subset W$, es decir, si la imagen $T(\vec{x})$ de todo vector $\vec{x} \in W$ es un elemento de W .

■ **Ejemplo 3.1** Sea $T \in L(\mathbb{R}^2)$ una aplicación lineal en \mathbb{R}^2 cuya matriz respecto de la base canónica $\{\vec{e}_1, \vec{e}_2\}$ de \mathbb{R}^2 está dada por

$$T = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

Entonces $W_1 = \{x_1\vec{e}_1, x_1 \in \mathbb{R}\}$ y $W_2 = \{x_2\vec{e}_2, x_2 \in \mathbb{R}\}$ son invariantes respecto de T .

En efecto, de la definición de matriz de una transformación lineal se tiene que $T(\vec{e}_1) = 2\vec{e}_1 + 0\vec{e}_2$ y $T(\vec{e}_2) = 0\vec{e}_1 + 1\vec{e}_2$.

Luego,

$$T(x_1\vec{e}_1) = x_1T(\vec{e}_1) = x_1(2\vec{e}_1) = (2x_1)\vec{e}_1 \in W_1$$

y

$$T(x_2\vec{e}_2) = x_2T(\vec{e}_2) = x_2\vec{e}_2 \in W_2$$

■

■ **Ejemplo 3.2** Sea R_α una rotación de ángulo $\alpha \neq 0$ en \mathbb{R}^3 con respecto al eje z . Geométricamente se observa que el plano xy y el eje z son invariantes con respecto a esta aplicación. Para comprobar algebraicamente que el plano xy es invariante se observa, en primer lugar, que la matriz de R_α con respecto a la base canónica de \mathbb{R}^3 es

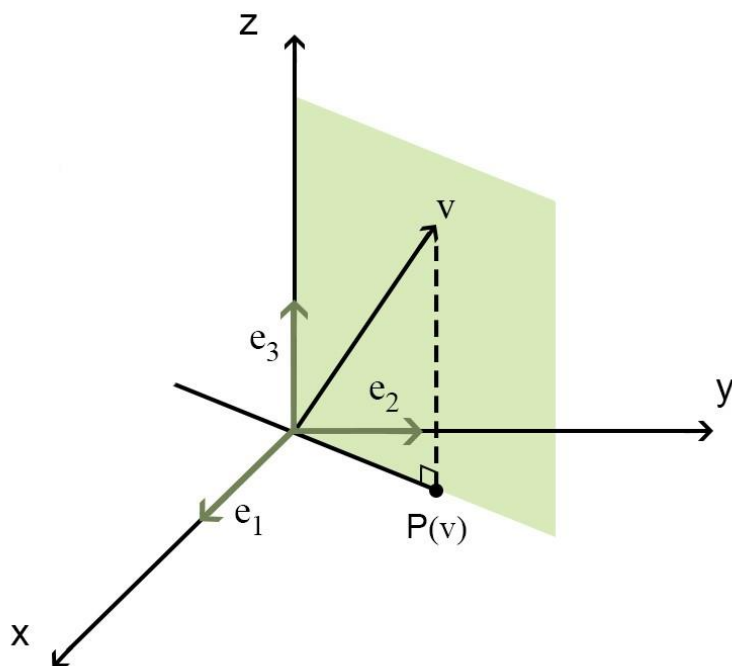
$$R = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Si $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2$ es un elemento del plano xy , se tiene que su rotación da el vector

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1\cos\alpha - x_2\sin\alpha \\ x_1\sin\alpha + x_2\cos\alpha \\ 0 \end{pmatrix}$$

y, por lo tanto, $R_\alpha(\vec{x}) = (x_1\cos\alpha - x_2\sin\alpha)\vec{e}_1 + (x_1\sin\alpha + x_2\cos\alpha)\vec{e}_2$ y es nuevamente un elemento del plano xy . ■

Figura 3.1: Todo plano que contiene al eje z es invariante por P



■ **Ejemplo 3.3** En la Figura 3.1 se muestra la proyección ortogonal P de \mathbb{R}^3 sobre el plano xy . (Ver Ejemplo 2.5). La matriz de P es

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

con respecto a la base canónica. Se puede ver que todo plano π que contiene al eje z es invariante: el plano π tiene ecuación $x_1x + x_2y + x_3z = 0$, donde $x_3 = 0$ ya que $(0, 0, 1) \in \pi$. Los vectores de ese plano son de la forma $\vec{x} = x_1\vec{e}_1 + \lambda x_1\vec{e}_2 + x_3\vec{e}_3$, $\lambda \in \mathbb{R}$, y se tiene que su imagen, $P(\vec{x})$, es de nuevo un elemento del plano π :

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \lambda x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ \lambda x_1 \\ 0 \end{pmatrix}$$

Otros subespacios invariantes de esta proyección ortogonal son el plano xy , el eje z y cualquier recta del plano xy que pase por el origen de coordenadas. ■

i Para cualquier transformación lineal $T \in L(V)$ (o endomorfismo), el subespacio $S = \{\vec{0}\}$, formado sólo por el elemento nulo, es invariante ya que $T(\vec{0}) = \vec{0}$ y el propio espacio vectorial V es también invariante ya que $T(\vec{x})$ para todo vector $\vec{x} \in V$ es un elemento de V .

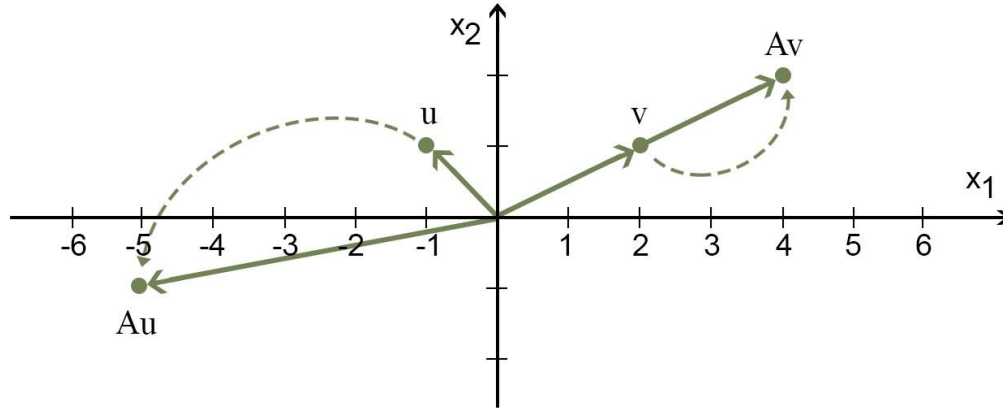
Proposición 3.2.1 La intersección y la suma de subespacios invariantes respecto de una aplicación lineal $T \in L(V)$ son subespacios invariantes respecto de T .

Se deja la demostración al lector.

Definición 3.2.1 Un vector $\vec{v} \neq \vec{0}$ de un espacio vectorial V sobre K se llama *autovector* o *vector propio* de una aplicación lineal $T \in L(V)$ si existe un escalar $\lambda \in K$ tal que $T(\vec{v}) = \lambda\vec{v}$. Este número λ se denomina *autovalor* o *valor propio* de la aplicación T correspondiente al vector \vec{v} .

i Si \vec{v} es un vector propio de T con autovalor λ , todo elemento no nulo del subespacio unidimensional generado por \vec{v} es un autovector de T con el mismo autovalor λ . Esto es porque $T(c\vec{v}) = cT(\vec{v}) = c\lambda\vec{v} = \lambda(c\vec{v})$.

Figura 3.2: El vector \vec{v} es autovector de A , mientras que el vector \vec{u} no es autovector, $A\vec{u}$ no es múltiplo de \vec{u}



Proposición 3.2.2 Una aplicación lineal $T \in L(V)$ es diagonalizable si y sólo si existe una base de V formada por vectores propios.

Demostración:

Supongamos que una aplicación lineal T en un espacio V de dimensión n tiene n vectores propios linealmente independientes, $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ con valores propios $\lambda_1, \lambda_2, \dots, \lambda_n$ respectivamente, tomando $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ como una base de V se tiene que

$$T(\vec{e}_1) = \lambda_1 \vec{e}_1, T(\vec{e}_2) = \lambda_2 \vec{e}_2, \dots, T(\vec{e}_n) = \lambda_n \vec{e}_n$$

y, por lo tanto, la matriz de T con respecto a esta base es la matriz diagonal

$$T = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \lambda_n \end{pmatrix}$$

Recíprocamente, toda aplicación lineal que tiene una matriz diagonal en una cierta base, tiene a los elementos de esta base como vectores propios.



i Una aplicación lineal $T \in L(V)$ es *diagonalizable* sí y solo sí existe una base de V en la cual la matriz de T es diagonal.

Definición 3.2.2 Una matriz $T \in K^{n \times n}$ se dice diagonalizable en K si la aplicación lineal $T : K^n \rightarrow K^n$ que la matriz que la representa es diagonalizable.

De esta definición se deduce que una matriz es diagonalizable en K si existe una matriz $C \in K^{n \times n}$ con determinante no nulo, tal que $T' = C^{-1}TC$ es una matriz diagonal.

En el ejemplo del inicio de la sección, para

$$T = \begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix}$$

se tiene que

$$\begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \end{pmatrix} = 3 \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

Se tiene que $T(\vec{v}_1) = 2\vec{v}_1$ y $T(\vec{v}_2) = 3\vec{v}_2$, siendo $\vec{v}_1 = (1, 2)$ y $\vec{v}_2 = (2, 3)$, con lo que \vec{v}_1 y \vec{v}_2 son vectores propios de T con sus correspondientes valores propios 2 y 3 respectivamente. Como \vec{v}_1 y \vec{v}_2 forman una base de \mathbb{R}^2 (son linealmente independientes y son 2), la matriz T es diagonalizable en \mathbb{R} y su matriz diagonal asociada es

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

Cálculo de autovalores y autovectores de una transformación lineal.

Supongamos que \vec{v} es un vector propio de una aplicación lineal T en un espacio vectorial V y que λ es su autovalor, es decir $T(\vec{v}) = \lambda\vec{v}$. Sea $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ una base de V y sean $v_j, j = 1, \dots, n$

las coordenadas de \vec{v} en esa base, es decir, $\vec{v} = \sum_{j=1}^n v_j \vec{e}_j$. Si (a_{ij}) es la matriz de T con respecto a la base tenemos que

$$\begin{aligned} \sum_{j=1}^n \lambda v_j \vec{e}_j &= \lambda \sum_{j=1}^n v_j \vec{e}_j = \lambda \vec{v} = T(\vec{v}) = T\left(\sum_{j=1}^n v_j \vec{e}_j\right) \\ &= \sum_{j=1}^n v_j T(\vec{e}_j) = \sum_{j=1}^n v_j \left(\sum_{i=1}^n a_{ij} \vec{e}_i\right) = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} v_j\right) \vec{e}_i \end{aligned}$$

Como $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ es una base de V , del primer y del último término de la igualdad anterior se tiene que:

$$\begin{cases} \lambda v_1 &= a_{11}v_1 + a_{12}v_2 + \dots + a_{1n}v_n \\ \lambda v_2 &= a_{21}v_1 + a_{22}v_2 + \dots + a_{2n}v_n \\ \dots & \\ \lambda v_n &= a_{n1}v_1 + a_{n2}v_2 + \dots + a_{nn}v_n \end{cases} \quad (3.5)$$

o en forma equivalente,

$$\begin{cases} (a_{11} - \lambda)v_1 + a_{12}v_2 + \dots + a_{1n}v_n &= 0 \\ a_{21}v_1 + (a_{22} - \lambda)v_2 + \dots + a_{2n}v_n &= 0 \\ \dots & \\ a_{n1}v_1 + a_{n2}v_2 + \dots + (a_{nn} - \lambda)v_n &= 0 \end{cases} \quad (3.6)$$

Como es un sistema homogéneo, para que exista una solución no nula debe ocurrir que

$$\text{Det}(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & \dots & \dots & a_{2n} \\ \dots & \dots & \ddots & \dots & \dots & \dots \\ \dots & \dots & \dots & \ddots & \dots & \dots \\ \dots & \dots & \dots & \dots & \ddots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn-1} & a_{nn} - \lambda \end{vmatrix} = 0 \quad (3.7)$$

donde I denota la matriz identidad. Esto es una ecuación de grado n en λ y sus soluciones en K (\mathbb{R} o \mathbb{C}) son los autovalores de T . Si V es un espacio vectorial complejo, por el teorema fundamental del álgebra la ecuación anterior tiene n soluciones complejas contando cada una con su multiplicidad. Si V es un espacio vectorial real, no podemos asegurar que la ecuación anterior tenga n soluciones reales.

Pasos para resolver $A\vec{v} = \lambda\vec{v}$

1. Se calcula $\text{Det}(A - \lambda I)$ (se anota en forma equivalente como $|A - \lambda I|$), restando λ de los elementos de la diagonal de la matriz A . Es un polinomio de grado n , con coeficiente $(-\lambda)^n$.
2. Se hallan las raíces de este polinomio. Las n raíces son los autovalores de la matriz A .
3. Para cada autovalor λ , se resuelve el sistema lineal $(A - \lambda I)\vec{v} = \vec{0}$. Como el determinante es cero, tendrá soluciones no nulas. Esos son los autovectores.

■ **Ejemplo 3.4** Se desea determinar los valores y vectores propios de la aplicación lineal de $T : V \rightarrow V$, $V = \mathbb{R}^2$, que tiene como matriz,

$$T = \begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}$$

1.

$$0 = |T - \lambda I| = \begin{vmatrix} 1-\lambda & 2 \\ 5 & 4-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda) - 10 = \lambda^2 - 5\lambda - 6$$

2. Las raíces son $\lambda_1 = 6$ y $\lambda_2 = -1$

3. Para $\lambda_1 = 6$ se resuelve el sistema

$$(T - 6I) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Los vectores propios correspondientes a $\lambda_1 = 6$ son de la forma $\alpha \begin{pmatrix} 2 \\ 5 \end{pmatrix}$.

Para $\lambda_2 = -1$ se resuelve el sistema

$$(T - (-1)I) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Los vectores propios correspondientes a $\lambda_2 = 6$ son de la forma $\beta \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Como $\begin{pmatrix} 2 \\ 5 \end{pmatrix}$ y $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ forman una base de \mathbb{R}^2 , por la Proposición 3.2.2, T es diagonalizable,

con matriz diagonal

$$\begin{pmatrix} 6 & 0 \\ 0 & -1 \end{pmatrix}$$

y la matriz de cambio de base (de la base de autovectores a la base canónica) está dada por la matriz

$$C = \begin{pmatrix} 2 & 1 \\ 5 & -1 \end{pmatrix}$$

■

William Gilbert Strang (1934)

Es un matemático estadounidense, actualmente Professor Mathworks de Matemáticas del Department of Mathematics del Massachusetts Institute of Technology (MIT). Ha contribuido a la teoría de elementos finitos, al cálculo de variaciones, al análisis wavelet y al álgebra lineal. Ha contribuido enormemente a la educación en matemáticas, en forma de libros técnicos y cursos online. En MIT enseña Álgebra Lineal, Ciencia Computacional e Ingeniería, Aprendiendo de los Datos. Sus clases están disponibles en la plataforma MIT OpenCourseWare (en inglés). Gilbert Strang nació en Chicago, Illinois. Cursó estudios en el propio MIT y en el Balliol College, en la Universidad de Oxford. Se doctoró en la Universidad de California, Los Ángeles (UCLA) y desde ese momento ha llevado a cabo su actividad docente en el MIT. Entre las publicaciones más notables del Professor Strang se destaca *An Analysis of the Finite Element Method*, conjuntamente con

George Fix, así como seis manuales: Introduction to Linear Algebra (1993, 1998, 2003), Linear Algebra and Its Applications (1976, 1980, 1988, 2005), Introduction to Applied Mathematics (1986), Calculus (1991), Wavelets and Filter Banks, con Truong Nguyen (1996), Linear Algebra, Geodesy, and GPS, con Kai Borre (1997).

Gilbert Strang fue Presidente de la SIAM (Society for Industrial and Applied Mathematics) durante los años 1999-2000. También ha sido Chairman of the US National Committee on Mathematics durante los años 2003-2004. Es Honorary Fellow, en el Balliol College de Oxford. También es Chairman, en la National Science Foundation (NSF) del Advisory Panel del área de Matemáticas.

Fue pionero al abrir sus clases y permitir que fueran grabadas en vídeo mientras explicaba matemáticas a sus alumnos del MIT para su difusión abierta y gratuita en Internet. [10]

3.2.1 Localización de autovalores

Teorema 3.2.3 Teorema de Gershgorin.

Los autovalores de una matriz A están en la unión de los discos D_1, D_2, \dots, D_n (del plano complejo) donde D_i es el disco centrado en el elemento de la diagonal a_{ii} :

$$|\lambda - a_{ii}| \leq r_i$$

Su radio $r_i = \sum_{j \neq i} |a_{ij}|$ es igual a la suma de los valores absolutos de los elementos del resto de la fila.

Demostración:

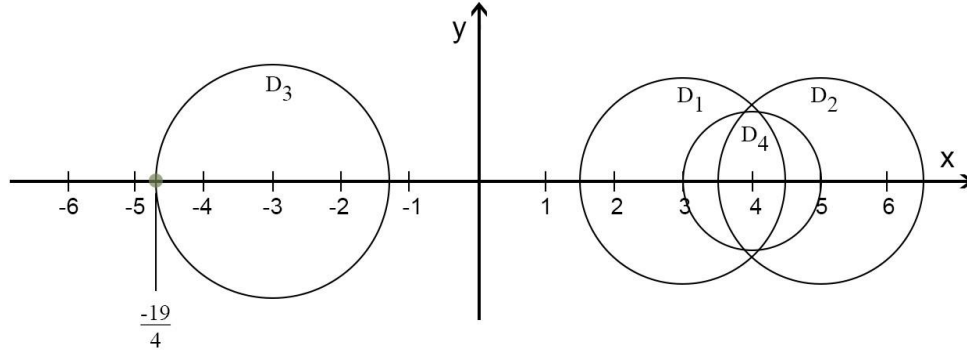
Supongamos v_i es la mayor componente en valor absoluto del autovector \vec{v} , $A\vec{v} = \lambda\vec{v}$. Entonces,

$$(\lambda - a_{ii})v_i = \sum_{j \neq i} a_{ij}v_j, \text{ de donde,}$$

$$|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \frac{|v_j|}{|v_i|} \leq \sum_{j \neq i} |a_{ij}| = r_i$$

□

Figura 3.3: Los autovalores de A se encuentran en la unión de los cuatro discos, D_1, D_2, D_3 y D_4



■ **Ejemplo 3.5** Dada la matriz,

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

Se muestra en la Figura 3.3 la localización de sus autovalores de acuerdo al Teorema 3.2.3. Los discos están centrados en los elementos de la diagonal de la matriz A y tienen radios $r_1 = 3/2$, $r_2 = 3/2$, $r_3 = 7/4$ y $r_4 = 1$. Están en el intervalo $[-19/4, 7]$, en el caso que sean números reales.

■

Para la matriz del Ejemplo 3.4 se tienen los discos $D_1 = |\lambda - 1| \leq 2$ y $D_2 = |\lambda - 4| \leq 5$. Sin calcular los autovalores, se sabe que si son números reales, estarán en el intervalo $[-1, 9]$.

3.3 Polinomio característico

Al polinomio $\text{Det}(T - \lambda I) \in P_K^{(n)}[\lambda]$ (Ec. (3.7)) se lo denomina *polinomio característico* de la aplicación T (o de la matriz $A \in K^{n \times n}$). Lo anotaremos $\mathbf{P}_T(\lambda)$.

Proposición 3.3.1 Sea $T \in K^{n \times n}$ y sea $\lambda \in K$. Entonces λ es autovalor de T sí y sólo sí λ es raíz del polinomio característico de T .

■ **Ejemplo 3.6** La matriz

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

es diagonalizable en $\mathbb{C}^{n \times n}$, pero no es diagonalizable en $\mathbb{Q}^{n \times n}$ ni en $\mathbb{R}^{n \times n}$. Sus autovalores son las raíces del polinomio

$$\text{Det} \begin{pmatrix} 0 - \lambda & 1 \\ -1 & 0 - \lambda \end{pmatrix} = \lambda^2 + 1 = 0$$

Se tiene, entonces, que los autovalores son $\lambda_{1-2} = \mp i$ ■

Proposición 3.3.2 El polinomio característico no depende de la base elegida en V para representar la aplicación lineal T .

Demostración:

Para demostrar esto, sea $\mathbf{P}_{T,B}(\lambda) = \text{Det}(T - \lambda I)$ el polinomio característico de la aplicación T en la base $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ como una base de V y sea $\mathbf{P}_{T,B'}(\lambda) = \text{Det}(T' - \lambda I)$ el polinomio característico de T en la base $B' = \{\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_m\}$; si C es la matriz del cambio de base, se sabe que $T = CT'C^{-1}$, y se tiene

$$\mathbf{P}_{T,B}(\lambda) = \mathbf{P}_{T,B'}(\lambda)$$

Se deja al lector completar esta demostración. □

■ **Ejemplo 3.7** La matriz de la proyección sobre el eje x , P_x , (Ver Ejemplo 2.15) en la base canónica $B = \{\vec{e}_1, \vec{e}_2\}$ es

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

Los autovalores son $\lambda_1 = 1$ y $\lambda_2 = 0$, ya que para los vectores \vec{v}_1 que están sobre el eje x se verifica $P_x(\vec{v}_1) = 1\vec{v}_1$, mientras que para los los vectores \vec{v}_2 que están sobre el eje y se verifica $P_x(\vec{v}_2) = \vec{0}$.

Si cambio de base, por ejemplo a la base $B' = \{\vec{e}'_1, \vec{e}'_2\}$ (rotando en $\phi = \pi/4$ los vectores de la base canónica B), donde los vectores de B' son las columnas de la matriz A del Ejemplo 1.28, se tendrá que

$$P_x(\vec{e}'_1) = (\sqrt{2}/2, 0) = 1/2\vec{e}'_1 - 1/2\vec{e}'_2$$

y

$$P_x(\vec{e}'_2) = (-\sqrt{2}/2, 0) = -1/2\vec{e}'_1 + 1/2\vec{e}'_2$$

entonces la matriz en esta nueva base es

$$\begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}$$

Ahora el polinomio característico es $\mathbf{P}_{B'} = (\lambda)(1/2 - \lambda)^2 - 1/4$, con las mismas raíces, $\lambda_1 = 1$ y $\lambda_2 = 0$, y los mismos autovectores que se obtuvieron con la base B , ya que

$$\vec{v}'_1 = (1, -1)_{B'} = 1(\sqrt{2}/2, \sqrt{2}/2) - 1(-\sqrt{2}/2, \sqrt{2}/2) = (1, 0) \text{ y}$$

$$\vec{v}'_2 = (1, 1)_{B'} = 1(\sqrt{2}/2, \sqrt{2}/2) + 1(-\sqrt{2}/2, \sqrt{2}/2) = (0, 1).$$

Se verifica, entonces, como se mencionó en la Observación i en 2.6 que las matrices de una misma transformación lineal en distintas bases son semejantes. Se tiene la relación,

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = P_{B,B'} \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} P_{B',B}$$

donde $P_{B,B'}$ y $P_{B',B}$ son las matrices del cambio de base de B' a B y de B a B' , respectivamente. ■

■ **Ejemplo 3.8** Se quieren determinar los valores y vectores propios de la aplicación lineal que corresponde a la rotación de ángulo α que tiene como matriz: (ver la matriz 1.8, Ejemplo 1.28)

$$R_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

con respecto a la base canónica de \mathbb{R}^2 .

$$0 = |R_\alpha - \lambda I| = \begin{vmatrix} \cos(\alpha) - \lambda & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) - \lambda \end{vmatrix} = (\cos(\alpha) - \lambda)^2 + \sin^2(\alpha) = \lambda^2 - 2\cos(\alpha)\lambda + 1$$

Las raíces son $\lambda_1 = \cos(\alpha) + i\sin(\alpha)$ y $\lambda_2 = \cos(\alpha) - i\sin(\alpha)$, que son números complejos a no ser que $\alpha = 2k\pi$ o $\alpha = (2k+1)\pi$, con $k \in \mathbb{Z}$.

Si $\alpha = 2k\pi$, se tiene la matriz identidad y todo vector de \mathbb{R}^2 es autovector, y corresponden a $\lambda = 1$.

Mientras que si $\alpha = (2k+1)\pi$, la matriz es $-I$ y también resulta autovector cualquier vector de \mathbb{R}^2 , y corresponden a $\lambda = -1$. En este caso, la transformación es una simetría respecto al origen de coordenadas. ■

■ **Ejemplo 3.9** La matriz correspondiente a una rotación de ángulo α en \mathbb{R}^3 con respecto al eje z , es,

$$R = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Su polinomio característico es $(\lambda^2 - 2\cos(\alpha)\lambda + 1)(1 - \lambda)$, cuyas raíces son $\lambda_1 = \cos(\alpha) + i\sin(\alpha)$, $\lambda_2 = \cos(\alpha) - i\sin(\alpha)$ y $\lambda_3 = 1$.

Los vectores propios correspondientes a $\lambda_3 = 1$ son las soluciones del sistema:

$$\begin{pmatrix} \cos(\alpha) - 1 & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) - 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (3.8)$$

Dado que

$$\begin{vmatrix} \cos(\alpha) - 1 & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) - 1 \end{vmatrix} = (\cos(\alpha) - 1)^2 + \sin^2(\alpha) = 2 - 2\cos(\alpha) = 4\sin^2(\alpha/2)$$

El sistema (3.8) tiene únicamente la solución $x_1 = x_2 = 0$ si $\alpha \neq 2k\pi$, con $k \in \mathbb{Z}$. En este caso los autovectores correspondientes a $\lambda_3 = 1$, son los vectores sobre el eje z , de la forma $(0, 0, x_3)$.

Si $\alpha = 2k\pi$, se trata de la identidad y todos los vectores de \mathbb{R}^3 son autovectores.

Si $\alpha = (2k+1)\pi$, $\lambda_1 = \lambda_2 = -1$, y los autovectores son las soluciones del sistema

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y son los vectores $(x_1, x_2, 0)$ o sea del plano xy , y la transformación es una simetría respecto al eje z . ■

3.4 Diagonalización

La Proposición 3.2.2 nos da una condición necesaria y suficiente para saber cuándo una aplicación lineal es diagonalizable, a saber, que exista una base del espacio vectorial V formada por vectores propios. En algunos casos puede resultar laborioso encontrar esta base. Una condición que es suficiente para poder asegurar la diagonalización de una matriz está contenida en la proposición siguiente:

Proposición 3.4.1 Los vectores propios de una aplicación T correspondientes a valores propios distintos dos a dos, son linealmente independientes.

Demostración:

Por inducción sobre la cantidad de vectores.

- Para $k = 2$. Supongamos se tienen \vec{v}_1 correspondiente a λ_1 y \vec{v}_2 correspondiente a λ_2 , con $\lambda_1 \neq \lambda_2$.

Si se tiene

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 = \vec{0} \tag{3.9}$$

$$T(\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2) = T(\vec{0}) = \vec{0}$$

$$\alpha_1 T(\vec{v}_1) + \alpha_2 T(\vec{v}_2) = \vec{0}$$

$$\alpha_1 \lambda_1 \vec{v}_1 + \alpha_2 \lambda_2 \vec{v}_2 = \vec{0}$$

De (3.9) se tiene que $\alpha_2 \vec{v}_2 = -\alpha_1 \vec{v}_1$. Si se reemplaza en la última ecuación, se tiene que

$$\begin{aligned}\alpha_1 \lambda_1 \vec{v}_1 - \alpha_1 \lambda_2 \vec{v}_1 &= \vec{0} \\ \alpha_1 \vec{v}_1 (\lambda_1 - \lambda_2) &= \vec{0}\end{aligned}\tag{3.10}$$

Como $\vec{v}_1 \neq \vec{0}$ y $\lambda_1 - \lambda_2 \neq 0$, resulta $\alpha_1 = 0$, de donde $\alpha_2 = 0$, pues $\vec{v}_2 \neq \vec{0}$.

Por lo tanto, \vec{v}_1 y \vec{v}_2 son linealmente independientes.

- Supongamos ahora el resultado es válido para $k - 1$ autovalores distintos.

Si

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_{k-1} \vec{v}_{k-1} + \alpha_k \vec{v}_k = \vec{0},$$

como en el caso $k = 2$, aplicamos T , despejamos de la igualdad anterior el término $\alpha_k \vec{v}_k$.

Entonces,

$$\alpha_1 \lambda_1 \vec{v}_1 + \alpha_2 \lambda_2 \vec{v}_2 + \cdots + \alpha_{k-1} \lambda_{k-1} \vec{v}_{k-1} + \alpha_k \lambda_k \vec{v}_k = \vec{0}$$

$$\alpha_1 \lambda_1 \vec{v}_1 + \alpha_2 \lambda_2 \vec{v}_2 + \cdots + \alpha_{k-1} \lambda_{k-1} \vec{v}_{k-1} + (-\alpha_1 \vec{v}_1 - \alpha_2 \vec{v}_2 - \cdots - \alpha_{k-1} \vec{v}_{k-1}) \lambda_k = \vec{0}$$

$$\alpha_1 (\lambda_1 - \lambda_k) \vec{v}_1 + \alpha_2 (\lambda_2 - \lambda_k) \vec{v}_2 + \cdots + \alpha_{k-1} (\lambda_{k-1} - \lambda_k) \vec{v}_{k-1} = \vec{0}$$

Como $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{k-1}$ son linealmente independientes y además, como

$$(\lambda_1 - \lambda_k) \neq 0, (\lambda_2 - \lambda_k) \neq 0, \dots, (\lambda_{k-1} - \lambda_k) \neq 0,$$

se tiene que $\alpha_1 = \alpha_2 = \cdots = \alpha_{k-1} = 0$, y también $\alpha_k = 0$, ya que $\vec{v}_k \neq \vec{0}$.

Por lo tanto $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{k-1}, \vec{v}_k$ son linealmente independientes.

□



- Una matriz T puede ser diagonalizable y tener autovalores múltiples. Un ejemplo es la matriz identidad, que tiene único autovalor 1 y es diagonalizable.
- Si una matriz $A \in K^{n \times n}$ tiene sus n autovalores distintos, sus autovectores son linealmente independientes y forman una base, por lo tanto, A es diagonalizable.

- $A\vec{v} = \lambda\vec{v}$ es una ecuación no lineal (λ multiplica \vec{v}). Si hallamos λ , sí la ecuación es lineal. Como $(A - \lambda I)\vec{v} = \vec{0}$, \vec{v} está en el espacio nulo de $(A - \lambda I)$, ($\vec{v} \in \text{Nul}(A - \lambda I)$).
- La condición $\text{Det}(A - \lambda I) = 0$ (donde A es la matriz que representa la aplicación T en alguna base) es equivalente a que la aplicación $T - \lambda I$ no es inyectiva, o sea $N(T - \lambda I) \neq \{\vec{0}\}$. O equivalentemente el núcleo de $T - \lambda I$ contiene no solo al vector nulo y es un subespacio de dimensión es mayor que 0.
- Si λ es autovalor con autovector \vec{v} de una matriz A no singular, $\frac{1}{\lambda}$ es autovalor de A^{-1} , con el mismo autovector. Ya que, si $A\vec{v} = \lambda\vec{v}$, multiplicando por la inversa, $\vec{v} = A^{-1}\lambda\vec{v}$, se tiene que

$$A^{-1}\vec{v} = \frac{1}{\lambda}\vec{v}.$$

- Si A es diagonalizable, $A = C^{-1}DC$, de donde, aplicando recursivamente, se tiene

$$A^n = C^{-1}D^nC \quad \forall n \in \mathbb{Z}.$$

- Dos matrices diagonalizables A y B comparten la matriz de autovectores S sí y solo si $AB = BA$. Para ver esto, supongamos $A = SD_1S^{-1}$ y $B = SD_2S^{-1}$. Si comparten la matriz de autovectores S , se tiene que

$$AB = SD_1S^{-1}SD_2S^{-1} = SD_1D_2S^{-1}$$

y

$$BA = SD_2S^{-1}SD_1S^{-1} = SD_2D_1S^{-1},$$

como $D_1D_2 = D_2D_1$ (las matrices diagonales siempre conmutan), entonces $AB = BA$. Y recíprocamente, si $AB = BA$, se puede demostrar que A y B comparten autovectores.

- Se puede demostrar que la suma de los n autovalores de una matriz A es igual a la traza de la matriz (suma de los elementos de la diagonal), es decir que

$$\text{Tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

y que el producto de los n autovalores es el determinante de A .

$$\text{Det}(A) = \prod_{i=1,2,\dots,n} \lambda_i$$

■ **Ejemplo 3.10** Se deja al lector estudiar si las matrices

$$A = \begin{pmatrix} 6 & -2 & 1 \\ 6 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

son diagonalizables.

Notar que la matriz A , por el Teorema 3.2.3, tiene un autovalor $\lambda_1 = 1$. Verifique que A es diagonalizable y los autovalores restantes son $\lambda_2 = 2$ y $\lambda_3 = 3$. Puede utilizar las sentencias Octave del programa presentado en la Introducción, Sección 3.1 $[U, D] = \text{eig}(A)$. Obtiene así la matriz U de autovectores y la matriz diagonal D con los autovalores. O bien, puede usar las sentencias en Python que están en recuadro.

En cuanto a la matriz T , ya fue estudiada en el Ejemplo 3.8. Es la matriz de una rotación en $\pi/2$ en sentido horario.

■



Sentencias en Python para diagonalizar una matriz

```
import numpy as np
A = np.array([[6, -2, 1], [6, -1, 1], [0, 0, 1]])
print(mat)
print()
D, U = np.linalg.eig(A)
%print('autovalores', D)
print(D)
%print('autovectores', U)
print(U)
```

■ **Ejemplo 3.11** Sea la matriz $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ \alpha & 0 & 0 \end{pmatrix}$.

Si se desea analizar para qué valores de $\alpha \in \mathbb{R}$ la matriz es diagonalizable, se tiene, en primer lugar que con solo observar los elementos de la matriz se sabe que tiene un autovalor 0 y otro α (Por

Teorema 3.2.3). A partir de esa información, hay que ver para qué valores de $\alpha \in \mathbb{R}$ existe una base de autovectores.

Se deja al lector verificar que solo es diagonalizable si $\alpha = 0$ y en ese caso todo vector de \mathbb{R}^3 es autovector. ■

Veamos a continuación una propiedad especial y muy útil de los autovectores de una matriz simétrica.

Teorema 3.4.2 Los autovectores de una matriz real simétrica, asociados a autovalores diferentes, son ortogonales.

Demostración:

Sea A una matriz simétrica y $\lambda_1 \neq \lambda_2$ autovalores con autovectores correspondientes \vec{v}_1 y \vec{v}_2 , es decir $A\vec{v}_1 = \lambda_1\vec{v}_1$ y $A\vec{v}_2 = \lambda_2\vec{v}_2$. Se verá que $\vec{v}_1 \cdot \vec{v}_2 = (\vec{v}_1)^t \vec{v}_2 = 0$

$$\begin{aligned} \lambda_1(\vec{v}_1 \cdot \vec{v}_2) &= \lambda_1(\vec{v}_1)^t \vec{v}_2 \\ &= (A\vec{v}_1)^t \vec{v}_2 \\ &= (\vec{v}_1)^t A^t \vec{v}_2 \\ &= (\vec{v}_1)^t A \vec{v}_2 \\ &= \lambda_2(\vec{v}_1)^t \vec{v}_2 \\ &= \lambda_2(\vec{v}_1 \cdot \vec{v}_2) \end{aligned}$$

de donde

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2) - \lambda_2(\vec{v}_1 \cdot \vec{v}_2) = (\lambda_1 - \lambda_2)(\vec{v}_1 \cdot \vec{v}_2) = 0,$$

y al ser $\lambda_1 \neq \lambda_2$, se tiene que $\vec{v}_1 \cdot \vec{v}_2 = 0$ y los vectores son ortogonales. □

■ **Ejemplo 3.12** Si las matrices A y B de $n \times n$ tienen autovalores λ y μ , podemos preguntarnos si la matriz producto AB tiene como autovalor a $\lambda\mu$. Es decir si se verifica para algún $\vec{x} \neq \vec{0}$ tal que,

$$AB\vec{x} = A\mu\vec{x} = \mu A\vec{x} = \mu\lambda\vec{x}.$$

Se deja al lector analizar este ejemplo:

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

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

y

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

■

3.4.1 Espacios propios

Sea $T \in L(V)$. El conjunto de los autovectores correspondientes a un autovalor λ no es un subespacio de K^n , puesto que $\vec{0}$ no es autovector de T . Sin embargo, podemos considerar el siguiente subespacio:

Definición 3.4.1 Sea $T \in K^{n \times n}$ y sea λ un autovalor de T . Se define el *espacio propio asociado* a λ y se lo anota \mathbf{E}_λ a

$$\mathbf{E}_\lambda = N(T - \lambda I) = \{\vec{v} \in K^n / T\vec{v} = \lambda \vec{v}\} \quad (3.11)$$

\mathbf{E}_λ es un subespacio de K^n , puesto que es el conjunto de soluciones de un sistema lineal homogéneo. Contiene todos los vectores propios correspondientes a λ junto con el vector $\vec{0}$.

De los resultados ya vistos en el Teorema 2.4.3, se tiene que

$$\dim(\mathbf{E}_\lambda) = \dim(V) - \dim(\text{Im}(T - \lambda I)) = \dim(V) - r(T - \lambda I)$$

Teorema 3.4.3 Sea $T \in L(V)$, V de $\dim < \infty$. Sean $\lambda_1, \lambda_2, \dots, \lambda_k$ los k , ($k \leq n$) autovalores distintos de T . Entonces $\mathbf{E}_{\lambda_1}, \mathbf{E}_{\lambda_2}, \dots, \mathbf{E}_{\lambda_k}$ están en suma directa.

Demostración:

Lo probaremos por inducción sobre la cantidad k de vectores considerados.

Para $k = 2$, sean λ_1 y λ_2 autovalores distintos de T . Si $\vec{v} \in \mathbf{E}_{\lambda_1} \cap \mathbf{E}_{\lambda_2}$, se tiene que $T\vec{v} = \lambda_1 \vec{v}$ y

$T\vec{v} = \lambda_2\vec{v}$, de donde $(\lambda_1 - \lambda_2)\vec{v} = 0$. Como $\lambda_1 - \lambda_2 \neq 0$, resulta que $\vec{v} = \vec{0}$. Luego $\mathbf{E}_{\lambda_1} \cap \mathbf{E}_{\lambda_2} = \vec{0}$ y la suma es directa.

Supongamos ahora que el resultado vale para el caso de $k - 1$ autovalores distintos, y sean $\lambda_1, \lambda_2, \dots, \lambda_k$ autovalores distintos de T .

Debemos probar que para cada $1 \leq i \leq k$, $\mathbf{E}_{\lambda_i} \cap (\sum_{j \neq i}^k \mathbf{E}_{\lambda_j}) = \vec{0}$ (ver Observación i al final de la Sección 1.4).

Supongamos que $i = k$, y sea $\vec{v} \in \mathbf{E}_{\lambda_k} \cap (\sum_{j=1}^{k-1} \mathbf{E}_{\lambda_j})$. Entonces, existen $\vec{v}_j \in \mathbf{E}_{\lambda_j}$, $(1 \leq j \leq k-1)$ tales que

$$\vec{v} = \vec{v}_1 + \vec{v}_2 + \dots + \vec{v}_{k-1}. \quad (3.12)$$

Multiplicando la igualdad (3.12) por la matriz T , como $\vec{v} \in \mathbf{E}_{\lambda_k}$, se tiene

$$\lambda_k \vec{v} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_{k-1} \vec{v}_{k-1} \quad (3.13)$$

y multiplicando ahora la igualdad Ec.(3.12) por λ_k , se tiene,

$$\lambda_k \vec{v} = \lambda_k \vec{v}_1 + \lambda_k \vec{v}_2 + \dots + \lambda_k \vec{v}_{k-1}$$

Restando las igualdades miembro a miembro,

$$\vec{0} = (\lambda_1 - \lambda_k) \vec{v}_1 + (\lambda_2 - \lambda_k) \vec{v}_2 + \dots + (\lambda_{k-1} - \lambda_k) \vec{v}_{k-1}$$

Como por hipótesis inductiva, los subespacios \mathbf{E}_{λ_j} $(1 \leq j \leq k-1)$ están en suma directa, el vector nulo se escribe de forma única como suma de vectores nulos, de donde $(\lambda_j - \lambda_k) \vec{v}_j = \vec{0}$ para cada $1 \leq j \leq k-1$ y por lo tanto $\vec{v}_j = \vec{0}$ para cada $1 \leq j \leq k-1$, con lo cual $\vec{v} = \vec{0}$. \square

Proposición 3.4.4 Sea $T \in K^{n \times n}$ y sea $\lambda \in K$ un autovalor de T . Sea r la *multiplicidad* de λ como raíz del polinomio característico \mathbf{P}_T y sea \mathbf{E}_λ su espacio propio. Entonces

$$\dim(\mathbf{E}_\lambda) \leq r$$

Proposición 3.4.5 Sea $T \in K^{n \times n}$ y sean $\lambda_1, \lambda_2, \dots, \lambda_k \in K$ los k autovalores distintos de T ($\lambda_i \neq \lambda_j$ si $i \neq j$). Son equivalentes:

1. T es diagonalizable
2. $K^n = \mathbf{E}_{\lambda_1} \oplus \mathbf{E}_{\lambda_2} \oplus \dots \oplus \mathbf{E}_{\lambda_k}$
3. El polinomio característico de T es

$$\mathbf{P}_T(\lambda) = (\lambda - \lambda_1)^{\alpha_1} (\lambda - \lambda_2)^{\alpha_2} \dots (\lambda - \lambda_k)^{\alpha_k}$$

y se tiene que $\alpha_i = \dim(\mathbf{E}_{\lambda_i})$, para $1 \leq i \leq k$.

■ **Ejemplo 3.13** Se quiere estudiar si la matriz

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

es diagonalizable. Sus autovalores son: $\lambda_1 = 2$, que es una raíz simple (multiplicidad algebraica 1) y $\lambda_2 = \lambda_3 = -2$ que es una raíz doble (multiplicidad algebraica 2).

$$\mathbf{E}_{\lambda_1} = N(T - 2I) = \{\vec{v} \in K^n / T\vec{v} = 2\vec{v}\} = \langle (1, 1, -1)^t \rangle$$

$$\mathbf{E}_{\lambda_2} = N(T + 2I) = \{\vec{v} \in K^n / T\vec{v} = -2\vec{v}\} = \langle (1, -1, 1)^t \rangle$$

La $\dim(\mathbf{E}_{\lambda_1}) = 1$ coincide con la multiplicidad algebraica mientras que $\dim(\mathbf{E}_{\lambda_2}) = 1$ es menor que 2, que es la multiplicidad algebraica. Por el Teorema 3.4.5, T no es diagonalizable. ■

Proposición 3.4.6 Sea $T \in L(V)$, V de $\dim < \infty$. Si T es diagonalizable y $\lambda_1, \lambda_2, \dots, \lambda_k$ son los autovalores distintos de T , entonces existen E_1, E_2, \dots, E_k aplicaciones lineales tales que

1. $E_1 + E_2 + \dots + E_k = I$
2. $T = \lambda_1 E_1 + \lambda_2 E_2 + \dots + \lambda_k E_k$
3. $E_i \circ E_j = \vec{0}$ $i \neq j$
4. $E_i^2 = E_i$, $i = 1, 2, \dots, k$

$$5. \operatorname{Im}(E_i) = \mathbf{E}_{\lambda_i}$$

Demostración:

1. Por ser T diagonalizable, de la Proposición 3.4.5 se tiene que

$$V = \mathbf{E}_{\lambda_1} \oplus \mathbf{E}_{\lambda_2} \cdots \oplus \mathbf{E}_{\lambda_k}$$

Para ver 1, se consideran las proyecciones E_i sobre cada espacio propio, \mathbf{E}_{λ_i} , asociadas a la descomposición anterior y $\vec{v} = \vec{w}_1 + \vec{w}_2 + \cdots + \vec{w}_k$, con $\vec{w}_i \in \mathbf{E}_{\lambda_i}$.

Entonces,

$$(E_1 + E_2 + \cdots + E_k)(\vec{v}) = (E_1 + E_2 + \cdots + E_k)(\vec{w}_1 + \vec{w}_2 + \cdots + \vec{w}_k) = E_1(\vec{w}_1 + \vec{w}_2 + \cdots + \vec{w}_k) + E_2(\vec{w}_1 + \vec{w}_2 + \cdots + \vec{w}_k) \cdots + E_k(\vec{w}_1 + \vec{w}_2 + \cdots + \vec{w}_k) = \vec{w}_1 + \vec{w}_2 + \cdots + \vec{w}_k = \vec{v}.$$

De donde se tiene, que $E_1 + E_2 + \cdots + E_k = I$

2. Para ver 2, se usa 1 y se toma la composición $T \circ I = T = T(E_1 + E_2 + \cdots + E_k)$. Luego, si $\vec{v} \in V$, $T(E_1 + E_2 + \cdots + E_k)(\vec{v}) = T(E_1(\vec{v})) + T(E_2(\vec{v})) \cdots T(E_k(\vec{v}))$

Como

$$T(E_i(\vec{v})) = T(\vec{w}_i) = \lambda_i \vec{w}_i = \lambda_i E_i(\vec{v})$$

se tiene, $T(E_1(\vec{v})) + T(E_2(\vec{v})) \cdots T(E_k(\vec{v})) = \lambda_1 E_1(\vec{v}) + \lambda_2 E_2(\vec{v}) + \cdots + \lambda_k E_k(\vec{v}) = (\lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_k E_k)(\vec{v})$, y por lo tanto,

$$T = \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_k E_k$$

La demostración de 3, 4 y 5 se dejan como ejercicio para el lector.

□

3.4.2 Polinomios minimales

Sea $P \in P_K[x]$, $P(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_r x^r$

Dada $T \in K^{n \times n}$ se define

$$P(T) = \alpha_0 I_n + \alpha_1 T + \alpha_2 T^2 + \cdots + \alpha_r T^r \in K^{n \times n}$$

Recordar que T^r es la composición de la aplicación lineal T r veces, y además, que si $P, Q \in P_K[x]$, y $T \in K^{n \times n}$, entonces $(P+Q)(T) = P(T) + Q(T)$ y $(P \cdot Q)(T) = P(T) \cdot Q(T)$.

Dada T una aplicación lineal cualquiera de un espacio vectorial V sobre K , interesa considerar polinomios que anulen a T , es decir

$$\left\{ P \in P_K[x] \mid P(T) = \vec{0} \right\}.$$

(Se pueden ver en [23] más detalles sobre este tema). El resultado que sigue asegura que para cualquier matriz existe un polinomio no nulo con esta propiedad.

Proposición 3.4.7 Sea $T \in K^{n \times n}$. Existe un polinomio $P \in P_K[x]$, $P \neq 0$, tal que $P(T) = \vec{0}$

Demostración:

Consideremos el conjunto $\{I, T, T^2, \dots, T^{n^2}\} \subseteq K^{n \times n}$ son $n^2 + 1$ transformaciones lineales de $L(K^n, K^n)$. Son linealmente dependientes porque ya vimos que la dimensión es la de $K^{n \times n} = n^2$. Luego existe una combinación lineal con escalares no todos nulos tales que

$$\alpha_0 I + \alpha_1 T + \alpha_2 T^2 + \dots + \alpha_{n^2} T^{n^2} = \vec{0}$$

Sea

$$P(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \dots + \alpha_{n^2} x^{n^2} = \vec{0}$$

$$P \in P_K[x], P \neq \vec{0} \text{ y } P(T) = \vec{0}$$

□

Es decir, para toda matriz, distinguimos un polinomio particular entre todos los polinomios que la anulan: el de grado mínimo y mónico. Siempre existe un polinomio con esas propiedades y además, es único.

Definición 3.4.2 Sea $T \in K^{n \times n}$. Se llama *polinomio minimal* de T al polinomio mónico de grado mínimo que anula a T . Lo simbolizamos m_T .



m_T se caracteriza por

1. $m_T(T) = \vec{0}$
2. m_T es mónico y es el de menor grado que anula a T
3. m_T / \mathbf{P}_T (divide al polinomio característico)

■ **Ejemplo 3.14** Dada la matriz, $T = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$,

el lector puede verificar que $\{I, T\}$ son linealmente independientes, es decir, no existe $P \in P_{\mathbb{R}}[x]$ de grado 1 y tal que $P(T) = \vec{0}$.

En cambio, $\{I, T, T^2\}$ es un conjunto linealmente dependiente, ya que

$$\begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}^2 - 3 \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix} + 2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \vec{0} \quad (3.14)$$

Se tiene que $T^2 - 3T + 2I = (T - 2I)(T - I) = \vec{0}$, luego el polinomio minimal es

$$m_T(\lambda) = \lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1).$$

m_T divide a \mathbf{P}_T ya que $\mathbf{P}_T = (\lambda - 2)^2(\lambda - 1)$ es el polinomio característico.

■



- En la Proposición 3.3.1 de la Sección 3.3 vimos que las raíces del polinomio característico de una matriz son sus autovalores. Lo mismo vale para el polinomio minimal.
- Dos matrices semejantes tienen el mismo polinomio minimal (y el mismo polinomio característico como se vio en la Proposición 3.3.2).

Proposición 3.4.8 Sea $T \in K^{n \times n}$, sea $\lambda \in K$ y m_T el polinomio minimal de T . Entonces λ es autovalor de T sí y sólo sí λ es raíz de m_T .

Criterio de diagonalización usando el polinomio minimal

Si el polinomio característico de $T \in K^{n \times n}$, $\mathbf{P}_T(\lambda)$ se factoriza linealmente en $P_K[x]$, con todas sus raíces $\lambda_1 \cdots \lambda_n$ simples (raíces distintas), entonces T es diagonalizable.

Esto sale de la Proposición 3.4.5 de la sección 3.4.1, ya que existe una base de K^n formada por autovectores y los espacios propios $E_{\lambda_1}, E_{\lambda_2}, \dots, E_{\lambda_n}$ están en suma directa. La recíproca no es cierta, es decir, $\mathbf{P}_T(\lambda)$ puede tener raíces múltiples y ser diagonalizable.

Proposición 3.4.9 Sea $T \in K^{n \times n}$. Entonces T es diagonalizable en $K^{n \times n}$ sí y sólo sí el polinomio minimal m_T tiene todas sus raíces en K y son simples. O en forma equivalente, sean $\lambda_1, \lambda_2, \dots, \lambda_k$ los autovalores distintos de T . T es diagonalizable sí y sólo sí

$$m_T(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_k)$$

Proposición 3.4.10 Sea T una aplicación lineal sobre un espacio vectorial de V de $\dim < \infty$. El polinomio característico y el polinomio minimal tienen las mismas raíces.

El Teorema que sigue fue enunciado por Arthur Cayley (1821-1895) en 1858. Lo demostró inicialmente para matrices de 2×2 .

Teorema 3.4.11 Teorema de Cayley-Hamilton:

Sea T una aplicación lineal sobre un espacio vectorial de V de $\dim < \infty$. Si \mathbf{P}_T es el polinomio característico de T , entonces $\mathbf{P}_T(T) = \vec{0}$.

Demostración:

Sea $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ una base de V y sea A la matriz que representa a T en la base dada. Entonces

$$T\vec{v}_i = \sum_{j=1}^n a_{ji}\vec{v}_j, \quad 1 \leq i \leq n$$

Estas ecuaciones, que son las mismas que las desarrolladas en 3.5, pueden escribirse en forma equivalente

$$\sum_{j=1}^n (\delta_{ij}T - a_{ji}I)\vec{v}_j = \vec{0}, \quad 1 \leq i \leq n$$

Sea $B \in K^{n \times n}$ con elementos $B_{ij} = \delta_{ij}T - a_{ji}I$

Cuando $n = 2$,

$$B = \begin{pmatrix} T - a_{11}I & -a_{21}I \\ -a_{12}I & T - a_{22}I \end{pmatrix}$$

y

$$\det(B) = (T - a_{11}I)(T - a_{22}I) - a_{21}a_{12}I$$

$$\det(B) = T^2 - (a_{11} + a_{22})T + (a_{11}a_{22} - a_{12}a_{21})I$$

$$\det(B) = \mathbf{P}_T(T)$$

donde \mathbf{P}_T es el polinomio característico correspondiente a T .

$$\mathbf{P}_T(\lambda) = \lambda^2 - \text{traza}(A)\lambda + \det(A)$$

Para $n > 2$, también se tiene $\det(B) = f(T)$, ya que \mathbf{P}_T es el determinante de la matriz $(\lambda I - A)$ cuyos elementos son los polinomios

$$(\lambda I - A)_{ij} = \delta_{ij}\lambda - a_{ji}$$

Se quiere demostrar que $\mathbf{P}_T(T) = \vec{0}$, y para eso es necesario y suficiente ver que $(\det(B)) \vec{v}_k = \vec{0}$ para $1 \leq k \leq n$. Por la definición de B , los vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ satisfacen las ecuaciones

$$\sum_{j=1}^n B_{ij} \vec{v}_j = \vec{0}, \quad 1 \leq i \leq n$$

Cuando $n = 2$,

$$\begin{pmatrix} T - a_{11}I & -a_{21}I \\ -a_{12}I & T - a_{22}I \end{pmatrix} \begin{pmatrix} \vec{v}_1 \\ \vec{v}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

la matriz adjunta de B transpuesta es

$$B^* = \begin{pmatrix} T - a_{22}I & a_{21}I \\ a_{12}I & T - a_{11}I \end{pmatrix}$$

y se tiene

$$B^*B = \begin{pmatrix} \det(B) & 0 \\ 0 & \det(B) \end{pmatrix}$$

Luego,

$$(\det(B)) I \begin{pmatrix} \vec{v}_1 \\ \vec{v}_2 \end{pmatrix} = B^*B \begin{pmatrix} \vec{v}_1 \\ \vec{v}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Para el caso general, como

$$\sum_{j=1}^n B_{ij}\vec{v}_j = \vec{0}, \quad 1 \leq i \leq n$$

$$\sum_{j=1}^n B_{ki}^* B_{ij}\vec{v}_j = \vec{0}$$

para todo par k, i , y sumando sobre i , se tiene $\sum_{i=1}^n \sum_{j=1}^n B_{ki}^* B_{ij}\vec{v}_j = \vec{0}$, y entonces,

$$\sum_{j=1}^n \left(\sum_{i=1}^n B_{ki}^* B_{ij} \right) \vec{v}_j = \vec{0}$$

Como $B^*B = \det(B)I$, se tiene que $\sum_{i=1}^n B_{ki}^* B_{ij} = \delta_{kj} \det(B)$

y, por lo tanto,

$$\sum_{j=1}^n \delta_{kj} \det(B) \vec{v}_j = \det(B) \vec{v}_j = \vec{0}$$

□



- Una utilidad de este teorema es que reduce la búsqueda del polinomio minimal. Si se conoce la matriz A que representa a T en cierta base, se calcula el polinomio característico, y se sabe que el polinomio minimal lo divide y que los dos tienen las mismas raíces.
- En algunos casos el teorema resulta útil para calcular la inversa de una matriz. Si existe T^{-1} y $\mathbf{P}_T(T) = \vec{0}$, entonces, $T^{-1}\mathbf{P}_T(T) = \vec{0}$.

Si

$$\mathbf{P}_T(\lambda) = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \cdots + \alpha_n\lambda^n$$

$$\mathbf{P}_T(T) = \alpha_0 I + \alpha_1 T + \alpha_2 T^2 + \cdots + T^n \in K^{n \times n},$$

entonces,

$$T^{-1}\mathbf{P}_T(T) = \alpha_0 T^{-1} + \alpha_1 I + \alpha_2 T + \cdots + T^{n-1} = \vec{0}$$

de donde, despejando,

$$T^{-1} = \frac{1}{\alpha_0}(-\alpha_1 I - \alpha_2 T - \cdots - T^{n-1})$$

$\alpha_0 \neq 0$, ya que $\alpha_0 = \mathbf{P}_T(\vec{0}) = \text{Det}(0I - T) = (-1)^n \text{Det}(T)$ y T es invertible.

Arthur Cayley (1821-1895)

Fue un matemático británico. Fue uno de los fundadores de la escuela británica moderna de matemáticas puras. Además de su predilección por las matemáticas, también era un ávido lector de novelas, le gustaba pintar, sentía pasión por la botánica y por la naturaleza en general, y era aficionado al alpinismo. Se educó en el Trinity College de Cambridge. Estudió durante algún tiempo la carrera de leyes con lo que trabajó de abogado durante 14 años, a la vez que publicaba un gran número de artículos. Luego pasó a ser profesor en Cambridge. Fue el primero que introdujo la multiplicación de las matrices. Es el autor del teorema de Cayley-Hamilton que dice que cualquier matriz cuadrada es solución de su polinomio característico. Dio la primera definición moderna de la noción de grupo. En combinatoria, su nombre está unido a la fórmula que cuenta los posibles árboles generadores con nodos etiquetados de orden n . Se llama a veces octavas de Cayley o números de Cayley a los octoniones. Es el tercer matemático más prolífico de la historia, sobrepasado tan solo por

Euler y Cauchy, con aportaciones a amplias áreas de la matemática. En 1889, Cambridge University Press le pidió que preparara sus artículos matemáticos en forma de colección. Siete volúmenes aparecieron con Cayley como editor, pero tras su fallecimiento, el resto de artículos fue editado por Andrew Forsyth, su sucesor en la cátedra de Cambridge. En total los Collected Mathematical Papers comprenden trece grandes volúmenes que contienen 967 artículos. [5]

3.5 Teorema de la descomposición prima

Es de interés, dada T una aplicación lineal sobre un espacio vectorial de V de $\dim < \infty$, la descomposición de V como suma directa de subespacios invariantes por T ,

$$V = W_1 \oplus W_2 \cdots \oplus W_k, W_i \text{ subespacio de } V \text{ y } T(W_i) \subseteq W_i$$

es decir, generalizar la Proposición 3.4.5 de la sección 3.4.1 para el caso que T no es diagonalizable. Esto lo muestra el Teorema que sigue.

Teorema 3.5.1 Teorema de la descomposición prima:

Sea T una aplicación lineal sobre un espacio vectorial de V de $\dim < \infty$. Sea $m_T(\lambda)$ el polinomio minimal de T cuya factorización prima es

$$m_T(\lambda) = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k} = (\lambda - \lambda_1)^{r_1} (\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}$$

los p_i son los polinomios primos (mónicos e irreducibles) en $P_K[\lambda]$, todos distintos y r_i son enteros positivos.

Sea $W_i = N(p_i^{r_i}(T)) = N(T - \lambda_i)^{r_i}$. Entonces

1. $V = W_1 \oplus W_2 \cdots \oplus W_k$
2. Cada W_i es invariante por T , o sea, $T(W_i) \subseteq W_i$

3. Si $T_i = T|_{W_i}$ (T restringido a W_i), entonces T_i tiene a $p_i^{r_i}$ como polinomio minimal

■ **Ejemplo 3.15** Dada la matriz

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

Las raíces del polinomio característico son $\lambda_1 = \lambda_2 = 2$ raíz doble y $\lambda_3 = 1$, raíz simple.

Hay dos opciones para el polinomio minimal,

$$m_T(\lambda) = (\lambda - 2)(\lambda - 1) \text{ o } m_T(\lambda) = (\lambda - 2)^2(\lambda - 1).$$

Puede verificarse que $(T - 2I)(T - I) \neq \vec{0}$, así que resulta $m_T(\lambda) = (\lambda - 2)^2(\lambda - 1)$ y T no es diagonalizable (por la Proposición 3.4.9).

Por el Teorema de la descomposición prima ,

$$p_1 = (\lambda - 2)^2 \text{ y } p_2 = (\lambda - 1), (r_1 = 2 \text{ y } r_2 = 1)$$

Teniendo en cuenta que

$$W_i = N(p_i^{r_i}(T)) = N(T - \lambda_i)^{r_i}, \text{ se tiene que } W_1 = N(p_1^{r_1}(T)) = N(T - 2I)^2.$$

Para hallar W_1 se calcula la matriz $(T - 2I)$,

$$T - 2I = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{pmatrix}$$

luego $(T - 2I)^2$ y se resuelve el sistema homogéneo con la matriz $(T - 2I)^2$:

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Se obtiene que $\{(1, 1, 2)\}$ y $\{(1, 1, 2), (1, 1, 0)\}$ son bases de $N(T - 2I)$ y de W_1 , respectivamente.

Por otro lado, $W_2 = N(p_2^{T_2}(T)) = N(T - 1I) = \mathbf{E}_{\lambda_3}$.

Al resolver este sistema homogéneo,

$$\begin{pmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

se tiene que una base de W_2 es $\{(1, 0, 2)\}$.

Es posible hallar la matriz de T en la base $B = \{(1, 1, 2), (1, 1, 0), (1, 0, 2)\}$.

Utilizando la matriz en la base canónica, se tiene que

■

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

de donde

$$T((1, 1, 2)) = 2(1, 1, 2) + 0(1, 1, 0) + 0(1, 0, 2).$$

$(1, 1, 2)$ es un autovector correspondiente a $\lambda_1 = \lambda_2 = 2$.

■

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

$$T((1, 1, 0)) = (4, 4, 4) = 2(1, 1, 2) + 2(1, 1, 0) + 0(1, 0, 2)$$

De la misma forma,

$$\blacksquare \quad T((1, 0, 2)) = (1, 0, 2) = 0(1, 1, 2) + 0(1, 1, 0) + 1(1, 0, 2),$$

La matriz de T en la nueva base es casi diagonal,

$$(T)_B = \begin{pmatrix} 2 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Para verificar que

$V = \mathbb{R}^3 = W_1 \oplus W_2$, se toma un vector (x_1, x_2, x_3) y se buscan sus proyecciones sobre los subespacios W_1 y W_2 .

A partir de la igualdad $(x_1, x_2, x_3) = \alpha(1, 1, 2) + \beta(1, 1, 0) + \gamma(1, 0, 2)$ se resuelve el sistema y se obtiene que $\alpha = -x_1 + x_2 + x_3/2$, $\beta = x_1 - x_3/2$ y $\gamma = x_1 - x_2$. Los dos primeros términos del lado derecho corresponden a la proyección sobre W_1 , que llamaremos E_1 y el tercero a la proyección sobre W_2 , E_2 . Es decir que

$$E_1((x_1, x_2, x_3)) = (x_2, x_2, x_3 - 2x_1 + 2x_2)$$

y

$$E_2((x_1, x_2, x_3)) = (x_1 - x_2, 0, 2x_1 - 2x_2)$$

Como ejemplo, las proyecciones del vector $(2, 1, 2)$ sobre W_1 y W_2 , son $E_1((2, 1, 2)) = (1, 1, 0)$ y $E_2((2, 1, 2)) = (1, 0, 2)$, respectivamente ya que $(2, 1, 2) = (1, 1, 0) + (1, 0, 2)$.

■

■ **Ejemplo 3.16** $T : V \rightarrow V$, $V = \mathbb{R}^5$ y $m_T(\lambda) = (\lambda - 1)(\lambda^2 + 1)(\lambda + 1)^2$

Los autovalores reales son $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = -1$

$$W_1 = N(p_1^{r_1}(T)) = N(T - 1I) = E_{\lambda_1}$$

$$W_2 = N(p_2^{r_2}(T)) = N(T^2 + I)$$

$$W_3 = N(p_3^{r_3}(T)) = N(T + I)^2$$

$$V = \mathbb{R}^5 = W_1 \oplus W_2 \oplus W_3$$

Se deja al lector verificar esta descomposición de V en suma directa de subespacios .

■



Si aplicamos el teorema anterior en el caso que T sea un operador diagonalizable se tiene que

$$m_T(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_k)$$

donde $\lambda_1, \lambda_2, \dots, \lambda_k$ son los autovalores distintos de T y entonces, $W_i = N(T - \lambda_i I) = E_{\lambda_i}$ y

$$V = E_{\lambda_1} \oplus E_{\lambda_2} \cdots \oplus E_{\lambda_k}$$

(ver Proposición 3.4.5).

Definición 3.5.1 Sea N una aplicación lineal sobre un espacio vectorial V . Se dice que N es *nilpotente* si existe algún entero positivo r tal que $N^r = \mathbf{0}$ (matriz nula).

Se tiene el siguiente resultado:

Teorema 3.5.2 Sea T una transformación lineal sobre un espacio vectorial de V de $\dim < \infty$. Supongamos que $m_T(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}$. Entonces existe una transformación lineal D diagonalizable y un operador lineal N nilpotente tal que

1. $T = D + N$
2. $DN = ND$

A la transformación lineal D se la llama parte diagonal de la transformación lineal T .

Demostración:

1. Supongamos que T es una transformación lineal tal que $m_T(\lambda) = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k} = (\lambda - \lambda_1)^{r_1} (\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}$.

Sean E_1, E_2, \dots, E_k las proyecciones tales que $\text{Im}(E_i) = N(T - \lambda_i I)^{r_i}$ y $D = \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_k E_k$.

Consideremos $N = T - D$. De la Proposición 3.4.6 se tiene que $E_1 + E_2 + \cdots + E_k = I$ y

$$T = TE_1 + TE_2 + \cdots + TE_k.$$

$$N = (TE_1 + TE_2 + \cdots + TE_k) - (\lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_k E_k)$$

$$N = (T - \lambda_1 I)E_1 + \cdots + (T - \lambda_k I)E_k$$

$$N^2 = ((T - \lambda_1 I)E_1 + \cdots + (T - \lambda_k I)E_k)((T - \lambda_1 I)E_1 + \cdots + (T - \lambda_k I)E_k)$$

$$N^2 = (T - \lambda_1 I)^2 E_1 + \cdots + (T - \lambda_k I)^2 E_k$$

$$\dots = \dots$$

$$N^r = (T - \lambda_1 I)^r E_1 + \cdots + (T - \lambda_k I)^r E_k$$

Si $r \geq r_i$, se tiene que $N^r = \vec{0}$, ya que, dado $\vec{v} \in V$, $\vec{v} = \vec{v}_1 + \vec{v}_2 + \cdots + \vec{v}_k$, con $\vec{v}_i \in N(T - \lambda_i I)^{r_i}$

$$N^r \vec{v} = (T - \lambda_1 I)^r E_1 \vec{v} + \cdots + (T - \lambda_k I)^r E_k \vec{v}$$

$$N^r \vec{v} = (T - \lambda_1 I)^r \vec{v}_1 + \cdots + (T - \lambda_k I)^r \vec{v}_k$$

y como $\vec{v}_i \in N(T - \lambda_i I)^{r_i}$ cada término da el vector nulo, $\vec{0}$.

2. ver en ([23])

□

■ **Ejemplo 3.17** Del Ejemplo 3.15

$$(T)_B = \begin{pmatrix} 2 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$D(x_1, x_2, x_3) = \lambda_1 E_1(x_1, x_2, x_3) + \lambda_2 E_2(x_1, x_2, x_3)$$

y de acuerdo a lo obtenido antes, si se reemplazan E_1 y E_2 , se tiene que la transformación D es

$$D((x_1, x_2, x_3)) = \lambda_1(x_2, x_2, x_3 - 2x_1 + 2x_2) + \lambda_2(x_1 - x_2, 0, 2x_1 - 2x_2)$$

Su matriz asociada, en la base B es

$$(D)_B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

mientras que

$$N = (T)_B - (D)_B = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

es nilpotente de orden 2 ($N^2 = \mathbf{0}$, es la matriz nula). Se cumple además que $DN = ND$.

■

3.6 Forma canónica de Jordan

Como ya vimos, existen transformaciones lineales (o matrices $A \in K^{n \times n}$) que no son diagonalizables, es decir, no existe una base de autovectores. En ese caso, aún es posible demostrar que la matriz es semejante a otra, una matriz más sencilla aunque no es diagonal. La matriz de la transformación es un poco más difícil de obtener que en el caso diagonalizable. Veremos a continuación el procedimiento para matrices de 2×2 .

Forma de Jordan de matrices de orden 2

Sea $T \in L(V)$, V espacio vectorial de dimensión 2, y supongamos su matriz, en cierta base es

$$T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Su polinomio característico es $\mathbf{P}_T(\lambda) = (a - \lambda)(d - \lambda) - cb$ con lo que $\mathbf{P}_T(\lambda) = 0$ es una ecuación de grado 2 en la variable λ y se tendrán dos casos diferentes según las dos soluciones sean iguales o distintas.

1. Caso I. Las raíces del polinomio característico son distintas $\lambda_1 \neq \lambda_2$.

En este caso la matriz T es diagonalizable (consecuencia de la Proposición 3.4.1). Su forma de Jordan es la matriz diagonal

$$J = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

y $T = CJC^{-1}$, donde la matriz C del cambio de base tiene en sus columnas las coordenadas de vectores \vec{v}_1 e $\vec{v}_2 \in N(T - \lambda_i I)$, $i = 1, 2$, respectivamente.

2. Caso II. Las raíces del polinomio característico coinciden $\lambda_1 = \lambda_2$

En este caso, si $N(T - \lambda_1 I)$ tiene dimensión 1, no podremos encontrar una base de autovectores en V .

Se tiene el siguiente resultado:

Lema: Suponga T es una matriz de 2×2 que tiene dos autovalores iguales λ . Sea \vec{v}_1 un autovector correspondiente a λ . Existe un vector \vec{v}_2 que satisface la ecuación

$$(T - \lambda I)\vec{v}_2 = \vec{v}_1$$

Definición 3.6.1 Se denomina *autovector generalizado* al vector \vec{v}_2 de la ecuación anterior.

Como \vec{v}_1 y \vec{v}_2 son linealmente independientes, forman una base de V . En esta base tenemos,


$$(T - \lambda I)\vec{v}_1 = \vec{0} \Leftrightarrow T\vec{v}_1 = \lambda\vec{v}_1$$

$$(T - \lambda I)\vec{v}_2 = \vec{v}_1 \Leftrightarrow T\vec{v}_2 = \vec{v}_1 + \lambda\vec{v}_2$$

con lo que la matriz de la transformación lineal en esta base es

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

y la matriz de cambio de base C tiene en sus columnas a los vectores \vec{v}_1 y \vec{v}_2 .

 A partir de $(T - \lambda I)\vec{v}_2 = \vec{v}_1$, se tiene que

$$(T - \lambda I)(T - \lambda I)\vec{v}_2 = (T - \lambda I)\vec{v}_1 = \vec{0}$$

y, entonces,

$$\vec{v}_2 \in N((T - \lambda I)^2)$$

Se pueden resumir estos resultados en la siguiente proposición, en donde K designa el cuerpo de los números reales o el de los complejos.

Proposición 3.6.1 Dada una matriz $T \in K^{2 \times 2}$ siempre puede encontrarse una matriz $J \in K^{2 \times 2}$ de una cualquiera de las formas

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \text{ o } \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

con $\lambda_1, \lambda_2 \in K$ y una matriz $C \in K^{2 \times 2}$ tal que $T = CJC^{-1}$.

La matriz J se denomina *matriz de Jordan* de T .

■ **Ejemplo 3.18** Dada la matriz,

$$T = \begin{pmatrix} 3 & -2 \\ 8 & -5 \end{pmatrix}$$

al tener el polinomio característico una raíz doble (Caso II), su forma de Jordan es:

$$T = CJC^{-1} = \begin{pmatrix} 1 & 1/4 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1/2 \\ 4 & -2 \end{pmatrix}$$

■

Forma de Jordan de matrices de orden 3

Ahora veremos para aplicaciones lineales entre espacios vectoriales de dimensión 3. Servirá para comprender los resultados teóricos necesarios para obtener la forma de Jordan en espacios vectoriales de cualquier dimensión.

■ **Ejemplo 3.19** Se quiere reducir la matriz

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

a su forma de Jordan. Sus autovalores son $\lambda_1 = 2$ (simple) y $\lambda_2 = \lambda_3 = -2$ (doble), y los espacios propios correspondientes,

$$\mathbf{E}_{\lambda_1} = N(T - 2I) = \left\{ \alpha \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \mid \alpha \in K \right\}$$

y

$$\mathbf{E}_{\lambda_2} = N(T + 2I) = \left\{ \alpha \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \mid \alpha \in K \right\}$$

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \text{ y } \vec{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

son dos autovectores linealmente independientes pero no forman una base. Se debe realizar un trabajo análogo al realizado en el caso de raíces iguales para matrices de orden 2.

Se halla \vec{v}_3 tal que $(T - \lambda I)\vec{v}_3 = \vec{v}_2$, o, equivalentemente, $(T - \lambda I)(T - \lambda I)\vec{v}_3 = (T - \lambda I)\vec{v}_2 = \vec{0}$, es decir,

$$\vec{v}_3 \in N((T - \lambda_2 I)^2) = N((T + 2I)^2)$$

$$\text{Sea } \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \text{ se tiene que } \vec{v}_3 \in N((T + 2I)^2)$$

$$\text{pues } (T + 2I)(T + 2I) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y

$$N((T + 2I)^2) = \left\{ \alpha \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mid \alpha, \beta \in K \right\}$$

Ahora es posible elegir una base de V con estos vectores, $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ de manera que la matriz de T en esta base sea sencilla.

$$T\vec{v}_1 = 2\vec{v}_1, T\vec{v}_2 = -2\vec{v}_2, T\vec{v}_3 = \vec{v}_2 - 2\vec{v}_3$$

$$T \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

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

y la matriz del cambio de base es

$$C = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix}$$

■

Lema: Suponga T es una matriz de 3×3 que tiene 3 autovalores iguales a λ y la dimensión de su espacio propio es 1. Sea \vec{v}_1 un autovector correspondiente a λ .

- Existe un vector \vec{v}_2 que satisface la ecuación

$$(T - \lambda I)\vec{v}_2 = \vec{v}_1,$$

tal que \vec{v}_1 y \vec{v}_2 son linealmente independientes.

- Con el \vec{v}_2 hallado en el punto anterior, existe un vector \vec{v}_3 solución del sistema

$$(T - \lambda I)\vec{v}_3 = \vec{v}_2$$

tal que \vec{v}_1 , \vec{v}_2 y \vec{v}_3 son linealmente independientes.

- **Ejemplo 3.20** Dada la matriz,

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

$\lambda = -2$ es raíz triple del polinomio característico. Se deja al lector verificar que su forma de Jordan es $T = CJC^{-1}$ donde

$$C = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Las columnas de C son vectores \vec{v}_1 , \vec{v}_2 y \vec{v}_3 que cumplen el lema anterior, y

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

■

Marie Ennemond Camille Jordan (1838 - 1922)

Fue un matemático francés, conocido tanto por su trabajo sobre la teoría de grupos, como por su influyente Curso de análisis (Cours d'analyse). Estudió en la Escuela Politécnica (promoción de 1855). Fue ingeniero de minas y, más tarde, ejerció como examinador en la misma escuela. En 1876 entró como profesor en el Colegio de Francia, sustituyendo a Joseph Liouville. Su nombre se asocia a un determinado número de resultados fundamentales: El teorema de la curva de Jordan: un resultado topológico recogido en análisis complejo. La forma canónica de Jordan en álgebra lineal. El teorema de Jordan-Hölder, que es el resultado básico de unas series de composiciones. El trabajo de Jordan incidió de manera sustancial en la introducción de la teoría de Galois en la corriente del pensamiento mayoritario. Investigó también los grupos de Mathieu, los primeros ejemplos de grupos esporádicos. Su Tratado de las sustituciones (Traité des substitutions) sobre las permutaciones de grupos fue publicado en 1870. El 4 de abril de 1881 fue elegido miembro de la Academia de la Ciencia. De 1885 a 1921 dirige la «Revista de matemáticas puras y aplicadas» (Journal de mathématiques pures et appliquées), fundado por Liouville. [7]

Teorema de clasificación de Jordan

Se denomina *matriz elemental de Jordan* de orden k y autovalor $\lambda \in \mathbb{C}$ a la matriz de orden k cuyos elementos son todos nulos, excepto los de la diagonal principal, que valen λ y los situados inmediatamente encima de la diagonal principal que son unos. Por ejemplo:

$$J_1(\lambda) = (\lambda) \quad J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad J_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \text{ y así sucesivamente.}$$

Se llama *matriz de Jordan* a cualquier matriz cuadrada formada por yuxtaposición de matrices elementales de Jordan a lo largo de la diagonal, de la forma

$$\begin{pmatrix} J_{1j} & 0 & 0 & 0 \\ 0 & J_{2j} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & J_{nj} \end{pmatrix}$$

donde, en el caso de orden 3, $J_{ij} = \begin{pmatrix} \lambda_i & 1 & 0 \\ 0 & \lambda_i & 1 \\ 0 & 0 & \lambda_i \end{pmatrix}$

Una transformación lineal T puede expresarse en la forma canónica de Jordan si sus polinomios característico y minimal se factorizan en polinomios lineales. Esto siempre es verdadero si el cuerpo K es \mathbb{C} . Análogamente toda matriz es semejante a una matriz en forma canónica de Jordan.

Teorema 3.6.2 Teorema de Jordan:

Sea $T \in L(V)$ cuyos polinomios característico y minimal son respectivamente,

$$\mathbf{P}_T(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_k)^{n_k}$$

$$m_T(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_k)^{m_k} \text{ donde los } \lambda_i \text{ son distintos.}$$

Entonces T tiene una representación matricial J que es diagonal por bloques.

Para cada λ_i los bloques correspondientes J_{ij} tienen las siguientes propiedades:

1. Existe al menos un J_{ij} de orden m_i , los demás J_{ij} son de orden $\leq m_i$.
2. La suma de los órdenes de los J_{ij} es n_i .
3. La cantidad de J_{ij} es igual a la multiplicidad geométrica de λ_i (dimensión de \mathbf{E}_{λ_i}).
4. La cantidad de J_{ij} de cada orden posible está determinado únicamente por T .



La matriz J se llama forma canónica de Jordan de la transformación lineal T . A J_{ij} se lo llama bloque de Jordan correspondiente al valor propio λ_i . Observar que $J_{ij} = \lambda_i I + N$, donde N es una matriz nilpotente. ▲

■ **Ejemplo 3.21** Supongamos

$$\mathbf{P}_T(\lambda) = (\lambda - 2)^4(\lambda - 3)^3 \text{ y } m_T(\lambda) = (\lambda - 2)^2(\lambda - 3)^2.$$

Se quiere hallar su matriz de Jordan aplicando el teorema anterior:

Como $m_1 = m_2 = 2$ existe al menos un bloque de orden 2 para cada λ , $\lambda_1 = 2$ y $\lambda_2 = 3$.

La suma de los órdenes de los bloques para $\lambda_1 = 2$ es $n_1 = 4$ y para $\lambda_2 = 3$ es $n_2 = 3$.

La cantidad de bloques es la dimensión del espacio propio correspondiente. Para $\lambda_2 = 3$, hay un bloque de orden 2 y uno de orden 1.

Para $\lambda_1 = 2$ hay 2 posibilidades, dependiendo de su multiplicidad geométrica:

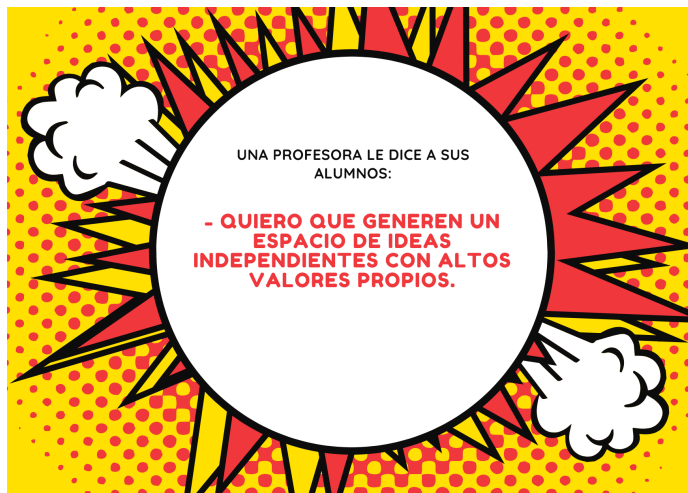
$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0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

En el primer caso, la dimensión del espacio propio es 2, hay 2 bloques de orden 2. En el segundo, como hay 3 autovectores, hay 3 bloques, 1 de orden 2 y 2 de orden 1.

Se puede observar que la cantidad de 1 en la matriz de Jordan corresponde a la resta: multiplicidades algebraicas - multiplicidades geométricas, $3 = 7 - 2 - 2$ en el primer caso y $2 = 7 - 3 - 2$ en el segundo.

■

Figura 3.4



3.7 Actividades propuestas

Problema de Aplicación 3 Un circuito eléctrico que describe el valor del voltaje de dos capacitores en función del tiempo, y en paralelo como se aprecia en la Figura 3.5 se resuelve con el siguiente sistema de ecuaciones diferenciales de primer orden:

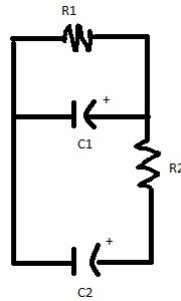
$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} -(1/R_1 + 1/R_2)/C_1 & 1/(R_2 C_1) \\ 1/(R_2 C_2) & -1/(R_2 C_2) \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

donde $x_1(t)$ y $x_2(t)$ son los voltajes en los dos capacitores al tiempo t . El voltaje es la magnitud que da cuenta de la diferencia en el potencial eléctrico entre dos puntos determinados. Aclaremos que un capacitor es un dispositivo capaz de almacenar energía a través de campos eléctricos. Los capacitores se utilizan principalmente como filtros de corriente continua, ya que evitan cambios bruscos y ruidos en las señales debido a su funcionamiento.

En la matriz puede observarse que esta corriente dependerá de las resistencias (R) del circuito en cuestión. Una resistencia es un dispositivo eléctrico que tiene la particularidad de oponerse al flujo de la corriente. Suponga que las resistencias valen para nuestro problema: $R_1 = 1$, $R_2 = 2$, las capacitancias involucradas tendrán el valor de $C_1 = 1$ y $C_2 = 0.5$. Con todos esos datos es posible obtener la matriz del problema, a la que deberá encontrarle sus autovectores y autovalores. De esa manera, haciendo uso de la matriz diagonal pertinente podrá desacoplar el sistema. Podrá construir una familia de soluciones. Para hallar la solución particular del problema necesita conocer los voltajes iniciales que se miden en los capacitores a un tiempo $t = 0$. Para este problema supondremos que las cargas iniciales han sido $x_1(0)=5$ y $x_2(0)=4$.

Su tarea consiste en encontrar las soluciones $x_1(t)$ y $x_2(t)$ que describan cómo los voltajes evolucionan en el tiempo. Tenga en cuenta que en general dado $\vec{y}' = A\vec{y}$ un sistema de ecuaciones diferenciales de primer orden donde la matriz A , de dimensión $n \times n$ es diagonalizable (los vectores $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$ asociados a los valores propios distintos $\lambda_1, \lambda_2, \dots, \lambda_n$ de A , respectivamente, son linealmente independientes). Entonces el conjunto $\{e^{\lambda_1 x} \vec{u}_1, e^{\lambda_2 x} \vec{u}_2, \dots, e^{\lambda_n x} \vec{u}_n\}$ es una base del espacio de soluciones de $\vec{y}' = A\vec{y}$.

Figura 3.5: Circuito



3.7.1 Ejercicios

Ejercicio 3.1

Halle el polinomio característico, autovalores y autovectores de las Matrices de Pauli:

$$\sigma_x) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Las sentencias Python a continuación dan los autovalores de la matriz. Investigue cómo puede hallar los autovectores en Python.

```
import numpy as np
a = np.array([[0, 1],
              [1, 0]])
LA.eigvals(a)
```

Ejercicio 3.2

Sea $T \in L(\mathbb{R}^2)$, dado por $T((x, y)) = (y, x)$. Halle el polinomio característico, autovalores y autovectores. Interprete geoméricamente.

Ejercicio 3.3

Demuestre que si $0 < \theta < \pi$, la matriz

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \text{ no tiene autovalores ni autovectores reales. Interprete geoméricamente.}$$

Ejercicio 3.4

Sea $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ la transformación lineal definida por:

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

Encuentre una base B de \mathbb{R}^3 tal que $(T)_B$ sea diagonal.

Ejercicio 3.5

$$\text{Sea } A = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

la matriz que representa la transformación lineal que proyecta cualquier vector $v \in \mathbb{R}^3$ sobre la recta de vector director $(1, 1, 0)$:

a) Analice si A es semejante sobre el cuerpo \mathbb{R} a una matriz diagonal. En caso afirmativo, halle la matriz diagonal correspondiente.

b) Interprete geoméricamente lo hallado en a).

Ejercicio 3.6

$$\text{Sea } A = \begin{pmatrix} \alpha & \beta & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Indique para qué valores de α y β la matriz es diagonalizable.

Ejercicio 3.7

Sea

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

Analice si A es semejante sobre el cuerpo \mathbb{R} a una matriz diagonal. Idem sobre el cuerpo \mathbb{C} . En caso afirmativo, hallar la matriz diagonal correspondiente.

Ejercicio 3.8

Halle A^{10} , donde $A = \begin{pmatrix} 1 & 3 \\ -3 & -1 \end{pmatrix}$

Deberá encontrar una matriz P que diagonalice a A .

Ejercicio 3.9

Conforme a que $A = T^{-1}BT$ con

$$A = \begin{pmatrix} -3 & -4 & 0 & -2 \\ 8 & 13 & 4 & 8 \\ 4 & 6 & 3 & 4 \\ -12 & -20 & -8 & -13 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & 4 & 1 \\ 1 & 1 & 2 & 1 \\ 2 & 3 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{pmatrix}.$$

Calcule A^6 .

Ejercicio 3.10

Encuentre la solución del sistema

$$\begin{cases} 2y_1 + 2y_2 + y_3 = y'_1 \\ y_1 + 3y_2 + y_3 = y'_2 \\ y_1 + 2y_2 + 2y_3 = y'_3 \end{cases}$$

con las condiciones iniciales $y_1(0) = 0$, $y_2(0) = 1$ y $y_3(0) = 1$.

Ejercicio 3.11

Resuelva la ecuación diferencial homogénea de tercer orden

$$y''' - y' = 0, \text{ con las condiciones iniciales}$$

$$y(0)=1, y'(0) = 0, \text{ e } y''(0) = 1$$

Realizando el cambio

$$z_1 = y, \quad z_2 = y', \quad z_3 = y''.$$

Ejercicio 3.12

Sea $T \in L(\mathbb{R}^3)$, definida por $T((x, y, z)) = (x, x + y, z)$

- Halle el polinomio característico, y el polinomio minimal.
- Halle autovalores y una base para cada espacio propio de T .
- Determine si T es o no diagonalizable.

Ejercicio 3.13

$$\text{Dada } A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

utilice el teorema de Cayley-Hamilton para hallar A^{-1} y A^3 .

Ejercicio 3.14

Utilice las propiedades del polinomio minimal para determinar si las matrices siguientes son diagonalizables o no (considerar sobre el cuerpo \mathbb{R} y sobre el cuerpo \mathbb{C}).

$$\text{a) } A = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}$$

$$\text{b) } A \text{ es una matriz cuadrada tal que } A \neq I \text{ y } A^3 - A^2 + A = I$$

Ejercicio 3.15

Encuentre la forma de Jordan de la matriz

$$\begin{pmatrix} -10 & -7 \\ 7 & 4 \end{pmatrix}$$

Ejercicio 3.16

Escriba todas las matrices de Jordan de 4×4 posibles.

Ejercicio 3.17

Determine las formas de Jordan posibles de una matriz de 4×4 cuyo polinomio característico es $(\lambda + 2)^3 \cdot (\lambda - 3)$.

Ejercicio 3.18

Determine las formas de Jordan posibles de una matriz de 5×5 cuyo polinomio minimal es $(\lambda - 2)^2$.

Ejercicio 3.19

Sea $T \in L(\mathbb{R}^4)$, tal que su polinomio característico es $(\lambda + 1)^2(\lambda - 2)\lambda$:

- Indique los polinomios minimales de T y describa en qué casos es diagonalizable.
- Si T no es diagonalizable, encuentre su forma de Jordan.

Ejercicio 3.20

Dada

$$N_6 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Demuestre que es nilpotente con índice de nilpotencia 6.

Ejercicio 3.21

La matriz

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

tiene polinomio característico $(\lambda + 1)^2(\lambda - 1)^2$, y polinomio minimal $(\lambda + 1)(\lambda - 1)^2$, por lo que no es diagonalizable. Encuentre su forma de Jordan y utilícela para encontrar A^{10} .

Nota: para hallar las potencias de los bloques de Jordan que son de la forma $(\lambda I_m + N)^k$ utilice el binomio de Newton y el hecho que N es una matriz nilpotente.

Ejercicio 3.22

Pruebe la proposición 3.2.1: La intersección y la suma de subespacios invariantes respecto de una aplicación lineal $T \in L(V)$ son subespacios invariantes respecto de T .

Ejercicio 3.23

Dado un cuerpo K , sea $A \in K^{n \times n}$ inversible. Pruebe que los autovalores de A^{-1} son los inversos de los autovalores de A , y que los autovectores correspondientes a autovalores inversos coinciden.

Ejercicio 3.24

Demuestre que dos matrices semejantes B y B' tienen el mismo polinomio característico $\mathbf{P}_{T,B}(\lambda) = \mathbf{P}_{T,B'}(\lambda)$

Ejercicio 3.25

$$\text{Sea } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

a) Demuestre que A es diagonalizable si $(a - d)^2 + 4bc > 0$.

b) Analice el caso que A sea simétrica ($b = c$).

Ejercicio 3.26

Sea D el operador derivación sobre el \mathbb{R} -espacio vectorial de las funciones derivables de \mathbb{R} en \mathbb{R} . Si $k \in \mathbb{Z}, k \neq 0$, demuestre que las funciones $\sin(kx)$ y $\cos(kx)$ son autovectores de D^2 . Indique cuáles son los autovalores correspondientes.

Ejercicio 3.27

Sea $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una transformación lineal con matriz asociada A respecto a la base canónica, \vec{u} y $\vec{v} \in \mathbb{R}^n$ autovectores asociados a los autovalores λ y μ . Indique justificando cuáles de las siguientes afirmaciones son verdaderas

- a) Para todo $\alpha \in \mathbb{R}$ el vector $\alpha\vec{u}$ es un autovector asociado a λ .
- b) Todo vector del núcleo es autovector.

- c) El vector $\vec{w} = \vec{v} + \vec{u}$ es autovector de T .
- d) λ^n es autovalor de T^n con autovector asociado \vec{u} .
- e) Una matriz diagonalizable es invertible.

Ejercicio 3.28

Dado un cuerpo K , sean A y $P \in K^{n \times n}$, P inversible. Demuestre que $(P^{-1}AP)^2 = P^{-1}A^2P$ y $(P^{-1}AP)^k = P^{-1}A^kP$ para k un entero positivo.

Ejercicio 3.29

Sea $T \in L(\mathbb{R}^2)$ la transformación lineal cuya matriz en la base canónica es : $A = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix}$

- a) Demuestre que los únicos subespacios de \mathbb{R}^2 invariantes por T son \mathbb{R}^2 y 0 .
- b) Si U es la misma transformación pero en \mathbb{C}^2 , cuya matriz en la base canónica es A , demuestre que U tiene algún subespacio unidimensional invariante.

Ejercicio 3.30

Sea $T \in L(\mathbb{R}^2)$ la transformación lineal cuya matriz en la base canónica es :

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

y sea W_1 el subespacio de \mathbb{R}^2 generado por $(1, 0)^t$:

- a) Pruebe que W_1 es T -invariante.
- b) Demuestre que no existe un subespacio W_2 que sea invariante tal que $\mathbb{R}^2 = W_1 + W_2$.

3.7.2 Autoevaluación**Verdadero o Falso**

1. Si A es invertible entonces cero no es un valor propio de A .
2. Los valores propios de una matriz triangular son los elementos en la diagonal de la matriz.
3. Si la matriz real $A \in \mathbb{R}^{3 \times 3}$ tiene tres valores propios distintos, entonces los vectores propios correspondientes a esos valores propios constituyen una base para \mathbb{R}^3 .
4. Si la matriz $A \in \mathbb{R}^{3 \times 3}$ tiene dos valores propios distintos, entonces A tiene a lo sumo dos vectores propios linealmente independientes.

5. Si A tiene elementos reales, entonces A puede tener exactamente un valor propio complejo.
6. Si $\text{Det}(A) = 0$, entonces 0 es un valor propio de A .
7. Si una matriz de $n \times n$ tiene n valores propios diferentes, se puede diagonalizar.
8. Si la matriz A de 5×5 tiene 3 valores propios diferentes, entonces A no puede ser semejante a la matriz diagonal.
9. El subespacio propio contiene todos los vectores propios asociados a λ y además al vector nulo.
10. El determinante de una matriz y el de su transpuesta son iguales, por lo tanto tienen el mismo polinomio característico, los mismos valores y vectores propios.
11. La matriz $\lambda I - A$ es invertible entonces λ es un valor propio de A .
12. Dos matrices semejantes tienen el mismo polinomio característico y los mismos valores propios con las mismas multiplicidades algebraicas.
13. Una matriz es diagonalizable si la multiplicidad algebraica de cada valor propio de la matriz, coincide con la dimensión del espacio propio correspondiente.
14. El determinante de una matriz es igual a la suma de todos sus autovalores (reales y complejos, y elevados a sus respectivas multiplicidades).
15. La traza de una matriz es igual al producto de todos sus autovalores (reales y complejos, y elevados a sus respectivas multiplicidades).
16. Las variables ángulo-acción están relacionadas con la diagonalización de matrices simétricas en la mecánica analítica, y corresponden a las coordenadas en el espacio de los autovectores de la matriz Hessiana.



4. Espacios vectoriales con producto interno

Los conceptos geométricos de longitud, distancia y perpendicularidad, que son bien conocidos para \mathbb{R}^2 y \mathbb{R}^3 , se definen en este capítulo para cualquier espacio vectorial euclídeo V . Estos conceptos proporcionan herramientas geométricas potentes para resolver muchos problemas aplicados, incluidos los problemas de mínimos cuadrados. Los tres conceptos se definen en términos del producto escalar o producto interior de dos vectores.

4.1 Producto interno. Ejemplos

Definición 4.1.1 Producto interno

Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} . Un *producto interno* sobre V es una función $\phi : V \times V \rightarrow \mathbb{R}$ (o \mathbb{C}) que cumple:

1. $\phi(\vec{x}, \vec{y}) = \overline{\phi(\vec{y}, \vec{x})}$ para todo $\vec{x}, \vec{y} \in V$
2. $\phi(\vec{x} + \vec{z}, \vec{y}) = \phi(\vec{x}, \vec{y}) + \phi(\vec{z}, \vec{y})$, para todos $\vec{x}, \vec{y}, \vec{z} \in V$
3. $\phi(\alpha\vec{x}, \vec{y}) = \alpha\phi(\vec{x}, \vec{y})$ para todo $\vec{x}, \vec{y} \in V$ y todo $\alpha \in \mathbb{R}$ o \mathbb{C}
4. $\phi(\vec{x}, \vec{x}) > 0$ para todo $\vec{x} \neq 0$



Consecuencias de 1, 2 y 3:

- De 1. y 2. se deduce

$$\phi(\vec{x}, \vec{y} + \vec{z}) = \phi(\vec{x}, \vec{y}) + \phi(\vec{x}, \vec{z}), \text{ para todos } \vec{x}, \vec{y}, \vec{z} \in V.$$

- De 3. y de 1. se deduce que

$$\phi(\vec{x}, \alpha \vec{y}) = \overline{\alpha} \phi(\vec{x}, \vec{y}) \text{ para todo } \vec{x}, \vec{y} \in V \text{ y todo } \alpha \in \mathbb{R}.$$

- De 2. se deduce que

$$\phi(\vec{0} + \vec{y}, \vec{x}) = \phi(\vec{y}, \vec{x}) = \phi(\vec{0}, \vec{x}) + \phi(\vec{y}, \vec{x}), \text{ sí y sólo sí } \phi(\vec{0}, \vec{x}) = 0.$$

Por la propiedad simétrica $\phi(\vec{x}, \vec{0}) = 0$ y, en particular, $\phi(\vec{x}, \vec{x}) = 0$ si $\vec{x} = \vec{0}$.

- **Ejemplo 4.1** Los productos internos en \mathbb{R}^n y \mathbb{C}^n son, respectivamente:

$$\phi(\vec{x}, \vec{y}) = \phi((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)) = x_1 y_1 + x_2 y_2 + \dots + x_n y_n \quad (\mathbb{R}^n)$$

$$\phi(\vec{x}, \vec{y}) = \phi((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)) = x_1 \bar{y}_1 + x_2 \bar{y}_2 + \dots + x_n \bar{y}_n \quad (\mathbb{C}^n)$$

Son los productos internos canónicos. Se deja al lector la verificación de las propiedades 1 – 4 de la Definición 4.1.1 en cada caso. ■



- A un espacio vectorial real (o complejo) V provisto de un producto interno se lo llama *espacio euclídeo*, \mathbb{E} , (respectivamente, *espacio unitario*).
- El producto interno generaliza el producto escalar de los vectores \vec{x} e $\vec{y} \in \mathbb{R}^n$ a un espacio vectorial V cualquiera.
- Si se tiene un producto interno, se anotará $\phi(\vec{x}, \vec{y}) = (\vec{x}, \vec{y})$

■ **Ejemplo 4.2** En el espacio vectorial de las funciones continuas en $[a, b]$, $C([a, b])$, a valores reales, se define

$$\phi : C([a, b]) \times C([a, b]) \rightarrow \mathbb{R}$$

$$\phi(f, g) = \int_a^b f(t)g(t)dt.$$

En el caso de funciones a valores complejos, se define $\phi(f, g) = \int_a^b f(t)\overline{g(t)}dt$ (similar a \mathbb{C}^2). ■



- El producto interno anterior es el que se utiliza para hallar los coeficientes de la serie de Fourier de una función $f(x)$ en $[0, 2\pi]$. Se calculan con el producto interno entre la función $f(x)$ y la base ortogonal $\{e^{inx}\}_{n \in \mathbb{Z}}$ o $\{1, \cos(nx), \sin(nx)\}_{n \in \mathbb{N}}$.
- La serie de Fourier tiene importantes aplicaciones, por ejemplo, en el procesamiento de señales. Permite la descomposición de la señal en una base ortonormal y obtener sus componentes frecuenciales. En señales de música, por ejemplo, posibilita separar los instrumentos.

Una vez fijada una base de V , si V es un espacio vectorial de dimensión finita con un producto interno, es posible construir una matriz asociada al producto interno y a dicha base.

4.2 Matriz de un producto Interno

Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} de dimensión finita con producto interno y sea $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ una base de V . Se define la *matriz del producto interno* (\cdot, \cdot) en la base B como la matriz $\in \mathbb{R}^{n \times n}$ (resp. $\in \mathbb{C}^{n \times n}$) tal que

$$P_{ij} = (\vec{v}_i, \vec{v}_j) \quad 1 \leq i, j \leq n$$

Esta matriz nos permite calcular el producto interno entre cualquier par de vectores. Si $\vec{x} = \sum_{i=1}^n x_i \vec{v}_i$ e $\vec{y} = \sum_{j=1}^n y_j \vec{v}_j$

$$(\vec{x}, \vec{y}) = \left(\sum_{i=1}^n x_i \vec{v}_i, \sum_{j=1}^n y_j \vec{v}_j \right) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j (\vec{v}_i, \vec{v}_j)$$

En particular, para $n = 3$, se tiene

$$(\vec{x}, \vec{y}) = (x_1, x_2, x_3) \begin{pmatrix} (\vec{v}_1, \vec{v}_1) & (\vec{v}_1, \vec{v}_2) & (\vec{v}_1, \vec{v}_3) \\ (\vec{v}_2, \vec{v}_1) & (\vec{v}_2, \vec{v}_2) & (\vec{v}_2, \vec{v}_3) \\ (\vec{v}_3, \vec{v}_1) & (\vec{v}_3, \vec{v}_2) & (\vec{v}_3, \vec{v}_3) \end{pmatrix} \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_3 \end{pmatrix}$$

i Si P es la matriz de un producto interno, entonces $P_{ij} = \bar{P}_{ji}$ para todo $i \neq j$. Sin embargo, esa condición no es suficiente para que P sea la matriz de un producto interno. Por ejemplo, la matriz

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

no puede ser la matriz de un producto interno en una base, ya que si \vec{v} es el primer vector de la base, se tendría $(\vec{v}, \vec{v}) = 0$ y sería el vector nulo.

4.3 Longitud, ángulos, distancia y ortogonalidad

A partir de la definición de un producto interno, es posible generalizar las nociones de longitud, ángulos, distancia y ortogonalidad ya vistas para vectores de \mathbb{R}^2 y \mathbb{R}^3 .

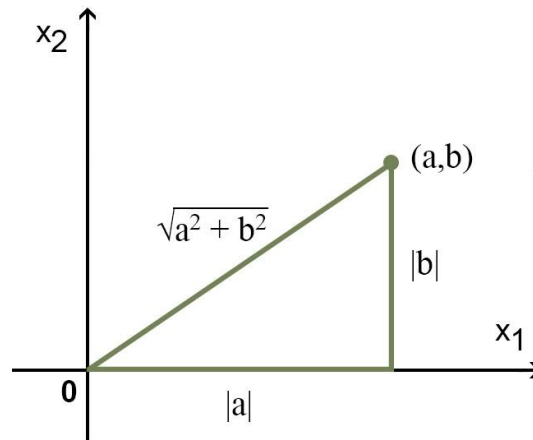
Definición 4.3.1 Longitud o norma de un vector \vec{x} en un espacio con producto interno se define como

$$\|\vec{x}\| = \sqrt{(\vec{x}, \vec{x})}, \quad \vec{x} \in \mathbb{E} \quad (4.1)$$

■ **Ejemplo 4.3** En \mathbb{R}^2 , si $\vec{v} = (a, b)$, $\|\vec{v}\|$ es la longitud del segmento que va desde el origen hasta \vec{v} , y es consecuencia del Teorema de Pitágoras (ver Figura 4.1). ■

■ **Ejemplo 4.4** Sea $\vec{x} = (1, -2, 2, 0)$. Como $\|\vec{x}\|^2 = (1)^2 + (-2)^2 + (2)^2 = 9$, su longitud o norma euclídea es $\|\vec{x}\| = \sqrt{9} = 3$. ■

Figura 4.1: La norma es la longitud del vector



- La definición de longitud tiene sentido por la propiedad 4. del producto interno (Definición 4.1.1).
- Se tiene la propiedad siguiente:

$$\|\alpha \vec{x}\| = \sqrt{(\alpha \vec{x}, \alpha \vec{x})} = \sqrt{\alpha^2 (\vec{x}, \vec{x})} = |\alpha| \sqrt{(\vec{x}, \vec{x})} = |\alpha| \|\vec{x}\|$$

- Todo vector de longitud 1 se dice *unitario*; todo vector \vec{x} no nulo de un espacio euclídeo puede *normalizarse*, es decir, hacerlo unitario multiplicándolo por $\frac{1}{\|\vec{x}\|}$.

■ **Ejemplo 4.5** El vector unitario $\vec{u} \in \mathbb{R}^4$ que tiene la misma dirección que el vector \vec{x} del Ejemplo 4.4 es

$$\vec{u} = \frac{\vec{x}}{\|\vec{x}\|} = \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}, 0\right) \quad (4.2)$$

■

Ángulo entre dos vectores.

En \mathbb{R}^2 el producto escalar verifica la expresión que sigue:

$$(\vec{x}, \vec{y}) = \|\vec{x}\| \|\vec{y}\| \cos \theta \quad (4.3)$$

La verificación para \mathbb{R}^3 es similar. Cuando $n > 3$, puede usarse la Ec.(4.3) para definir el ángulo entre dos vectores de \mathbb{R}^n , o en espacios vectoriales cualesquiera.

Dados dos vectores \vec{x} e \vec{y} de un espacio euclídeo, definimos el *coseno* del ángulo entre ellos como

$$\cos(\theta) = \frac{(\vec{x}, \vec{y})}{\|\vec{x}\| \|\vec{y}\|} \quad (4.4)$$

Para que tenga sentido la definición anterior, es necesario demostrar que el valor absoluto del cociente

$$\frac{(\vec{x}, \vec{y})}{\|\vec{x}\| \|\vec{y}\|}$$

sea menor que o igual que 1.



En Estadística, el valor de $\cos(\theta)$ definido mediante la Ec.(4.4) para los vectores \vec{x} e \vec{y} es llamado *coeficiente de correlación entre los vectores \vec{x} e \vec{y}* , y mide de alguna forma la similitud entre ambos.

Augustin Louis Cauchy (1789 - 1857)

Fue un matemático francés, miembro de la Academia de Ciencias de Francia y profesor en la Escuela politécnica. Cauchy ha sido uno de los matemáticos más prolíficos de todos los tiempos, solo superado por Leonhard Euler, Paul Erdős y Arthur Cayley con cerca de 800 publicaciones y siete trabajos; su investigación cubre el conjunto de áreas matemáticas de la época. Fue pionero en análisis donde se le debe la introducción de las funciones holomorfas, los criterios de convergencia de series y las series de potencias. Sus trabajos sobre permutaciones fueron precursores de la teoría de grupos, contribuyendo de manera medular a su desarrollo. En óptica se le atribuyen trabajos sobre la propagación de ondas electromagnéticas. [6]

Proposición 4.3.1 Desigualdad de Cauchy-Schwartz:

En todo espacio vectorial con producto interno,

$$|(\vec{x}, \vec{y})| \leq \|\vec{x}\| \|\vec{y}\|$$

para todo $\vec{x}, \vec{y} \in \mathbb{E}$.

Demostración:

Si $\vec{y} = \vec{0}$ vale, pues $|\langle \vec{x}, \vec{0} \rangle| = 0 \leq \|\vec{x}\| \cdot 0 = 0$.

Si $\vec{y} \neq \vec{0}$, y llamamos $a = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{y}\|^2}$ donde $\|\vec{y}\|^2 = \langle \vec{y}, \vec{y} \rangle$, se tiene que

$$0 \leq \langle \vec{x} - a\vec{y}, \vec{x} - a\vec{y} \rangle$$

$$0 \leq \langle \vec{x}, \vec{x} - a\vec{y} \rangle - a\langle \vec{y}, \vec{x} - a\vec{y} \rangle$$

$$0 \leq \langle \vec{x}, \vec{x} \rangle - \overline{a\langle \vec{x}, \vec{y} \rangle} - a\langle \vec{y}, \vec{x} \rangle + a\overline{a\langle \vec{y}, \vec{y} \rangle}$$

como

$$\overline{a\langle \vec{x}, \vec{y} \rangle} = \frac{\overline{\langle \vec{y}, \vec{x} \rangle}}{\|\vec{y}\|^2} \langle \vec{x}, \vec{y} \rangle$$

$$a\langle \vec{y}, \vec{x} \rangle = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{y}\|^2} \langle \vec{y}, \vec{x} \rangle$$

y

$$a\overline{a\langle \vec{y}, \vec{y} \rangle} = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{y}\|^2} \frac{\overline{\langle \vec{y}, \vec{x} \rangle}}{\|\vec{y}\|^2} \langle \vec{y}, \vec{y} \rangle = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{y}\|^2} \overline{\langle \vec{y}, \vec{x} \rangle}$$

se cancelan el segundo y el cuarto término de la desigualdad y queda

$$0 \leq \langle \vec{x}, \vec{x} \rangle - a\langle \vec{y}, \vec{x} \rangle$$

$$0 \leq \langle \vec{x}, \vec{x} \rangle - \overline{a\langle \vec{x}, \vec{y} \rangle}$$

$$0 \leq \langle \vec{x}, \vec{x} \rangle - \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{y}\|^2} \overline{\langle \vec{x}, \vec{y} \rangle}$$

Entonces,

$$0 \leq \|\vec{x}\|^2 - \frac{|\langle \vec{x}, \vec{y} \rangle|^2}{\|\vec{y}\|^2}$$

que equivale a

$$0 \leq \frac{\|\vec{x}\|^2 \|\vec{y}\|^2 - |\langle \vec{x}, \vec{y} \rangle|^2}{\|\vec{y}\|^2}$$

En el numerador se tiene la desigualdad que se quería demostrar.

□



Se acredita a Cauchy la desigualdad para vectores y a Schwarz para los productos escalares

con integrales. Sin embargo, fue Bunyakovsky quien demostró y publicó la desigualdad de Schwarz en una monografía, 25 años antes que Schwarz.

4.4 Distancia entre vectores

Definición 4.4.1 Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} con producto interno. Se define la distancia $d, d : V \times V \rightarrow \mathbb{R}$ como:

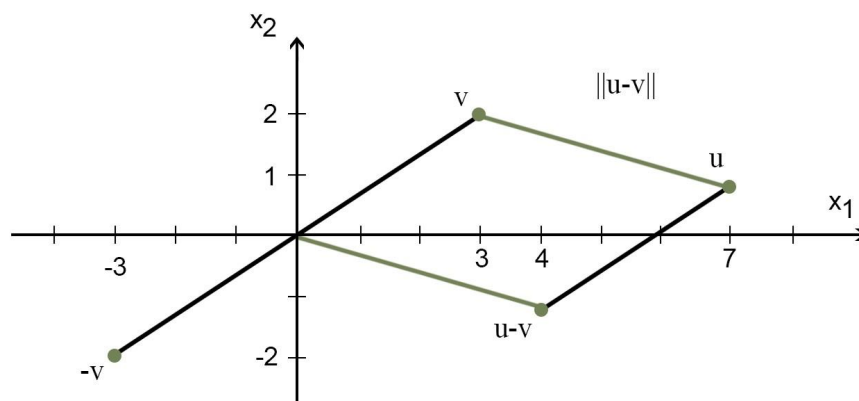
$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|.$$

Usando las propiedades de la norma, se puede verificar que d satisface:

1. $d(\vec{x}, \vec{y}) \geq 0$ para todo $\vec{x}, \vec{y} \in E$
2. $d(\vec{x}, \vec{y}) = 0$ si y sólo si $\vec{x} = \vec{y}$
3. $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$ para todo $\vec{x}, \vec{y} \in E$
4. $d(\vec{x}, \vec{z}) \leq d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z})$ para todo $\vec{x}, \vec{y}, \vec{z} \in E$.

Karl Herman Amandus Schwarz (1843 - 1921)

Fue un matemático alemán conocido por su trabajo en análisis complejo. Schwarz inicialmente estudio química en Berlín pero Kummer y Weierstrass lo persuadieron para que se hiciera matemático. Entre 1867 y 1869 trabajó en Halle, después en Zürich. Desde 1875 trabajó en el universidad de Gotinga, tratando los temas de teoría de funciones, geometría diferencial y cálculo de variaciones. Su memoria en ocasión del 70 aniversario de Weierstrass contiene, entre otros temas importantes, la desigualdad para integrales que hoy se conoce como desigualdad de Schwarz. [14]

Figura 4.2: Distancia entre los vectores \vec{u} y \vec{v} 

Viktor Yakovlevich Bunyakovsky (1804 - 1889)

Nació en Ucrania. Estudió matemáticas en la Sorbona, en la que se doctoró en 1825 bajo la tutoría de Augustin Cauchy. En 1826 volvió a San Petersburgo donde ejerció como profesor de la Escuela de Cadetes de la Academia Naval y del Instituto de Comunicaciones. De 1846 a 1880 fue profesor en la Universidad de San Petersburgo.

Entre otros campos de las matemáticas, Buniakovski trabajó sobre todo en teoría de números, análisis matemático y en teoría de la probabilidad. Son relevantes las aportaciones que llevan su nombre como la conjetura de Buniakovski (nunca demostrada) y la desigualdad de Cauchy-Buniakovski-Schwarz. Sus aportaciones más originales son en teoría de la probabilidad, acerca de la cual publicó numerosos artículos sobre el estudio de problemas estadísticos de la población de Rusia. [17]

■ **Ejemplo 4.6** En la Figura 4.2 se muestran los vectores \vec{u} , \vec{v} y $\vec{u} - \vec{v}$,

$$\vec{u} - \vec{v} = \begin{pmatrix} 7 \\ 1 \end{pmatrix} - \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$$

$$\|\vec{u} - \vec{v}\| = \sqrt{4^2 + (-1)^2} = \sqrt{17}.$$

Puede observarse que la distancia de \vec{v} a \vec{u} es la misma que la de $\vec{u} - \vec{v}$ a $\vec{0}$, y también que si se suma el vector $\vec{u} - \vec{v}$ a \vec{v} se obtiene el vector \vec{u} . ■



- Dados dos vectores \vec{x} e \vec{y} , se dice que $d(\vec{x}, \vec{y})$ es la *distancia* entre \vec{x} e \vec{y} .
- Una distancia es una función que verifica las 4 propiedades anteriores. Puede no provenir de ninguna norma.

Con la definición que sigue se generaliza la noción de perpendicularidad entre vectores de un espacio vectorial.

Definición 4.4.2 Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} con producto interno. Dos vectores \vec{x}, \vec{y} se dicen *ortogonales* (o *perpendiculares*), si

$$(\vec{x}, \vec{y}) = 0 \quad (4.5)$$



Por las propiedades vistas en las observaciones i en la Sección 4.1, el vector nulo, es ortogonal a todo vector de V .

Teorema 4.4.1 Teorema de Pitágoras

Dos vectores \vec{x} e \vec{y} son ortogonales, sii

$$\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$$

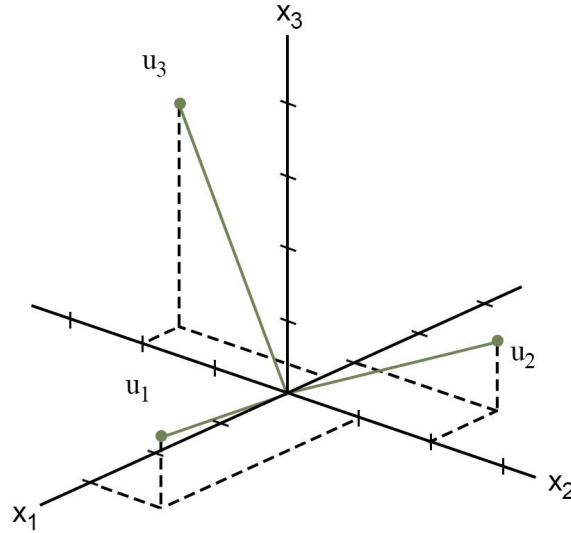
La demostración se deja al lector.

Definición 4.4.3 Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} de dimensión finita con producto interno. Se dice que $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\} \subset V$ es un conjunto ortogonal si $(\vec{v}_i, \vec{v}_j) = 0$ para todo $i \neq j$.

■ **Ejemplo 4.7** El conjunto $S = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, donde $\vec{u}_1 = (3, 1, 1)^T$, $\vec{u}_2 = (-1, 2, 1)^T$ y $\vec{u}_3 = (-1/2, -2, 7/2)^T$, es un conjunto ortogonal, ya que al considerar los tres pares posibles de vectores,

$\{\vec{u}_1, \vec{u}_2\}$, $\{\vec{u}_1, \vec{u}_3\}$, y $\{\vec{u}_2, \vec{u}_3\}$, se tiene

Figura 4.3: $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es un conjunto ortogonal de \mathbb{R}^3




$$\vec{u}_1 \cdot \vec{u}_2 = 3(-1) + 1(2) + 1(1) = 0$$

$$\vec{u}_1 \cdot \vec{u}_3 = 3(-1/2) + 1(-2) + 1(7/2) = 0$$

$$\vec{u}_2 \cdot \vec{u}_3 = -1(-1/2) + 2(-2) + 1(7/2) = 0$$

Cada par de vectores distintos es ortogonal, así que $S = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, es un conjunto ortogonal como se muestra en la Figura 4.3. ■

 Un conjunto de r vectores se dice *ortonormal* si es ortogonal y $\|\vec{u}_i\| = 1$ para cada $1 \leq i \leq r$.

Proposición 4.4.2 Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} de dimensión finita con producto interno y sea $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\} \subset V$ un conjunto ortogonal de V con $\vec{v}_i \neq 0$ para $1 \leq i \leq r$. Entonces $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\}$ es un conjunto de vectores linealmente independiente.

Demostración:

Supongamos que $\sum_{i=1}^r \alpha_i \vec{v}_i = \vec{0}$. Entonces, para cada j , $1 \leq j \leq r$,

$$0 = (\vec{0}, \vec{v}_j) = \left(\sum_{i=1}^r \alpha_i \vec{v}_i, \vec{v}_j \right) = \sum_{i=1}^r \alpha_i (\vec{v}_i, \vec{v}_j) = \alpha_j \|\vec{v}_j\|^2,$$

y como $\vec{v}_j \neq \vec{0}$, se tiene que $\alpha_j = 0$ para $1 \leq j \leq r$ y entonces, $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\}$ es un conjunto de vectores linealmente independiente. □

En el teorema siguiente se muestra por qué una base ortogonal es más conveniente que otras bases ya que las coordenadas de un vector en esa base pueden calcularse muy fácilmente.

Proposición 4.4.3 Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} de dimensión finita con producto interno y sea $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\} \subset V$ es un conjunto ortogonal de V con $\vec{v}_i \neq 0$ para $1 \leq i \leq r$. Sea $\vec{v} \in \langle \vec{v}_1, \vec{v}_2, \dots, \vec{v}_r \rangle$. Entonces

$$\vec{v} = \sum_{j=1}^r \frac{(\vec{v}, \vec{v}_j)}{\|\vec{v}_j\|^2} \vec{v}_j \quad (4.6)$$

Demostración:

Si $\vec{v} = \sum_{i=1}^r \alpha_i \vec{v}_i$, para cada j , $1 \leq j \leq r$, se tiene que

$$(\vec{v}, \vec{v}_j) = \left(\sum_{i=1}^r \alpha_i \vec{v}_i, \vec{v}_j \right) = \sum_{i=1}^r \alpha_i (\vec{v}_i, \vec{v}_j) = \alpha_j (\vec{v}_j, \vec{v}_j) = \alpha_j \|\vec{v}_j\|^2,$$

y como $\vec{v}_j \neq \vec{0}$, se tiene entonces que

$$\alpha_j = \frac{(\vec{v}, \vec{v}_j)}{\|\vec{v}_j\|^2}$$
□

■ **Ejemplo 4.8** El conjunto $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ del Ejemplo 4.7 es una base ortogonal para \mathbb{R}^3 . Si se desea expresar el vector $\vec{y} = (6, 1, -8)^T$ como una combinación lineal de los vectores en S , de acuerdo a la Ec.(4.6) se tiene que,

$$\vec{y} = \sum_{j=1}^3 \frac{(\vec{y}, \vec{u}_j)}{\|\vec{u}_j\|^2} \vec{u}_j$$

Para hallar las coordenadas de \vec{y} en la base ortogonal, se calculan los productos escalares

$$\vec{y} \cdot \vec{u}_1 = 11, \vec{y} \cdot \vec{u}_2 = -12, \vec{y} \cdot \vec{u}_3 = -33$$

$$\text{y } \vec{u}_1 \cdot \vec{u}_1 = 11, \vec{u}_2 \cdot \vec{u}_2 = 6, \vec{u}_3 \cdot \vec{u}_3 = 33/2$$

$$\frac{(\vec{y}, \vec{u}_1)}{\|\vec{u}_1\|^2} \vec{u}_1 = \frac{11}{11} (3, 1, 1) = (3, 1, 1)$$

$$\frac{(\vec{y}, \vec{u}_2)}{\|\vec{u}_2\|^2} \vec{u}_2 = \frac{-12}{6} (-1, 2, 1) = 2(-1, 2, 1)$$

$$\frac{(\vec{y}, \vec{u}_3)}{\|\vec{u}_3\|^2} \vec{u}_3 = \frac{-33}{33/2} (-1/2, -2, 7/2) = 2(-1/2, -2, 7/2)$$

Y se obtiene,

$$\vec{y} = 1\vec{u}_1 - 2\vec{u}_2 - 2\vec{u}_3$$

que puede verificarse fácilmente,

$$\vec{y} = (6, 1, -8) = (3, 1, 1) - 2(-1, 2, 1) - 2(-1/2, -2, 7/2). \quad \blacksquare$$



- Como se vio en el Ejemplo 4.8, es muy fácil calcular las coordenadas de un vector \vec{y} en una base ortogonal. En otro caso, se debe que resolver un sistema de ecuaciones lineales para hallarlas.
- Si el conjunto además, es ortonormal, se tiene

$$\vec{v} = \sum_{j=1}^r (\vec{v}, \vec{v}_j) \vec{v}_j$$

La proposición que sigue asegura que en todo espacio vectorial de dimensión finita con producto interno tiene bases ortonormales. Más aún, en la demostración se da un procedimiento recursivo conocido como Gram-Schmidt que permite obtener una base ortonormal del espacio vectorial a partir de una base cualquiera del mismo.

Proposición 4.4.4 Método de ortonormalización de Gram-Schmidt

Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} de dimensión finita con producto interno y sea $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ una base de V . Existe una base ortonormal $B = \{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_n\}$ de V tal que

$$\langle \vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \rangle = \langle \vec{w}_1, \vec{w}_2, \dots, \vec{w}_k \rangle$$

para todo $1 \leq k \leq n$

Demostración:

Se construyen los vectores $\{\vec{z}_1, \vec{z}_2, \dots, \vec{z}_n\}$ de una base ortogonal, recursivamente

1. Se toma $\vec{z}_1 = \vec{v}_1$
2. Se busca $\vec{z}_2 \in V$ tal que $(\vec{z}_2, \vec{z}_1) = 0$ y tal que $\langle \vec{z}_1, \vec{z}_2 \rangle = \langle \vec{v}_1, \vec{v}_2 \rangle$

La segunda condición vale sí y sólo sí $\vec{z}_2 = a\vec{v}_1 + b\vec{v}_2$ con $b \neq 0$. Es posible considerar $b = 1$ y buscar a para que se cumpla la primera condición:

$$0 = (\vec{z}_2, \vec{z}_1) = (a\vec{v}_1 + b\vec{v}_2, \vec{z}_1) = a(\vec{v}_1, \vec{v}_1) + (\vec{v}_2, \vec{v}_1),$$

lo que implica

$$a = \frac{-(\vec{v}_2, \vec{v}_1)}{\|\vec{v}_1\|^2}.$$

Luego, el vector,

$$\vec{z}_2 = \vec{v}_2 - \frac{(\vec{v}_2, \vec{v}_1)}{\|\vec{v}_1\|^2} \vec{v}_1 = \vec{v}_2 - \frac{(\vec{v}_2, \vec{v}_1)}{\|\vec{v}_1\|^2} \vec{z}_1$$

satisface las condiciones.

Supongamos contruídos $\vec{z}_1, \vec{z}_2, \dots, \vec{z}_r \in V$ tales que

1. $(\vec{z}_i, \vec{z}_j) = 0$ cuando $i \neq j$
2. $\langle \vec{z}_1, \vec{z}_2, \dots, \vec{z}_r \rangle = \langle \vec{v}_1, \vec{v}_2, \dots, \vec{v}_r \rangle$ con $1 \leq k \leq r$

consideramos el vector

$$\vec{z}_{r+1} = \vec{v}_{r+1} - \sum_{i=1}^r \frac{(\vec{v}_{r+1}, \vec{z}_i)}{\|\vec{z}_i\|^2} \vec{z}_i$$

Se tiene que

- $\langle \vec{z}_1, \vec{z}_2, \dots, \vec{z}_r, \vec{z}_{r+1} \rangle = \langle \vec{v}_1, \vec{v}_2, \dots, \vec{v}_r, \vec{v}_{r+1} \rangle$ con $1 \leq k \leq r$
- para cada $j \leq r$, reemplazando \vec{z}_{r+1} y teniendo en cuenta 1.,

$$(\vec{z}_{r+1}, \vec{z}_j) = (\vec{v}_{r+1} - \sum_{i=1}^r \frac{(\vec{v}_{r+1}, \vec{z}_i)}{\|\vec{z}_i\|^2} \vec{z}_i, \vec{z}_j)$$

$$(\vec{z}_{r+1}, \vec{z}_j) = (\vec{v}_{r+1}, \vec{z}_j) - \frac{(\vec{v}_{r+1}, \vec{z}_j)}{\|\vec{z}_j\|^2} (\vec{z}_j, \vec{z}_j) = 0$$

Luego \vec{z}_{r+1} satisface las condiciones requeridas.

De esta manera, al concluir el n -ésimo paso, se obtiene una base ortogonal $\{\vec{z}_1, \vec{z}_2, \dots, \vec{z}_n\}$ de V que además satisface

$$\langle \vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \rangle = \langle \vec{z}_1, \vec{z}_2, \dots, \vec{z}_k \rangle$$

para todo $1 \leq k \leq n$.

Finalmente, para cada $1 \leq i \leq n$ consideramos el vector $\vec{w}_i = \frac{\vec{z}_i}{\|\vec{z}_i\|}$. Luego, el conjunto $B = \{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_n\}$ resulta una base de V que cumple lo pedido.

□

Corolario Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} de dimensión finita con producto interno y sea S un subespacio de V , $S \neq \vec{0}$. Entonces existe una base ortonormal de V que contiene una base ortonormal de S . Se demuestra tomando una base de S , completando a una base de V y aplicando a esta base el procedimiento de Gram-Schmidt. ■

■ Ejemplo 4.9 Aplicación del Método de Gram-Schmidt

Dada la base $B = \{(1, 0, i), (1, 1, 2 + i), (0, 0, 1)\}$ de \mathbb{C}^3 se desea hallar una base ortonormal con G-S.

$$\vec{v}_1 = (1, 0, i), \vec{v}_2 = (1, 1, 2 + i) \text{ y } \vec{v}_3 = (0, 0, 1)$$

$$\vec{z}_1 = \vec{v}_1$$

$$\vec{z}_2 = \vec{v}_2 - \frac{(\vec{v}_2, \vec{z}_1)}{\|\vec{z}_1\|^2} \vec{z}_1$$

$$\vec{z}_2 = (1, 1, 2 + i) - \frac{((1, 1, 2 + i), (1, 0, i))}{\|(1, 0, i)\|^2} (1, 0, i)$$

$$\vec{z}_2 = (1, 1, 2 + i) - (1 - i)(1, 0, i) = (i, 1, 1)$$

y luego,

$$\vec{z}_3 = \vec{v}_3 - \frac{(\vec{v}_3, \vec{z}_1)}{\|\vec{z}_1\|^2} \vec{z}_1 - \frac{(\vec{v}_3, \vec{z}_2)}{\|\vec{z}_2\|^2} \vec{z}_2$$

$$\vec{z}_3 = (i/6, -1/3, 1/6)$$

$\{\vec{z}_1, \vec{z}_2, \vec{z}_3\}$ resulta una base de ortogonal de \mathbb{C}^3

Dividiendo por su norma queda una base ortonormal $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$

$$\text{donde } \vec{w}_1 = (1/\sqrt{2}, 0, i/\sqrt{2})$$

$$\vec{w}_2 = (i/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$$

$$\vec{w}_3 = (\sqrt{6}i/6, -\sqrt{6}/3, \sqrt{6}/6)$$

■

La existencia de bases ortogonales para subespacios de dimensión finita de un espacio con producto interior puede establecerse por medio del proceso Gram-Schmidt, de igual forma que en \mathbb{R}^n . Al aplicar este proceso, es posible plantear ciertas bases ortogonales que surgen con frecuencia en las aplicaciones y construir la proyección ortogonal de un vector sobre un subespacio S . La proyección no depende de la selección de la base ortogonal y tiene muy buenas propiedades que se describirán más adelante. En el teorema que sigue se ve cómo es posible escribir la matriz de una transformación lineal usando producto interno para escribir las coordenadas de la imagen de cada vector de la base ortogonal.

Teorema 4.4.5 Si T es una transformación lineal sobre V donde V es un espacio vectorial con producto interno y de dimensión finita, entonces

$$(T)_B = (T(\vec{u}_j), \vec{u}_i)_{ij}$$

siendo $B = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ cualquier base ortonormal de V .

Demostración:

$$T(\vec{u}_1) = k_1\vec{u}_1 + k_2\vec{u}_2 + \dots + k_n\vec{u}_n.$$

En la primera columna de la matriz deben ir las coordenadas de $T(\vec{u}_1)$, o sea k_1, \dots, k_n .

y resulta que las coordenadas son


$$\begin{aligned} (T(\vec{u}_1), \vec{u}_i) &= (k_1\vec{u}_1 + k_2\vec{u}_2 + \dots + k_n\vec{u}_n, \vec{u}_i) \\ &= k_1(\vec{u}_1, \vec{u}_i) + k_2(\vec{u}_2, \vec{u}_i) + \dots + k_n(\vec{u}_n, \vec{u}_i) = k_i \end{aligned}$$

□

4.4.1 Complemento Ortogonal

Definición 4.4.4 Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} con producto interno y sea S un subespacio de V . Se define el complemento ortogonal de S como

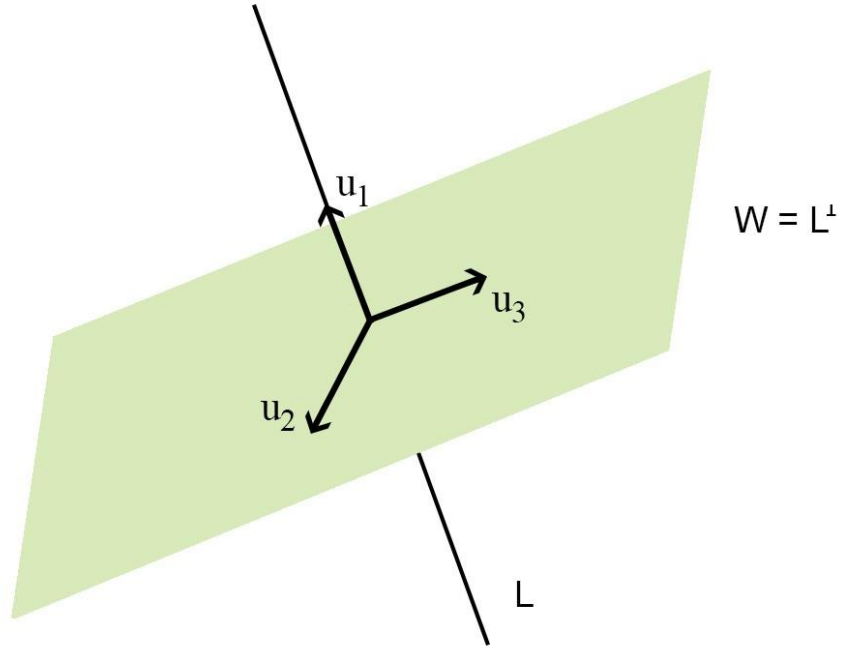
$$S^\perp = \{\vec{v} \in V \mid (\vec{v}, \vec{s}) = 0 \quad \forall \vec{s} \in S\}$$

 S^\perp es un subespacio de V .

■ **Ejemplo 4.10** Para el subespacio de \mathbb{R}^2 generado por el vector $(1, 1)$, su complemento ortogonal es $\langle (1, 1) \rangle^\perp = \{(x, y) \in \mathbb{R}^2, ((x, y), (1, 1)) = 0\} = \{(x, y) \in \mathbb{R}^2, x + y = 0\} = \langle (1, -1) \rangle$. ■

■ **Ejemplo 4.11** Sea W un plano que pasa por el origen en \mathbb{R}^3 , y sea L la recta que pasa por el origen y es perpendicular a W . Si \vec{u}_1 y \vec{u}_2 son diferentes de $\vec{0}$, \vec{u}_1 está sobre L , y \vec{u}_2 está en W , $\vec{u}_1 \cdot \vec{u}_2 = 0$, como se muestra en la Figura 4.4. Así que cada vector sobre L es ortogonal a cada vector \vec{w} en W . De hecho, L consiste en todos los vectores que son ortogonales a los \vec{w} en W , y W consiste en todos los vectores ortogonales a los vectores en L . Es decir, $L = W^\perp$ y $W = L^\perp$. ■

Figura 4.4: Complemento ortogonal



■ **Ejemplo 4.12** En \mathbb{C}^3 hallar el complemento ortogonal de $\langle (1, i, 1+i) \rangle$.

$$\begin{aligned} \langle (1, i, 1+i) \rangle^\perp &= \{ (x_1, x_2, x_3) \in \mathbb{C}^3, (x_1, x_2, x_3) \cdot (\alpha, \alpha i, \alpha(1+i)) = 0 \quad \forall \alpha \in \mathbb{C} \} \\ &= \{ (x_1, x_2, x_3) \in \mathbb{C}^3, x_1 \bar{\alpha} + x_2 \bar{\alpha} i + x_3 \bar{\alpha} (1+i) = 0 \quad \forall \alpha \in \mathbb{C} \} \end{aligned}$$

De donde

$$x_1 \bar{\alpha} - x_2 \bar{\alpha} i + x_3 \bar{\alpha} (1-i) = 0$$

o

$$\bar{\alpha} (x_1 - x_2 i + x_3 (1-i)) = 0.$$

Se tiene, entonces, $x_1 = x_2 i - x_3 (1-i)$ y resulta

$$\langle (1, i, 1+i) \rangle^\perp = \{ (x_2 i - x_3 (1-i), x_2, x_3) = x_2 (i, 1, 0) + x_3 (i-1, 0, 1) \}$$

■

Teorema 4.4.6 Sea A una matriz de $m \times n$. El complemento ortogonal del espacio fila de A , $FilA$, (subespacio de \mathbb{R}^n que generan los vectores filas de A) es el espacio nulo de A , $Nul(A)$ y el complemento ortogonal del espacio columna de A , $ColA$, (subespacio de \mathbb{R}^m que generan los vectores columnas de A) es el espacio nulo de la matriz A^T . Es decir,

$$(FilA)^\perp = Nul(A) \text{ y } (ColA)^\perp = Nul(A^T)$$

■ **Ejemplo 4.13** Dada la matriz

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{pmatrix}, \quad (4.7)$$

■

se describen los espacios $FilA$, $Nul(A)$ y $ColA$:

$Nul(A)$ son los vectores de \mathbb{R}^3 soluciones del sistema homogéneo. Esos vectores son perpendiculares a las filas de la matriz A , y pertenecen entonces a subespacio $(FilA)^\perp$.

Si ahora se resuelve el sistema homogéneo con la matriz A^T , da el vector nulo de \mathbb{R}^2 . El subespacio que generan las columnas de A , $ColA$ es todo \mathbb{R}^2 , pues hay en las columnas 2 vectores linealmente independientes. De ahí que el subespacio ortogonal, $(ColA)^\perp = Nul(A^T) = \vec{0}$.

Proposición 4.4.7 Sea V un espacio vectorial de dimensión finita con producto interno y sea $S \subseteq V$ un subespacio. Entonces

1. $S \cap S^\perp = \vec{0}$
2. $\dim(S) + \dim(S^\perp) = \dim(V)$

En consecuencia, $S \oplus S^\perp$.

Demostración:

- Sea $\vec{v} \in S \cap S^\perp$. Como $\vec{v} \in S^\perp$, $(\vec{v}, \vec{s}) = 0$, $\forall \vec{s} \in S$. En particular para $\vec{s} = \vec{v}$, entonces $(\vec{v}, \vec{v}) = \|\vec{v}\|^2 = 0$, de donde $\vec{v} = \vec{0}$.
- Sea $\{\vec{s}_1, \vec{s}_2, \dots, \vec{s}_r\}$ una base de S . Existen $\vec{v}_{r+1}, \dots, \vec{v}_n$ tales que $B = \{\vec{s}_1, \vec{s}_2, \dots, \vec{s}_r, \vec{v}_{r+1}, \dots, \vec{v}_n\}$ es una base de V . Aplicando Gram-Schmidt se obtiene una base ortonormal de V , $B' = \{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_r, \vec{w}_{r+1}, \dots, \vec{w}_n\}$ tal que

$$\langle \vec{w}_1, \vec{w}_2, \dots, \vec{w}_r \rangle = \langle \vec{s}_1, \vec{s}_2, \dots, \vec{s}_r \rangle = S.$$

Sea $j > r$. Veamos que $\vec{w}_j \in S^\perp$. Dado $\vec{s} \in S$, existen escalares $\alpha_1, \dots, \alpha_r$, tales que $\vec{s} = \sum_{i=1}^r \alpha_i \vec{w}_i$, entonces

$$(\vec{w}_j, \vec{s}) = (\vec{w}_j, \sum_{i=1}^r \alpha_i \vec{w}_i) = \sum_{i=1}^r \alpha_i (\vec{w}_j, \vec{w}_i) = 0.$$

como la base es ortonormal y $j > r$, $(\vec{w}_j, \vec{w}_i) = 0$ para $1 \leq i \leq r$. De donde, $\vec{w}_j \in S^\perp$, y se tiene que,

$$\{\vec{w}_{r+1}, \dots, \vec{w}_n\} \in S^\perp,$$

y, por lo tanto,

$$\langle \vec{w}_{r+1}, \dots, \vec{w}_n \rangle \subseteq S^\perp,$$

por ser S^\perp un subespacio.


$\dim(S^\perp) \geq \dim(\langle \vec{w}_{r+1}, \dots, \vec{w}_n \rangle) = n - r = n - \dim(S)$. Entonces, $\dim(S^\perp) + \dim(S) \geq n$.

Por otro lado como $S \cap S^\perp = \{\vec{0}\}$

$$\dim(S^\perp) + \dim(S) = \dim(S^\perp + S) \leq \dim(V) = n.$$

Entonces $\dim(S^\perp) + \dim(S) = \dim(V)$

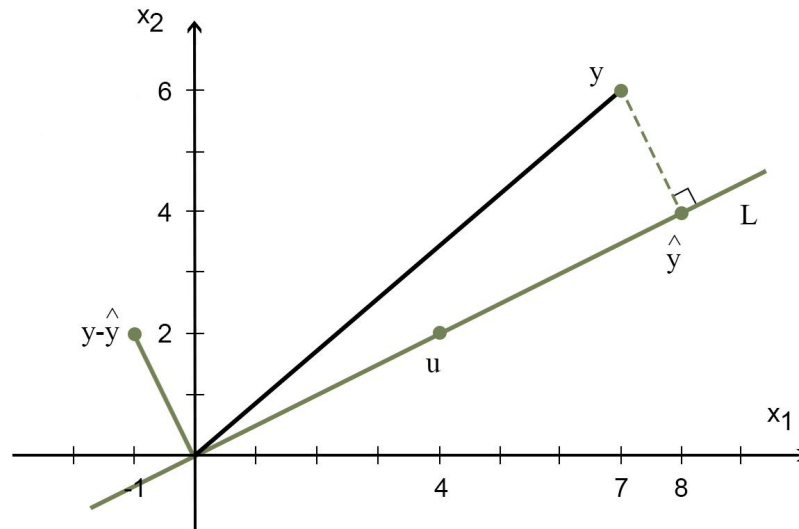
□

 Del teorema sale cómo generar el subespacio S^\perp , a partir de la base ortonormal de V .

$$S^\perp = \langle \vec{w}_{r+1}, \dots, \vec{w}_n \rangle$$

Proposición 4.4.8 Sea V un espacio vectorial de dimensión finita con producto interno y sea S un subespacio de V . Entonces $(S^\perp)^\perp = S$

Figura 4.5: El vector \hat{y} es la proyección del vector y sobre la recta L



Demostración:

Por definición, $(S^\perp)^\perp = \{\vec{v} \in V / (\vec{v}, \vec{w}) = 0 \quad \forall \vec{w} \in S^\perp\}$. Veamos que $S \subseteq (S^\perp)^\perp$. Sea $\vec{s} \in S$. Para cada $\vec{w} \in S^\perp$ se tiene que $(\vec{s}, \vec{w}) = \overline{(\vec{w}, \vec{s})} = 0$, de donde se deduce que $\vec{s} \in (S^\perp)^\perp$. \square

4.5 Proyección ortogonal

Dado un subespacio S de un espacio vectorial V de dimensión finita con producto interno. Como $S \oplus S^\perp = V$ se puede considerar el proyector $P_S : V \rightarrow V$ cuya imagen es S y cuyo núcleo es S^\perp .

Definición 4.5.1 Sea V un espacio vectorial de dimensión finita con producto interno y sea $S \subseteq V$ un subespacio. Se define la *proyección ortogonal* sobre S como la transformación lineal $P_S : V \rightarrow V$ que satisface

1. $P_S(\vec{s}) = \vec{s} \quad \forall \vec{s} \in S$
2. $P_S(\vec{t}) = \vec{0} \quad \forall \vec{t} \in S^\perp$

■ **Ejemplo 4.14** En \mathbb{R}^2 se desea hallar la proyección ortogonal de un vector \vec{y} , \hat{y} sobre el subespacio

S generado por otro vector \vec{u} , o sea sobre la recta L generada por \vec{u} que pasa por el origen. Esto se muestra en la Figura 4.5.

$$P_S(\vec{y}) = \vec{\hat{y}} = c\vec{u},$$

y se debe cumplir que \vec{u} sea ortogonal al vector $\vec{y} - \vec{\hat{y}}$, es decir, $(\vec{y} - c\vec{u}) \cdot \vec{u} = 0$, entonces,

$$\vec{y} \cdot \vec{u} - c\vec{u} \cdot \vec{u} = 0$$

de donde

$$c = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}$$

y entonces, la proyección sobre L es ,

$$\vec{\hat{y}} = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u} \quad (4.8)$$

En el caso de la Figura 4.5, se tiene que $\vec{y} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$ y $\vec{u} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$.

Usando Ec.(4.8), como $\vec{y} \cdot \vec{u} = 40$ y $\vec{u} \cdot \vec{u} = 20$, se obtiene,

$$\vec{\hat{y}} = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u} = \frac{40}{20} \vec{u} = 2\vec{u} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$$

y la componente ortogonal a \vec{u} es $\vec{y} - \vec{\hat{y}} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

La descomposición de \vec{y} , como suma de proyecciones sobre S y sobre S^\perp , es

$$\begin{pmatrix} 7 \\ 6 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

■



- Si $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r, \vec{v}_{r+1}, \dots, \vec{v}_n\}$ una base ortonormal de V tal que $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\}$ es una base de S y $\{\vec{v}_{r+1}, \dots, \vec{v}_n\}$ una base de S^\perp , la proyección ortogonal sobre S es la única transformación lineal $P_S : V \rightarrow V$ que satisface

1. $P_S(\vec{v}_i) = \vec{v}_i \quad \forall 1 \leq i \leq r$
2. $P_S(\vec{v}_i) = \vec{0} \quad \forall r+1 \leq i \leq n$

En consecuencia, para $\vec{v} \in V$, recordando que $\vec{v} = \sum_{j=1}^n (\vec{v}, \vec{v}_j) \vec{v}_j$ resulta,

$$P_S(\vec{v}) = P_S\left(\sum_{j=1}^n (\vec{v}, \vec{v}_j) \vec{v}_j\right) = \sum_{j=1}^n (\vec{v}, \vec{v}_j) P_S(\vec{v}_j) = \sum_{j=1}^r (\vec{v}, \vec{v}_j) \vec{v}_j, \quad (4.9)$$

que es una expresión para $P_S(\vec{v})$ en términos de los vectores de la base ortonormal de S .

- Sea V un espacio vectorial de dimensión finita con producto interno y sea $S \subseteq V$ un subespacio. Entonces $P_S + P_{S^\perp} = id_V$,

donde, $P_S(\vec{v}) = \sum_{j=1}^r (\vec{v}, \vec{v}_j) \vec{v}_j$ y $P_{S^\perp}(\vec{v}) = \sum_{j=r+1}^n (\vec{v}, \vec{v}_j) \vec{v}_j$.

- **Ejemplo 4.15** Si $V = \mathbb{R}^4$ y W es el subespacio

$$W = \{\vec{x} \in \mathbb{R}^4, x_1 - 2x_2 + x_3 = x_1 - 3x_2 + x_4 = 0\},$$

se desea hallar la proyección ortogonal del vector $\vec{v} = (4, 8, -4, 12)$ sobre el subespacio W .

Para usar la expresión de la Ec.(4.9) debemos hallar una base ortonormal de W . En primer lugar calculamos una base de W resolviendo el sistema por eliminación de Gauss:

$$\begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & -3 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{pmatrix} \quad (4.10)$$

A partir de la matriz escalonada, al resolver el sistema homogéneo, quedan como variables independientes x_3 y x_4 y una base de W es $\{(-3, -1, 1, 0), (2, 1, 0, 1)\}$. Una base ortogonal de W a partir de aplicar Gram-Schmidt es,

$$\left\{ (-3, -1, 1, 0), \left(\frac{1}{11}, \frac{4}{11}, \frac{7}{11}, 1 \right) \right\}.$$

Si llamamos $\vec{w}_1 = (-3, -1, 1, 0)$ y $\vec{w}_2 = (\frac{1}{11}, \frac{4}{11}, \frac{7}{11}, 1)$, los vectores de la base ortonormal \vec{v}_i se obtienen dividiéndolos por su norma, $\vec{v}_i = \frac{\vec{w}_i}{\|\vec{w}_i\|}$.

Entonces, como $P_W(\vec{v}) = \sum_{j=1}^2 (\vec{v}, \vec{v}_j) \vec{v}_j$, se calculan los productos escalares (\vec{v}, \vec{v}_j) , y se obtiene que

$$P_W((4, 8, -4, 12)) = \left(\frac{124}{17}, \frac{88}{17}, \frac{52}{17}, \frac{140}{17}\right)$$

es el vector del subespacio W más cercano a $\vec{v} = (4, 8, -4, 12)$. ■

Teorema 4.5.1 Teorema de la proyección ortogonal.

Sea S un subespacio de un espacio vectorial con producto interno, V . Entonces, para cada $\vec{v} \in V$, el vector de S a menor distancia de \vec{v} es $P_S(\vec{v})$.

Demostración:

Si $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r, \vec{v}_{r+1}, \dots, \vec{v}_n\}$ una base ortonormal de V tal que $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\}$ es una base de S .

Sea $\vec{v} \in V$. Se tiene que $\vec{v} = \sum_{j=1}^n (\vec{v}, \vec{v}_j) \vec{v}_j$ y $P_S(\vec{v}) = \sum_{j=1}^r (\vec{v}, \vec{v}_j) \vec{v}_j$. Por otro lado, si $\vec{s} \in S$, $\vec{s} = \sum_{j=1}^r (\vec{s}, \vec{v}_j) \vec{v}_j$.

Entonces

$$\vec{v} - \vec{s} = \sum_{j=1}^n (\vec{v}, \vec{v}_j) \vec{v}_j - \sum_{j=1}^r (\vec{s}, \vec{v}_j) \vec{v}_j = \sum_{j=1}^r (\vec{v} - \vec{s}, \vec{v}_j) \vec{v}_j + \sum_{j=r+1}^n (\vec{v}, \vec{v}_j) \vec{v}_j$$

de donde,

$$\|\vec{v} - \vec{s}\|^2 = \sum_{j=1}^r |(\vec{v} - \vec{s}, \vec{v}_j)|^2 + \sum_{j=r+1}^n |(\vec{v}, \vec{v}_j)|^2 \geq \sum_{j=r+1}^n |(\vec{v}, \vec{v}_j)|^2 = \|\vec{v} - P_S(\vec{v})\|^2$$

(el primer término de la desigualdad se anula cuando $\vec{s} = P_S(\vec{v})$) □



- El teorema anterior es conocido también como el teorema de la mejor aproximación.
- En la Figura 4.5, \vec{y} es el punto de L más cercano a \vec{y} , en el sentido que

$$\|\vec{y} - \vec{y}\| \leq \|\vec{y} - \vec{v}\|,$$

para todo \vec{v} en L distinto de \vec{y} .

4.5.1 Problema de cuadrados mínimos

El problema de hallar la proyección de un vector \mathbf{b} sobre un subespacio surge cuando se tiene el problema $A\mathbf{x} = \mathbf{b}$, con A una matriz de $m \times n$, donde m es la cantidad de observaciones y es mucho mayor que la cantidad de incógnitas n , de forma tal que se espera que el sistema $A\mathbf{x} = \mathbf{b}$ sea incompatible. En otras palabras, el vector \mathbf{b} no es combinación lineal de los vectores columna de A (no está en el espacio columna de A). Se trata entonces de hallar $\hat{\mathbf{x}}$ que minimice el error, y esto se realizará en el sentido de los cuadrados mínimos. El error es $E = \|A\mathbf{x} - \mathbf{b}\|$, y es la distancia de \mathbf{b} al vector $A\mathbf{x}$ en el espacio columna. El vector \mathbf{p} del espacio columna más próximo a \mathbf{b} que cualquier otro es la *proyección de \mathbf{b} sobre el espacio columna*. El error $\mathbf{e} = \mathbf{b} - A\hat{\mathbf{x}}$ es perpendicular al espacio columna. Recordando que el espacio nulo de la matriz A es el conjunto de vectores de \mathbb{R}^n que son perpendiculares a todas las filas de A , el error \mathbf{e} pertenecerá al espacio nulo de la matriz A^T (es perpendicular al espacio columna). Es decir que el error \mathbf{e} es perpendicular a cada columna de A (ver Figura 4.6). Entonces se tiene que,

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0 \quad \text{o} \quad A^T A\hat{\mathbf{x}} = A^T \mathbf{b} \quad (4.11)$$

Las ecuaciones (4.11) se conocen como *ecuaciones normales*. Pueden obtenerse a partir de buscar \mathbf{x} que minimiza E^2 , tomando derivadas parciales de $E^2 = (A\mathbf{x} - \mathbf{b})^T (A\mathbf{x} - \mathbf{b})$. Al igualar a cero, se tiene $2A^T A\mathbf{x} - 2A^T \mathbf{b} = 0$.

Si $A^T A$ tiene inversa (esto ocurre cuando las columnas de A son linealmente independientes), entonces

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b} \quad (4.12)$$

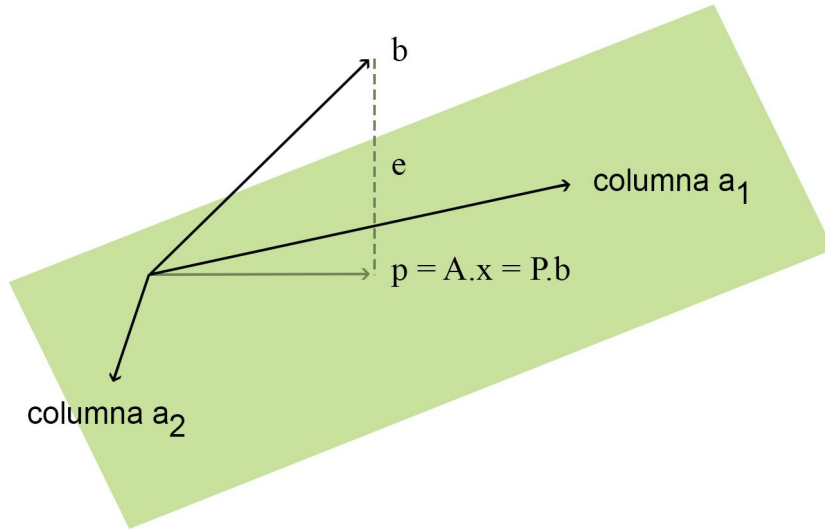
y la proyección \mathbf{p} de \mathbf{b} sobre el espacio columna, es el vector $A\hat{\mathbf{x}}$

$$\mathbf{p} = A\hat{\mathbf{x}} = A(A^T A)^{-1} A^T \mathbf{b} \quad (4.13)$$

Un ejemplo de ajuste de datos por cuadrados mínimos

Supongamos realizamos un experimento en el que se espera que la salida \mathbf{b} sea una función lineal de la entrada \mathbf{t} . Se buscará la recta $\mathbf{b} = \alpha + \beta \mathbf{t}$. Por ejemplo, si a diferentes tiempos medimos la distancia a un satélite en su recorrido a Marte. En este caso \mathbf{t} es el tiempo y \mathbf{b} la distancia y el satélite se moverá con una velocidad casi constante ($\mathbf{b} = \mathbf{b}_0 + \mathbf{v}\mathbf{t}$). ¿Es posible calcular α y β ? Si

Figura 4.6: Proyección sobre el espacio columna



no hay errores experimentales dos mediciones determinan la recta $\mathbf{b} = \alpha + \beta \mathbf{t}$. Pero si hay error se deberá promediar los experimentos y hallar la *mejor* recta. Al realizar m mediciones,

$$\alpha + \beta t_1 = b_1$$

$$\alpha + \beta t_2 = b_2$$

...

$$\alpha + \beta t_m = b_m$$

Se tendrá un sistema *sobredeterminado*, con m ecuaciones y solo 2 incógnitas. Si las mediciones tienen error, el sistema no tiene solución. En este caso la matriz A tiene dos columnas y $\mathbf{x} = (\alpha, \beta)^T$:

$$\begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ \dots \\ 1 & t_m \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \dots \\ b_m \end{pmatrix} \quad (4.14)$$

o $A\mathbf{x} = \mathbf{b}$. La mejor recta se tendrá con $\hat{\mathbf{x}} = (\hat{\alpha}, \hat{\beta})$ que minimizan

$$E^2 = \|A\mathbf{x} - \mathbf{b}\|^2 = (b_1 - \alpha - \beta t_1)^2 + (b_2 - \alpha - \beta t_2)^2 + \dots + (b_m - \alpha - \beta t_m)^2$$

El vector $\mathbf{p} = A\hat{\mathbf{x}}$ es el más cercano a \mathbf{b} . De todas las rectas $\mathbf{b} = \alpha + \beta t$ estamos eligiendo la que mejor ajusta los datos. Los errores son las distancias verticales $\mathbf{b} - \alpha - \beta t$ (no perpendiculares). Estas distancias verticales se elevan al cuadrado, se suman y se minimizan.

■ Ejemplo 4.16 ■

Si se considera como ejemplo, que se tienen tres mediciones: $b_1 = 1$ en $t_1 = -1$, $b_2 = 1$ en $t_2 = 1$ y $b_3 = 3$ en $t_3 = 2$ se tiene el sistema $Ax = b$

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \quad (4.15)$$

El sistema es incompatible porque los puntos no están sobre una misma recta. Se resuelve entonces, por cuadrados mínimos, $A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$

$$\begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix} \quad (4.16)$$

La solución es $\hat{\alpha} = \frac{9}{7}$, $\hat{\beta} = \frac{4}{7}$ y la mejor recta es $\frac{9}{7} + \frac{4}{7}t$.

Como se muestra en la Figura 4.7, el vector \mathbf{b} no es combinación lineal de las columnas $(1, 1, 1)^T$ y $(-1, 1, 2)^T$. Con mínimos cuadrados se reemplaza \mathbf{b} que no está en la recta por el vector $\mathbf{p} = A\hat{\mathbf{x}}$ que sí está, al no poder resolver $A\mathbf{x} = \mathbf{b}$, se resuelve $A\hat{\mathbf{x}} = \mathbf{p}$. El vector $\mathbf{p} = (\frac{5}{7}, \frac{13}{7}, \frac{17}{7})$ está en el espacio columna, es la proyección en ese subespacio. Restando \mathbf{p} de \mathbf{b} , los errores son $\mathbf{e} = \mathbf{b} - \mathbf{p} = (\frac{2}{7}, -\frac{6}{7}, \frac{4}{7})$. Son los errores verticales en la Figura 4.8. Ese vector \mathbf{e} , como se muestra en la Figura 4.7, es ortogonal a las columnas de A (está en el espacio nulo de A^T).

Para el caso de m mediciones b_1, b_2, \dots, b_m en puntos distintos t_1, t_2, \dots, t_m , la recta $\alpha + \beta t$ que minimiza E^2 , surge de resolver el sistema lineal

$$A^T A \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = A^T b \quad (4.17)$$

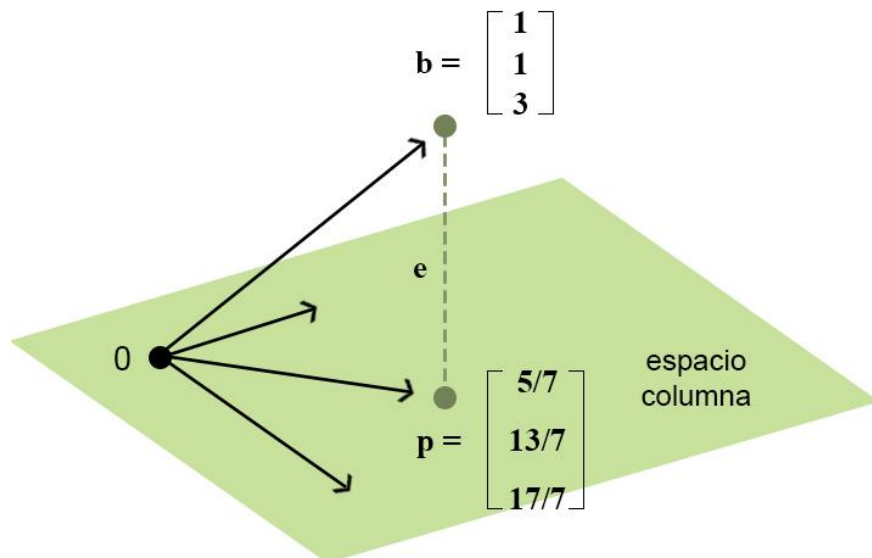
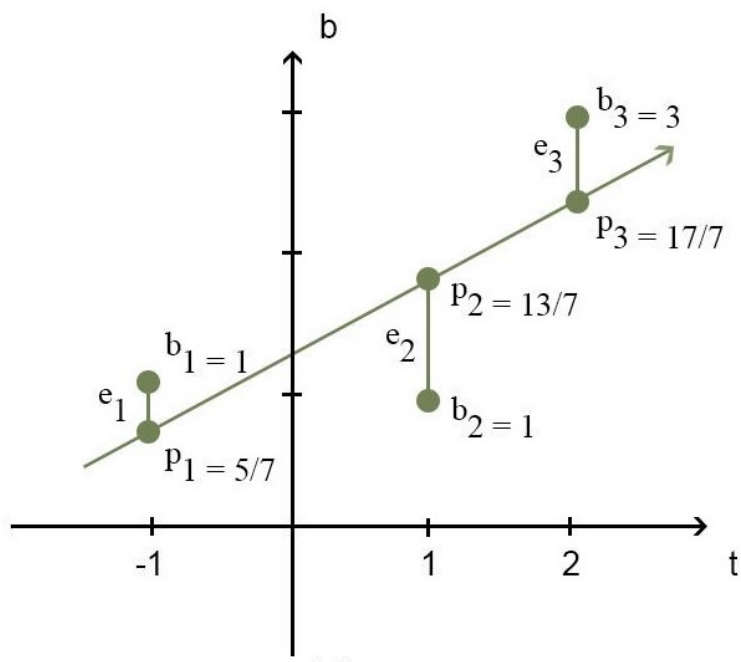


Figura 4.7: Proyección del vector \mathbf{b} en el espacio columna

Figura 4.8: Recta que ajusta por mínimos cuadrados los datos del ejemplo



$$\begin{pmatrix} m & \sum_{i=1}^m t_i \\ \sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2 \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m b_i \\ \sum_{i=1}^m t_i b_i \end{pmatrix} \quad (4.18)$$

-  Es importante notar que el método de cuadrados mínimos no está limitado a ajustar datos con una recta. En muchos casos interesan otros ajustes, con polinomios de grado más alto o con otras funciones como es el caso de ajuste exponencial o el ajuste con senos y cosenos. Pueden conducir a problemas lineales o a problemas no lineales de cuadrados mínimos, siendo estos últimos más complejos de abordar.
-  El método por mínimos cuadrados fue inventado por Karl Friedrich Gauss, y lo usó para resolver un problema de astronomía. En 1801 el asteroide Ceres se había observado mucho más brillante durante más de un mes antes de desaparecer cuando se acercó al Sol. Con base en las observaciones disponibles, los astrónomos deseaban aproximar la órbita de Ceres para observarlo de nuevo cuando se alejara del sol. Gauss empleó los mínimos cuadrados e impactó a la comunidad científica al predecir la hora y el lugar correctos (unos 10 meses después) para localizar el asteroide.

Karl Friedrich Gauss (1777 - 1855)

Fue un matemático, astrónomo y físico alemán que contribuyó significativamente en muchos ámbitos, incluida la teoría de números, el análisis matemático, la geometría diferencial, la estadística, el álgebra, la geodesia, el magnetismo y la óptica.

Gauss pronto fue reconocido como un niño prodigio, pese a provenir de una familia campesina de padres con poca cultura: su madre sabía leer, aunque no escribir; su padre sí, pero en cuanto a las matemáticas, no pasaba de la aritmética más elemental. De Carl Friedrich Gauss existen muchas anécdotas acerca de su asombrosa precocidad. Completó su magnum opus, *Disquisitiones arithmeticae*, a los veintiún años (1798), aunque la obra no se publicó hasta 1801. Constituye un trabajo fundamental como consolidación de la teoría de los números y ha moldeado esta área hasta los días presentes. [8]

4.6 Endomorfismos de espacios vectoriales con producto interno

Vamos a asociar ahora a cada endomorfismo f de un espacio vectorial V de dimensión finita y con producto interno otra transformación lineal $f^*, f^* : V \rightarrow V$.

Definición 4.6.1 Sea V un espacio vectorial con producto interno y sea f una transformación lineal. Se llama *adjunta* de f y se anota f^* a una transformación lineal $f^* : V \rightarrow V$ tal que

$$(f(\vec{v}), \vec{w}) = (\vec{v}, f^*(\vec{w})) \quad \forall \vec{v}, \vec{w} \in V \quad (4.19)$$

■ **Ejemplo 4.17** Sea $f : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ con el producto interno canónico, dada por $f(x, y) = ((x + iy, 2x - (1 + i)y)$

Se tiene que

$$(f(x, y), (z, w)) = ((x + iy, 2x - (1 + i)y), (z, w)) = ((x, y), (z + 2w, -iz + (-1 + i)w))$$

de donde, $f^* : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ definida por $f^*(z, w) = (z + 2w, -iz + (-1 + i)w)$ satisface $(f(x, y), (z, w)) = ((x, y), f^*(z, w))$ para todo par de vectores $(x, y), (z, w)$ en \mathbb{C}^2 . ■

El resultado que sigue prueba que en espacios vectoriales con producto interno de dimensión finita, la transformación lineal adjunta existe y es única.

Proposición 4.6.1 Sea V un espacio vectorial de dimensión finita con producto interno y sea f una transformación lineal, $f : V \rightarrow V$. Entonces existe una única transformación lineal $f^* : V \rightarrow V$ tal que

$$(f(\vec{v}), \vec{w}) = (\vec{v}, f^*(\vec{w}))$$

Demostración:

Unicidad

Para ver la unicidad supongamos existen transformaciones lineales $g^* : V \rightarrow V$ y $h^* : V \rightarrow V$ tales que, para \vec{w} fijo y para cada \vec{v}

$$(f(\vec{v}), \vec{w}) = (\vec{v}, g^*(\vec{w})) \quad (f(\vec{v}), \vec{w}) = (\vec{v}, h^*(\vec{w}))$$

entonces

$(\vec{v}, g(\vec{w})) = (\vec{v}, h(\vec{w}))$ o equivalentemente, $(\vec{v}, g(\vec{w}) - h(\vec{w})) = \vec{0}$, para todo $\vec{v} \in V$; tomando $\vec{v} = g(\vec{w}) - h(\vec{w})$, se tiene $(g(\vec{w}) - h(\vec{w}), g(\vec{w}) - h(\vec{w})) = 0$ y por la propiedad del producto escalar $g(\vec{w}) - h(\vec{w}) = 0$, entonces $g(\vec{w}) = h(\vec{w})$, para cada $\vec{w} \in V$, con lo cual g y h coinciden.

Existencia

Sea $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ una base ortonormal de V . Si existe $f^* : V \rightarrow V$ con las condiciones del enunciado, debe cumplirse, para cada $\vec{w} \in V$

$$\begin{aligned} f^*(\vec{w}) &= \sum_{i=1}^n (f^*(\vec{w}), \vec{v}_i) \vec{v}_i \\ &= \sum_{i=1}^n \overline{(\vec{v}_i, f^*(\vec{w}))} \vec{v}_i \\ &= \sum_{i=1}^n \overline{(f(\vec{v}_i), \vec{w})} \vec{v}_i = \sum_{i=1}^n (\vec{w}, f(\vec{v}_i)) \vec{v}_i \end{aligned}$$

Se define entonces, $f^* : V \rightarrow V$

$$f^*(\vec{w}) = \sum_{i=1}^n (\vec{w}, f(\vec{v}_i)) \vec{v}_i \quad (4.20)$$

f^* es una transformación lineal

- Usando la definición Ec.(4.20),

$$\begin{aligned} f^*(\vec{w}_1 + \vec{w}_2) &= \sum_{i=1}^n (\vec{w}_1 + \vec{w}_2, f(\vec{v}_i)) \vec{v}_i \\ &= \sum_{i=1}^n (\vec{w}_1, f(\vec{v}_i)) \vec{v}_i + (\vec{w}_2, f(\vec{v}_i)) \vec{v}_i \\ &= \sum_{i=1}^n (\vec{w}_1, f(\vec{v}_i)) \vec{v}_i + \sum_{i=1}^n (\vec{w}_2, f(\vec{v}_i)) \vec{v}_i \\ &= f^*(\vec{w}_1) + f^*(\vec{w}_2) \end{aligned}$$

- Para $\lambda \in \mathbb{C}$ (o \mathbb{R}) $\vec{w} \in V$

$$f^*(\lambda \vec{w}) = \sum_{i=1}^n (\lambda \vec{w}, f(\vec{v}_i)) \vec{v}_i$$

$$= \lambda \sum_{i=1}^n (\vec{w}, f(\vec{v}_i)) \vec{v}_i = \lambda f^*(\vec{w})$$

Veamos que para todo $\vec{v}, \vec{w} \in V$ vale $(f(\vec{v}), \vec{w}) = (\vec{v}, f^*(\vec{w}))$

Sean $\vec{v}, \vec{w} \in V$. Se tiene que $\vec{v} = \sum_{i=1}^n (\vec{v}, \vec{v}_i) \vec{v}_i$ y entonces, $f(\vec{v}) = \sum_{i=1}^n (\vec{v}, \vec{v}_i) f(\vec{v}_i)$

$$\begin{aligned} (f(\vec{v}), \vec{w}) &= \left(\sum_{i=1}^n (\vec{v}, \vec{v}_i) f(\vec{v}_i), \vec{w} \right) \\ &= \sum_{i=1}^n (\vec{v}, \vec{v}_i) (f(\vec{v}_i), \vec{w}) \end{aligned}$$

Por otro lado

$$\begin{aligned} (\vec{v}, f^*(\vec{w})) &= \left(\sum_{i=1}^n (\vec{v}, \vec{v}_i) \vec{v}_i, \sum_{j=1}^n (\vec{w}, f(\vec{v}_j)) \vec{v}_j \right) \\ &= \sum_{i=1}^n (\vec{v}, \vec{v}_i) \left(\vec{v}_i, \sum_{j=1}^n (\vec{w}, f(\vec{v}_j)) \vec{v}_j \right) \\ &= \sum_{i=1}^n (\vec{v}, \vec{v}_i) \sum_{j=1}^n \overline{(\vec{w}, f(\vec{v}_j))} (\vec{v}_i, \vec{v}_j) \\ &= \sum_{i=1}^n (\vec{v}, \vec{v}_i) \overline{(\vec{w}, f(\vec{v}_i))} = \sum_{i=1}^n (\vec{v}, \vec{v}_i) (f(\vec{v}_i), \vec{w}) \end{aligned}$$

Concluimos entonces que se cumple $(f(\vec{v}), \vec{w}) = (\vec{v}, f^*(\vec{w}))$

□

A partir de la matriz de una transformación lineal $f : V \rightarrow V$ en una base ortonormal de V , puede obtenerse fácilmente la matriz de su adjunta en la misma base.

Proposición 4.6.2 Sea V un espacio vectorial de dimensión finita con producto interno y sea f una transformación lineal, $f : V \rightarrow V$. Sea B una base ortonormal de V . Entonces, la matriz que representa la transformación adjunta es la conjugada y transpuesta de la matriz de la transformación f , es decir,

$$(f^*)_B = ((f)_B)^* \quad (4.21)$$

Demostración:

Supongamos que $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ es una base ortonormal de V . Entonces para cada $1 \leq i, j \leq n$,

$$((f^*)_B)_{ij} = (f^*(\vec{v}_j), \vec{v}_i)$$

como en cada columna j van las coordenadas de $f^*(\vec{v}_j)$ en la base B ,

$$= \overline{(\vec{v}_i, f^*(\vec{v}_j))} = \overline{(f(\vec{v}_i), \vec{v}_j)} = \overline{((f)_B)_{ji}} = ((f)_B)_{ij}^*$$

□

■ **Ejemplo 4.18** En el caso de la transformación lineal adjunta del Ejemplo 4.17, si B es la base canónica de \mathbb{C}^2 , se tiene, de acuerdo a la proposición anterior, que la matriz que la representa es:

$$(f)_B = \begin{pmatrix} 1 & i \\ 2 & -1-i \end{pmatrix} \quad \text{y} \quad (f^*)_B = \begin{pmatrix} 1 & 2 \\ -i & -1+i \end{pmatrix}$$

$(f^*)_B$ es la matriz transpuesta y conjugada de $(f)_B$.

■

Existe el caso particular de transformaciones lineales $f: V \rightarrow V$ cuya adjunta f^* coincide con f .

Definición 4.6.2 Sea V un espacio vectorial con producto interno y sea $f: V \rightarrow V$ una transformación lineal. Se dice que f es *autoadjunta* si $f = f^*$ o sea, tal que

$$(f(\vec{v}), \vec{w}) = (\vec{v}, f(\vec{w})) \quad \forall \vec{v}, \vec{w} \in V \quad (4.22)$$

Definición 4.6.3 Una matriz $A \in \mathbb{R}^{n \times n}$ se dice *simétrica* si $A_{ij} = A_{ji} \forall 1 \leq i, j \leq n$, o equivalentemente, si $A = A^T$. Una matriz $A \in \mathbb{C}^{n \times n}$ se dice *hermitiana* si $A_{ij} = \overline{A_{ji}} \forall 1 \leq i, j \leq n$, o equivalentemente, si $A = A^*$.

Si A es la matriz de una transformación lineal f en una base ortonormal, sabemos que A^* es la matriz de la transformación adjunta en la misma base. Si f es autoadjunta se tiene $A = A^*$, por lo tanto la matriz de una transformación lineal autoadjunta es simétrica (hermítica).

Si f es autoadjunta, entonces, es una transformación lineal diagonalizable. Más aún, existe una base ortonormal de V formada por autovectores de f y todos sus autovalores son reales.

Proposición 4.6.3 Sea V un espacio vectorial de dimensión finita con producto interno. Sea f una transformación lineal autoadjunta. Entonces el polinomio característico de f tiene todas sus raíces reales.



Si $A \in \mathbb{C}^{n \times n}$ es una matriz hermitiana, entonces todas las raíces del polinomio característico de A son reales.

El que sigue es un resultado importante sobre diagonalización de transformaciones lineales autoadjuntas.

Proposición 4.6.4 Sea V un espacio vectorial de dimensión finita con producto interno. Sea $f : V \rightarrow V$ una transformación lineal autoadjunta. Entonces existe una base ortonormal B de V tal que la matriz $(f)_B$ es diagonal real.



- Sea $A \in \mathbb{R}^{n \times n}$ es una matriz simétrica. Si se considera el producto interno canónico en \mathbb{R}^n , la transformación lineal $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $f_A(\vec{x}) = A\vec{x}$ es autoadjunta. Por la Proposición 4.6.4 existe una base ortonormal B de \mathbb{R}^n tal que $(f_A)_B = D$, donde D es una diagonal real. En este caso, si E es la base canónica de \mathbb{R}^n , $(f_A)_E = A$, $D = (P_{B,E})^{-1}A(P_{B,E})$, y $(P_{B,E})^{-1} = (P_{B,E})^t$.
- Análogamente, si $A \in \mathbb{C}^{n \times n}$ es una matriz hermitiana. Si se considera \mathbb{C}^n con el producto interno canónico, la transformación lineal $f_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$, definida por $f_A(\vec{x}) = A\vec{x}$ es autoadjunta, y si E es la base canónica de \mathbb{C}^n , $(f_A)_E = A$. Por la proposición anterior existe una base ortonormal B de \mathbb{C}^n tal que $(f_A)_B = D$, donde D es una diagonal real. Entonces $(P_{E,B})^{-1}A(P_{E,B})$, donde, por lo anterior $(P_{E,B})^{-1} = (P_{E,B})^*$.

Esto anterior nos lleva a la siguiente definición

Definición 4.6.4 Una matriz O se dice *ortogonal* si es invertible y $O^{-1} = O^t$. Una matriz $U \in \mathbb{C}^{n \times n}$ se dice *unitaria* si es invertible y $U^{-1} = U^*$.

■ **Ejemplo 4.19** Dada la matriz A ,

$$A = \begin{pmatrix} 1 & 1-i \\ 1+i & 0 \end{pmatrix}, \quad (4.23)$$

se desea hallar una matriz unitaria U tal que U^*AU sea diagonal ■


Si en el programa Octave escribimos $[U, D] = \text{eig}(A)$ nos devuelve las matrices U y D (con edición de 4 dígitos):

$$U = \begin{pmatrix} 0.4082 - 0.4082i & 0.5774 - 0.5774i \\ -0.8165 & 0.5774 \end{pmatrix}, \quad D = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \quad (4.24)$$

Se verifican $U^*U = I$ y $U^*AU = D$

Resumiendo

1. Sea $A \in \mathbb{R}^{n \times n}$ una matriz simétrica. Entonces existe una matriz ortogonal $O \in \mathbb{R}^{n \times n}$ tal que O^tAO es diagonal real.
2. Sea $A \in \mathbb{C}^{n \times n}$ una matriz hermitiana. Entonces existe una matriz unitaria $C \in \mathbb{C}^{n \times n}$ tal que C^*AC es diagonal real.

 Toda transformación lineal autoadjunta en un espacio euclídeo de dimensión finita tiene sus autovalores reales

■ **Ejemplo 4.20** Se quiere diagonalizar ortogonalmente la matriz

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}, \quad (4.25)$$

La ecuación característica de A es $\det(a - \lambda I) = \begin{vmatrix} 1-\lambda & -2 \\ -2 & 3-\lambda \end{vmatrix} = \lambda^2 - 4\lambda - 1 = 0$.

Tiene dos raíces, $\lambda_1 = 2 - \sqrt{5}$ y $\lambda_2 = 2 + \sqrt{5}$

y los vectores propios correspondientes son $\vec{v}_1 = \begin{pmatrix} 2 \\ -1 + \sqrt{5} \end{pmatrix}$ y $\vec{v}_2 = \begin{pmatrix} 1 - \sqrt{5} \\ 2 \end{pmatrix}$.

Para obtener vectores ortonormales los dividimos por su longitud, entonces,

$$\begin{aligned} \vec{u}_1 &= \frac{1}{\sqrt{10-2\sqrt{5}}} \begin{pmatrix} 2 \\ -1 + \sqrt{5} \end{pmatrix} \text{ y } \vec{u}_2 = \frac{1}{\sqrt{10-2\sqrt{5}}} \begin{pmatrix} 1 - \sqrt{5} \\ 2 \end{pmatrix} \\ O &= \frac{1}{\sqrt{10-2\sqrt{5}}} \begin{pmatrix} 2 & 1 - \sqrt{5} \\ -1 + \sqrt{5} & 2 \end{pmatrix}, \end{aligned} \quad (4.26)$$

$$\text{y } O^T A O = \begin{pmatrix} 2 - \sqrt{5} & 0 \\ 0 & 2 + \sqrt{5} \end{pmatrix}$$

■

■ **Ejemplo 4.21** Sea

$$A = \begin{pmatrix} 2 & 3 - 3i \\ 3 + 3i & 5 \end{pmatrix}, \quad (4.27)$$

una matriz hermitiana $\in \mathbb{C}^{2 \times 2}$.

Es posible diagonalizarla con la matriz unitaria

$$U = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 + i & 1 \\ 1 & 1 + i \end{pmatrix}, \quad (4.28)$$

y se tiene que $U^* A U = \begin{pmatrix} -1 & 0 \\ 0 & 8 \end{pmatrix}$ es una matriz diagonal real.

■

Transformaciones ortogonales

Veremos ahora endomorfismos de un espacio vectorial con producto interno que preservan el producto interno y, en particular, las distancias entre vectores.

Definición 4.6.5 Sea V un espacio Euclídeo. Una transformación lineal f se llama ortogonal si $(f(\vec{v}), f(\vec{w})) = (\vec{v}, \vec{w}) \quad \forall \vec{v}, \vec{w} \in V$. Es decir, cuando f conserva el producto escalar.

Proposición 4.6.5 Toda transformación lineal f en un espacio euclídeo que conserve la longitud de los vectores es una transformación ortogonal.

Demostración:

$$\begin{aligned} \|f(\vec{v} + \vec{w})\|^2 &= (f(\vec{v} + \vec{w}), f(\vec{v} + \vec{w})) \\ &= \|f(\vec{v})\|^2 + 2(f(\vec{v}), f(\vec{w})) + \|f(\vec{w})\|^2 \end{aligned}$$

y, por otro lado,

$$\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + 2(\vec{v}, \vec{w}) + \|\vec{w}\|^2$$

Como f conserva la longitud de los vectores, $\|f(\vec{v} + \vec{w})\| = \|\vec{v} + \vec{w}\|$, $\|f(\vec{v})\| = \|\vec{v}\|$, $\|f(\vec{w})\| = \|\vec{w}\|$, entonces, igualando términos en las expresiones anteriores, se tiene que,

$$(f(\vec{v}), f(\vec{w})) = (\vec{v}, \vec{w}),$$

y se prueba que f es ortogonal. □

■ **Ejemplo 4.22** $f = id_V$ es una transformación lineal ortogonal ■

■ **Ejemplo 4.23** Una rotación en \mathbb{R}^2 (con centro en el origen de coordenadas) es una transformación lineal ortogonal.

Ya vimos en la Sección 2.2, (Ec.(2.2)) que la matriz de una rotación en un ángulo θ en sentido antihorario, respecto de una base ortonormal es

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \quad (4.29)$$

Se verifica que $\forall \vec{v} \in \mathbb{R}^2$, $\|R\vec{v}\|^2 = \|\vec{v}\|^2$ ■

■ **Ejemplo 4.24** En \mathbb{R}^2 cualquier simetría con respecto a un subespacio vectorial unidimensional es una transformación lineal ortogonal.

En el caso de simetría respecto del eje x , en la base canónica, la matriz es

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (4.30)$$

■



- En \mathbb{R}^2 las únicas transformaciones ortogonales son las rotaciones y las simetrías, mientras que en \mathbb{R}^3 hay más posibilidades de tener transformaciones lineales ortogonales que en \mathbb{R}^2 .
- Para estudiar todas las posibles transformaciones lineales ortogonales se deben analizar los autovalores.
- Puede demostrarse que toda transformación lineal f que transforma al menos una base ortonormal en una base ortonormal, es ortogonal.

■ **Ejemplo 4.25** En \mathbb{R}^3 la simetría con respecto a una recta, como se muestra en la Figura 4.9. La matriz de la transformación en la base $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es

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

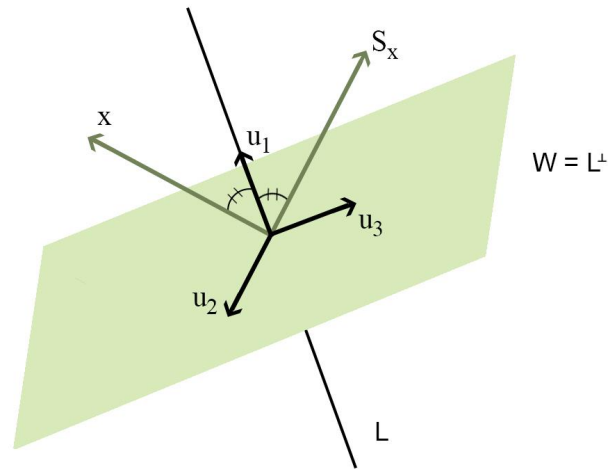
■

Proposición 4.6.6 Los autovalores reales de una transformación lineal ortogonal son iguales a 1 o a -1 .

Demostración:

Si λ es un autovalor real de una transformación ortogonal, con autovector \vec{v} , se tiene

Figura 4.9: Simetría con respecto a una recta



$$(\vec{v}, \vec{v}) = (f(\vec{v}), f(\vec{v})) = (\lambda \vec{v}, \lambda \vec{v}) = \lambda^2 (\vec{v}, \vec{v})$$

Por lo tanto $\lambda^2 = 1$ y $\lambda = + - 1$.

□



- Una transformación lineal que cumple las condiciones anteriores se dice unitaria si V es un \mathbb{C} espacio vectorial y ortogonal si V es un \mathbb{R} espacio vectorial.

f es unitaria (ortogonal) $\longleftrightarrow (f)_B$ es unitaria (ortogonal)

- Cuando A es simétrica y no demasiado grande, los algoritmos de computadora modernos que se usan actualmente calculan con gran precisión vectores y valores propios. Esos algoritmos aplican a A una sucesión de transformaciones de semejanza en las que intervienen matrices ortogonales. El uso de matrices ortogonales evita que los errores numéricos se acumulen durante el proceso. Cuando A es simétrica, la sucesión de matrices ortogonales se combina para formar una matriz ortogonal cuyas columnas son vectores propios de A . Una matriz no simétrica no puede tener un conjunto completo de vectores propios ortogonales, por lo que se necesitan técnicas no ortogonales para calcular los vectores propios.

- Cuando una matriz A tiene n vectores propios ortogonales se llama *descomposición espectral* de A a la expresión

$$A = \lambda_1 \vec{u}_1 \vec{u}_1^t + \lambda_2 \vec{u}_2 \vec{u}_2^t + \cdots + \lambda_n \vec{u}_n \vec{u}_n^t.$$

La matriz A queda dividida en partes determinadas por el espectro, y cada término es una matriz de rango 1. Entre las aplicaciones de esta descomposición está la compresión de imágenes, que se realiza considerando k términos en lugar de n , con $k < n$.

4.7.1 Ejercicios

Ejercicio 4.1

Sea \mathbb{E} un espacio Euclídeo, si \vec{x}, \vec{y} y \vec{z} son vectores de \mathbb{E} , desarrolle la siguiente expresión: $(\vec{x} + \vec{z}, \vec{x} - \vec{z} + \vec{y})$

Ejercicio 4.2

Calcule la distancia entre los vectores $\vec{u} = (2, i, 1 - i)$ y $\vec{v} = (-i, 0, 4i)$ en \mathbb{C}^3 con el producto interno canónico.

Ejercicio 4.3

Pruebe que las siguientes funciones definen productos internos sobre los espacios vectoriales considerados:

a) $(\cdot, \cdot): C[0, 1] \times C[0, 1] \rightarrow \mathbb{R}, (f(x), g(x)) = \int_0^1 f(x)g(x)dx.$

b) $(\cdot, \cdot): K^{n \times n} \times K^{n \times n} \rightarrow K, (A, B) = \text{tr}(A \cdot B^*),$ con $K = \mathbb{R}$ y $K = \mathbb{C}$ (B^* es la matriz traspuesta conjugada de B).

Ejercicio 4.4

Determine para qué valores de $\alpha \in \mathbb{R}$:

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

es un producto interno en \mathbb{R}^2 .

Ejercicio 4.5

Sean $\vec{u}_1 = (-2, -1, 1), \vec{u}_2 = (0, -1, 0)$ y $\vec{u}_3 = (1, -1, 0)$ tres vectores linealmente independientes de \mathbb{R}^3 . Si definimos el producto escalar en \mathbb{R}^3 afirmando que $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es una base ortonormal. ¿Cuál sería la expresión analítica de este producto escalar en la base canónica de \mathbb{R}^3 ?

Ejercicio 4.6

Demuestre que $Q \in \mathbb{R}^{2 \times 2}$ es una matriz de rotación puesto que Q es ortogonal y además su determinante vale 1. $Q = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}$

Ejercicio 4.7

En $P_{\mathbb{R}}^{(2)}[x]$ se define el producto escalar: $\phi(p(x), q(x)) = \int_{-1}^1 p(x)q(x)dx$

Pruebe que el conjunto $\{1, x, \frac{1}{3}(3x^2 - 1)\}$ es ortogonal.

Ejercicio 4.8

Dada la matriz simétrica $A = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$ construya una matriz ortogonal T tal que $(T^{-1}AT)$ sea una matriz diagonal. Verifique también que $T^{-1} = T^t$ y que el determinante de T es igual a 1. Obtenga asimismo la matriz diagonal $(T^{-1}AT)$.

Ejercicio 4.9

Halle una base ortogonal de \mathbb{R}^3 -con el producto interno canónico- que contenga al vector $\vec{u} = (1, -1, 2)$.

Ejercicio 4.10

Sea

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

- Halle una base ortonormal de autovectores de A .
- Halle una matriz P ortogonal tal que P^tAP sea diagonal.

Ejercicio 4.11

Sea $B = \{(1, 0, 1), (2, 0, 1), (1, 1, 0)\}$ una base de \mathbb{R}^3 . Considere el producto interno canónico y utilice Gram-Schmidt para hallar a partir de B una base B' que sea ortonormal. Calcule las coordenadas de $\vec{v} = (2, -1, 3)$ en la base B' . Utilice sus resultados para encontrar la factorización $A = QR$. Puede chequear sus resultados utilizando el siguiente programa Python:

```
import numpy as np
# Definimos la matriz
A = np.array([[4, 3, 1],
              [2, 1, 3],
              [1, 1, 1]])
# Realizamos la descomposición QR
Q, R = np.linalg.qr(A)
# Imprimimos las matrices Q y R
print("Matriz Q:")
print(Q)
print("Matriz R:")
```

```
print(R)
```

Ejercicio 4.12

Demuestre que $S = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es un conjunto ortogonal donde

$$\vec{u}_1 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} -\frac{1}{2} \\ -2 \\ \frac{7}{2} \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} 6 \\ 1 \\ -8 \end{pmatrix}$$

Expresé el vector \vec{y} como combinación lineal del conjunto S . Recuerde que las coordenadas se calculan como $c_j = \frac{\vec{y} \cdot \vec{u}_j}{\vec{u}_j \cdot \vec{u}_j}$ con $j = 1, 2, 3$ por ser una base ortogonal.

Ejercicio 4.13

Calcule la distancia de un punto \vec{y} en \mathbb{R}^3 a un subespacio W generado por $\{\vec{u}_1, \vec{u}_2\}$ sabiendo que el punto más cercano se calcula como $\|\vec{y} - \hat{y}\|$, donde $\hat{y} = \text{proy}_W \vec{y}$.

$$\vec{y} = \begin{pmatrix} -1 \\ -5 \\ 10 \end{pmatrix}, \quad \vec{u}_1 = \begin{pmatrix} 5 \\ -2 \\ -1 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

Ejercicio 4.14

Halle el complemento ortogonal para el subespacio de V :

$$V = \mathbb{R}^3, S = \{(x_1, x_2, x_3) \in \mathbb{R}^3, 2x_1 - x_2 = 0\} \text{ con el producto interno canónico.}$$

Ejercicio 4.15

Demuestre que el conjunto $S = \{\cos(nx), \sin(mx)\}_{n,m \in \mathbb{N}}$ es linealmente independiente en $C([0, 2\pi])$.

Sugerencia: observar que

$$\int_0^{2\pi} \cos^2(nx) dx \neq 0, \quad \int_0^{2\pi} \sin^2(mx) dx \neq 0, \quad n, m \in \mathbb{N}$$

y

$$\int_0^{2\pi} \cos(nx) \cos(mx) dx = \int_0^{2\pi} \cos(nx) \sin(mx) dx = \int_0^{2\pi} \sin(nx) \sin(mx) dx = 0$$

si $n \neq m, n, m \in \mathbb{N}$

Ejercicio 4.16

Halle T^* para cada una de las transformaciones lineales siguientes:

a) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $T((x_1, x_2)) = (3x_1 + x_2, -x_1 + x_2)$.

b) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, tal que $[T]_B = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$, donde $B = \{(1, 2, -1), (1, 0, 0), (0, 1, 1)\}$.

Ejercicio 4.17

Determine si los siguientes endomorfismos definidos sobre \mathbb{R}^3 son autoadjuntos:

a) $T((x, y, z)) = (x + y, x, -z)$

b) $S((x, y, z)) = (-2x + 2z, y, 2x)$

Ejercicio 4.18

Encuentre en cada caso una matriz $O \in \mathbb{R}^{n \times n}$ ortogonal tal que $O.A.O^t$ sea diagonal

a) $A = \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix}$

b) $A = \begin{pmatrix} 5 & 0 & -2 \\ 0 & 7 & -2 \\ -2 & -2 & 6 \end{pmatrix}$

Ejercicio 4.19

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

encuentre una matriz $U \in \mathbb{C}^{n \times n}$ ortogonal tal que UAU^* sea diagonal.

Ejercicio 4.20

Halle la matriz en la base canónica de las siguientes transformaciones ortogonales

a) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, rotación de un ángulo de $\frac{\pi}{4}$.

b) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, simetría respecto de la recta $x_1 = x_2$.

Ejercicio 4.21

Sea V un espacio vectorial y sea (\cdot, \cdot) un producto interno sobre V . Pruebe

- a) $(\vec{x}, \vec{y} + \vec{z}) = (\vec{x}, \vec{y}) + (\vec{x}, \vec{z})$
- b) $(\vec{x}, c\vec{y}) = \bar{c}(\vec{x}, \vec{y})$
- c) $(\vec{x}, \vec{y}) = (\vec{x}, \vec{z}) \quad \forall \vec{x} \in V \Rightarrow \vec{y} = \vec{z}$

Ejercicio 4.22

Sea V un espacio vectorial con producto interno (\cdot, \cdot) . Pruebe que $|(\vec{x}, \vec{y})| = \|\vec{x}\| \|\vec{y}\|$ sí y sólo sí $\{\vec{x}, \vec{y}\}$ es un conjunto linealmente dependiente.

Ejercicio 4.23

Pruebe que dos vectores \vec{x} e \vec{y} son ortogonales, si

$$\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$$

Ejercicio 4.24

Sea V un espacio vectorial sobre K de dimensión finita con producto interno (\cdot, \cdot) . Sea $T \in L(V)$ biyectiva. Considerar la aplicación $(\cdot, \cdot)_T : V \times V \rightarrow K$, $(\vec{x}, \vec{y})_T = (T(\vec{x}), T(\vec{y}))$, $\forall \vec{x}, \vec{y} \in V$

Pruebe que $(\cdot, \cdot)_T$ también es un producto interno sobre V .

Ejercicio 4.25

Sea $A \in \mathbb{R}^{2 \times 2}$. Sea $\phi : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $\phi(x, y) = y^t A x$. Pruebe que ϕ es un producto interno sobre \mathbb{R}^2 sí y sólo sí $A = A^t$, $a_{11} \geq 0$ y $\text{Det}(A) \geq 0$.

Ejercicio 4.26

Sea V un \mathbb{C} -espacio vectorial con producto interno (\cdot, \cdot) y sea $T \in L(V)$ sobre \mathbb{C} . Pruebe que si λ es autovalor de T , entonces $\bar{\lambda}$ es un autovalor de T^* .

Ejercicio 4.27

Sea V un \mathbb{C} -espacio vectorial con producto interno (\cdot, \cdot) y sea $T \in L(V)$ sobre \mathbb{C} autoadjunta. Pruebe que:

- a) si λ es autovalor de T , entonces $\lambda \in \mathbb{R}$.
- b) Si \vec{v}_i es autovector asociado al autovalor λ_i de T (para $i = 1, 2$) y $\lambda_1 \neq \lambda_2$, entonces $(\vec{v}_1, \vec{v}_2) = 0$.

Ejercicio 4.28

Sea V un espacio vectorial de dimensión finita con producto interno y sean S y $T \in L(V)$. Si $k \in K$, pruebe:

- a) $(S + T)^* = S^* + T^*$

b) $(kT)^* = \bar{k}T^*$

c) $(ST)^* = T^*S^*$

4.7.2 Autoevaluación

Verdadero o Falso

1. El subespacio imagen de una transformación lineal es ortogonal a su núcleo.
2. El vector nulo es ortogonal a todo vector en \mathbb{R}^n .
3. Sean W un plano a través del origen en \mathbb{R}^3 y L la recta que pasa por el origen y es perpendicular a W . Entonces $L^\perp = W$ y $W^\perp = L$.
4. $T(\vec{x}) = \left(\frac{\vec{x} \cdot \vec{v}}{\vec{v} \cdot \vec{v}}\right)\vec{v}$ es una transformación de proyección.
5. Una matriz cuadrada U tal que $U^{-1} = U^t$ se denomina ortogonal.
6. Si U es ortonormal tanto las filas como las columnas de U son ortonormales.
7. En la factorización QR el hecho de que R sea invertible es consecuencia directa de que las columnas de A sean linealmente independientes.
8. Si $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es una base ortogonal para W , entonces la multiplicación de \vec{v}_3 por un escalar, da una nueva base ortogonal $\{\vec{v}_1, \vec{v}_2, 3\vec{v}_3\}$.
9. Si $A = QR$, donde Q tiene columnas ortonormales, entonces $R = Q^t A$.
10. Si \vec{x} no está en un subespacio W , entonces $\vec{x} - P_w \vec{x}$ no es el vector nulo.
11. Un espacio vectorial con un producto escalar se dice que es un espacio vectorial euclídeo.
12. Si Q es ortogonal se cumple que la norma de \vec{x} es igual a la norma de $Q\vec{x}$.
13. Todo conjunto ortogonal de un espacio euclídeo es linealmente dependiente.
14. El determinante de una matriz ortogonal es 1 ó -1 .
15. El producto de dos matrices ortogonales es la matriz identidad.
16. Descomponer un vector \vec{y} en una suma de proyecciones ortogonales sobre espacios unidimensionales es la esencia del proceso de Gram-Smith.
17. $\dim V = \dim W + \dim W^\perp$ (W un subespacio de V).
18. $d(\vec{u}, -\vec{v})^2 = \|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2(\vec{u}, \vec{v})$.



5. Formas bilineales y cuadráticas

En este capítulo se dará una breve introducción al tema enfocada a mostrar aplicaciones de la diagonalización de las matrices de las formas bilineales y cuadráticas en el estudio de secciones cónicas y superficies cuádricas.

5.1 Formas bilineales y cuadráticas

La ecuación general de una cónica está dada por una ecuación de segundo grado de la forma

$$a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2 + a_1x_1 + a_2x_2 + a = 0, \quad (5.1)$$

donde a_{ij} , a_i ($i, j = 1, 2$) y a son números reales y al menos uno de los números a_{ij} no es cero. La parte principal es:

$$a_{11}x^2 + a_{12}xy + a_{22}y^2 \quad (5.2)$$

y puede escribirse

$$P(x, y) = a_{11}x^2 + a_{12}xy + a_{22}y^2 = (x, y) A \begin{pmatrix} x \\ y \end{pmatrix}, \quad (5.3)$$

donde A es la matriz simétrica,

$$A = \begin{pmatrix} a_{11} & a_{12}/2 \\ a_{12}/2 & a_{22} \end{pmatrix}. \quad (5.4)$$

Y su generalización a \mathbb{R}^n , dado $\vec{x} = (x_1, x_2, x_3, \dots, x_n) \in \mathbb{R}^n$, es

$$P(\vec{x}) = (x_1, x_2, x_3, \dots, x_n) A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{pmatrix} \quad (5.5)$$

donde A es una matriz simétrica $\in \mathbb{R}^{n \times n}$.


Análogamente al caso $n = 2$, para $n = 3$ se obtienen superficies de segundo grado.

Los dos ejemplos anteriores corresponden a *formas cuadráticas* (en \mathbb{R}^2 y en \mathbb{R}^n), y son casos particulares de *formas bilineales*, las que se definen a continuación.

Definición 5.1.1 Forma bilineal

Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} . Una aplicación $\mathbf{A} : V \times V \rightarrow \mathbb{R}$ (o \mathbb{C}) se dice que es una *forma bilineal* si y solo si satisface:

1. $\mathbf{A}(\vec{x} + \vec{z}, \vec{y}) = \mathbf{A}(\vec{x}, \vec{y}) + \mathbf{A}(\vec{z}, \vec{y}), \forall \vec{x}, \vec{y}, \vec{z} \in V$
2. $\mathbf{A}(\alpha \vec{x}, \vec{y}) = \alpha \mathbf{A}(\vec{x}, \vec{y}) \forall \vec{x}, \vec{y} \in V \text{ y } \forall \alpha \in \mathbb{R} \text{ o } \mathbb{C}$
3. $\mathbf{A}(\vec{x}, \vec{y} + \vec{z}) = \mathbf{A}(\vec{x}, \vec{y}) + \mathbf{A}(\vec{x}, \vec{z}), \forall \vec{x}, \vec{y}, \vec{z} \in V$
4. $\mathbf{A}(\vec{x}, \beta \vec{y}) = \beta \mathbf{A}(\vec{x}, \vec{y}) \forall \vec{x}, \vec{y} \text{ y } \forall \beta \in \mathbb{R} \text{ o } \mathbb{C}$

 Los productos internos reales son formas bilineales.

■ **Ejemplo 5.1** \mathbb{E} un espacio euclídeo de dimensión finita, y T una aplicación lineal de \mathbb{E} en \mathbb{E} . Puede demostrarse que $\mathbf{A}(\vec{x}, \vec{y}) = (\vec{x}, T\vec{y})$ es una forma bilineal. ■

■ **Ejemplo 5.2** Dados

$\vec{x} = (x_1, x_2, x_3, \dots, x_n)$ e $\vec{y} = (y_1, y_2, y_3, \dots, y_n) \in \mathbb{R}^n$, el producto

$$(x_1, x_2, x_3, \dots, x_n) A \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_n \end{pmatrix}$$

donde A es una matriz simétrica $\in \mathbb{R}^{n \times n}$, es una forma bilineal.

La forma cuadrática en \mathbb{R}^n de la Ec.(5.5) presentada antes sale de tomar para este caso $\vec{x} = \vec{y}$. ■



- Una forma bilineal se dice *simétrica* si $A(\vec{x}, \vec{y}) = A(\vec{y}, \vec{x})$
- Una forma bilineal se dice *antisimétrica* si $A(\vec{x}, \vec{y}) = -A(\vec{y}, \vec{x})$

Definición 5.1.2 Matriz de una forma bilineal

Sea \mathbf{A} una forma bilineal en un espacio V y sea $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ una base de V .

Si $\vec{x} = \sum_{i=1}^n x_i \vec{e}_i$, e $\vec{y} = \sum_{j=1}^n y_j \vec{e}_j$,

$$\begin{aligned} \mathbf{A}(\vec{x}, \vec{y}) &= \mathbf{A}\left(\sum_{i=1}^n x_i \vec{e}_i, \sum_{j=1}^n y_j \vec{e}_j\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i y_j \mathbf{A}(\vec{e}_i, \vec{e}_j) \end{aligned}$$

Se define la matriz de la forma bilineal \mathbf{A} en la base B como la matriz $A \in \mathbb{R}^{n \times n}$ tal que $A_{ij} = \mathbf{A}(\vec{e}_i, \vec{e}_j)$ para $1 \leq i, j \leq n$



- Si $\mathbf{A}(\vec{x}, \vec{y}) = \mathbf{A}(\vec{y}, \vec{x})$ (es decir \mathbf{A} es una forma bilineal simétrica), entonces $\mathbf{A}(\vec{e}_i, \vec{e}_j) = \mathbf{A}(\vec{e}_j, \vec{e}_i)$ para cualquier base $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ de V , es decir la matriz de una forma bilineal A simétrica en cualquier base es simétrica. Vale también la recíproca: si la matriz de

una forma bilineal es simétrica en alguna base $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ de V , entonces la forma bilineal es simétrica, pues

$$\begin{aligned}
 \mathbf{A}(\vec{y}, \vec{x}) &= \sum_{i,j=1}^n \mathbf{A}(\vec{e}_i, \vec{e}_j) y_i x_j \\
 &= \sum_{i,j=1}^n a_{ij} y_i x_j \\
 &= \sum_{i,j=1}^n a_{ji} x_j y_i \\
 &= \sum_{i,j=1}^n \mathbf{A}(\vec{e}_j, \vec{e}_i) x_j y_i \\
 &= \mathbf{A}(\vec{x}, \vec{y})
 \end{aligned}$$

- Si \mathbf{A} es una forma bilineal antisimétrica, $\mathbf{A}(\vec{x}, \vec{y}) = -\mathbf{A}(\vec{y}, \vec{x})$ para cualquier base $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ de V . La matriz satisface $a_{ij} = -a_{ji}$, de donde $a_{ii} = 0$, $i = 1, \dots, n$.
- Si A es la matriz de una forma bilineal respecto a la base $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ y \tilde{A} con respecto a la base $\tilde{B} = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$, entonces $A = C^T \tilde{A} C$, donde C es la matriz de cambio de base de B a \tilde{B} . Tienen el mismo rango \tilde{A} y A , ya que $\det(C) \neq 0$.
- El rango de una forma bilineal es el rango que tiene la matriz de la forma bilineal en cualquier base.

Proposición 5.1.1 Toda forma bilineal puede ser representada como la suma de una forma bilineal simétrica y una forma bilineal antisimétrica.

Demostración:

Sea \mathbf{A} una forma bilineal definida en V y sea $\mathbf{B} : V \times V \rightarrow K$ definida así

$$\mathbf{B}(\vec{x}, \vec{y}) = \mathbf{A}(\vec{x}, \vec{y}) + \mathbf{A}(\vec{y}, \vec{x})$$

y sea $\mathbf{C} : V \times V \rightarrow K$

$$\mathbf{C}(\vec{x}, \vec{y}) = \mathbf{A}(\vec{x}, \vec{y}) - \mathbf{A}(\vec{y}, \vec{x})$$

$$2\mathbf{A}(\vec{x}, \vec{y}) = \mathbf{B}(\vec{x}, \vec{y}) + \mathbf{C}(\vec{x}, \vec{y})$$

$$\mathbf{A}(\vec{x}, \vec{y}) = \frac{\mathbf{B}(\vec{x}, \vec{y})}{2} + \frac{\mathbf{C}(\vec{x}, \vec{y})}{2}$$

donde

$$\mathbf{B}(\vec{y}, \vec{x}) = \mathbf{A}(\vec{y}, \vec{x}) + \mathbf{A}(\vec{x}, \vec{y}) = \mathbf{B}(\vec{x}, \vec{y})$$

$$\mathbf{C}(\vec{y}, \vec{x}) = \mathbf{A}(\vec{y}, \vec{x}) - \mathbf{A}(\vec{x}, \vec{y}) = -\mathbf{C}(\vec{x}, \vec{y})$$

Esto es similar a lo que ocurre con una matriz A , que puede expresarse,

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2}$$

donde $\frac{A+A^T}{2}$ es simétrica y $\frac{A-A^T}{2}$ antisimétrica.

□

Definición 5.1.3 Forma cuadrática

Dada $\mathbf{A} : V \times V \rightarrow \mathbb{R}$ una forma bilineal, se define una *forma cuadrática*

$$\mathbf{Q} : V \rightarrow \mathbb{R}, \mathbf{Q}(\vec{x}) = \mathbf{A}(\vec{x}, \vec{x}).$$

■ Ejemplo 5.3

$$\mathbf{Q}(x_1, x_2) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2$$

■

Usando la matriz P definida antes, Ec.(5.3) una forma cuadrática se escribe,

$$\mathbf{Q}(\vec{x}) = (x_1, x_2, x_3, \dots, x_n)P(x_1, x_2, x_3, \dots, x_n)^T.$$


En general, toda expresión de la forma

$$\mathbf{Q}(\vec{x}) = \sum_{j=1}^n \sum_{i \leq j}^n a_{ij} x_i x_j$$

en un espacio vectorial define una forma cuadrática, ya que alcanza con tomar la forma bilineal

$$\mathbf{A}(\vec{x}, \vec{y}) = \sum_{i=1}^n a_{ii} x_i y_i + \sum_{j=1}^n \sum_{i < j}^n \frac{a_{ij}}{2} x_i y_j + \sum_{j=1}^n \sum_{i < j}^n \frac{a_{ij}}{2} x_j y_i$$

con $\vec{x} = \sum_{i=1}^n x_i \vec{e}_i$ e $\vec{y} = \sum_{j=1}^n y_j \vec{e}_j$.

-  Un ejemplo de forma bilineal es el tensor de inercia, $I(\vec{x}, \vec{y})$. Gran parte de su interés radica en que $I(\omega, \omega)$ da la energía de rotación cuando la velocidad angular es ω .

5.1.1 Formas cuadráticas. Aplicación a las secciones cónicas

La Ec.(5.1) se reescribe,

$$Q(x_1, x_2) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2$$

y también en forma matricial,

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12}/2 \\ a_{12}/2 & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + a = 0 \quad (5.6)$$

$$\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{K} \mathbf{x} + \mathbf{a} = 0 \quad (5.7)$$

donde

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \quad (5.8)$$

Con esta notación, la forma cuadrática asociada a la Ec.(5.7) es $\mathbf{x}^T \mathbf{A} \mathbf{x}$. La matriz simétrica A se denomina *matriz de la forma cuadrática* $\mathbf{x}^T \mathbf{A} \mathbf{x}$.

■ **Ejemplo 5.4** En la ecuación,

$$3x_1^2 + 5x_1x_2 - 7x_2^2 + 2x_1 + 7 = 0, \quad (5.9)$$

la matriz de la forma cuadrática es,

$$\begin{pmatrix} 3 & 5/2 \\ 5/2 & -7 \end{pmatrix}.$$

mientras que para la ecuación

$$8x_1^2 - 4x_2^2 = 0, \quad (5.10)$$

la matriz de la forma cuadrática es,

$$\begin{pmatrix} 8 & 0 \\ 0 & -4 \end{pmatrix}.$$

■

Para una cónica C con ecuación (5.7) es posible hacer girar los ejes de coordenadas x_1x_2 de modo que la ecuación de la cónica, en el sistema de coordenadas $x'_1x'_2$, no tenga término con producto cruzado.

- Se halla una matriz O que diagonaliza ortogonalmente a A , $A = ODO^t$. Se intercambian sus columnas en el caso que el $\text{Det}(O) \neq 1$ para asegurar que la transformación de coordenadas sea una rotación.

$$\mathbf{x} = O\mathbf{x}'$$

- Para obtener la ecuación en el sistema $x'_1x'_2$ se sustituye la ecuación anterior en la Ec.(5.7).

$$(\mathbf{x}')^t \mathbf{O}^t \mathbf{A} \mathbf{O} \mathbf{x}' + \mathbf{K} \mathbf{O} \mathbf{x}' + \mathbf{a} = \mathbf{0} \quad (5.11)$$

Como O diagonaliza ortogonalmente a la matriz A , $O^t A O = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, donde λ_1 y λ_2 son los autovalores de A . Por lo tanto, la ecuación (5.11) reescribir como

$$\begin{pmatrix} x'_1 & x'_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} + \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} o_{11} & o_{12} \\ o_{21} & o_{22} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} + a = 0 \quad (5.12)$$

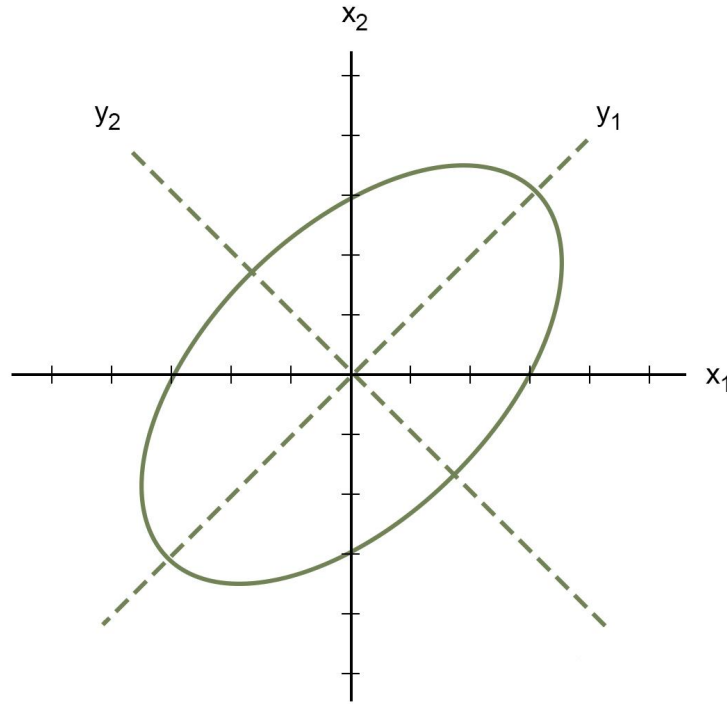
o bien

$$\lambda_1 (x'_1)^2 + \lambda_2 (x'_2)^2 + a'_1 x'_1 + a'_2 x'_2 + a = 0$$

donde $a'_1 = a_1 o_{11} + a_2 o_{21}$ y $a'_2 = a_1 o_{12} + a_2 o_{22}$.

El análisis anterior se resume en el Teorema siguiente:

Figura 5.1: Elipse

**Teorema 5.1.2 Teorema de los ejes principales para \mathbb{R}^2 ..**

Sea

$$a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2 + a_1x_1 + a_2x_2 + a = 0, \quad (5.13)$$

la ecuación de una cónica C , y supongamos que

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{a}_{11}\mathbf{x}_1^2 + \mathbf{a}_{12}\mathbf{x}_1\mathbf{x}_2 + \mathbf{a}_{22}\mathbf{x}_2^2$$

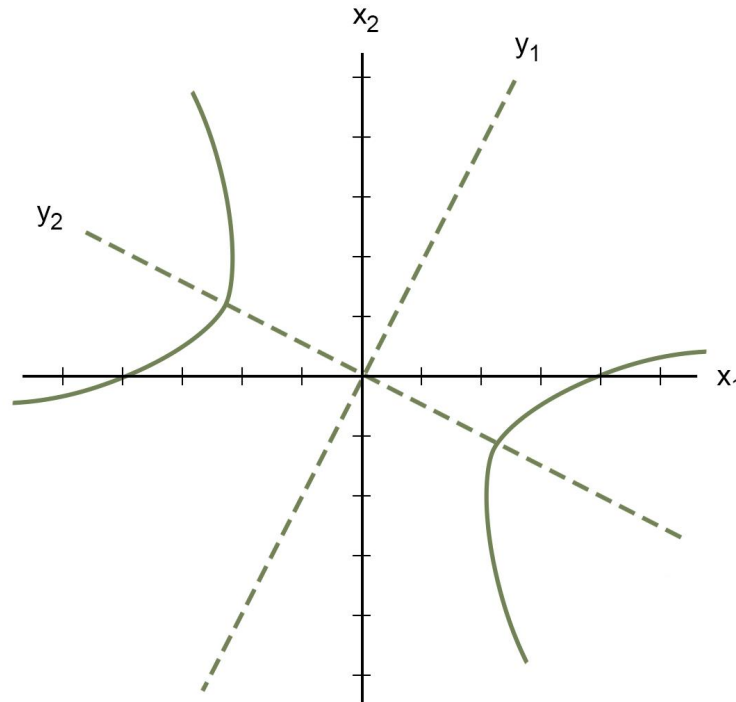
es la forma cuadrática asociada. Entonces, es posible girar los ejes de coordenadas de modo que la ecuación para C en el nuevo sistema de coordenadas $x'y'$ tenga la forma

$$\lambda_1 x_1'^2 + \lambda_2 x_2'^2 + a_1'x_1' + a_2'y_2' + a = 0$$

donde λ_1 y λ_2 son los autovalores de A . Se puede llevar a cabo la rotación por medio de la sustitución $\mathbf{x} = O\mathbf{x}'$, donde O diagonaliza a A y $\text{Det}(O) = 1$.

Si A es una matriz no diagonal, la gráfica está girada hasta salirse de la posición estándar, como

Figura 5.2: Hipérbola



en las Figuras 5.1 y 5.2. Encontrar los ejes principales (determinados por los vectores propios de A) equivale a encontrar un nuevo sistema de coordenadas con respecto al cual la gráfica está en posición estándar.

5.1.2 Formas cuadráticas: aplicación a las superficies cuádricas

Sea

$$a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{22}x_2^2 + a_{23}x_2x_3 + a_{33}x_3^2 + a_1x_1 + a_2x_2 + a_3x_3 + a = 0 \quad (5.14)$$

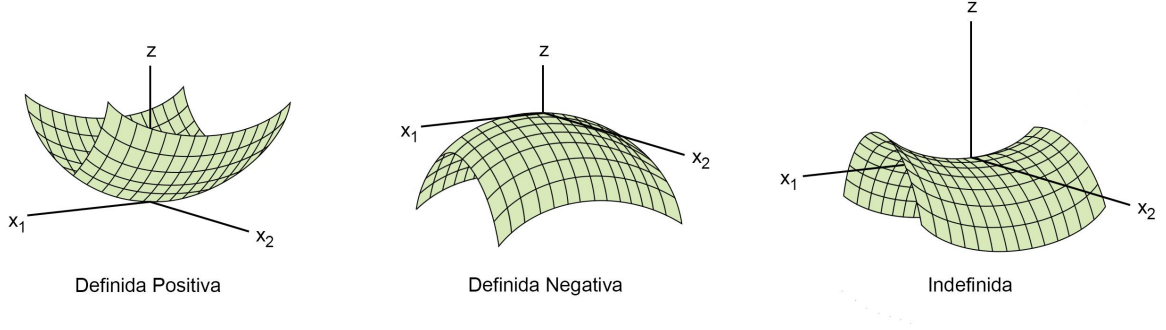
donde a_{ij} , a_i ($i, j = 1, 3$) y a son números reales y al menos uno de los números a_{ij} no es cero. La parte principal

$$a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{22}x_2^2 + a_{23}x_2x_3 + a_{33}x_3^2$$

es la forma cuadrática asociada.

La Ec.(5.14) puede escribirse en forma matricial

Figura 5.3: Formas cuadráticas. a) Definida Positiva. b) Definida Negativa. c) Indefinida



$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} a_{11} & a_{12}/2 & a_{13}/2 \\ a_{12}/2 & a_{22} & a_{23}/2 \\ a_{13}/2 & a_{23}/2 & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + a = 0 \quad (5.15)$$

Se tiene el Teorema siguiente:

Teorema 5.1.3 Teorema de los ejes principales para \mathbb{R}^3 .

Sea

$$a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{22}x_2^2 + a_{23}x_2x_3 + a_{33}x_3^2 + a_1x_1 + a_2x_2 + a_3x_3 + a = 0 \quad (5.16)$$

la ecuación de una cónica C , y supongamos que

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{22}x_2^2 + a_{23}x_2x_3 + a_{33}x_3^2$$

es la forma cuadrática asociada. Entonces se puede hacer girar los ejes de coordenadas de modo que la ecuación para C en el nuevo sistema de coordenadas $x'_1x'_2x'_3$ tenga la forma

$$\lambda_1x_1'^2 + \lambda_2x_2'^2 + \lambda_3x_3'^2 + a'_1x'_1 + a'_2x'_2 + a'_3x'_3 + a = 0$$

donde λ_1, λ_2 y λ_3 son los autovalores de A . Como en el caso de \mathbb{R}^2 se puede llevar a cabo la rotación por medio de la sustitución $\mathbf{x} = O\mathbf{x}'$, donde O diagonaliza a A y $\text{Det}(O) = 1$.

Este teorema sugiere el procedimiento para eliminar los términos de productos cruzados de una ecuación cuadrática en x_1, x_2 y x_3 . Y lo veremos con un ejemplo.

■ **Ejemplo 5.5** Se desea describir la superficie cuádrica cuya ecuación es

$$4x_1^2 + 4x_1x_2 + 4x_1x_3 + 4x_2^2 + 4x_2x_3 + 4x_3^2 - 3 = 0 \quad (5.17)$$

La forma matricial de la ecuación cuadrática anterior es

$$\mathbf{x}^T \mathbf{A} \mathbf{x} - 3 = 0 \quad (5.18)$$

donde

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix} \quad (5.19)$$

Los autovalores de \mathbf{A} son $\lambda_1 = \lambda_2 = 2$ y $\lambda_3 = 8$, y \mathbf{A} es diagonalizada ortogonalmente por la matriz

$$\mathbf{O} = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} \quad (5.20)$$

donde las dos primeras columnas de \mathbf{O} son los autovectores correspondientes a $\lambda_1 = \lambda_2 = 2$ mientras que la tercer columna es un autovector correspondiente a $\lambda_3 = 8$. Se puede verificar que $\text{Det}(\mathbf{O}) = 1$ por lo que la transformación de coordenadas $\mathbf{x} = \mathbf{O}\mathbf{x}'$ es una rotación.

Al sustituir en la Ec.(5.18) se obtiene

$$\mathbf{x}^T \mathbf{O} \mathbf{D} \mathbf{O}^T \mathbf{x} - 3 = \mathbf{x}'^T \mathbf{D} \mathbf{x}' - 3 = 0 \quad (5.21)$$

Y como

$$\mathbf{D} = \mathbf{O}^t \mathbf{A} \mathbf{O} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix} \quad (5.22)$$

se tiene

$$\begin{pmatrix} x' & y' & z' \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} - 3 = 0 \quad (5.23)$$


o bien

$$2(x'_1)^2 + 2(x'_2)^2 + 8(x'_3)^2 - 3 = 0 \quad (5.24)$$

que puede reescribirse

$$\frac{(x'_1)^2}{3/2} + \frac{(x'_2)^2}{3/2} + \frac{(x'_3)^2}{3/8} = 1 \quad (5.25)$$

y es la ecuación de un elipsoide. ■

 Las superficies cuádricas han sido representadas en varios edificios contemporáneos. Algunos de ellos son: Puente Juscelino Kubitschek, Brasilia (Brasil), Centro Nacional de las Artes Escénicas, Pekín (China), L'Oceanogràfic, Valencia (España).

Formas cuadráticas y valores propios

Cuando A es una matriz de $n \times n$, la forma cuadrática $\mathbf{Q}(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ es una función de valores reales con dominio \mathbb{R}^n . Se distinguen varias clases importantes de formas cuadráticas por el tipo de valores que asumen para diversos \mathbf{x} .

En la Figura 5.3 se muestran las gráficas de tres formas cuadráticas. Para cada punto $\mathbf{x} = (x_1, x_2)$ del dominio de una forma cuadrática \mathbf{Q} , se traza un punto (x_1, x_2, z) , donde $z = \mathbf{Q}(\mathbf{x})$. Observe que excepto en $\mathbf{x} = \mathbf{0}$, todos los valores de $\mathbf{Q}(\mathbf{x})$ son positivos en la Figura 5.3(a) y negativos en la Figura 5.3(b). En la Figura 5.3(c), en cambio, toma valores positivos y negativos. De acuerdo a los autovalores de A se tiene lo siguiente:

Sea A una matriz simétrica de $n \times n$. Entonces una forma cuadrática $\mathbf{x}^T A \mathbf{x}$ es:

- *definida positiva* si, y sólo si, todos los valores propios de A son positivos,
- *definida negativa* si, y sólo si, todos los valores propios de A son negativos, o
- *indefinida* si, y sólo si, A tiene valores propios tanto positivos como negativos


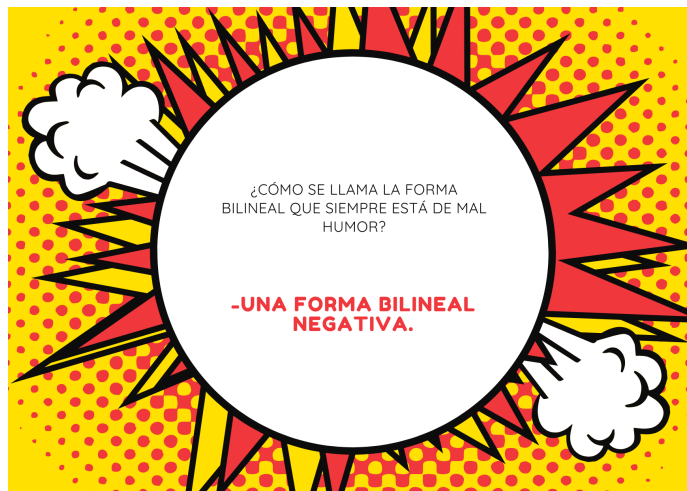
-  Una de las aplicaciones más conocidas es el estudio de extremos relativos de funciones de varias variables. En ese caso se calculan los autovalores de la matriz Hessiana en los puntos estacionarios. Corresponde a un mínimo local en caso de ser definida positiva, a un máximo local en caso de definida negativa y a un punto silla si es indefinida.

Figura 5.4



5.2 Actividades propuestas

Problema de Aplicación 5 Realice un cuadro conceptual donde describa las diferentes superficies cuádricas a partir de la diagonalización de las matrices de las formas bilineales y cuadráticas correspondientes. Indique cuáles de ellas tienen centro, cuáles no, cuáles son degeneradas, y qué significa ese término. Investigue además, qué característica de los paraboloïdes hace que los radiotelescopios usen esa forma para sus antenas. Complemente con imágenes de antenas de algún radiotelescopio y sus características físicas. Se propone la presentación oral del trabajo con el fin de contribuir al desarrollo de habilidades y capacidades del estudiante (15 minutos máximo).

5.2.1 Ejercicios

Ejercicio 5.1

Encuentre la matriz asociada a la forma bilineal $A(\vec{x}, \vec{y}) = x_1y_1 - x_1y_2 + 2x_2y_1 + 6x_2y_2 - 3x_1y_3 + x_3y_3$ y calcule su rango.

Ejercicio 5.2

Convierta la forma bilineal del ejercicio anterior en una forma cuadrática reemplazando $(\vec{x} = \vec{y})$. Calcule su nueva matriz asociada.

Ejercicio 5.3

Dado que la matriz asociada a una forma cuadrática es simétrica, haga una lista de todas las propiedades de las matrices simétricas.

Figura 5.5: Beletsky alma. https://www.eso.org/public/images/beletsky_alma_15-cc2/



Ejercicio 5.4

Encuentre la forma canónica de la siguiente forma cuadrática: $\mathbf{Q}(\vec{x}) = x_2^2 - 3x_3^2 + 2x_1x_2 + x_1x_3$

Ejercicio 5.5

Para la elipse $5x_1^2 + 5x_2^2 - 4x_1x_2 = 48$, encuentre un cambio de variables por medio de calcular sus valores y vectores propios unitarios tal que se elimine el producto cruzado de la ecuación.

Ejercicio 5.6

Especifique a qué cónica corresponden las siguientes ecuaciones y especifique su centro.

a) $(x - x_0)^2 + (y - y_0)^2 = r^2$

b) $(x - 2)^2 - (y - 3)^2 = 1$

c) $x^2 + y^2 + 4x = 1$

Ejercicio 5.7

Dada la ecuación de una cónica: $Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$, encuentre su forma matricial. Ayuda: si el elemento a_{11} es F , ¿cómo están relacionados los otros elementos de la matriz con la ecuación de la cónica?

$$\begin{pmatrix} 1 & x & y \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} = 0$$

Datos extras: piense que las cónicas describen las órbitas. Ejemplo de órbitas elípticas pueden ser el asteroide numerado 433 conocido con el nombre de Eros, en la página de NEODyS (objetos cercanos al planeta Tierra) podrá encontrar muchos más. Como ejemplo de órbita hiperbólica podría ser el cometa C/2002 E2 Snyder-Murakami, y como ejemplo de órbita parabólica (excentricidad = 1) la del cometa C/2002 B2 LINEAR.

Ejercicio 5.8

Usando la matriz del ejercicio anterior, encuentre la forma de sus invariantes y especifique de qué tipo de cónica estamos hablando si $B^2 - 4AC = 0$

Ejercicio 5.9

Responda cómo están los ejes de las cónicas con respecto a los ejes coordenados según:

- a) $B = 0$
- b) $B \neq 0$

Ejercicio 5.10

¿ $Q(\vec{x}) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$ es definida positiva?

5.2.2 Autoevaluación

Verdadero o Falso

Dada una matriz simétrica:

1. A es definida positiva si y solo si todos los valores propios de A son positivos.
2. A es definida negativa si y solo si los valores propios van alternando entre positivos y negativos.
3. A es indefinida si y solo si alguno de los valores propios es 0.
4. Es posible clasificar A por medio de su determinante.
5. Siempre existe un cambio ortogonal de la variable $\mathbf{x} = P\mathbf{y}$ tal que $Q(\mathbf{X}) = \mathbf{x}^t A \mathbf{x} = \mathbf{y}^t D \mathbf{y} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_n x_n^2$, con $\lambda_1, \lambda_2, \dots, \lambda_n$ los autovalores de A.



6. Cálculo tensorial

Este capítulo trata de una introducción al estudio de tensores (escalar, vector, tensor de segundo orden y de orden superior). Se expone la notación indicial por su simplicidad y facilidad de uso en las expresiones matemáticas. Se hace una revisión de las operaciones entre vectores, y de los sistemas de coordenadas rectangulares. Luego se plantean los sistemas de coordenadas curvilíneas, y la construcción de bases adecuadas. Para facilitar la comprensión de los temas se presentan ejemplos y aplicaciones. El objetivo de incorporar el cálculo tensorial es brindar al estudiante de Astronomía herramientas matemáticas que le resulten de utilidad para los cursos superiores de la carrera.

6.1 Invariancia y representación

Dado un espacio vectorial, la elección de la base es arbitraria. Una vez elegida la base, lo que se tiene es una *representación* del vector en una determinada base y por lo tanto, se tienen sus coordenadas.

Así para el vector $\mathbf{x} = (1, 2, 3)$, con las bases canónica, B y la base $B' = \{(1/2, 0, 0), (0, 0, -2), (0, -1, 0)\}$ se tendrán dos representaciones,

$$\mathbf{x} = 1 \mathbf{e}_1 + 2 \mathbf{e}_2 + 3 \mathbf{e}_3$$

y

$$\mathbf{x} = 2 \mathbf{e}'_1 + 3/2 \mathbf{e}'_2 - 1 \mathbf{e}'_3$$

En general, en un espacio vectorial de dimensión N , una vez elegida la base $B = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$, cada vector \mathbf{x} estará representado por un conjunto de n coordenadas representadas con λ^j

$$\sum_{j=1}^n \lambda^j \mathbf{e}_j$$

pero el vector \mathbf{x} es un *invariante*, ya que no depende de la base. De esta forma, dadas dos bases $B = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ y $B' = \{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\}$ de un espacio vectorial V de dimensión n , para un vector \mathbf{x} se satisface la igualdad

$$\sum_{j=1}^n \lambda^j \mathbf{e}_j = \sum_{i=1}^n \beta^i \mathbf{e}'_i$$



A diferencia de los capítulos anteriores, en este capítulo los vectores se indicarán con trazo resaltado.

6.2 Convenio de suma de Einstein

Albert Einstein, en 1916, propuso un criterio que permite escribir las sumas sin escribir los símbolos de sumatoria, dando origen a la notación indicial. Introdujo los dos convenios siguientes:

1. Para un espacio vectorial de dimensión N , los índices usados, ya sea como subíndices o como supraíndices pueden tomar todos los valores de 1 a N , a no ser que se especifique lo contrario.
2. Si se repite un índice en un término, esto implica una suma con respecto a aquel índice desde 1 a N . El índice repetido se llama índice *mudo*.

Usando los convenios anteriores, un vector \mathbf{x} puede expresarse, entonces,

$$\mathbf{x} = \lambda^j \mathbf{e}_j = \beta^i \mathbf{e}'_i$$

En las expresiones anteriores i y j son índices mudos.

Así,

$$a_i^i = a_1^1 + a_2^2 + a_3^3 + \dots + a_n^n$$

Si los elementos a_j^i son los de una matriz $A \in \mathbb{R}^{n \times n}$, (el supraíndice i corresponde a la fila y el subíndice j a la columna), a_i^i es la *traza* de la matriz.

■ **Ejemplo 6.1** En la expresión

$$A_{ij}x_i x_j \quad i, j = 1, 2, 3$$

no hay índice libre, tanto i como j son índices mudos, por lo que al sumar en i y en j , el resultado da un escalar. ■



Una de las ventajas de la notación indicial es que se tiene una expresión muy concisa. Así, un sistema lineal de 3 ecuaciones con 3 incógnitas usando el convenio de suma se escribe:

$$a_{ij}x_j = b_i \quad i, j = 1, 2, 3.$$

6.3 Notación indicial

El sistema de coordenadas cartesianas rectangulares está definido por tres vectores, \mathbf{i} , \mathbf{j} , \mathbf{k} que constituyen una base *ortonormal*. Es decir que se satisfacen dos propiedades: son vectores unitarios (longitud 1) y son ortogonales entre sí. El producto vectorial cumple la regla de la mano derecha: $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$ y $\mathbf{k} \times \mathbf{i} = \mathbf{j}$.

La representación de un vector \mathbf{P} en un sistema de coordenadas rectangulares es:

$$\mathbf{P} = P_x \mathbf{i} + P_y \mathbf{j} + P_z \mathbf{k} \quad (6.1)$$

que puede reescribirse de la forma

$$\mathbf{P} = P_1 \mathbf{e}_1 + P_2 \mathbf{e}_2 + P_3 \mathbf{e}_3 \quad (6.2)$$

donde hemos considerado que $P_1 \equiv P_x$, $P_2 \equiv P_y$, $P_3 \equiv P_z$, $\mathbf{e}_1 \equiv \mathbf{i}$, $\mathbf{e}_2 \equiv \mathbf{j}$, $\mathbf{e}_3 \equiv \mathbf{k}$, como se indica en la Figura 6.1.

La representación del vector \mathbf{P} de la Ec.(6.2) se expresa con la notación indicial de la forma siguiente ([18]):

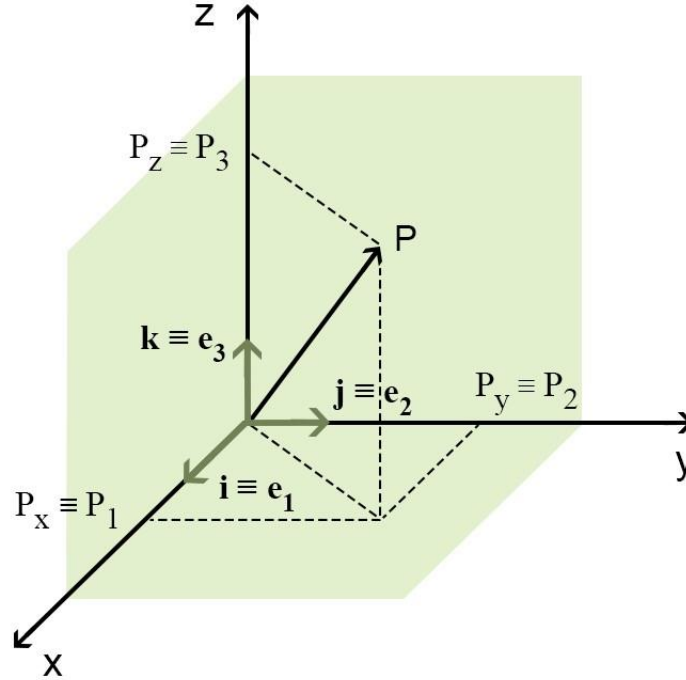
$$\mathbf{P} = P_i \mathbf{e}_i \quad (i = 1, 2, 3) \quad (6.3)$$

6.3.1 Delta de Kronecker

El símbolo *delta de Kronecker*, está definido por

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases} \quad (6.4)$$

Figura 6.1: Vector en el sistema de coordenadas cartesianas



y coincide con el resultado de hacer el producto escalar entre los vectores de la base ortonormal \mathbf{e}_i , es decir que $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$. Exponiendo esto en forma explícita se tiene:

$$\mathbf{e}_i \cdot \mathbf{e}_j = \begin{pmatrix} \mathbf{e}_1 \cdot \mathbf{e}_1 & \mathbf{e}_1 \cdot \mathbf{e}_2 & \mathbf{e}_1 \cdot \mathbf{e}_3 \\ \mathbf{e}_2 \cdot \mathbf{e}_1 & \mathbf{e}_2 \cdot \mathbf{e}_2 & \mathbf{e}_2 \cdot \mathbf{e}_3 \\ \mathbf{e}_3 \cdot \mathbf{e}_1 & \mathbf{e}_3 \cdot \mathbf{e}_2 & \mathbf{e}_3 \cdot \mathbf{e}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \delta_{ij} \quad (6.5)$$

Este símbolo δ_{ij} es llamado *operador de sustitución*, por la propiedad interesante que mostramos con un ejemplo.

Sea \mathbf{v} un vector de componentes v_i , entonces

$$\delta_{ij}v_i = \delta_{1j}v_1 + \delta_{2j}v_2 + \delta_{3j}v_3,$$

como j es un índice libre, se tiene:

$$\text{Si } j = 1, \delta_{i1}v_i = \delta_{11}v_1 + \delta_{21}v_2 + \delta_{31}v_3 = v_1$$

Si $j = 2$, $\delta_{i2}v_i = \delta_{12}v_1 + \delta_{22}v_2 + \delta_{32}v_3 = v_2$

Si $j = 3$, $\delta_{i3}v_i = \delta_{13}v_1 + \delta_{23}v_2 + \delta_{33}v_3 = v_3$

de donde

$$\delta_{ij}v_i = v_j$$

es decir, por la presencia de δ_{ij} se reemplaza en la componente v_i , el índice i por el j . De ahí el nombre de operador de sustitución.

6.3.2 Símbolo de Levi-Civita

El símbolo de permutación es llamado también de *Levi-Civita* y está definido por:

$$e_{i_1 i_2 \dots i_N} = e^{i_1 i_2 \dots i_N} = \begin{cases} 1 & \text{índices distintos e } i_1 i_2 \dots i_N \text{ es una permutación par de } 1, 2, 3, \dots, N \\ -1 & \text{índices distintos e } i_1 i_2 \dots i_N \text{ es una permutación impar de } 1, 2, 3, \dots, N \\ 0 & \text{en otro caso} \end{cases} \quad (6.6)$$

Este símbolo es utilizado en la definición de determinante de una matriz de $N \times N$:

$$|A| = \sum_{i_1 i_2 \dots i_N} e_{i_1 i_2 \dots i_N} a_{1i_1} a_{2i_2} \dots a_{Ni_N}$$

El determinante de una matriz $N \times N$ consiste en la suma de todos los productos posibles de N elementos que pertenecen a distintas filas y columnas multiplicados por 1 o -1 de acuerdo a si la permutación de los segundos índices es par o impar.

En el caso $N = 3$ se tiene lo siguiente:

$$e_{ijk} = e^{ijk} = [\mathbf{e}_i \mathbf{e}_j \mathbf{e}_k]$$

$$e_{ijk} = e^{ijk} = \begin{cases} 1 & \text{si } ijk \text{ es una permutación par de } 123 \\ -1 & \text{si } ijk \text{ es una permutación impar de } 123 \\ 0 & \text{en otro caso} \end{cases} \quad (6.7)$$

Y al calcular el determinante utilizando la definición anterior, se tiene la suma de seis términos, que son todos los posibles productos de a tres elementos de la matriz

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

$$= e_{123} a_{11} a_{22} a_{33} + e_{231} a_{12} a_{23} a_{31} + e_{312} a_{13} a_{21} a_{32} + e_{321} a_{13} a_{22} a_{31} + e_{132} a_{11} a_{23} a_{32} + e_{213} a_{12} a_{21} a_{33}$$

Luego, reemplazando los símbolos de permutación e_{ijk} de acuerdo a (6.7), resulta

$$|A| = a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{13} a_{22} a_{31} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33}$$

Por otro lado, si se expresa el símbolo de permutación en función de la delta de Kronecker (u operador de sustitución), obtenemos

$$e_{ijk} = e_{lmn} \delta_{li} \delta_{mj} \delta_{nk} \quad (6.8)$$

que es igual al resultado del determinante

$$e_{ijk} = \begin{vmatrix} \delta_{1i} & \delta_{1j} & \delta_{1k} \\ \delta_{2i} & \delta_{2j} & \delta_{2k} \\ \delta_{3i} & \delta_{3j} & \delta_{3k} \end{vmatrix}$$

■ Ejemplo 6.2

$$e_{321} = \begin{vmatrix} \delta_{13} & \delta_{12} & \delta_{11} \\ \delta_{23} & \delta_{22} & \delta_{21} \\ \delta_{33} & \delta_{32} & \delta_{31} \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = -1$$

■

6.4 Operaciones con vectores

Producto escalar

Dados $\mathbf{u} = \beta^i \mathbf{e}_i$ y $\mathbf{v} = \lambda^j \mathbf{e}_j$ con $i, j = 1, 2, \dots, n$ el producto escalar es

$$\mathbf{u} \cdot \mathbf{v} = \beta^1 \lambda^1 + \beta^2 \lambda^2 + \dots + \beta^n \lambda^n = \beta^l \lambda^l$$

Esto se obtiene al reemplazar los vectores \vec{u} y \vec{v} ,

$$\mathbf{u} \cdot \mathbf{v} = (\beta^i \mathbf{e}_i) \cdot (\lambda^j \mathbf{e}_j) = (\beta^i \lambda^j) (\mathbf{e}_i \cdot \mathbf{e}_j) \quad (6.9)$$

Usando (6.4), como $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$, la Ec.(6.9) queda

$$\mathbf{u} \cdot \mathbf{v} = (\beta^i \lambda^j) \delta_{ij} = \beta^i \lambda^i \quad (6.10)$$

La longitud de \mathbf{u} puede escribirse

$$\|\mathbf{u}\| = \sqrt{\beta^i \beta^i}$$

La multiplicación de dos matrices $A \in \mathbb{R}^{m \times k}$ y $B \in \mathbb{R}^{k \times n}$ da por resultado una matriz $C \in \mathbb{R}^{m \times n}$. Si indicamos con el supraíndice la fila y con subíndice la columna, los elementos de la matriz C son

$$c_j^i = \sum_{l=1}^k a_l^i b_j^l$$

que se simplifica usando el convenio de Einstein a

$$c_j^i = a_l^i b_j^l.$$



Note que c_j^i tiene dos índices libres. No es un escalar, porque no todos los índices están afectados a sumatorias.

Producto vectorial

Si $\mathbf{u}_i = \beta_i^j \mathbf{e}_j$ son vectores de \mathbb{R}^3 , $i = 1, 3$, el producto vectorial de \mathbf{u}_1 y \mathbf{u}_2 da como resultado un vector:

$$\mathbf{u}_1 \times \mathbf{u}_2 = e_{ijk} \beta_1^j \beta_2^k \mathbf{e}_i \quad (6.11)$$

donde e_{ijk} es el símbolo de permutación. Esa expresión se obtiene a partir de calcular el determinante

$$\mathbf{u}_1 \times \mathbf{u}_2 = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \beta_1^1 & \beta_1^2 & \beta_1^3 \\ \beta_2^1 & \beta_2^2 & \beta_2^3 \end{vmatrix} = c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3,$$

donde

$$c_1 = \beta_1^2 \beta_2^3 - \beta_2^2 \beta_1^3 \quad c_2 = -\beta_1^1 \beta_2^3 + \beta_2^1 \beta_1^3 \quad c_3 = \beta_1^1 \beta_2^2 - \beta_2^1 \beta_1^2$$

Teniendo en cuenta la definición de los símbolos de permutación (6.7), pueden reescribirse

$$c_1 = e_{123} \beta_1^2 \beta_2^3 + e_{132} \beta_2^2 \beta_1^3 \quad c_2 = e_{231} \beta_1^1 \beta_2^3 + e_{213} \beta_2^1 \beta_1^3 \quad c_3 = e_{312} \beta_1^1 \beta_2^2 + e_{321} \beta_2^1 \beta_1^2$$

luego, en forma reducida, usando la notación de Einstein,

$$c_i = e_{ijk} \beta_1^j \beta_2^k$$

y de ahí se obtiene la expresión (6.11).

■ **Ejemplo 6.3** Si se desarrolla la expresión $c_i = e_{ijk} A^j B^k$ para $i = 1$, se tiene:

$$\begin{aligned} c_1 &= e_{1jk} A^j B^k \\ &= e_{11k} A^1 B^k + e_{12k} A^2 B^k + e_{13k} A^3 B^k \\ &= e_{111} A^1 B^1 + e_{112} A^1 B^2 + e_{113} A^1 B^3 + e_{121} A^2 B^1 + e_{122} A^2 B^2 \\ &\quad + e_{123} A^2 B^3 + e_{131} A^3 B^1 + e_{132} A^3 B^2 + e_{133} A^3 B^3 \end{aligned}$$

De acuerdo a la definición de los símbolos de permutación (6.7) son nulos los términos con índices repetidos, entonces resulta

$$c_1 = e_{123}A^2B^3 + e_{132}A^3B^2 = A^2B^3 - A^3B^2$$

■

Producto mixto

Si $\mathbf{u}_l = \beta_l^i \mathbf{e}_i$ y anotamos como $[\mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3]$ al producto mixto,

$$[\mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3] = \mathbf{u}_1 \cdot (\mathbf{u}_2 \times \mathbf{u}_3)$$

se tiene que

$$[\mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3] = [\beta_1^i \mathbf{e}_i \beta_2^j \mathbf{e}_j \beta_3^k \mathbf{e}_k]$$

$$= \beta_1^i \beta_2^j \beta_3^k [\mathbf{e}_i \mathbf{e}_j \mathbf{e}_k]$$

$$= \beta_1^i \beta_2^j \beta_3^k e_{ijk}$$

o sea, es el determinante de la matriz de las coordenadas β_j^i ,

$$[\mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3] = |\beta_j^i|$$

Producto tensorial

El producto tensorial o diádico entre dos vectores \mathbf{u} y \mathbf{v} está definido de la forma siguiente:

$$\mathbf{u} \otimes \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \otimes \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & u_1 v_3 \\ u_2 v_1 & u_2 v_2 & u_2 v_3 \\ u_3 v_1 & u_3 v_2 & u_3 v_3 \end{bmatrix} \quad (6.12)$$

Se obtiene una matriz \mathbf{A} , $\mathbf{A} = \mathbf{u} \otimes \mathbf{v}$, donde, por ejemplo, $u_3 v_2$ es la componente de la fila 3, y columna 2. En este caso particular, el producto tensorial es el producto matricial usual de los vectores \mathbf{u} (de $n \times 1$) y \mathbf{v}^t , (de $1 \times n$).

Una matriz \mathbf{A} puede escribirse en términos de los productos tensoriales de los vectores de la base \mathbf{e}_1 , \mathbf{e}_2 , y \mathbf{e}_3 , de la forma siguiente:

$$\mathbf{A} = \mathbf{u} \otimes \mathbf{v} = u_1 v_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + u_1 v_2 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + u_1 v_3 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdots + u_3 v_3 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (6.13)$$

$$\mathbf{A} = \mathbf{u} \otimes \mathbf{v} = u_1 v_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + u_1 v_2 \mathbf{e}_1 \otimes \mathbf{e}_2 + \cdots + u_3 v_3 \mathbf{e}_3 \otimes \mathbf{e}_3 \quad (6.14)$$

que, utilizando el convenio de Einstein, puede reescribirse

$$\mathbf{A} = \mathbf{u} \otimes \mathbf{v} = u_i v_j \mathbf{e}_i \otimes \mathbf{e}_j \quad (6.15)$$

6.5 Transformaciones lineales

Sea $T : V \rightarrow W$ una transformación lineal entre dos espacios vectoriales V y W (ver Definición 2.1.1). Sean $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N\}$ una base de V y $\bar{B} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ una base de W . Si aplicamos T a un vector arbitrario $\mathbf{v} \in V$,

$$\mathbf{v} = \alpha^j \mathbf{v}_j$$

$$T(\mathbf{v}) = T(\alpha^j \mathbf{v}_j) = \alpha^j T(\mathbf{v}_j)$$

Como $T(\mathbf{v})$ es un elemento de W , se puede escribir como combinación lineal de los vectores de la base de W ,

$$T(\mathbf{v}) = \beta^i \mathbf{w}_i$$

Por otro lado, si se aplica T a los vectores de la base de V , se obtiene la expresión

$$T(\mathbf{v}_j) = a_j^i \mathbf{w}_i$$

donde los escalares a_i^j son las coordenadas $T(\mathbf{v}_j)$ en la base de W .

En la expresión anterior el término a_i^j tiene dos índices: j está asociado al espacio V e i al espacio W . T es la matriz de la transformación lineal, en cada columna j están las coordenadas de $T(\mathbf{v}_j)$ en la base de W .

$$T(\mathbf{v}) = \alpha^j T(\mathbf{v}_j) = \alpha^j a_i^j \mathbf{w}_i$$

$$T(\mathbf{v}) = \beta^i \mathbf{w}_i = a_i^j \alpha^j \mathbf{w}_i$$

de donde,

$$(\beta^i - a_i^j \alpha^j) \mathbf{w}_i = \mathbf{0}$$

y por ser los \mathbf{w}_i linealmente independientes (forman una base de W), resulta

$$\beta^i = a_i^j \alpha^j$$

que da la relación entre las coordenadas de \mathbf{v} y de $T(\mathbf{v})$. Corresponde al producto matricial

$$\begin{pmatrix} \beta^1 \\ \beta^2 \\ \beta^3 \\ \dots \\ \beta^n \end{pmatrix} = \begin{pmatrix} a_1^1 & a_1^2 & \dots & a_1^n \\ a_2^1 & a_2^2 & \dots & a_2^n \\ \dots & \dots & \dots & \dots \\ a_n^1 & a_n^2 & \dots & a_n^n \end{pmatrix} \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \alpha^3 \\ \dots \\ \alpha^n \end{pmatrix} \quad (6.16)$$

En los elementos de la matriz el supraíndice i y el subíndice j del elemento a_i^j corresponden a la fila y a la columna respectivamente. Es la matriz asociada a la transformación lineal T definida en la Sección 2.2).

■ **Ejemplo 6.4** Si la transformación lineal es la identidad, usamos dos bases distintas, $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ y $B' = \{\mathbf{u}_1, \mathbf{u}_2\}$, y las coordenadas en cada base son $\mathbf{v} = \beta^j \mathbf{e}_j = \alpha^j \mathbf{u}_j$, la relación entre las coordenadas es

$$\beta^i = a_i^j \alpha^j.$$

$$\begin{pmatrix} \beta^1 \\ \beta^2 \end{pmatrix} = \begin{pmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{pmatrix} \begin{pmatrix} \alpha^1 \\ \alpha^2 \end{pmatrix}$$

Estas expresiones se corresponden con las Ecs.(1.4) y (1.5), donde se vio la relación entre las coordenadas del vector \vec{x} en la antigua base B y en la nueva base B' , respectivamente, x^i y x'^i , con $i = 1, 2$. En forma matricial, $X = AX'$, o

$$\begin{cases} x^1 &= a_1^1 x'^1 + a_2^1 x'^2 \\ x^2 &= a_1^2 x'^1 + a_2^2 x'^2 \end{cases} \quad (6.17)$$

Si $\mathbf{u}_1 = (2, 3)$ y $\mathbf{u}_2 = (1, 4)$ usando lo anterior, para el vector

$$\begin{pmatrix} 7 \\ 7 \end{pmatrix}_B = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix}_{B'} = P_{B, B'} \begin{pmatrix} 3 \\ 1 \end{pmatrix}_{B'}$$

La relación está dada por la matriz de cambio de base de B' a B , $P_{B, B'}$ vista en la Sección 1.5.

Por otro lado, La relación entre los vectores de las bases B' y B está dada por la matriz transpuesta (Ver Ec.(1.1) en la Sección (1.5)):

$$\begin{cases} \mathbf{u}_1 &= a_1^1 \mathbf{e}_1 + a_1^2 \mathbf{e}_2 = 2\mathbf{e}_1 + 3\mathbf{e}_2 \\ \mathbf{u}_2 &= a_2^1 \mathbf{e}_1 + a_2^2 \mathbf{e}_2 = 1\mathbf{e}_1 + 4\mathbf{e}_2 \end{cases}$$

Usando el convenio de suma, la relación entre los vectores de las bases tiene la expresión

$$\mathbf{u}_j = a_j^l \mathbf{e}_l$$

■

6.6 Definición de tensor

El concepto de *tensor* tiene su origen en la evolución de la geometría diferencial de Gauss, Riemann y Christoffel. La necesidad del cálculo tensorial, como rama sistemática de la matemática, se debe a Ricci y a su discípulo Levi-Civita, que publicaron en colaboración el primer trabajo sobre esta materia: *Métodos del cálculo diferencial absoluto y sus aplicaciones*, en *Mathematische Annalen*, vol. 54 (1901).

El objeto principal del cálculo tensorial es la investigación de las relaciones que permanecen invariantes cuando se cambia de un sistema de coordenadas a otro. Las leyes de la física no pueden depender del sistema de referencia que elija el físico con fines descriptivos. Por eso es, estéticamente deseable y muchas veces conveniente, utilizar el cálculo tensorial como fundamento matemático en que se puedan formular tales leyes. Einstein, en particular, lo consideró un excelente instrumento para la presentación de su teoría general de la relatividad. El cálculo tensorial alcanzó gran importancia y es hoy en día inapreciable en sus aplicaciones en la mayoría de las ramas de la física teórica; es también indispensable en geometría diferencial.

6.6.1 Cambio de coordenadas

Si se considera un espacio vectorial V de dimensión N con el sistema de coordenadas $x^1, x^2, x^3, \dots, x^n$, las N ecuaciones

$$x'^i = x'^i(x^1, x^2, x^3, \dots, x^n) = \varphi^i(x^1, x^2, x^3, \dots, x^n), \quad i = 1, \dots, N \quad (6.18)$$

donde φ^i son funciones continuas y diferenciables de las coordenadas definen un nuevo sistema de coordenadas x'^1, x'^2, \dots, x'^N .

La condición necesaria y suficiente para que las Ec.(6.18) definan una transformación de coordenadas es que el Jacobiano formado por las derivadas parciales $\frac{\partial x'^i}{\partial x^j}$ no se anule. En ese caso se pueden resolver para las x^i como funciones de x'^i y se obtiene,

$$x^i = \psi^i(x'^1, x'^2, \dots, x'^N), \quad i = 1, \dots, N \quad (6.19)$$

■ **Ejemplo 6.5** En el caso particular de un cambio de base en \mathbb{R}^3 la relación entre las coordenadas está dada por el sistema lineal (1.4) para $n = 3$

$$\begin{cases} x^1 &= a_1^1 x'^1 + a_2^1 x'^2 + a_3^1 x'^3 \\ x^2 &= a_1^2 x'^1 + a_2^2 x'^2 + a_3^2 x'^3 \\ x^3 &= a_1^3 x'^1 + a_2^3 x'^2 + a_3^3 x'^3 \end{cases} \quad (6.20)$$

Corresponde a relaciones como las de la Ec.(6.19). El determinante de la matriz Jacobiana (cuyos elementos son $\frac{\partial x^i}{\partial x'^j} = a_j^i$) no se anula, porque la matriz de cambio de base A tiene inversa. Con su matriz inversa pueden escribirse las ecuaciones (6.18) y así tener las coordenadas x'^i como funciones de las x^i . ■

6.6.2 Tensores de orden 0, 1 y 2

Es importante tener presente la expresión (1.5) de la relación ya vista entre las coordenadas x'^i y x^i ante un cambio de base. Porque los tensores se definen en función de sus propiedades de transformación ante un cambio de coordenadas ([26]).:

$$x^i \rightarrow x'^i \quad i = 1, \dots, N \quad (6.21)$$

dado por las relaciones de las Ecs.(6.18) y (6.19).

Se tiene lo siguiente:

- **Tensor de orden cero o escalar** es una cantidad ϕ que permanece invariante al cambiar al sistema primado,

$$\phi' = \phi$$

Ejemplos La masa, la energía, la temperatura.

- **Tensor de orden uno o vector** son N cantidades

- **Vectores contravariantes**

Las funciones v^j de las N coordenadas x^i se dice que son las componentes de un *vector contravariante* si se transforman según la ecuación:

$$v'^i = \sum_{j=1}^n \frac{\partial x'^i}{\partial x^j} v^j \quad v^j = \frac{\partial x'^i}{\partial x^j} v^j \quad i = 1, \dots, n \quad (6.22)$$

en un cambio de coordenadas de x^i a x'^i .

Ejemplo Los diferenciales dx'^i ,

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j$$

forman las componentes de un vector contravariante, ya que se transforman de la misma forma que la expresión (6.22).

- **Vectores covariantes**

Las funciones v_j de las N coordenadas x^i se dice que son las componentes de un *vector covariante* si se transforman según la ecuación:

$$v'_i = \frac{\partial x^j}{\partial x'^i} v_j \quad (6.23)$$

en un cambio de coordenadas de x^i a x'^i . Ejemplo. El vector gradiente de una función f

$$\frac{\partial f}{\partial x'^i} = \frac{\partial f}{\partial x^j} \frac{\partial x^j}{\partial x'^i} = \frac{\partial x^j}{\partial x'^i} \frac{\partial f}{\partial x^j} \quad (6.24)$$

De acuerdo a la Ec.(6.23) las magnitudes $\frac{\partial f}{\partial x^j}$ son las componentes de un vector covariante (el índice j es considerado un subíndice).

Ejemplos de tensores de orden 1: \mathbf{r} , vector posición y \mathbf{v} vector velocidad.

- **Tensor de segundo orden: son N^2 cantidades:**

- t^{ij} ($i, j = 1, \dots, N$) son las componentes de un tensor dos veces contravariante si se transforman según

$$t'^{ij} = \frac{\partial x'^i}{\partial x^l} \frac{\partial x'^j}{\partial x^m} t^{lm} \quad (6.25)$$

- t_{ij} ($i, j = 1, \dots, N$) son las componentes de un tensor dos veces covariante si se transforman según

$$t'_{ij} = \frac{\partial x^l}{\partial x'^i} \frac{\partial x^m}{\partial x'^j} t_{lm} \quad (6.26)$$

- t_j^i ($i, j = 1, \dots, N$) son las componentes de un tensor una vez contravariante y otra covariante si se transforman según

$$t_j^i = \frac{\partial x^i}{\partial x^l} \frac{\partial x^m}{\partial x'^j} t_m^l \quad (6.27)$$

o,

$$t_i'^j = \frac{\partial x^l}{\partial x'^i} \frac{\partial x'^j}{\partial x^m} t_l^m \quad (6.28)$$

Los tensores de segundo orden están asociados a matrices:

$$t_{ij} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{15} \\ t_{21} & t_{22} & & \\ t_{31} & & & \\ t_{41} & & & \\ t_{51} & \cdots & & t_{55} \end{pmatrix}, \quad t_j^i = \begin{pmatrix} t_1^1 & t_2^1 & \cdots & t_5^1 \\ t_1^2 & t_2^2 & & \\ t_1^3 & & & \\ t_1^4 & & & \\ t_1^5 & \cdots & & t_5^5 \end{pmatrix}$$

Como ejemplo, en (6.12) se obtuvo un tensor de segundo orden \mathbf{A} , a partir del producto tensorial de dos vectores.

6.6.3 Suma. Contracción de índices

La suma de tensores de igual orden es un tensor del mismo orden, y el producto de un escalar por un tensor de orden q da un tensor de orden q . El producto de componentes de un tensor por las componentes de otro da las componentes de un tensor de orden suma de los órdenes originales.

■ **Ejemplo 6.6** Si u_i son las componentes de un vector y t_{ij} son las componentes de un tensor de orden 2,

$$u_i t_{lm}$$

son las componentes de un tensor de orden 3, ya que tiene 3 índices libres. ■

Una operación importante entre tensores es la llamada *contracción de índices*. Es la operación de multiplicar 2 tensores de orden n y m y hacer la suma sobre uno de los índices (de 1 a N). Se obtiene un tensor de orden $n + m - 2$.

También puede realizarse incluso sobre el mismo tensor:

$$T_i^i$$

obteniéndose un nuevo tensor de rango $n - 2$ (es un escalar o tensor de orden 0 para $n = 2$, como se vió en el ejemplo de la traza en la Sección 6.2).

■ **Ejemplo 6.7** Si se tiene un tensor de orden 3, de componentes t_{ijk} , contrayendo el segundo y tercer índice se obtiene un tensor de orden 1:

$$v_i = t_{ijj}$$

■

■ **Ejemplo 6.8** Dados 2 tensores de orden 2, en este caso $n = m = 2$, al sumar sobre el índice j :

$$T^{ij}H_{jm} = T^{i1}H_{1m} + T^{i2}H_{2m} + \cdots = S_m^i \quad (6.29)$$

se tiene como resultado un tensor de rango $n + m - 2 = 2 + 2 - 2 = 2$.

■



El producto escalar de 2 vectores ($n = m = 1$) es un caso particular y el resultado es un escalar o tensor de orden 0.

■ **Ejemplo 6.9** Aparece con frecuencia la contracción de uno de los índices de un tensor de orden 2 con el índice de un vector (corresponde al producto escalar del tensor por el vector) y da como resultado un vector:

$$v_i = t_{ij}u_j$$

■

Un tensor es *simétrico* respecto a dos de sus índices si al permutarlos se obtiene el mismo valor, por ejemplo

$$t_{ij} = t_{ji}$$

o

$$h_{ijkl} = h_{ilkj}$$


(respecto al segundo y cuarto índice). Será *antisimétrico* si cambia de signo:

$$t_{ij} = -t_{ji}$$

o

$$h_{ijkl} = -h_{ilkj}$$

Se dice que es *totalmente simétrico* o *totalmente antisimétrico* cuando se cumple lo anterior respecto de cualquier par de índices.

-  Se llama *tensor isotrópico* a un tensor cuyas componentes son las mismas en cualquier sistema de coordenadas. Todo escalar es un tensor isotrópico pero no hay vector no nulo que sea isotrópico. Se puede mostrar que todo tensor isotrópico de orden 2 es un escalar por la delta de Kronecker δ_{ij} y todo tensor isotrópico de orden 3 es un escalar por los símbolos de Levi-Civita.

Tullio Levi-Civita (1873 - 1941)

Fue un matemático italiano, famoso por su trabajo sobre cálculo tensorial, pero que también hizo contribuciones significativas en otras áreas de las matemáticas. Era discípulo de Gregorio Ricci-Curbastro, el inventor (algunos dicen co-inventor con Levi-Civita) del cálculo tensorial. Su trabajo incluye artículos fundamentales en matemáticas puras y aplicadas, la mecánica celeste (notable en el problema de los tres cuerpos) e hidrodinámica. Levi-Civita personalmente ayudó a Albert Einstein a aprender el cálculo tensorial, en el cual Einstein basaría su relatividad general, y que había luchado por dominar. Su libro de texto en cálculo tensorial *El Cálculo Diferencial Absoluto* (originalmente un conjunto de notas de la conferencia en italiano de coautoría con Ricci-Curbastro) sigue siendo uno de los textos estándares más de un siglo después de su primera publicación, con varias traducciones disponibles. [16]

6.6.4 Transformaciones ortogonales

Si se aplica una transformación lineal la relación entre las coordenadas en distintas bases está dada por la expresión, (ver Sección 1.5 y el Ejemplo 6.4)

$$x^j = a_i^j x^i \quad (6.30)$$

donde a_i^j son los elementos de la matriz cambio de base.

A partir de la Ec.(6.30), si $B = A^{-1}$ se tiene que

$$x'^i = b_j^i x^j \quad (6.31)$$

entonces las coordenadas se transforman mediante una ley contravariante .

En la Ec.(6.22), las cantidades $\frac{\partial x'^i}{\partial x^j}$ son,

$$\frac{\partial x'^i}{\partial x^j} = b_j^i \quad (6.32)$$

así que, se tiene

$$x'^i = \frac{\partial x'^i}{\partial x^j} x^j = b_j^i x^j \quad i = 1, \dots, n. \quad (6.33)$$

De la misma manera, un tensor de segundo rango dos veces contravariante se transforma de la forma siguiente:

$$T'^{ij} = a_l^i a_m^j T^{lm} \quad (6.34)$$

Las transformaciones ortogonales son un caso particular de transformaciones lineales, son aquellas que transforman un sistema de coordenadas cartesianas ortogonales en otro similar también ortogonal y son tales que la inversa de la matriz que la representa es igual a su transpuesta. Corresponden a rotaciones o reflexiones (ver matriz ortogonal Definición 4.6.4, y el Ejemplo 2.6 de la Sección 2.2). Es decir

$$A^{-1} = A^t$$

o bien

$$a_j^i (a^t)^j_l = \delta_l^i$$

o

$$a_j^i a^l_j = \delta^{il}$$

■ **Ejemplo 6.10** Las ecuaciones que corresponden a una rotación de un ángulo ϕ alrededor del eje z son (ver Ec.(6.18) y el Ejemplo 6.5).

$$x'^1 = \cos(\phi)x^1 + \sin(\phi)x^2 \quad (6.35)$$

$$x'^2 = -\sin(\phi)x^1 + \cos(\phi)x^2 \quad (6.36)$$

$$x'^3 = x^3 \quad (6.37)$$

Es una transformación ortogonal ya que la inversa de la matriz que la representa es igual a su transpuesta, o sea $A^t A = I$ (ver Sección 4.6).

Como se vió en la Proposición 4.6.6, si el determinante es 1 corresponde a una rotación y si es -1 corresponde a una reflexión (o a una composición de una simetría y una rotación). ■

■ **Ejemplo 6.11 Una transformación no ortogonal**

Las transformaciones de Lorentz relacionan las coordenadas en dos sistemas de referencia inerciales (x^0, x^1, x^2, x^3) y (x'^0, x'^1, x'^2, x'^3) .

Son las que dejan invariante $s^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$. y su relación en forma matricial es la siguiente:

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

donde $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ es el factor de Lorentz y $\beta = \frac{v}{c}$ es la velocidad relativa respecto de la luz (V es la velocidad del movimiento uniforme y c es la velocidad de la luz en el vacío).

La matriz inversa se obtiene cambiando β por $-\beta$, y por lo tanto se tiene que $A^{-1} \neq A^t$. No es una transformación ortogonal. ■

6.7 Tensores cartesianos

Como se mencionó en la Sección 6.6.2 los tensores están definidos por las propiedades de transformación de sus componentes ante cambios de coordenadas.

Se llaman *tensores cartesianos* a los tensores que están definidos por sus propiedades ante transformaciones entre sistemas de coordenadas cartesianos ortogonales (ver [25]). Esto los diferencia de los tensores en general, en los que se consideran transformaciones más generales de coordenadas.

En los tensores cartesianos no es necesario diferenciar entre componentes covariantes y contravariantes ya que se transforman igual:

En este caso, la Ec.(6.30) puede escribirse

$$x'_i = a_{ij}x_j \quad (6.38)$$

donde a_{ij} son los elementos de la matriz cambio de base.

Los tensores de orden 1 se transforman con la ley

$$v'_i = a_{ij}v_j \quad (6.39)$$

Ya que si se utilizan las relaciones (6.22) y (6.23), como $A^t.A = I$, se obtiene:

$$v'^i = a^i_j v^j \quad (6.40)$$

$$v'_i = (a^{-1})^j_i v_j = (a^T)^j_i v_j = a^j_i v_j \quad (6.41)$$

Análogamente, en el caso de tensores de orden 2, la expresión de la Ec.(6.34) se reescribe

$$t'_{ij} = a_{il}a_{jm}t_{lm} \quad (6.42)$$

Si se asocian las componentes t_{ij} a una matriz T , esta ley de transformación de (6.42) corresponde a la transformación

$$T' = ATA^t \quad (6.43)$$

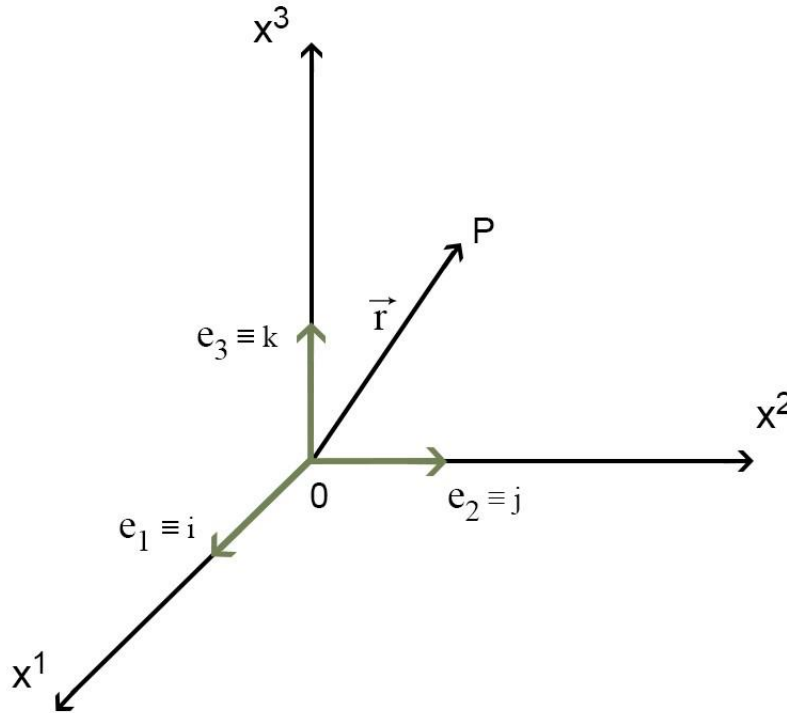
Ejemplos de tensores cartesianos: el vector posición \mathbf{r} , el vector velocidad \mathbf{v} , el tensor de inercia I_{ij} , y el tensor de tensiones τ_{ij} .

6.8 Sistema de coordenadas curvilíneas

Las transformaciones de coordenadas se presentaron en las Ecs. (6.18) y (6.19) de la Sección 6.6.1. Consideremos ahora, en particular, una región del espacio de tres dimensiones referida a un sistema de ejes cartesianos ortogonales, caracterizados con supraíndices x^1 , x^2 , y x^3 .

Sean, además

Figura 6.2: Bases en coordenadas cartesianas



$$\bar{x}^i = \bar{x}^i(x^1, x^2, x^3), \quad i = 1, 2, 3 \quad (6.44)$$

funciones continuas con derivadas parciales primeras continuas y tal que el jacobiano formado por las derivadas parciales $\frac{\partial \bar{x}^i}{\partial x^j}$ no se anule.

Entonces las ecuaciones anteriores pueden resolverse en las x^i , esto es

$$x^j = x^j(\bar{x}^1, \bar{x}^2, \bar{x}^3), \quad j = 1, 2, 3 \quad (6.45)$$

Las variables \bar{x}^i introducidas son tales que a cada punto P le corresponde una única terna de valores de ellas y recíprocamente: las denominamos coordenadas *curvilíneas* de P .

Uno de los sistemas de coordenadas curvilíneas más usados en el espacio son las coordenadas cilíndricas (r, φ, z) , donde, $\bar{x}^1 = r$, $\bar{x}^2 = \varphi$ y $\bar{x}^3 = z$. y tales que las relaciones con las cartesianas (x^1, x^2, x^3) son

$$r = \sqrt{(x^1)^2 + (x^2)^2}$$

$$\varphi = \arctan\left(\frac{x^2}{x^1}\right)$$

$$z = x^3$$

Es decir que la Ec.(6.44) expresa la transformación entre coordenadas cartesianas y coordenadas cilíndricas. En otro caso, puede expresar la transformación entre coordenadas cartesianas y coordenadas esféricas, otro sistema que también es utilizado con frecuencia.

Es importante notar que cuando se pasa de un sistema de coordenadas cartesianas a otras cartesianas, la transformación es lineal, y la relación entre las coordenadas de un mismo punto en los dos sistemas diferentes se obtiene multiplicando por una matriz.

Igualando a una constante la Ec.(6.44)

$$\bar{x}^i = \bar{x}^i(x^1, x^2, x^3) = C_i = \text{constante},$$

obtenemos la ecuación de una superficie para cada valor de la constante: es decir, la última ecuación representa para cada $i = 1, 2, 3$ tres familias de superficies, que se denominan *superficies coordenadas* y la condición de que el jacobiano no se anule, significa geoméricamente que tres de ellas (una de cada familia) se intersecan en uno y sólo un punto P .

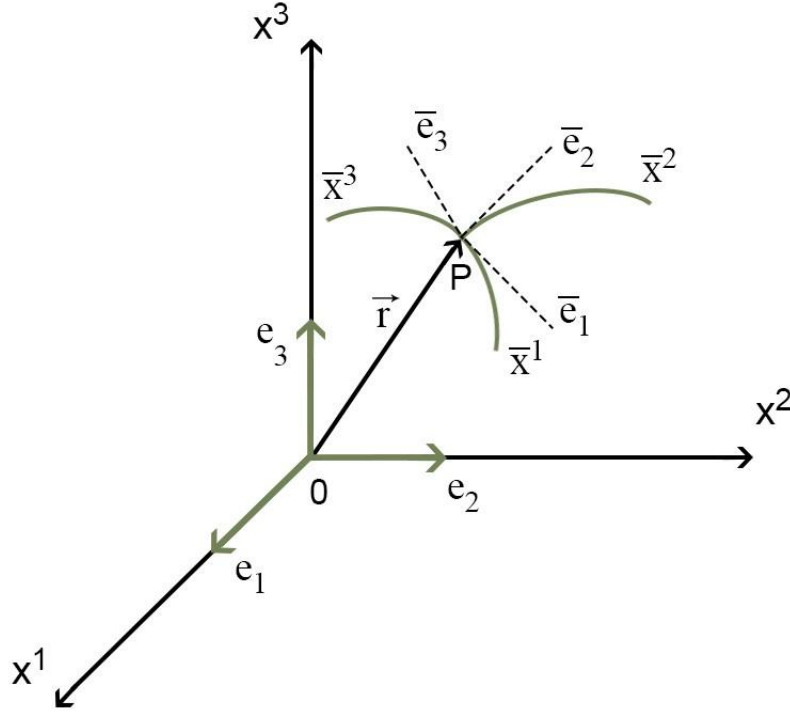
La intersección de las tres superficies que pasan por P determina tres líneas, a lo largo de las cuales sólo una coordenada \bar{x}^i es variable: se denominan líneas coordenadas.

6.8.1 Tensor fundamental (o Tensor Métrico)

Supongamos que se requiere calcular la longitud de un vector \mathbf{v} dado de R^3 , por ejemplo $\mathbf{v} = (7, 4, -1)$. Entonces:

- Al estar dadas sus coordenadas en la base canónica $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, (sino se hubiera anotado $\mathbf{v} = (7, 4, -1)_B$ para hallar la longitud del vector $\mathbf{v} = (7, 4, -1)$ se calcula $\|\mathbf{v}\|^2 = (7)^2 + (4)^2 + (-1)^2 = 49 + 16 + 1 = 66$. Se obtiene que su longitud es $\|\mathbf{v}\| = \sqrt{66}$.
- Si se tienen las coordenadas de \mathbf{v} en la base $B' = \{(1, 1, 0), (4, 2, 1), (2, 1, -2)\}$, $\mathbf{v} = (1, 1, 1)_B$, se deben transformar sus coordenadas a la base canónica para calcular la longitud.

Figura 6.3: Bases en coordenadas curvilíneas



Es deseable, entonces, una definición de longitud de un vector invariante ante un cambio de base. Con el llamado *tensor métrico* se redefine la longitud de un vector según la expresión:

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} \bar{\mathbf{e}}_1 \cdot \bar{\mathbf{e}}_1 & \bar{\mathbf{e}}_1 \cdot \bar{\mathbf{e}}_2 & \bar{\mathbf{e}}_1 \cdot \bar{\mathbf{e}}_3 \\ \bar{\mathbf{e}}_2 \cdot \bar{\mathbf{e}}_1 & \bar{\mathbf{e}}_2 \cdot \bar{\mathbf{e}}_2 & \bar{\mathbf{e}}_2 \cdot \bar{\mathbf{e}}_3 \\ \bar{\mathbf{e}}_3 \cdot \bar{\mathbf{e}}_1 & \bar{\mathbf{e}}_3 \cdot \bar{\mathbf{e}}_2 & \bar{\mathbf{e}}_3 \cdot \bar{\mathbf{e}}_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (6.46)$$

donde $\bar{\mathbf{e}}_i$ son los elementos de la base B' , y (x, y, z) sus coordenadas en esa base.

Las componentes del tensor métrico (representado por los elementos ij de la matriz) son los productos escalares de los vectores de la base B' , es decir, $\bar{\mathbf{e}}_i \cdot \bar{\mathbf{e}}_j$.

Así, para el vector $\mathbf{v} = (1, 1, 1)_B$, se tiene

$$\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}_B \begin{pmatrix} 2 & 6 & 3 \\ 6 & 21 & 8 \\ 3 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}_B = \begin{pmatrix} 11 & 35 & 20 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 66 \quad (6.47)$$

Con esta definición nueva, (que da la longitud del vector elevada al cuadrado), el tensor métrico da la matriz identidad (si la base es la canónica) y se tiene el resultado presentado al inicio:

$$\begin{pmatrix} 7 & 4 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 7 \\ 4 \\ -1 \end{pmatrix} = \begin{pmatrix} 7 & 4 & -1 \end{pmatrix} \begin{pmatrix} 7 \\ 4 \\ -1 \end{pmatrix} = 66 \quad (6.48)$$

De lo anterior surge que es posible introducir el concepto de distancia en un espacio V de dimensión N cualquiera, y que la distancia ds entre dos puntos próximos de coordenadas x^i y $x^i + dx^i$, está dada por la expresión:

$$ds^2 = g_{ij} dx^i dx^j = g_{11}(dx^1)^2 + g_{12} dx^1 dx^2 + \cdots + g_{1N} dx^1 dx^N + \cdots + g_{NN}(dx^N)^2 \quad (6.49)$$

donde g_{ij} son funciones de x^i , con la restricción que $g = |g_{ij}| \neq 0$.

Cuando se cumple esta definición de longitud se dice que el espacio es un *espacio de Riemann*.



- Se postula que la distancia entre dos puntos próximos es independiente del sistema de coordenadas, es decir que ds es un *invariante*.
- A la forma cuadrática $g_{ij} dx^i dx^j$ se la llama *métrica*. g_{ij} es un tensor simétrico covariante de segundo orden llamado *tensor fundamental*. Sus componentes contravariantes están dadas por los elementos de la matriz inversa.

$$g^{ij} g_{jk} = \delta_k^i \quad (6.50)$$

Los coeficientes g_{ij} son funciones de las coordenadas, y se obtienen a partir de los vectores de la base, $\bar{\mathbf{e}}_i$, ya que

$$g_{ij} = \bar{\mathbf{e}}_i \cdot \bar{\mathbf{e}}_j \quad (6.51)$$

- $g_{ij} = g_{ji}$. $g_{ij}dx^i dx^j$ es una forma cuadrática. Se llama *métrica*, y es el cuadrado del elemento de línea ds .
- La longitud de los vectores de la base viene dada por:

$$|\bar{\mathbf{e}}_i| = \sqrt{g_{ii}} \quad (6.52)$$

donde i no se suma en la última expresión. Además, si el sistema es ortogonal, $g_{ij} = 0$ para $i \neq j$.

- En un espacio euclídeo de tres dimensiones, referido a un sistema de ejes cartesianos rectangulares se tiene que la expresión de la Ec.(6.49) para tensores cartesianos para $N = 3$, es

$$ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \quad (6.53)$$

Teniendo en cuenta que,

$$\delta_{ij} = \delta_i^j = \begin{cases} 1 & \text{cuando } i = j, \\ 0 & \text{cuando } i \neq j \end{cases}$$

y desarrollando las sumas,

$$\begin{aligned} \delta_{ij}x^i x^j &= \delta_{1j}x^1 x^j + \delta_{2j}x^2 x^j + \delta_{3j}x^3 x^j \\ &= \delta_{11}x^1 x^1 + \delta_{12}x^1 x^2 + \delta_{13}x^1 x^3 + \delta_{21}x^2 x^1 + \delta_{22}x^2 x^2 + \delta_{23}x^2 x^3 + \dots + \delta_{33}x^3 x^3 \end{aligned}$$

$$\delta_{ij}x^i x^j = (x^1)^2 + (x^2)^2 + (x^3)^2 = ds^2$$



Las expresiones $\delta_{ij}x^i x^j = \delta_{ml}x^m x^l = \delta_{\beta\alpha}x^\beta x^\alpha$ son equivalentes por ser los índices mudos.

El tensor métrico en este caso es

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Las componentes del tensor fundamental son cero, excepto $g_{11} = g_{22} = g_{33} = 1$.

- La métrica en un espacio euclídeo es positiva. Será cero solo cuando $dx^1 = dx^2 = dx^3 = 0$.
- En la teoría especial de la relatividad la métrica no siempre es positiva. Su expresión está dada por

$$ds^2 = -(dx^1)^2 - (dx^2)^2 - (dx^3)^2 + c^2(dx^4)^2 \quad (6.54)$$

- Otro ejemplo de métrica en un espacio euclídeo es la referida a coordenadas polares esféricas $x^1 = r$, $x^2 = \theta$ y $x^3 = \psi$. La métrica está dada por

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\psi^2 \quad (6.55)$$

Otras métricas:

De acuerdo con la teoría de la relatividad general en presencia de materia, la geometría del espacio-tiempo no es plana. La métrica de Schwarzschild describe como se curva el espacio-tiempo a causa de un cuerpo esférico, aislado y estático que no gira sobre sí mismo (r : distancia, G : constante gravitatoria, M : Masa, c : velocidad de la luz y θ = ángulo):

$$\text{Métrica de Schwarzschild: } g_{ij} = \begin{pmatrix} -c^2(1 - \frac{2GM}{c^2 r}) & 0 & 0 & 0 \\ 0 & (1 - \frac{2GM}{c^2 r})^{-1} & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2(\theta) \end{pmatrix}$$

Por otro lado la métrica de Friedman-Lamaitre-Roberson-Waller nos describe la expansión de universo en términos del parámetro k . Si $k > 0$ el universo es cerrado y volverá a plegarse sobre sí mismo generando un nuevo bigbang (teoría del bigcrush), mientras que si $k \leq 0$ el universo se expande sin límites. ($a(t)$ representa la aceleración del universo):

$$\text{Métrica de FLRW: } g_{ij} = \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & a(t)(\frac{1}{1-kr^2}) & 0 & 0 \\ 0 & 0 & a(t)(\frac{1}{1-kr^2}) & 0 \\ 0 & 0 & 0 & a(t)r^2 \sin^2(\theta)(\frac{1}{1-kr^2}) \end{pmatrix}$$

6.8.2 Bases en coordenadas curvilíneas

De acuerdo a la Figura 6.3,

$$\mathbf{r} = x^1 \mathbf{e}_1 + x^2 \mathbf{e}_2 + x^3 \mathbf{e}_3 \quad (6.56)$$

donde

$$\mathbf{e}_j = \frac{\partial \mathbf{r}}{\partial x^j}$$

Para cada punto P del espacio se tienen tres líneas coordenadas \bar{x}^i ; es posible definir tres vectores base para P como:

$$\bar{\mathbf{e}}_i = \frac{\partial \mathbf{r}}{\partial \bar{x}^i} \quad (6.57)$$

que se llaman vectores tangentes a las líneas coordenadas ([22]) \bar{x}^i .

La base $\bar{\mathbf{e}}_i$, representada en la Figura 6.3 es, en general, variable punto a punto y sus versores no necesariamente tienen longitud unitaria. Se trata de una base *local*; cada punto P del espacio tiene su propia base.

En un sistema de coordenadas curvilíneas se tiene, en cada punto P , una base local dada por los vectores $\bar{\mathbf{e}}_i$ de la Ec.(6.57) como se muestra en la Figura 6.3. De ahora en adelante a los vectores de esta base los denotaremos \mathbf{g}_i .

■ **Ejemplo 6.12** Coordenadas cilíndricas

$$\vec{r} = x^1 \mathbf{e}_1 + x^2 \mathbf{e}_2 + x^3 \mathbf{e}_3 \quad (6.58)$$

donde $x^1 = r \cos(\varphi)$, $x^2 = r \sin(\varphi)$ y $x^3 = z$

Es una transformación entre las coordenadas x^i y, \bar{x}^i

$$\bar{x}^1 = r, \bar{x}^2 = \varphi \text{ y } \bar{x}^3 = z$$

Los vectores tangentes (o base covariante, Ec.(6.57)), son

$$\mathbf{g}_1 = \frac{\partial \mathbf{r}}{\partial \bar{x}^1} = \cos(\varphi) \mathbf{e}_1 + \sin(\varphi) \mathbf{e}_2 + 0 \cdot \mathbf{e}_3 \quad (6.59)$$

$$\mathbf{g}_2 = \frac{\partial \mathbf{r}}{\partial \bar{x}^2} = -r \sin(\varphi) \mathbf{e}_1 + r \cos(\varphi) \mathbf{e}_2 + 0 \cdot \mathbf{e}_3 \quad (6.60)$$

$$\mathbf{g}_3 = \frac{\partial \mathbf{r}}{\partial \bar{x}^3} = 0 \mathbf{e}_1 + 0 \mathbf{e}_2 + 1 \mathbf{e}_3 \quad (6.61)$$

■

Existe además, para cada punto P , otro conjunto de tres direcciones que puede ser adoptado para definir otra base local de vectores que denotaremos por \mathbf{g}^i .

Estos vectores \mathbf{g}^i constituyen la denominada *base recíproca o dual* de la \mathbf{g}_i , en virtud de las relaciones:

$$\mathbf{g}^i \cdot \mathbf{g}_j = \delta_j^i \quad (6.62)$$

Pueden obtenerse de la forma siguiente

$$\mathbf{g}^1 = \frac{\mathbf{g}_2 \times \mathbf{g}_3}{[\mathbf{g}_1 \mathbf{g}_2 \mathbf{g}_3]} \quad \mathbf{g}^2 = \frac{\mathbf{g}_3 \times \mathbf{g}_1}{[\mathbf{g}_1 \mathbf{g}_2 \mathbf{g}_3]} \quad \mathbf{g}^3 = \frac{\mathbf{g}_1 \times \mathbf{g}_2}{[\mathbf{g}_1 \mathbf{g}_2 \mathbf{g}_3]} \quad (6.63)$$

donde $[\mathbf{g}_1 \mathbf{g}_2 \mathbf{g}_3] = \mathbf{g}_1 \times \mathbf{g}_2 \cdot \mathbf{g}_3 = E$

En la Figura 6.4 se muestra un ejemplo en un sistema cartesiano.

Se tiene que,

$$\mathbf{g}_1 = \frac{\mathbf{g}^2 \times \mathbf{g}^3}{[\mathbf{g}^1 \mathbf{g}^2 \mathbf{g}^3]} \quad \mathbf{g}_2 = \frac{\mathbf{g}^3 \times \mathbf{g}^1}{[\mathbf{g}^1 \mathbf{g}^2 \mathbf{g}^3]} \quad \mathbf{g}_3 = \frac{\mathbf{g}^1 \times \mathbf{g}^2}{[\mathbf{g}^1 \mathbf{g}^2 \mathbf{g}^3]} \quad (6.64)$$

y $[\mathbf{g}^1 \mathbf{g}^2 \mathbf{g}^3] = E_R$

De la misma forma que en Ec.(6.51), se tiene

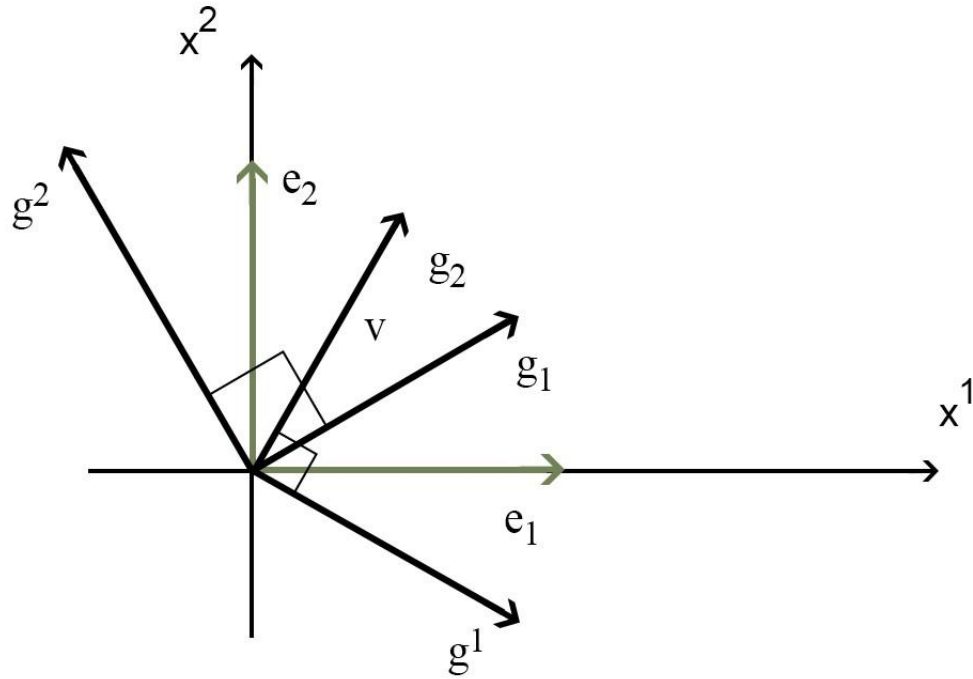
$$g^{ij} = \mathbf{g}^i \cdot \mathbf{g}^j \quad (6.65)$$

y naturalmente, $g^{ij} = g^{ji}$

Decimos que estas dos bases, que son mutuamente recíprocas, y se pueden demostrar las siguientes propiedades

$$\blacksquare [\mathbf{g}_1 \mathbf{g}_2 \mathbf{g}_3] [\mathbf{g}^1 \mathbf{g}^2 \mathbf{g}^3] = 1 \quad E_R = E^{-1}$$

$$\blacksquare \det(g_{ij}) = g = E^2 \quad E = \sqrt{g}$$

Figura 6.4: Ejemplo de una base \mathbf{g}_i y su base recíproca \mathbf{g}^j 

$$\blacksquare g^{ik} g_{kj} = \delta_j^i$$

$$\blacksquare \mathbf{e}^j \times \mathbf{e}_j = 0$$

En un sistema cartesiano ortogonal, la base recíproca coincide con aquella que la genera, es decir, $\mathbf{g}^i = \mathbf{g}_i$.

■ Ejemplo 6.13 Coordenadas cilíndricas.

La base recíproca, usando ec.(6.63)

$$\mathbf{g}^1 = \frac{\mathbf{g}_2 \times \mathbf{g}_3}{[\mathbf{g}_1 \mathbf{g}_2 \mathbf{g}_3]} \quad (6.66)$$

$$\mathbf{g}^1 = (r \cos(\varphi) \mathbf{e}_1 + r \sin(\varphi) \mathbf{e}_2 + 0 \mathbf{e}_3) / r \quad (6.67)$$

$$\mathbf{g}^2 = (-\sin(\varphi) \mathbf{e}_1 + \cos(\varphi) \mathbf{e}_2 + 0 \mathbf{e}_3) / r \quad (6.68)$$

$$\mathbf{g}^3 = (r \cos^2(\varphi) \mathbf{e}_1 + r \sin^2(\varphi) \mathbf{e}_2 + 0 \mathbf{e}_3) / r \quad (6.69)$$

En la Figura 6.5 se muestran las coordenadas cilíndricas en el espacio \mathbb{R}^3 y los vectores \bar{g}_i .

Como por la Ec.(6.51), $g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j$

$$g_{11} = \mathbf{g}_1 \cdot \mathbf{g}_1 = \cos^2(\varphi) + \sin^2(\varphi) = 1$$

$$g_{12} = \mathbf{g}_1 \cdot \mathbf{g}_2 = -r \cos(\varphi) \sin(\varphi) + r \cos(\varphi) \sin(\varphi) = 0$$

$$g_{13} = \mathbf{g}_1 \cdot \mathbf{g}_3 = 0$$

$$g_{22} = \mathbf{g}_2 \cdot \mathbf{g}_2 = r^2 \sin^2(\varphi) + r^2 \cos^2(\varphi) = r^2$$

$$g_{23} = \mathbf{g}_2 \cdot \mathbf{g}_3 = 0$$

$$g_{33} = \mathbf{g}_3 \cdot \mathbf{g}_3 = 1$$

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

y su inversa

$$g^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

■

Relaciones entre versores de base

Dados dos sistemas de coordenadas curvilíneas x^i, \bar{x}^i , y considerando para cada uno de ellos las bases anteriormente introducidas, existen entre sus versores las relaciones:

$$\begin{aligned} \mathbf{g}_j &= \frac{\partial \bar{x}^i}{\partial x^j} \bar{\mathbf{g}}_i, \\ \bar{\mathbf{g}}_i &= \frac{\partial x^j}{\partial \bar{x}^i} \mathbf{g}_j \end{aligned} \quad (6.70)$$

$$\mathbf{g}_i = g_{ij} \mathbf{g}^j \quad \mathbf{g}^i = g^{ij} \mathbf{g}_j \quad (6.71)$$

Si

$$\mathbf{g}^k = \frac{\partial x^k}{\partial \bar{x}^j} \bar{\mathbf{g}}^j, \quad \bar{\mathbf{g}}^k = \frac{\partial \bar{x}^k}{\partial x^j} \mathbf{g}^j \quad (6.72)$$

$$\mathbf{g}_m = \frac{\partial x^i}{\partial \bar{x}^j} g_{mi} \bar{\mathbf{g}}^j, \quad (6.73)$$

se tiene que esta expresión se demuestra de la forma siguiente:

usando Ec.(6.70)

$$\mathbf{g}_m = \frac{\partial \bar{x}^i}{\partial x^m} \bar{\mathbf{g}}_i$$

$$\mathbf{g}_m = \frac{\partial \bar{x}^i}{\partial x^m} \frac{\partial x^j}{\partial \bar{x}^i} \mathbf{g}_j$$

Reemplazando la Ec.(6.71)

$$\mathbf{g}_m = \frac{\partial \bar{x}^i}{\partial x^m} \frac{\partial x^j}{\partial \bar{x}^i} g_{jk} \mathbf{g}^k$$

Teniendo en cuenta la Ec.(6.72)

$$\mathbf{g}_m = \frac{\partial x^j}{\partial \bar{x}^i} \frac{\partial \bar{x}^i}{\partial x^m} g_{jk} \frac{\partial x^k}{\partial \bar{x}^j} \bar{\mathbf{g}}^j$$

y como $\frac{\partial x^j}{\partial \bar{x}^i} \frac{\partial \bar{x}^i}{\partial x^m} = \delta_m^j$,

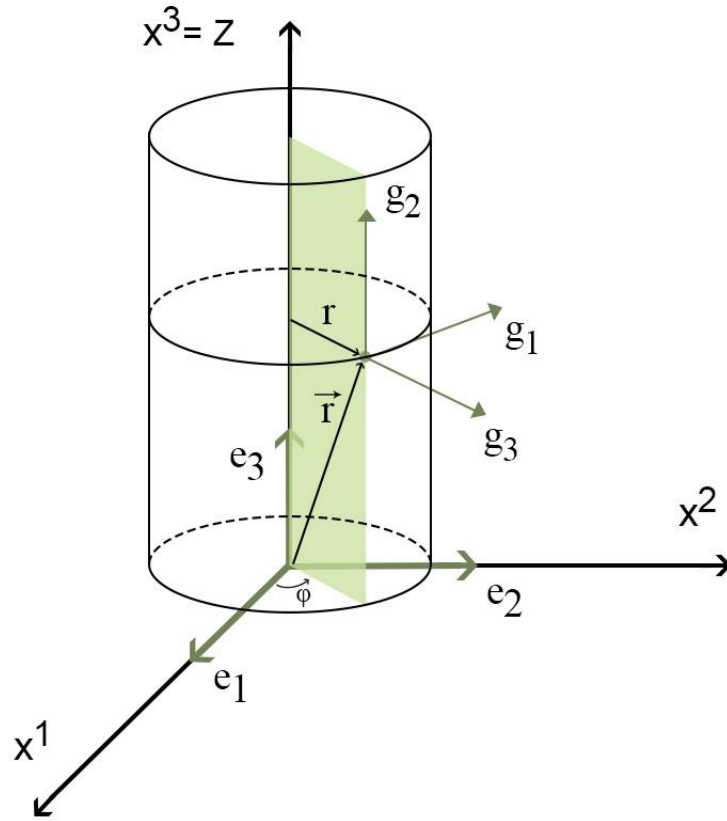
se tiene que

$$\mathbf{g}_m = \delta_m^j g_{jk} \frac{\partial x^k}{\partial \bar{x}^j} \bar{\mathbf{g}}^j$$

$$\mathbf{g}_m = \frac{\partial x^k}{\partial \bar{x}^j} g_{mk} \bar{\mathbf{g}}^j,$$

que coincide con Ec.(6.73) ya que k es un índice mudo.

Figura 6.5: Coordenadas cilíndricas



Componentes contravariantes y covariantes de un tensor

De lo anterior surge que dado un tensor \mathbf{A} se dispone de dos bases aptas para su expresión: la \mathbf{g}_i y su recíproca \mathbf{g}^j . Las componentes de \mathbf{A} en \mathbf{g}_i las denominamos *contravariantes* y las indicaremos A^i , mientras que las componentes en \mathbf{g}^j las llamaremos *covariantes*, designándolas A_j [24]. Entonces:

$$\mathbf{A} = A^i \mathbf{g}_i = A_j \mathbf{g}^j \quad (6.74)$$

Como se mencionó antes, si el sistema es cartesiano ortogonal, ambas componentes son indistinguibles: $A^i = A_i$. Cuando \mathbf{A} esté dado mediante A^i , diremos *vector contravariante*; o vector covariante A_k si se nos presenta mediante esas componentes. Entre ellas se cumple:

$$A_j = g_{ij}A^i \quad A^j = g^{ij}A_i \quad (6.75)$$

así, g_{ij} y g^{ij} bajan y suben índices, respectivamente.

Para demostrar la primera igualdad, partimos de la expresión

$$\mathbf{A} = A_k \mathbf{g}^k = A^l \mathbf{g}_l$$

Multiplicando escalarmente por \mathbf{g}_j a ambos lados,

$$A_k \mathbf{g}^k \cdot \mathbf{g}_j = A^l \mathbf{g}_l \cdot \mathbf{g}_j$$

y teniendo en cuenta la relación entre los vectores de la base y de su base recíproca, Ec.(6.62),

$$A_k \delta_j^k = A^i \mathbf{g}_i \cdot \mathbf{g}_j$$

y se obtiene

$$A_j = g_{ij}A^i$$

Entre las componentes contravariantes (o duales) de \mathbf{A} en dos sistemas x^i, \bar{x}^i se verifica:

$$A^i = \frac{\partial x^i}{\partial \bar{x}^j} \bar{A}^j \quad \bar{A}^j = \frac{\partial \bar{x}^j}{\partial x^i} A^i \quad (6.76)$$

y entre las covariantes

$$A_i = \frac{\partial \bar{x}^j}{\partial x^i} \bar{A}_j \quad \bar{A}_k = \frac{\partial x^j}{\partial \bar{x}^k} \bar{A}_j \quad (6.77)$$

y además,

$$A^j = \mathbf{A} \cdot \mathbf{g}^j \quad A_i = \mathbf{A} \cdot \mathbf{g}_i \quad (6.78)$$

De lo anterior sale que una vez definido el tensor métrico, las componentes covariantes y contravariantes de un tensor están relacionadas por el tensor métrico, así por ejemplo,

$$\begin{aligned}v^i &= g^{ij}v_j \\v_l &= g_{lm}v^m \\T^{ij} &= g^{il}T_l^j = g^{il}g^{jm}T_{lm}\end{aligned}$$

Es importante notar que bases definidas (\mathbf{g}_i y su recíproca \mathbf{g}^i) cumplen la relación de la Ec.(6.62) y que esta relación se mantiene al realizar una transformación de coordenadas. Para demostrarlo, se utilizan los vectores contravariantes y covariantes, $\bar{\mathbf{g}}^i$ y $\bar{\mathbf{g}}_j$, y sus relaciones con los vectores \mathbf{g}^i y \mathbf{g}_j , respectivamente (Ecs.(6.22) y (6.23)). Se verá que $\bar{\mathbf{g}}^i \cdot \bar{\mathbf{g}}_j = \delta^i_j$.

$$\begin{aligned}\bar{\mathbf{g}}^i \cdot \bar{\mathbf{g}}_j &= \frac{\partial \bar{x}^i}{\partial x^l} \mathbf{g}^l \cdot \frac{\partial x^k}{\partial \bar{x}^j} \mathbf{g}_k \\&= \frac{\partial \bar{x}^i}{\partial x^l} \frac{\partial x^k}{\partial \bar{x}^j} \mathbf{g}^l \cdot \mathbf{g}_k \\&= \frac{\partial \bar{x}^i}{\partial x^l} \frac{\partial x^k}{\partial \bar{x}^j} \delta^l_k \\&= \frac{\partial \bar{x}^i}{\partial x^l} \frac{\partial x^l}{\partial \bar{x}^j} \\&= \delta^i_j\end{aligned}$$

Componentes físicas de un vector En ciertos contextos son importantes las componentes físicas de un vector. Si $\mathbf{A} = A^l \mathbf{g}_l$, están dadas por

$$|A^l| |\mathbf{g}_l| = |A^l| \sqrt{g_{jj}}$$

6.9 Diagonalización de tensores de segundo orden. Invariantes

Dado un tensor de segundo orden simétrico y real siempre existe un sistema de coordenadas en el cual las únicas componentes no nulas del tensor son las que tienen los dos índices iguales, $t_{ij} = 0$ si $i \neq j$. Si T es la matriz de los t_{ij} , T es una matriz diagonal:

$$T = \begin{pmatrix} t_{11} & 0 & 0 \\ 0 & t_{22} & 0 \\ 0 & 0 & t_{33} \end{pmatrix}$$

(para $N=3$). El sistema de coordenadas en el que el tensor es diagonal se llama de *ejes principales*.

Para demostrar que todo tensor simétrico y real es diagonalizable, se hace el producto escalar del tensor de componentes t_{ij} por un vector \mathbf{v}_i , con lo que se obtiene otro vector \mathbf{w}_i :

$$\mathbf{w}_i = t_{ij}\mathbf{v}_j.$$

Se trata de hallar los vectores \mathbf{v}_i tales que el vector resultante sea un múltiplo, o sea

$$\mathbf{w}_i = \lambda \mathbf{v}_i \quad i = 1, 2, \dots, N.$$

donde λ es un escalar. En ese caso a \mathbf{v}_i se lo llama *autovector* y a λ *autovalor*.

Para que esto suceda los vectores \mathbf{v}_i deben satisfacer

$$t_{ij}\mathbf{v}_j = \lambda \mathbf{v}_i$$

o

$$(t_{ij} - \lambda \delta_{ij})\mathbf{v}_j = \vec{0}$$

Para el caso que $N = 3$, se tiene un sistema de ecuaciones algebraicas:

$$\begin{cases} (t_{11} - \lambda)\mathbf{v}_1 + t_{12}\mathbf{v}_2 + t_{13}\mathbf{v}_3 = 0 \\ t_{21}\mathbf{v}_1 + (t_{22} - \lambda)\mathbf{v}_2 + t_{23}\mathbf{v}_3 = 0 \\ t_{31}\mathbf{v}_1 + t_{32}\mathbf{v}_2 + (t_{33} - \lambda)\mathbf{v}_3 = 0 \end{cases} \quad (6.79)$$

Como es un sistema homogéneo, para que tenga solución no trivial el determinante del sistema debe ser nulo. A la ecuación

$$\text{Det}(T - \lambda I) = 0 \quad (6.80)$$

se la llama ecuación característica del tensor \mathbf{T} .

El determinante de la Ec.(6.80) es un polinomio de grado 3 con respecto a las potencias de λ :

$$P_T(\lambda) = \lambda^3 + \lambda^2 \mathbf{I}_T + \lambda \mathbf{II}_T - \mathbf{III}_T = 0$$

llamado polinomio característico del tensor \mathbf{T} . \mathbf{I}_T , \mathbf{II}_T , \mathbf{III}_T son los *invariantes principales* del tensor \mathbf{T} , definidos en función de sus componentes t_{ij} por (ver [18]):

$$\mathbf{I}_T = \text{tr}(\mathbf{T}) = t_{ii}$$

$$\mathbf{II}_T = \frac{1}{2}[tr(\mathbf{T})^2 - tr(\mathbf{T}^2)]$$

$$\mathbf{III}_T = Det(\mathbf{T})$$

Si \mathbf{T} es un tensor simétrico, los invariantes principales se resumen de la forma:

$$\mathbf{I}_T = t_{11} + t_{22} + t_{33}$$

$$\mathbf{II}_T = t_{11}t_{22} + t_{11}t_{33} + t_{22}t_{33} - t_{12}^2 - t_{13}^2 - t_{23}^2$$

$$\mathbf{III}_T = t_{11}t_{22}t_{33} + t_{12}t_{13}t_{23} + t_{13}t_{12}t_{23} - t_{12}^2t_{33} - t_{23}^2t_{11} - t_{13}^2t_{22}$$



- Encontrar los autovalores o valores principales es equivalente a encontrar unas direcciones principales (autovectores) tales que $t_{ij} = 0$ para $i \neq j$.
- Una vez obtenidos los autovalores, los autovectores se obtiene resolviendo las ecuaciones $(t_{ij} - \lambda_1 \delta_{ij})\mathbf{n}_j^{(1)} = \vec{0}$, $(t_{ij} - \lambda_2 \delta_{ij})\mathbf{n}_j^{(2)} = \vec{0}$ y $(t_{ij} - \lambda_3 \delta_{ij})\mathbf{n}_j^{(3)} = \vec{0}$.
- Si \mathbf{T} es un tensor simétrico el espacio de los autovectores está definido por una base ortonormal y los autovalores son todos reales.
- Cuando un tensor presenta los tres autovalores iguales, $\lambda_1 = \lambda_2 = \lambda_3$ se denomina *tensor esférico*.

6.10 Tensores de mayor orden

Hemos visto tensores de órdenes 0, 1 y 2. De la misma forma es posible definir tensores de orden mayor.

Un conjunto de N^{s+p} funciones $A_{q_1 q_2 \dots q_p}^{t_1 t_2 \dots t_s}$ de las N coordenadas x^i se dice que son las componentes de un *tensor mixto* de orden $(s+p)$ contravariante de orden s y covariante de orden p si se transforman según la ecuación

$$\bar{A}_{r_1 r_2 \dots r_p}^{u_1 u_2 \dots u_s} = \frac{\partial \bar{x}^{u_1}}{\partial x^{r_1}} \dots \frac{\partial \bar{x}^{u_s}}{\partial x^{r_s}} \frac{\partial x^{q_1}}{\partial \bar{x}^{r_1}} \dots \frac{\partial x^{q_p}}{\partial \bar{x}^{r_p}} A_{q_1 q_2 \dots q_p}^{t_1 t_2 \dots t_s}$$

con el cambio de coordenadas x^i en \bar{x}^i .

Aunque esta expresión parece complicada es simplemente una combinación de la Ec.(6.22) con respecto a los índices contravariantes y de la Ec.(6.23) en cuanto a los índices covariantes.

Un ejemplo de orden 4 es el *tensor constitutivo* \mathbf{C} que relaciona las componentes de dos tensores de orden 2, el *tensor de deformaciones*, $\boldsymbol{\varepsilon}$ con el *tensor de tensiones*, $\boldsymbol{\tau}$.

$$\tau_{ij} = C_{ijkl} \varepsilon_{kl}$$

El tensor constitutivo C_{ijkl} es de orden 4 y sus componentes, considerando dos bases ortogonales e_i y \bar{e}_j del sistema cartesiano, se transforman de la forma siguiente

$$\bar{C}_{ijkl} = p_{im} p_{jn} p_{kr} p_{ls} C_{mnr s} \varepsilon_{kl}$$



En las ecuaciones de elasticidad para el estudio de tensiones y deformaciones en cáscaras y láminas es muy útil expresar el tensor constitutivo en las componentes de una base local no ortogonal. Se utilizan las bases de vectores covariantes y contravariantes.

Juan Martín Maldacena

Nacido en Buenos Aires, 10 de septiembre de 1968. Entre sus muchos aportes al campo de la teoría de supercuerdas —o Teoría M—, se encuentra la denominada «conjetura de Maldacena», «dualidad de Maldacena» o correspondencia AdS/CFT, que propone la equivalencia entre ciertas teorías de gravedad cuántica y cualquier teoría conforme de campos bajo determinadas condiciones que satisfacen el principio holográfico. En 1997 se unió a la Universidad de Harvard como profesor asociado —entonces el profesor asociado vitalicio más joven de la historia de Harvard—. Ahí en 1999 ascendió a profesor titular. En 2012 fue honrado con el nuevo Premio Yuri Milner a la física fundamental. La distinción le dotó con tres millones de dólares. En ese momento sus investigaciones estaban orientadas a la relación entre espacio y tiempo cuánticos y a las teorías de partículas. En 2018 recibió la Medalla Lorentz siendo así el único científico de habla hispana y de Iberoamérica en haberla recibido. [13]

Figura 6.6



6.11 Actividades propuestas

Problema de Aplicación 6 Para el sistema de coordenadas esféricas

$$\vec{r}(\varphi, \phi, r) = r \sin(\varphi) \sin(\phi) \mathbf{e}_1 + r \cos(\phi) \mathbf{e}_2 + r \cos(\varphi) \sin(\phi) \mathbf{e}_3$$

halle:

- Los vectores base covariantes.
- Los vectores base contravariantes.
- La métrica del sistema de coordenadas.
- La matriz Jacobiana de la transformación de coordenadas.
- Calcule el ds^2 del sistema de coordenadas (φ, ϕ, r) .

6.11.1 Ejercicios

Ejercicio 6.1

Responda:

- ¿Qué información dan los subíndices libres?
- ¿Cómo se identifican los subíndices mudos?

- c) ¿Por qué se llama frecuentemente a la delta de Kronecker operador de sustitución?
- d) ¿Qué valor toma la terna 132 para el símbolo de permutación?

Ejercicio 6.2

Reescriba usando notación indicial las siguientes expresiones:

1. $a_1x_1x_3 + a_2x_2x_3 + a_3x_3x_3$
2. $x_1x_1 + x_2x_2$
- 3.

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z &= b_x \\ a_{21}x + a_{22}y + a_{23}z &= b_y \\ a_{31}x + a_{32}y + a_{33}z &= b_z \end{cases}$$

Ejercicio 6.3

Desarrolle las expresiones siguientes para $n = 3$:

1. $\delta_j^i a^j$,
2. $\delta_{ij} x^i x^j$,
3. δ_i^i ,
4. $\frac{\partial f_i}{\partial x^j} dx^j$

Ejercicio 6.4

Verifique en \mathbb{R}^3 las siguientes igualdades:

1. $\delta^{ij} e_{ijk} = 0$
2. $e^{ikm} e_{jkm} = 2 \delta_j^i$
3. $e^{ijk} e_{ijk} = 3!$
4. $e^{ijm} e_{klm} = \delta_k^i \delta_l^j - \delta_l^i \delta_k^j$

Ejercicio 6.5

Utilice el convenio de la suma de Einstein para escribir de forma tensorial:

1. Multiplicación de dos matrices $A \in \mathbb{R}^{n \times m}$ y $B \in \mathbb{R}^{m \times k}$, $C = A.B$ con elementos c_i^j (el supraíndice indica fila y el subíndice indica columna).
2. La traza de una matriz $A \in \mathbb{R}^{n \times n}$.
3. El determinante de una matriz $A \in \mathbb{R}^{n \times n}$.
4. El polinomio característico en función de los invariantes de un tensor.

Ejercicio 6.6

Calcule el producto tensorial de los versores en \mathbb{R}^3 , dos a dos y entre ellos mismos.

Ejercicio 6.7

¿Cuál es el orden de los tensores representados por sus componentes:

$$v_i, \varphi_{ijk}, F_{ijj}, \varepsilon_{ij}, C_{ijkl}, \sigma_{ij}?$$

¿Cuántas componentes tiene cada uno si los índices toman los valores 1, 2, 3 ?

Ejercicio 6.8

Dada la transformación,

$$\bar{x}^1 = 6x^1$$

$$\bar{x}^2 = \frac{-3}{\sqrt{3}}x^1 + 3x^2$$

$$\bar{x}^3 = x^3$$

y su inversa,

$$x^1 = \bar{x}^1$$

$$x^2 = \frac{1}{6\sqrt{3}}\bar{x}^1 + 1/3\bar{x}^2$$

$$x^3 = \bar{x}^3$$

verifique, luego de determinar los versores, $\bar{\mathbf{g}}_i$, y la base recíproca $\bar{\mathbf{g}}^j$, que

$$g_{ij} = \begin{pmatrix} \frac{1}{36} + \frac{1}{36 \cdot 3} & \frac{1}{18\sqrt{3}} & 0 \\ \frac{1}{18\sqrt{3}} & \frac{1}{9} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad g^{ij} = \begin{pmatrix} 36 & -\frac{18}{\sqrt{3}} & 0 \\ -\frac{18}{\sqrt{3}} & 12 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Ejercicio 6.9

Responda:

- 1) ¿A qué se denomina tensor esférico?
- 2) ¿Qué expresión toman los invariantes de una matriz que está en su espacio principal?
- 3) ¿Qué expresión toman los invariantes de un tensor esférico?
- 4) ¿Cómo queda la ecuación característica de un tensor antisimétrico?

- 5) ¿Cualquier combinación de los invariantes principales será un invariante?
- 6) ¿A qué se llama representación espectral de una matriz?
- 7) ¿Cómo se calcula la matriz inversa usando el teorema de Cayley-Hamilton y los invariantes de un tensor?
- 8) ¿La norma de un tensor es también un invariante?

6.11.2 Autoevaluación

Verdadero o Falso

1. Si V es de dimensión finita n , entonces los hiperplanos vectoriales de V son de dimensión $n + 1$.
2. Un hiperplano de V es el núcleo de un funcional lineal no nulo sobre el espacio V .
3. Si un espacio vectorial es suma directa de dos espacios vectoriales, la suma directa de los espacios duales de esos espacios conforman el espacio dual del espacio vectorial original.
4. La distancia de un punto depende de la forma o métrica donde se mide.
5. El teorema de Pitágoras se cumple por igual en un plano o sobre una esfera.
6. El valor de Curvatura Gaussiana o Función K en el Espacio Euclídeo habitual es igual a 1.
7. La distancia más corta entre dos puntos sobre una esfera, se llama geodésica y no es una línea recta.
8. En Radioastronomía se suele utilizar el término cubo de datos para nombrar la imagen de una región del espacio pero a diferentes velocidades.



7. Aplicaciones

En este capítulo se abordan aplicaciones del Álgebra Lineal en dos temas considerados de mucho interés, como son la resolución de ecuaciones diferenciales y la aproximación de funciones. Los sistemas de ecuaciones diferenciales surgieron para analizar cuantitativamente determinados sistemas físicos. En el campo de la astronomía, y contemplando los principios físicos como las leyes del movimiento de Newton y la ley de gravitación, el problema matemático al estudiar el movimiento de dos o más cuerpos, (moviéndose cada uno bajo la acción gravitatoria de los otros) es el de resolver un sistema de ecuaciones diferenciales ordinarias. Por otro lado, el estudio de la teoría de aproximación de funciones también es de importancia fundamental. Comprende dos tipos generales de problemas: uno se refiere a la búsqueda de la función óptima que pueda utilizarse para representar un conjunto de datos y fue tratado en el Capítulo 4, en la aproximación por mínimos cuadrados. En este capítulo nos ocuparemos del problema que se presenta cuando una función se da de manera explícita, pero se quiere encontrar un tipo más simple de ella, por ejemplo un polinomio, que sirva para determinar valores aproximados de la función dada.

7.1 Ecuaciones diferenciales

Muchas leyes de la física, química, biología y economía se expresan en términos de *ecuaciones diferenciales*, es decir, en ecuaciones que comprenden funciones y sus derivadas. En esta sección veremos que es posible aplicar el álgebra lineal para resolver ciertos sistemas de ecuaciones diferenciales. Una de las ecuaciones diferenciales más simples es

$$y' = ay \tag{7.1}$$

donde $y = f(x)$ es una función desconocida que se debe determinar, $y' = \frac{dy}{dx}$ es su derivada y a es una constante. La Ec.(7.1) tiene infinitas soluciones, las cuales son funciones de la forma

$$y = ce^{ax} \quad (7.2)$$

donde c es una constante arbitraria. Estas funciones son soluciones de $y' = ay$, dado que

$$y' = cae^{ax} = ay \quad (7.3)$$

A la Ec.(7.2) se le da el nombre de *solución general* de $y' = ay$

Con frecuencia el problema físico que genera una ecuación diferencial impone alguna condición que permite hallar una solución única a partir de la solución general. Por ejemplo, si se requiere que la solución de $y' = ay$ satisfaga que $y = 3$ cuando $x = 0$, entonces al sustituir en la solución general Ec.(7.3), se obtiene un valor para c :

$$3 = ce^0 = c \quad (7.4)$$

Por lo tanto, $y = 3e^{ax}$ es la única solución de $y' = ay$ que satisface la condición agregada, que se conoce como *condición inicial*. Al problema de resolver una ecuación diferencial sujeta a una condición inicial se denomina *problema de valor inicial*.

Dado ahora un sistema de ecuaciones diferenciales, por ejemplo,

$$\begin{aligned} y_1' &= 3y_1 \\ y_2' &= -2y_2 \\ y_3' &= 5y_3 \end{aligned} \quad (7.5)$$

se desea hallar la solución del sistema que satisface las condiciones iniciales $y_1(0) = 1$, $y_2(0) = 4$ y $y_3(0) = -2$. En forma matricial, se tiene

$$Y' = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 5 \end{pmatrix} Y \quad (7.6)$$

donde $Y = (y_1, y_2, y_3)^T$. Debido a que cada ecuación comprende sólo una función desconocida, se puede resolver cada una de las ecuaciones por separado. De la Ec.(7.2) se obtiene

$$\begin{aligned} y_1 &= c_1 e^{3x} \\ y_2 &= c_2 e^{-2x} \\ y_3 &= c_3 e^{5x} \end{aligned}$$

A partir de las condiciones iniciales dadas, se obtiene

$$\begin{aligned} 1 &= y_1(0) &= c_1 e^0 &= c_1 \\ 4 &= y_2(0) &= c_2 e^0 &= c_2 \\ -2 &= y_3(0) &= c_3 e^0 &= c_3 \end{aligned} \tag{7.7}$$

de modo que la solución que satisface las condiciones iniciales es

$$y_1 = e^{3x}, \quad y_2 = 4e^{-2x}, \quad y_3 = -2e^{5x}.$$

El sistema de este ejemplo fue fácil de resolver porque cada ecuación comprendía solo una función desconocida, y fue ese el caso porque la matriz de coeficientes (7.6) es diagonal. Para resolver el caso cuando la matriz no es diagonal es posible hacer una sustitución para Y , $Y = SU$ que conduzca a un nuevo sistema con una matriz diagonal de los coeficientes; y una vez resuelto este sistema más sencillo, se usa esa solución para determinar la del sistema original. Si se hacen las sustituciones $Y = PU$ e $Y' = PU'$ en el sistema original

$$Y' = AY$$

y se supone que S tiene inversa, se obtiene

$$SU' = A(SU)$$

o bien,

$$U' = (S^{-1}AS)U$$

o bien,

$$Y' = DY$$

donde $D = S^{-1}AS$. Está claro cómo elegir S si se desea que la nueva matriz de los coeficientes D sea diagonal. Se debe elegir S como la matriz que diagonalice a A . El procedimiento para resolver un sistema

$$Y' = AY$$

con una matriz A diagonalizable lo veremos con un ejemplo.

Suponga que una partícula se mueve en un campo de fuerzas plano y que su vector de posición X satisface $X' = AX$ y $X(0) = X_0$, donde

$$Y' = \begin{pmatrix} 4 & -5 \\ -2 & 1 \end{pmatrix} Y$$

las condiciones iniciales $x_1(0) = 2.9$, $x_2(0) = 2.6$. Se desea resolver este problema de valor inicial, y trazar la trayectoria de la partícula para $t \geq 0$. La matriz A de los coeficientes del sistema es

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

Como se analizó en la Sección 3.2 a partir de $\text{Det}(A - \lambda I) = 0$, se obtienen los autovalores de la matriz, $\lambda_1 = 6$ y $\lambda_2 = -1$. Los autovectores correspondientes son $\vec{v}_1 = (-5, 2)^T$ y $\vec{v}_2 = (1, 1)^T$.

De ahí que la matriz

$$S = \begin{pmatrix} -5 & 1 \\ 2 & 1 \end{pmatrix}$$

diagonaliza a A y

$$D = S^{-1}AS = \begin{pmatrix} 6 & 0 \\ 0 & -1 \end{pmatrix}.$$

Por lo tanto, la sustitución $X = SU$ y $X' = SU'$ conduce al nuevo sistema diagonal

$$U' = DU = \begin{pmatrix} 6 & 0 \\ 0 & -1 \end{pmatrix} U.$$

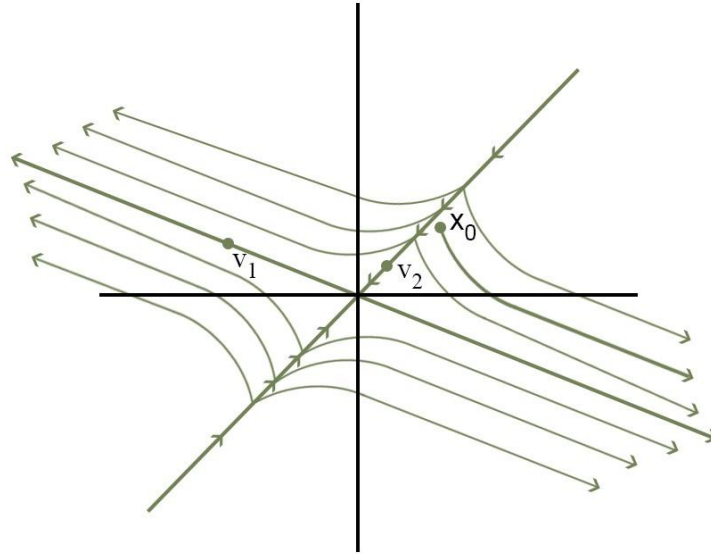
De acuerdo a (7.2), si $U = (u_1, u_2)^t$, la solución de este sistema es

$$\begin{aligned} u_1 &= c_1 e^{6t} \\ u_2 &= c_2 e^{-t} \end{aligned}$$

y la ecuación $X = SU$ proporciona la solución para X

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -5 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} c_1 e^{6t} \\ c_2 e^{-t} \end{pmatrix}$$

Figura 7.1: El origen es un punto silla



o bien

$$x_1 = -5c_1e^{6t} + c_2e^{-t}$$

$$x_2 = 2c_1e^{6t} + c_2e^{-t}$$

Si se sustituyen las condiciones iniciales, se obtiene $c_1 = -3/70$ y $c_2 = 188/70$, de modo que la solución que satisface las condiciones iniciales es

$$x_1 = 15/70e^{6t} + 188/70e^{-t}$$

$$x_2 = -6/70e^{6t} + 188/70e^{-t}$$

Las trayectorias de X se muestran en la Figura 7.1. Al origen se le llama punto silla del sistema dinámico porque algunas trayectorias se aproximan primero al origen y luego cambian de dirección y se alejan de él. Se presenta un punto silla siempre que la matriz A tiene valores propios tanto positivos como negativos. La dirección de mayor repulsión es la línea que pasa por \vec{v}_1 y $\vec{0}$ correspondiente al valor propio positivo. La dirección de mayor atracción es la línea que pasa por \vec{v}_2 y $\vec{0}$, correspondiente al valor propio negativo.

Luis Ángel Caffarelli

Nacido el 8 de diciembre de 1948 en Buenos Aires. Es el principal experto mundial en problemas de frontera libre para ecuaciones diferenciales en derivadas parciales no lineales. También es famoso por sus contribuciones a la ecuación Monge-Ampere y más en general ecuaciones completamente no lineales. Recientemente se ha interesado por los problemas de homogeneización. En 2023, la Academia Noruega de Ciencias le concedió el Premio Abel, el cual es semejante al Nobel en matemáticas, puesto que este último no cuenta con distinciones para esta rama del conocimiento. [15]

7.2 Problemas de aproximación de funciones

En muchas aplicaciones se tiene interés en encontrar la mejor aproximación posible sobre un intervalo, para una función f , por medio de otra función que pertenece a alguna clase especificada; por ejemplo:

- la mejor aproximación posible para e^x en $[0, 1]$ por medio de un polinomio de la forma $a_0 + a_1x + a_2x^2$.
- la mejor aproximación posible para $\text{sen}(\pi x)$ en $[-1, 1]$ por medio de una función de la forma $a_0 + a_1e^x + a_2e^{2x} + a_3e^{3x}$.
- la mejor aproximación posible para $|x|$ en $[0, 2\pi]$ por medio de una función de la forma $a_0 + a_1\text{sen}(x) + a_2\text{sen}(2x) + b_1\cos(x) + b_2\cos(2x)$.

En cada uno de esos ejemplos las funciones de aproximación pertenecen a un subespacio del espacio vectorial $C[a, b]$ (funciones continuas en $[a, b]$), es decir que se está buscando la *mejor aproximación* utilizando funciones de un subespacio W de $C[a, b]$. Intuitivamente, la mejor aproximación posible en $[a, b]$ será aquella que produzca el menor error. Si se desea aproximar en un solo punto x_0 , el error al aproximar $f(x)$ con $g(x)$ estaría dado por $|f(x_0) - g(x_0)|$ (desviación entre f y g en x_0). Si se desea la mejor aproximación en un intervalo, se necesita medir el error global de la aproximación $g(x)$. Una medida posible se obtiene integrando la desviación $|f(x) - g(x)|$ sobre todo el intervalo; es decir,

$$\int_a^b |f(x) - g(x)| dx \quad (7.8)$$

Geométricamente (7.8) es el área entre las gráficas de $f(x)$ y $g(x)$ sobre el intervalo $[a, b]$; mayor el área, mayor será el error global. Aunque es natural, y geoméricamente atractiva, la presencia del

valor absoluto hace que se utilice más frecuentemente otra medida del error, conocida como *error cuadrático medio*, definido por

$$ECM = \int_a^b |f(x) - g(x)|^2 dx \quad (7.9)$$

La ventaja adicional del ECM es que puede escribirse a partir de la teoría de espacios vectoriales con producto interno (ver Ejemplo 4.2).

Considerando el producto interior

$$(f, g) = \int_a^b f(x)g(x)dx \quad (7.10)$$

sobre el espacio vectorial $C[a, b]$, el error cuadrático medio

$$ECM = \|f - g\|^2 = (f - g, f - g) = \int_a^b |f(x) - g(x)|^2 dx \quad (7.11)$$

es el cuadrado de la distancia entre f y g . La aproximación g en W que minimiza el error cuadrático medio es el vector g en W *más próximo* a f con el producto interno (7.10). Por lo que vimos en el Teorema 4.5.1, g es la proyección ortogonal de f sobre el subespacio W . Entonces, si f es una función continua sobre $[a, b]$ y W es un espacio con dimensión finita de $C[a, b]$, la función g en W que minimiza el error cuadrático medio Ec.(7.9) es $g = P_W f$, que se conoce como *aproximación de los mínimos cuadrados para f en W* .

7.2.1 Series de Fourier

Una función de la forma

$$\begin{aligned} f(x) = & c_0 + c_1 \cos(x) + c_2 \cos(2x) + \cdots + c_n \cos(nx) \\ & + d_1 \sin(x) + d_2 \sin(2x) \cdots + c_n \sin(nx) \end{aligned} \quad (7.12)$$

se conoce como *polinomio trigonométrico*; si c_n y d_n no son ambos nulos, entonces se dice que $f(x)$ tiene *orden* n .

■ Ejemplo 7.1

$$f(x) = 5 + \cos(x) - 3\cos(2x) + 7\sin(4x)$$

es un polinomio trigonométrico. $c_0 = 5$, $c_1 = 1$, $c_2 = -3$, $d_1 = d_2 = d_3 = 0$ y $d_4 = 7$ de orden 4. ■

Por la expresión (7.12) los polinomios trigonométricos de orden n o menos son combinaciones lineales de

$$1, \quad \cos(x) \quad \cos(2x), \quad \cos(nx) \quad \sin(x), \quad \sin(2x) \cdots \sin(nx) \quad (7.13)$$

y forman un subespacio W del espacio vectorial de funciones continuas; generado por las $2n + 1$ funciones de (7.13). Se puede demostrar que estas funciones son linealmente independientes y, como consecuencia, forman una base para W .

Si se desea hallar una aproximación para una función continua $f(x)$ sobre el intervalo $[-\pi, \pi]$ o $[0, 2\pi]$ por medio de un polinomio trigonométrico de orden n o menor se deberá calcular la proyección ortogonal de f sobre W . Para hallar esa proyección ortogonal (Teorema 4.5.1), se deberá encontrar una base ortonormal $g_0, g_1, g_2, \dots, g_{2n}$ para W y luego utilizar la fórmula

$$P_W(f) = (f, g_0)g_0 + (f, g_1)g_1 + \dots + (f, g_{2n})g_{2n} \quad (7.14)$$

Es posible obtener una base ortonormal para W aplicando el método de Gram-Schmidt a la base (7.13) usando el producto interno (7.8). Esto conduce a la base ortonormal.

$$\begin{aligned} g_0 &= \frac{1}{\sqrt{2\pi}}, & g_1 &= \frac{1}{\sqrt{\pi}}\cos(x), & g_2 &= \frac{1}{\sqrt{\pi}}\cos(2x), \\ \dots, & & g_n &= \frac{1}{\sqrt{\pi}}\cos(nx), & g_{n+1} &= \frac{1}{\sqrt{\pi}}\sen(x), \\ & & & & \dots & g_{n+1} = \frac{1}{\sqrt{\pi}}\sen(nx) \end{aligned} \quad (7.15)$$

Si se introduce la notación

$$\begin{aligned} a_0 &= \frac{2}{\sqrt{2\pi}}(f, g_0), & a_1 &= \frac{1}{\sqrt{\pi}}(f, g_1), \dots & a_n &= \frac{2}{\sqrt{2\pi}}(f, g_n), \\ & & b_1 &= \frac{1}{\sqrt{\pi}}(f, g_{n+1}), \dots & b_n &= \frac{2}{\sqrt{2\pi}}(f, g_{2n}), \end{aligned} \quad (7.16)$$

Reemplazando en la Ec.(7.14), se obtiene

$$P_W(f) = \frac{a_0}{2} + [a_1\cos(x) + \dots + a_n\cos(nx)] + [b_1\sen(x) + \dots + b_n\sen(nx)] \quad (7.17)$$

Los números $a_0, a_1, \dots, a_n, b_1, \dots, b_n$ se denominan *coeficientes de Fourier de f* .

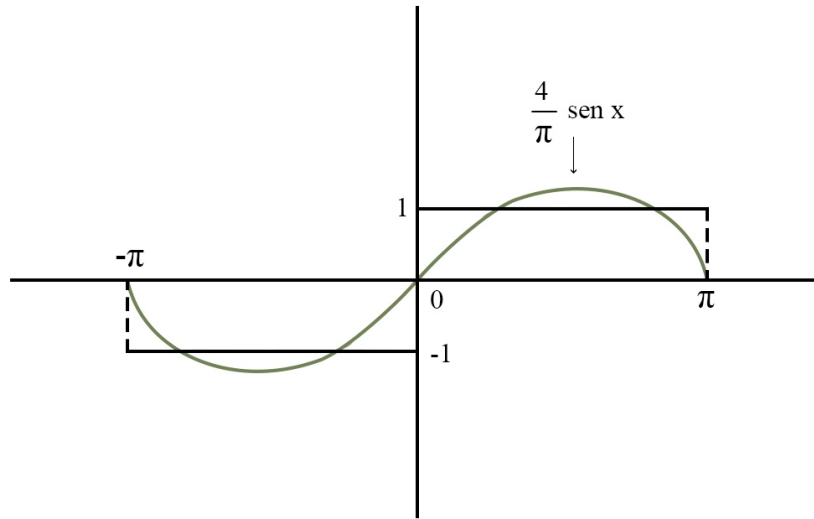
■ **Ejemplo 7.2** Se desea hallar la aproximación de mínimos cuadrados de $f(x)$,

$$f(x) = \begin{cases} -1 & \text{cuando } -\pi \leq x \leq 0, \\ 1 & \text{cuando } 0 < x \leq \pi \end{cases}$$

en $[-\pi, \pi]$ por medio de un polinomio trigonométrico de orden 2 o menor.

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)dx = \frac{1}{2\pi} \left(\int_{-\pi}^0 -1dx + \int_0^{\pi} 1dx \right) = \frac{1}{2\pi} (-\pi + \pi) = 0$$

Figura 7.2: Aproximación mediante un polinomio trigonométrico



$$b_1 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(x) dx = \frac{1}{\pi} \left(\int_{-\pi}^0 -\sin(x) dx + \int_0^{\pi} \sin(x) dx \right) = \frac{4}{\pi}$$

El resto de los coeficientes son nulos, así que la mejor aproximación a $f(x)$ (por medio de un polinomio trigonométrico de orden 2 o menor) es,

$$f(x) \approx \frac{4}{\pi} \sin(x)$$

La función y su aproximación se muestran en la figura 7.2. ■

■ **Ejemplo 7.3** Se desea hallar la aproximación de mínimos cuadrados de $f(x) = x$ en $[0, 2\pi]$ por medio de un polinomio trigonométrico de orden 2 o menor.

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) dx = \frac{1}{\pi} \int_0^{2\pi} x dx = 2\pi \quad (7.18)$$

Para $k = 1, 2, \dots, n$ se puede verificar que, realizando integración por partes, se obtiene

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx = \frac{1}{\pi} \int_0^{2\pi} x \cos(kx) dx = 0$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx = \frac{1}{\pi} \int_0^{2\pi} x \sin(kx) dx = \frac{-2}{k}$$

Por lo tanto, la aproximación de mínimos cuadrados por medio de un polinomio trigonométrico de orden 2 o menor es

$$x \approx \pi - 2\text{sen}(x) - \text{sen}(2x)$$

■

De lo anterior se desprende que la aproximación de mínimos cuadrados de $f(x) = x$ en $[0, 2\pi]$ por medio de un polinomio trigonométrico de orden n o menor, teniendo en cuenta (7.19), es

$$x \approx \pi - 2\left(\text{sen}(x) + \frac{\text{sen}(2x)}{2} + \frac{\text{sen}(3x)}{3} + \dots + \frac{\text{sen}(nx)}{n}\right)$$

y resulta obvio esperar que el error cuadrático medio disminuya a medida que se aumenta el número de términos en la aproximación de mínimos cuadrados

$$f(x) \approx \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \text{sen}(kx))$$

Se puede probar que el error cuadrático medio tiende a 0 cuando $n \rightarrow \infty$, esto se denota escribiendo

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \text{sen}(kx)) \quad (7.19)$$

El segundo miembro de esta ecuación se denomina *serie de Fourier* para f . Las series de este tipo tienen importancia primordial en ingeniería, ciencias y matemáticas.

7.2.2 Series de Haar. Bases de wavelets ortogonales

En la sección anterior se describió el sistema trigonométrico.

$$\{1, \cos(nx), \text{sen}(nx)\}_{n \in \mathbb{N}} \quad (7.20)$$

de período 2π a partir del cual se halla la serie de Fourier de una función $f(x)$. Puede reescribirse con exponenciales complejas

$$\{e^{inx}\}_{n \in \mathbb{Z}} \quad (7.21)$$

ya que por la fórmula de Euler, $e^{ix} = \cos(x) + i\text{sen}(x)$.

Los sistemas (7.20) y (7.21) pueden obtenerse uno del otro mediante combinaciones lineales simples. En particular, para $n \in \mathbb{Z}$,

$$e^{inx} = \begin{cases} \cos(nx) + i\text{sen}(nx) & \text{si } n \neq 0, \\ 1 & \text{si } n = 0 \end{cases}$$

y para $n \in \mathbb{N}$

$$\cos(nx) = \frac{e^{inx} + e^{-inx}}{2}$$

y

$$\sin(nx) = \frac{e^{inx} - e^{-inx}}{2i}.$$

La serie de Fourier (7.19) puede escribirse, entonces,

$$f(x) = \sum_{k \in \mathbb{Z}} c(k) e^{inx} \quad (7.22)$$

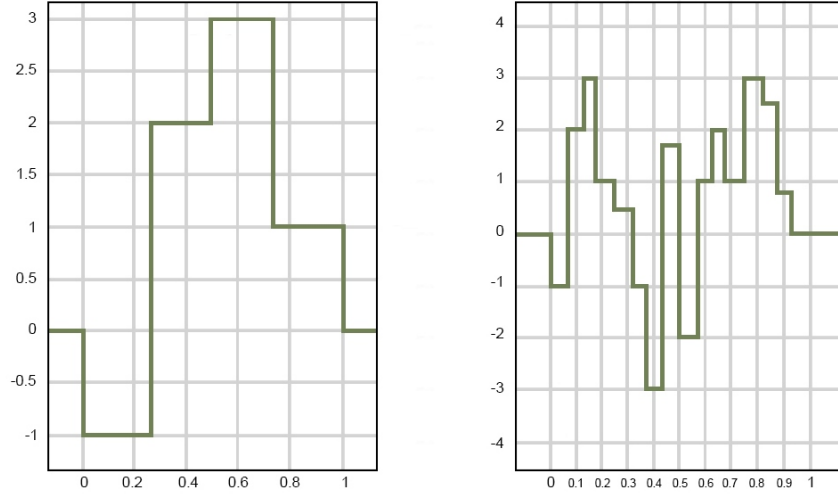
con ciertos coeficientes $c(k)$. Vamos a presentar en esta sección un ejemplo de sistema ortogonal en $[0, 1]$ conocido como *el sistema de Haar*. Es la más simple e históricamente el primer ejemplo de una base *wavelet ortogonal*. Muchas de sus propiedades contrastan con las propiedades del sistema trigonométrico (7.21):

- tienen soporte en pequeños subintervalos de $[0, 1]$, mientras que las funciones base de Fourier son no nulas en todo el intervalo $[0, 1]$.
- son escalonadas, con discontinuidades, mientras que las funciones base de Fourier con C^∞ en $[0, 1]$.
- las bases de Haar tienen un índice que indica la escala j que reemplaza a la frecuencia n de las bases de Fourier.
- las bases de Haar proveen una representación eficiente para funciones que son suaves en algunos segmentos y con picos y discontinuidades en otros, mientras que las bases de Fourier dan buenas representaciones para funciones con comportamiento oscilatorio en intervalos largos.

Alfréd Haar (1885 - 1933)

Matemático húngaro de origen judío, nacido en 1885. En 1904 comenzó a estudiar en la Universidad de Gotinga. Su doctorado fue supervisado por David Hilbert. La medida de Haar, la ondícula de Haar y la transformación de Haar reciben su nombre. Entre 1912 y 1919 enseñó en la Universidad Francisco José de Kolozsvár. Junto con Frigyes Riesz, hizo de la Universidad de Szeged un centro de las matemáticas. También fundó la revista *Acta Scientiarum Mathematicarum* junto con Riesz. [2]

Figura 7.3: Bases de Haar. Escala $j = 2$ a la izquierda y $j = 4$ a la derecha



Para cada par de enteros $j, k \in \mathbb{Z}$, se definen el intervalo diádico $I_{j,k}$:

$$I_{j,k} = [2^{-j}k, 2^{-j}(k+1)) \quad (7.23)$$

y la función de Haar:

$$h_{j,k}(x) = 2^{j/2}(\chi_{I_{j,k}^l}(x) - \chi_{I_{j,k}^r}(x)) \quad (7.24)$$

De esta forma, $h_{j,k}(x)$ está soportada en el intervalo $I_{j,k}$ (no se anula en ese intervalo). Decimos que la función de Haar $h_{j,k}(x)$ está asociada a ese intervalo.

La longitud del intervalo $I_{j,k}$ es 2^{-j} . Si j es grande, la longitud es pequeña. Se dice, entonces, que la función $h_{j,k}(x)$ está bien localizada en el tiempo. Esta propiedad contrasta con las bases de Fourier que tienen todas módulo 1 para todo $x \in [0, 1)$ y por lo tanto no se anulan para ningún x de ese intervalo.

En la Figura 7.3 se muestra un ejemplo de aproximación de una función mediante bases de Haar, para dos escalas o niveles de resolución diferentes.

Las bases de Haar, creadas por Alfred Haar en 1909, fueron el primer registro histórico de lo que hoy se denomina familias de funciones *ondículas* o *wavelets* desarrolladas en los últimos 40 años

para poder analizar señales que no se comportan en forma estacionaria o que presentan cambios bruscos en intervalos pequeños. Esas señales de interés provienen de distintas áreas como la medicina, sismología, geología, electrónica y también astronomía. Así, la Teoría Wavelet, caracterizada por una base matemática compleja, constituye una potente herramienta en el procesamiento de señales y de imágenes digitales. Permite la reducción de ruido, la compresión de señales (muy importante tanto para la transmisión de grandes cantidades de datos como en su almacenamiento) o la detección de determinados objetos en imágenes o en irregularidades locales, por ejemplo en un electrocardiograma (ECG).

El concepto de wavelets como lo conocemos fue propuesto por Jean Morlet y el equipo del Centro de Física Teórica de Marsella, Francia. Con el fin de descomponer y estudiar ciertas señales sísmicas, diseñaron la wavelet que se muestra en la Figura 7.4. Cabe señalar que los métodos del análisis wavelet fueron desarrollados principalmente por Yves Meyer y sus colegas y que recién en 1988 apareció el primer algoritmo de cálculo y su autor fue Stéphane Mallat. Desde entonces la investigación acerca del análisis Wavelet captó mucho interés y se destacan científicos como Ingrid Daubechies, quien en 1988 creó una familia de ondículas o wavelets ortogonales con soporte compacto. En la Figura 7.5 se muestra la wavelet de Daubechies de orden 6, utilizada con frecuencia, por su similitud, para analizar electrocardiogramas.

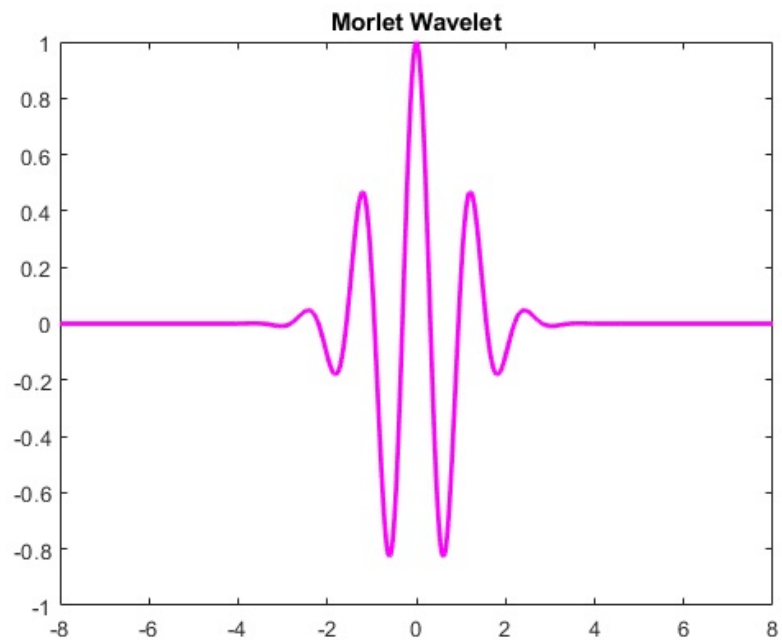
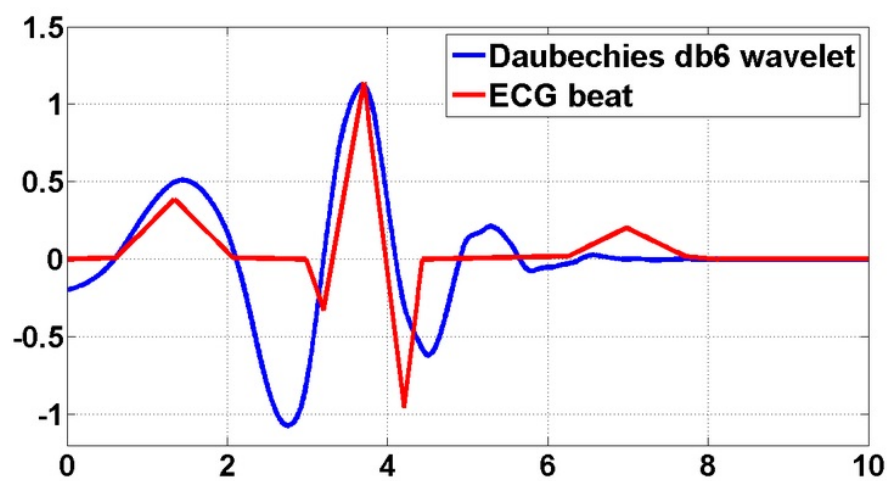
Su aplicación se extiende a campos muy diversos. En cuanto a las aplicaciones en medicina, el análisis con Wavelets permite interpretar los resultados de exámenes médicos, facilitando el diagnóstico de enfermedades.

En Astronomía, algunos ejemplos del uso de Wavelets son:

- para el procesado de imágenes planetarias.
- en el estudio de la actividad solar.
- para la detección de períodos en curvas de luz.

Ingrid Daubechies

Es una matemática y física belga. Nació en 1954. Estudió física en la Vrije Universiteit Brussel (la universidad de Bruselas en lengua flamenca), en la que también se doctoró en física teórica en 1980 y estuvo investigando hasta 1987. Ese año se trasladó a Estados Unidos con su marido, el también matemático Robert Calderbank, recién casados. Daubechies trabajó en los Laboratorios Bell de Nueva

Figura 7.4: Wavelet Morlet*Figura 7.5: Wavelet de Daubechies y latido de un ECG*

Jersey y en varias universidades estadounidenses. En 1993 se convirtió en profesora de matemática computacional en la Universidad de Princeton hasta 2011, cuando trasladó a la Universidad Duke como catedrática de matemáticas. En 2012, el rey Alberto II de Bélgica la concedió el título de Baronesa en reconocimiento de su trayectoria profesional. Es miembro de numerosas instituciones. Fue la primera mujer matemática en presidir la Unión Matemática Internacional (desde 2011). En 1993 fue admitida en la Academia Estadounidense de las Artes y las Ciencias, en 1998 en la Academia Nacional de Ciencias de Estados Unidos y en 2012 en la Sociedad Estadounidense de Matemática. Además, ha sido invitada a participar en numerosas ocasiones en el Congreso Internacional de Matemáticas. Daubechies ha recibido numerosos premios, entre ellos destacan el Premio Nemmers en Matemáticas de 2012 y el Premio Fundación BBVA Fronteras del Conocimiento en Ciencias Básicas 2012 junto a David Mumford. En 2020 fue reconocida, junto a Emmanuel Candès, Yves Meyer y Terence Tao, con el Premio Princesa de Asturias de Investigación Científica y Técnica por «haber realizado contribuciones pioneras y trascendentales a las teorías y técnicas modernas del procesamiento matemático de datos y señales». En 2023 recibió el Premio Wolf en Matemáticas, por sus investigaciones sobre ondículas y análisis armónico aplicado. Daubechies es la primera mujer que ha recibido este reconocimiento.

Ingrid Daubechies ha trabajado en el campo de las ondículas, herramientas que permiten el análisis de señales para entregar información temporal y frecuencial de manera casi simultánea. En 1988, Daubechies propuso la ondícula ortogonal con soporte compacto (conocida como ondícula Daubechies), y en 1992 la ondícula biortogonal, también conocida como ondícula CDF (Cohen-Daubechies-Feauveau), empleada para el formato de compresión de imágenes JPEG 2000. Estas herramientas matemáticas permiten el avance e investigación tanto en matemática teórica como aplicada, pues sirve en la demostración tanto de teoremas como en el desarrollo de las telecomunicaciones, tanto en audio como vídeo, y hasta el ámbito biosanitario, con transmisión de datos de imágenes sanitarias. [11]



Apéndice: Ejercicios preliminares

¿Puede un escalar no ser un número?

- 1) A.B.C.D.E.F.G.H.I.J.K.L.M.N.P.Q.R.S. T.U.V.W.X.Y.Z
- 2) AA, B; AAA, C; AAAA, D; AAAAA, E; AAAAAA, F; AAAAAAA, G; AAAAAAAA, H; AAAAAAAAA, I; AAAAAAAAAA, J;
- 3) AKALB; AKAKALC; AKAKAKALD; AKALB; BKALC; CKALD; DKALE; BKELG; GLEKB; FKDLJ; JLFKD.
- 4) CMALB; DMALC; IMGLB.
- 5) CKNLC; HKNLH; DMDLN; EMELN.
- 6) JLAN; JKALAA; JKBLAB; AAKALAB; JKJLBN; JKJKJKJLCN; FNKGLFG.
- 7) BPCLF; EPBLJ; FPJLFN.
- 8) FQBLC; JQBLE; FNQFLJ.
- 9) CRBLI; BRELCB.
- 10) JPJLJRBLSLANN; JPJPJLRCLTLANN; JPSLT; JPTLJRD.
- 11) AQJLU; UQJLAQSLV.
- 12) ULWA; UPBLWB; AWDMALWDLDP; VLWNA; VPCLWNC; VQJLWNA; VQSLWNNNA; JPEWFGHLEFWGH; SPEWFGHLEFGWH.
- 13) GIWIHYHN; TKCYT; ZYCWADAF.

14) DPZPWNNIBRCQC.

Mensaje interplanetario

Por Adrián Paenza

Supongamos que uno quisiera mandar un mensaje al espacio de manera tal de que en el caso de que fuera interceptado por algún ser “inteligente”, éste pudiera leerlo e interpretarlo. ¿Cómo hacer para escribir algo en “ningún idioma” en particular, pero lo suficientemente explícito como para que cualquiera que pueda “razonar” lo pueda entender? Por otro lado, una vez superado el obstáculo del “medio” o sistema de comunicación que se va a utilizar y que suponga que “el otro” va a entender, ¿qué escribirle?, ¿qué decirle?

Ahora quiero contar una historia que apareció en un diario japonés. Alicia Dickenstein, una de las mejores matemáticas argentinas de la historia y actual profesora en Exactas (UBA), volvía de un viaje por Oriente y me comentó lo que había leído en la revista El Correo de la Unesco, correspondiente al mes de enero de 1966. Me tomo el atrevimiento de reproducirlo textualmente ya que el texto circula por Internet desde hace muchísimo tiempo:

“En 1960, Ivan Bell, un profesor de inglés en Tokio, oyó hablar del ‘Project Ozma’, un plan de escucha de los mensajes que por radio pudieran venirnos desde el espacio. Bell redactó entonces un mensaje interplanetario de 24 símbolos, que el diario japonés Japan Times publicó en su edición del 22 de enero de 1960, pidiendo a sus lectores que lo descifrarán. El diario recibió cuatro respuestas. De las cuatro, una correspondió a una lectora norteamericana que contestó usando el mismo código que había sido utilizado para escribir el mensaje, agregando que ella vivía en Júpiter.”

Acá usted se va a encontrar con el mensaje de Ivan Bell que, como dice el artículo original, es “extraordinariamente fácil de descifrar y mucho más sencillo de lo que parece a simple vista”.

Es un ejemplo muy disfrutable y original de lo que puede hacer el intelecto humano, cualquiera sea el idioma que hable: sólo se requiere tener voluntad de pensar. Acá va la lista de 14 frases. La numeración corre por cuenta mía, pero piense que cada línea es una parte del mensaje.

Ejercicios

Producto escalar. Norma. Distancia y ángulo entre vectores

1. Calcule $\vec{u} \cdot \vec{v}$, siendo \cdot el producto escalar, para $\vec{u} = (2, -5, -1)$ y $\vec{v} = (3, 2, -3)$.

```
import numpy as np
a=np.array([2,-5,-1])
b=np.array([3,2,-3])
print (a.dot(b))
```

2. Sea $\vec{v} = (1, -2, 2, 0)$, encuentre un vector unitario \vec{u} en la misma dirección que \vec{v} .
3. Demuestre que \vec{c} es ortogonal a \vec{d} siendo $\vec{c} = (4/3, -1, 2/3)$ y $\vec{d} = (5, 6, -1)$.
4. Determine el coseno del ángulo entre los vectores $\vec{u} = (2, -5, -1)$ y $\vec{v} = (3, 2, -3)$, en estadística este valor recibe el nombre de coeficiente de correlación. Si el valor está cercano a 1 o a -1 los datos están relacionados, de lo contrario si el valor es cercano a 0 no existe ninguna relación entre ellos.
5. Encuentre la distancia entre $\vec{x} = (1, -1, 2)$ y $\vec{y} = (3, 4, -5)$.

Sistemas de ecuaciones lineales

6. Dado el sistema

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 - 2x_3 = 3 \end{cases}$$

- a) Compruebe que la terna $(2t + 3, -3t - 2, t)$ es solución de dicho sistema, $\forall t \in \mathbb{R}$.
 - b) Justifique por qué este sistema tiene infinitas soluciones.
 - c) Indique cómo se clasifican los sistemas de ecuaciones lineales.
7. Dados los sistemas

$$\begin{cases} x_1 - x_2 = 3 \\ 2x_1 - x_2 = 5 \end{cases}$$

y

$$\begin{cases} x_1 - x_2 = 3 \\ x_2 = -1 \end{cases}$$

- a) Compruebe que tienen el mismo conjunto solución.

- b) Comente la relación entre los dos sistemas.
 c) Mencione cuáles son las llamadas operaciones elementales.
8. Encuentre los valores de b que hacen que el sistema

$$\begin{cases} x_1 + bx_2 - 2x_3 = 2 \\ -x_1 + (b-2)x_2 + 2x_3 = -2 \\ 2x_1 + 2x_2 + (b-4)x_3 = 3 \end{cases}$$

Tenga:

- 1) Una solución.
- 2) Infinitas soluciones.
- 3) Ninguna solución.
- 4) Describa el conjunto solución para a) y b)

Nota: utilice el algoritmo de eliminación Gauss prestando atención a la notación de las operaciones.

Matrices. Matrices semejantes. Matrices elementales.

El mismo Arthur Cayley relató en 1894 que lo condujo a las matrices, el ser estas un modo conveniente de expresar las ecuaciones

$$\begin{cases} x' = ax + by \\ y' = cx + dx \end{cases}$$

Simbolizando esta transformación lineal con dos variables independientes por medio de la disposición en cuadro.

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

9. Halle las matrices X e Y sabiendo que:

$$X + Y = \begin{pmatrix} 3 & 1 & -4 \\ 7 & 1 & 5 \end{pmatrix}$$

$$2X - 3Y = \begin{pmatrix} -4 & -3 & 2 \\ -6 & 2 & -5 \end{pmatrix}$$

10. Si A es una matriz de tamaño $m \times n$ y B es una matriz de tamaño $n \times p$, el producto de las matrices A y B es la matriz de tamaño $m \times p$, cuyo elemento (i, k) es el producto escalar de la fila i de la matriz A por la columna k de B , o sea:

$$(A.B)_{i,k} = \sum_{j=1}^n a_{ij}b_{jk}$$

Dadas las matrices: $\begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 2 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix}$, $\begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix}$

Analice en qué casos es posible calcular: $A.B - B.A$, a ésta diferencia se la conoce como conmutador.

```
import numpy as np
a = np.array([[1, 0],
              [0, 1]])
b = np.array([[4, 1],
              [2, 2]])
a @ b
```

11. Dadas las matrices

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

Es de destacar que se simplifica el producto $A.B$ si se utilizan submatrices.

12. Calcule A^6 siendo:

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

13. Dados los tres pares de datos $(0, -1)$, $(1, 1)$ y $(2, 0)$, halle el polinomio de grado menor o igual que 2: $p(x) = a_0 + a_1x + a_2x^2$, que pasa por dichos pares de datos. Evalúe el valor de $p(x)$ si $x = 2/3$. Al polinomio resultante se lo denomina: Polinomio de Interpolación.

Determinantes y matrices inversibles. Rango de una matriz.

14. ¿Es invertible la siguiente matriz?:

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

Justifique de varias maneras su conclusión.

15. Sea la matriz:

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

Encuentre los valores del parámetro b para que A sea invertible.

16. Piense una forma conveniente para calcular el valor del determinante de una matriz cuadrada y úselo para calcular el determinante de:

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

17. Encuentre el rango de la matriz en función del parámetro a :

$$A = \begin{pmatrix} 1 & a & -1 & 2 \\ 2 & -1 & a & 5 \\ 1 & 10 & -6 & 1 \end{pmatrix}$$

18. Demuestre que si

- 1) $D = (d_{ij})$ una matriz diagonal de tamaño $n \times n$. Entonces $\text{Det}(D) = \prod_i^n d_{ii}$.
- 2) Si A es una matriz triangular del tamaño $n \times n$, entonces $\text{Det}(A)$ es producto de los elementos diagonales de A .
- 3) Una matriz A de tamaño $n \times n$, el $\text{Det}(tA) = t^n \text{Det}(A)$ siendo t un escalar cualquiera.

19. Demuestre el siguiente enunciado:

Supongamos que \vec{u}_0 es una solución particular del sistema de ecuaciones lineales $A\vec{x} = \vec{b}$. Si \vec{v} es una solución cualquiera del sistema homogéneo asociado $A\vec{x} = \vec{0}$, entonces $\vec{u}_0 + \vec{v}$ es solución de $A\vec{x} = \vec{b}$.

Aplique lo demostrado para el siguiente sistema:

$$\begin{cases} x + z + w = 4 \\ 2x + y - w = -2 \\ 3x + y + z = 7 \end{cases}$$

Autoevaluación

Verdadero o Falso.

1. Si la distancia entre \vec{u} y \vec{v} es igual a la distancia de \vec{u} y $-\vec{v}$, entonces \vec{u} y \vec{v} son ortogonales.
2. Sean las matrices $A, B \in \mathbb{R}^{n \times n}$ se verifica que $AB^t = A^t B$.
3. Consideremos la matriz $A \in \mathbb{R}^{3 \times 3}$, donde

$$A = \begin{pmatrix} x & 0 & 1 \\ 0 & x & 0 \\ 1 & 0 & x \end{pmatrix}$$

La ecuación $\text{Det}(A) = 0$, no tiene solución.

4. Sean $A, B \in \mathbb{R}^{n \times n}$, se verifica la relación $(A + B)^2 = A^2 + 2AB + B^2$.
5. Sea $A \in \mathbb{R}^{n \times n}$ una matriz invertible se verifica que $\text{Det}(A^{-1}) = -\text{Det}(A) = 1$.
6. Sea $A\vec{x} = \vec{b}$ un sistema de m ecuaciones con n incógnitas, si se verifica rango de A es n , entonces el sistema es compatible determinado.
7. Sea $A \in \mathbb{R}^{n \times n}$ de forma tal que $A^2 = A$, se verifica que $\text{Det}(A) = 0$ o $\text{Det}(A) = 1$.



Glosario

$C([a, b])$ espacio vectorial de las funciones continuas en $[a, b]$

$ColA$ espacio vectorial generado por los vectores columna de la matriz A .

$Det(A)$ o $|A|$ determinante de la matriz A .

$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$ distancia entre los vectores \vec{x} e \vec{y} .

$dim(V)$ dimensión del espacio vectorial V .

\mathbb{E} espacio euclídeo

E_k proyecciones tales que $Im(E_i) = N(T - \lambda_i I)$.

E_{λ_i} espacio propio correspondiente a λ_i .

f^* transformación adjunta de una transformación lineal f .

$FilA$ espacio vectorial generado por los vectores fila de la matriz A .

$A(\vec{x}, \vec{y})$ forma bilineal

$Im(T)$ imagen de la transformación lineal T .

$\langle \vec{z} \rangle$ subespacio generado por el vector \vec{z}

$L(S)$ subespacio vectorial de V generado por S .

$L(V)$ espacio vectorial de transformaciones lineales de V en V (endomorfismos).

$L(V, W)$ espacio vectorial de transformaciones lineales de V en W .

m_T polinomio minimal.

$N(T)$ núcleo de la transformación lineal T .

$Nul(A)$ espacio nulo de la matriz A .

$P_K[x]$ polinomios en x , con coeficientes en K .

$P_K^{(n)}[x]$ polinomios en x , con coeficientes en K de grado $\leq n$.

$\mathbf{P}_T(\lambda)$ polinomio característico de T .

$P_{B,B'}$ matriz del cambio de base de B' a B .

$P_S(\vec{v})$ proyección ortogonal de \vec{v} sobre el subespacio S .

$\phi(\vec{x}, \vec{y}) = (\vec{x}, \vec{y})$ producto interno

$r(T)$ rango de la matriz T .

$S = \langle \vec{v} \rangle$ subespacio generado por el vector \vec{v} .

S^\perp complemento ortogonal de S en un espacio vectorial (V con producto interno).

$Tr(A)$ traza de la matriz A

V^* espacio dual de V .



Índice alfabético

A

Ajuste por cuadrados mínimos.....	185
Angulo entre dos vectores.....	165
Anulador de un subespacio.....	90
Aplicación transpuesta.....	93
Aproximación de funciones.....	272
Autovalor (o valor propio).....	112
Autovector (o vector propio).....	112

B

Base canónica.....	26
Base de un espacio vectorial.....	26
Bunyakovsky, Viktor Yakovlevich.....	168

C

Caffarelli, Luis Ángel.....	271
Cambio de coordenadas celestes.....	47
Cauchy, Augustin Louis.....	166

Cayley, Arthur.....	137
Circuito eléctrico.....	152
Coefficientes de Fourier.....	274
Combinación lineal.....	20
Complemento ortogonal.....	177
Composición de transformaciones lineales.....	68
Contracción de índices.....	240
Coordenadas.....	28
Coordenadas cilíndricas.....	254
Coordenadas esféricas.....	263
Cuadrados mínimos.....	185

D

Daubechies, Ingrid.....	279
Delta de Kronecker.....	227
Descomposición espectral de una matriz.....	199
Desigualdad de Cauchy-Schwartz.....	166
Diagonalización.....	123, 195
Dickestein, Alicia.....	95
Dimensión de un espacio vectorial.....	32

Distancia entre vectores 168

Doble dual 92

E

Ecuaciones diferenciales 267

Ecuaciones normales 185

Ejes principales \mathbb{R}^2 216

Endomorfismos 190

Epimorfismo 69

Espacio columna de una matriz 178

Espacio dual 84

Espacio fila de una matriz 25

Espacio nulo de una matriz 21

Espacio propio 128

Espacio Vectorial 16

Esperanza 97

F

Forma bilineal 210

Forma cuadrática 213

Forma de Jordan 144

Fourier, Jean-Baptiste 30

G

Gauss, Karl Friedrich 189

González, Gabriela 77

Goobar, Ariel 45

H

Haar, Alfréd 277

Hamilton, William R. 22

Hiperplano 32

I

Imagen de una transformación lineal 57

Independencia lineal 24

Intersección de subespacios 34

Isomorfismo 69

J

Jordan, Marie Ennemond Camille 149

L

Levi-Civita, Tullio 242

Localización de autovalores 118

Longitud de un vector 164

M

Maldacena, Juan Martín 262

Matrices equivalentes 83

Matrices semejantes 83

Matriz de cambio de base 39

Matriz de la forma cuadrática 214

Matriz de un producto interno 163

Matriz de una forma bilineal 211

Matriz de una transformación lineal 61

Matriz diagonalizable 114

Matriz ortogonal 194

Matriz unitaria 194

Monomorfismo 69

Método de Gram-Schmidt 174

N

Nilpotente 142

Nulidad de una matriz 32

Núcleo de una transformación lineal 70

O

Orbita de un cometa 201

P

Polinomio minimal 131

Producto escalar 231

Producto interno 161

Producto mixto 233

Producto tensorial 233

Producto vectorial 232

Proyección ortogonal 55, 181

R

Radiotelescopios 222

Rango de una matriz 73

Reflexión 78, 79

Rotación 63

S

Schwarz, Karl Herman 168

Secciones cónicas 214

Sistema de generadores 24

Strang, William Gilbert 117

Subespacio vectorial 18

Subespacios invariantes 110

Suma de subespacios 34

Suma directa 37

Superficies cuádricas 217

Símbolo de Levi-Civita 229

T

Tensor esférico 261

Tensor Métrico 247

Teorema de Cayley-Hamilton 134

Teorema de Gershgorin 118

Teorema de la proyección ortogonal 184

Transformación ortogonal 197

Transformación lineal 55

Transformación lineal autoadjunta 193

Transformación lineal inyectiva 69

Transformación lineal suryectiva 69

Transformación ortogonal 196

Traza de una matriz 56



La autora

Victoria Vampa, autora

Nació en La Plata, Argentina, en 1960. Obtuvo la Licenciatura en Matemática Aplicada en la Facultad de Ciencias Exactas, UNLP, la Maestría en Simulación Numérica y Control en la Facultad de Ingeniería de la Universidad de Buenos Aires y luego su Doctorado en Matemática en la Facultad de Ciencias Exactas de la UNLP. Su área de investigación es la Matemática Aplicada. Trabajó en la resolución numérica de ecuaciones diferenciales mediante el método de elementos finitos; en particular, con las ecuaciones de elasticidad, en problemas que describen el comportamiento membranar de estructuras laminares. Luego estudió el análisis wavelet y desarrolló aportes en la solución de esos problemas de ecuaciones diferenciales. Trabajó, además, en el uso del análisis multirresolución wavelet diseñando un método eficiente para resolver problemas de contorno. Ha participado de una gran cantidad de congresos y reuniones científicas, y ha publicado numerosos trabajos. Los trabajos considerados de mayor relevancia son:

- *Mejoras en el comportamiento membranar del elemento de lámina MITC4*. (ISBN 978-3-8484-6763-1), Editorial Académica Española, 2012.
<https://www.eae-publishing.com/catalog/details/store/es/book/978-3-8484-6763-1/mejoras-en-el-comportamiento-membranal-del-elemento-de-lamina-mitc4>

- *A new refinement Wavelet-Galerkin method in a spline local multiresolution analysis scheme for boundary value problems*. Autores: V. Vampa, M. T. Martín y E. Serrano, Journal on Wavelets, Multiresolution and Information Processing, IJWMIP-120605, Vol. 11, 2, 1350015-1-19, 2013.
<https://doi.org/10.1142/S021969131350015X>
- *Entropy-Based Informational Study of the COVID-19 series of data*. Autores: A. M. Kowalski, M. Portesi, V. Vampa, M. Losada and F. Holik. Mathematics, 10, 4590, 2022.
<https://doi.org/10.3390/math10234590>
- *Spatio-Temporal Analysis of Acute Myocardial Ischaemia Based on Entropy-Complexity Plane*. Autores E. R. Valverde, V. Vampa, O. A. Rosso, P. D. Arini. Entropy, 27(1), 8, 2025.
<https://doi.org/10.3390/e27010008>

Actualmente es Profesora Adjunta Dedicación Exclusiva en la Facultad de Ingeniería, UNLP, en la cátedra Matemática C. Coordinó la Unidad de Investigación, Desarrollo, Extensión y Transferencia, UIDET Matemática Aplicada del Departamento de Ciencias Básicas entre 2021 y 2023. Actualmente es Miembro del Centro de Matemática de La Plata (CMALP), Facultad de Ciencias Exactas, UNLP. Es directora de proyectos acreditados por la UNLP sobre la aplicación de la Transformada Wavelet en el estudio de sistemas dinámicos y es profesora del curso de posgrado Análisis Wavelet para el Procesamiento de Señales. En la Facultad de Ciencias Astronómicas y Geofísicas, UNLP, es Profesora Adjunta Dedicación Simple, en la cátedra Álgebra Lineal, cargo en el que se desempeña desde el año 2013. Otras aficiones de la autora, además de la docencia y la matemática, son el ciclismo, la música, la repostería y la fotografía.

Lucía Rizzo Buschiazzo, colaboradora

Lucía Rizzo Buschiazzo nació en Rosario en 1987. Graduada de Licenciada en Astronomía por la Facultad de Ciencias Astronómicas y Geofísicas de La Plata en el año 2013 y de Doctora en Astronomía por la Facultad de Ciencias Astronómicas y Geofísicas de La Plata en el año 2023. Recibió el premio Joaquín González del año 2013 en mérito a los mejores promedios de la UNLP, entregado por la Municipalidad de esa ciudad. Al año siguiente comenzó funciones de Jefa de trabajos Prácticos de Álgebra Lineal, cargo en el que se encuentra activa hasta el presente. Becaria del CONICET entre 2013-2018. En su tema de Doctorado investigó la Estructura de la Vía Láctea. Docente en las cátedras de Sistemas Estelares en la Carrera de Astronomía y de Álgebra y Análisis Matemático I y II en la Facultad de Ciencias Exactas e Informática. Ha brindado charlas informativas sobre astronomía en diversas instituciones y en diferentes niveles curriculares. Es miembro de la Asociación Argentina de Astronomía desde 2013. Ha participado en numerosos congresos y en algunas publicaciones científicas de jerarquía internacional.

- *A new look at SCO OB1 association with Gaia DR2..* (Revista: MNRAS. Año: 2020), Autores: L. Yalyalieva, G. Carraro, R. Vazquez, L. Rizzo and E. Glushkova.
- *A multiwavelength study of the H II region G347.600 + 00.211 and its effects on the surrounding medium.* (Revista: MNRAS Año:2022), Cárdenas. S.B, Cichowolski. S, Suad. L.A, Molina Lera J.A, Gamen. R, Rizzo. L..



Referencias

- [1] <https://www.lavanguardia.com/cultura/20210520/7467794/bagdad-matematico-persa>.
- [2] Alfred Haar. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Alfred_Haar. 2023.
- [3] Alicia Dickenstein. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Alicia_Dickenstein. 2023.
- [4] Ariel Goobar. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://sv.wikipedia.org/wiki/Ariel_Goobar y Extractos de la nota realizada por Soledad Vallejos para el Pagina 12, publicada el el 10 de abril de 2016. 2023.
- [5] Arthur Cayley. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Arthur_Cayley. 2023.
- [6] Augustin Louis Cauchy. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Augustin_Louis_Cauchy. 2023.
- [7] Camille Jordan. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Camille_Jordan. 2023.
- [8] Carl Friedrich Gauss. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Carl_Friedrich_Gauss. 2023.

- [9] Gabriela Gonzalez. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Gabriela_Gonzalez. 2023.
- [10] Gilbert Strang. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Gilbert_Strang. 2023.
- [11] Ingrid Daubechies. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Ingrid_Daubechies. 2023.
- [12] Joseph Fourier. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Joseph_Fourier. 2023.
- [13] Juan Martin Maldacena. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Juan_Martin_Maldacena. 2023.
- [14] Karl Herman Amandus Schwartz. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Hermann_Schwarz. 2023.
- [15] Luis Caffarelli. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Luis_Caffarelli. 2023.
- [16] Tullio Levi Civita. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Tullio_Levi-Civita. 2023.
- [17] Viktor Buniakovski. (s.f.). En Wikipedia. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/Viktor_Buniakovskiy. 2023.
- [18] Eduardo W. V. Chaves. *Mecánica del Medio Continuo. Conceptos básicos-3ª Edición*. CIMNE, 2012.
- [19] Vampa V. et al. Ajustes del modelo sir a los datos iniciales de la pandemia del covid-19 en argentina. *Revista Ciencia y Tecnología. Facultad de Ingeniería, Universidad de Palermo*, pages 21–36, 2022.
- [20] S.I. Grossman. *Álgebra Lineal*. McGraw-Hill, 2008.
- [21] Eugenio Hernández. *Álgebra y Geometría*. Addison Wesley-Universidad Autónoma de Madrid, 1994.
- [22] Mikhail Itskov. *Tensor Algebra and Tensor Analysis for Engineers: With Applications to Continuum Mechanics*. Mathematical Engineering. Springer Berlin Heidelberg, 2012.

-
- [23] R. Kunze K. Hoffman. *Introduction to Linear Algebra*. Wellesley-Cambridge Press, 1961.
- [24] Ana Maria Platzeck. *Apuntes de Analisis Tensorial*. FGAGLP. UNLP.
- [25] Luis A. Santaló. *VECTORES Y TENSORES*. EUDEBA. Buenos Aires., 1973.
- [26] Barry Spain. *CÁLCULO TENSORIAL*. Editorial DOSSAT, Madrid-Buenos Aires, 1952.
- [27] Gilbert Strang. *Introduction to Linear Algebra*. Wellesley-Cambridge Press, 2016.
- [28] Wikipedia contributors. *William Rowan Hamilton*. (s.f.). En *Wikipedia*. Recuperado el 15 de abril de 2023 de https://es.wikipedia.org/wiki/William_Rowan_Hamilton. 2023.



Bibliografía recomendada

- [1] Lay, D.C. (2003). *Linear Algebra and Its Applications*. Pearson Education. ISBN: 9788177583335. URL: <https://books.google.com.ar/books?id=v8Zls26Y0wkC>.
- [2] Lay, D.C. & Murrieta, J.E.M. (2007). *Algebra Lineal Y Sus Aplicaciones*. Pearson Educación. ISBN: 9789702609063. Series: Área: Matemáticas. URL: <https://books.google.com.ar/books?id=lTIVrKT9CMIC>.
- [3] Lang, S. (1990). *Introducción al álgebra lineal*. Addison-Wesley Iberoamericana. ISBN: 9780201629125. URL: <https://books.google.com.ar/books?id=NBqnAAAACAAJ>.
- [4] Anton, H. (2003). *Introducción al álgebra lineal*. Ed. Limusa.
- [5] Grossman, S.I. (2008). *Álgebra Lineal*. McGraw-Hill.
- [6] Díaz, J. (2016). Ejemplos básicos de álgebra lineal con Python. *Innovación Educativa*, 16(73), 109-124. URL: <https://www.redalyc.org/journal/6079/607963609005/html/>.



Respuestas a las autoevaluaciones

Apéndice: 1. V, 2. F, 3. F, 4. F, 5. F, 6. F, 7. V

1.6.2: 1. V, 2. F, 3. F, 4. V, 5. V, 6. F, 7. V, 8. V, 9. F

2.8.2: 1. F, 2. V, 3. F, 4. V, 5. F, 6. V, 7. V, 8. V, 9. V, 10. V, 11. V

3.7.2: 1. V, 2. V, 3. V, 4. F, 5. F, 6. V, 7. V, 8. F, 9. V, 10. F, 11. F,
12. V, 13. V, 14. F, 15. F, 16. V

4.7.2: 1. V, 2. V, 3. V, 4. V, 5. V, 6. V, 7. V, 8. V, 9. V, 10. V, 11. V,
12. V, 13. F, 14. V, 15. F, 16. V, 17. V, 18. V

5.2.2: 1. V, 2. F, 3. F, 4. F, 5. V

6.11.2: 1. F, 2. V, 3. V, 4. V, 5. F, 6. V, 7. V, 8. V

First release, 8 de septiembre de 2025

Libros de Cátedra

Vampa, Victoria

Álgebra lineal aplicada / Victoria Vampa; Contribuciones de Lucía Rizzo Buschiazzi; Ilustrado por Claudia Contreras. - 1a ed. - La Plata: Universidad Nacional de La Plata; La Plata: EDULP, 2025.
Libro digital, PDF - (Libros de cátedra)

Archivo Digital: descarga
ISBN 978-950-34-2611-1

1. Álgebra Lineal. I. Rizzo Buschiazzi, Lucía, colab. II. Contreras, Claudia, illus. III. Título.
CDD 515.14

Diseño de tapa: Dirección de Comunicación Visual de la UNLP

Universidad Nacional de La Plata – Editorial de la Universidad de La Plata
48 N.º 551-599 / La Plata B1900AMX / Buenos Aires, Argentina
+54 221 644 7150
edulp.editorial@gmail.com
www.editorial.unlp.edu.ar

Edulp integra la Red de Editoriales Universitarias Nacionales (REUN)

Primera edición, 2025
ISBN 978-950-34-2611-1
© 2025 - Edulp

FACULTAD DE
CIENCIAS ASTRONÓMICAS Y GEOFÍSICAS

e
exactas


EDITORIAL DE LA UNLP



UNIVERSIDAD
NACIONAL
DE LA PLATA