Álgebra lineal

Leandro Vendramin
Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
http://mate.dm.uba.ar/~lvendram/
lvendramin@dm.uba.ar

Índice general

Capítulo 1

Sistemas lineales y matrices

Dos clases. Temas: Repaso de sistemas lineales y matrices.

1.1. Cuerpos

1.1.1. Un **cuerpo** $\mathbb{K}$ es un conjunto con dos operaciones binarias en $\mathbb{K}$, una llamada **suma** y denotada por $(x, y) \mapsto x + y$, y otra llamada **producto** y denotada por $(x, y) \mapsto xy$, que satisfacen las siguientes propiedades:

1) $(x + y) + z = x + (y + z)$ para todo $x, y, z \in \mathbb{K}$,
2) existe un único $0 \in \mathbb{K}$ tal que $x + 0 = 0 + x = x$ para todo $x \in \mathbb{K}$,
3) para cada $x \in \mathbb{K}$ existe un único $-x \in \mathbb{K}$ tal que
$$x + (-x) = (-x) + x = 0.$$
4) $x + y = y + x$ para todo $x, y \in \mathbb{K}$,
5) $(xy)z = x(yz)$ para todo $x, y, z \in \mathbb{K}$,
6) existe un único $1 \in \mathbb{K} \setminus \{0\}$ tal que $x1 = 1x = x$ para todo $x \in \mathbb{K}$,
7) para cada $x \in \mathbb{K} \setminus \{0\}$ existe un único $x^{-1} \in \mathbb{K}$ tal que
$$xx^{-1} = x^{-1}x = 1,$$
8) $xy = yx$ para todo $x, y \in \mathbb{K}$,
9) $x(y + z) = xy + xz$ para todo $x, y, z \in \mathbb{K}$.

1.1.2. Ejemplos.

1) $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son cuerpos.
2) $\mathbb{Z}_p$ es cuerpo si y sólo si p es un número primo.
3) $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$ es un cuerpo. Determine el inverso de un elemento de la forma $a + b\sqrt{2}$, donde $a, b \in \mathbb{Q}$.

1.2. Sistemas lineales y matrices

1.2.1. Estudiaremos **sistemas de ecuaciones lineales** con m ecuaciones y n incógnitas. Dados a_{ij}, donde $i \in \{1, \dots, m\}$ y $j \in \{1, \dots, n\}$, el problema consiste en encontrar n elementos de $\mathbb{K}$, digamos $x_1, \dots, x_n$, que satisfagan las condiciones

eq:sistema

$$\left\{\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned}\right. \tag{1.2.2}$$

Toda tupla $(x_1, \dots, x_n)$ de elementos de $\mathbb{K}$ que satisface (1.2.2) se denomina **solución** del sistema. Si $b_1 = \cdots = b_m = 0$ el sistema (1.2.2) se denomina sistema **homogéneo**.

1.2.3. Un sistema lineal es **compatible determinado** cuando admite una única solución, **compatible indeterminado** cuando admite más de una solución e **incompatible** si no tiene soluciones. Observemos que un sistema homogéneo siempre tiene solución: basta tomar la **solución trivial** $x_1 = x_2 = \cdots = x_n = 0$.

1.2.4. Ejemplo. Muchos problemas no lineales pueden resolverse mediante un sistema lineal. Como ejemplo demostraremos que existe una única terna (a, b, c) que satisface las ecuaciones

:sistema_nolineal
$$abc = 1, \quad a^2b^2c^3 = 8, \quad a(bc)^{-1} = 16, \tag{1.2.5}$$

En efecto, si aplicamos $\log_2$, las ecuaciones (1.2.5), se transforman en el sistema lineal

$$\begin{cases} x + y + z = 0, \\ 2x + 2y + 3z = 3, \\ x - y - z = 4, \end{cases}$$

que tiene a $(x, y, z) = (2, -5, 3)$ como única solución. Luego la única solución del sistema (1.2.5) es $(a, b, c) = (4, 2^{-5}, 8)$.

1.2.6. En ocasiones resulta conveniente escribir al sistema (1.2.2) de la siguiente forma:

$$\begin{cases} F_1(x_1, \dots, x_n) = b_1, \\ \quad\vdots \\ F_m(x_1, \dots, x_n) = b_m, \end{cases} \tag{1.2.7}$$

donde $F_1(x_1, \dots, x_n), \dots, F_m(x_1, \dots, x_n)$ son las funciones definidas por

$$F_i(x_1, \dots, x_n) = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$$

para $i \in \{1, \dots, m\}$. Observemos que las F_i satisfacen

$$\begin{aligned} &F_i(x_1 + y_1, \dots, x_n + y_n) = F_i(x_1, \dots, x_n) + F_i(y_1, \dots, y_n), \\ &F_i(\lambda x_1, \dots, \lambda x_n) = \lambda F_i(x_1, \dots, x_n), \end{aligned}$$

para todo $\lambda \in \mathbb{K}$.

1.2.8. Proposición. *Supongamos que $\mathbb{K}$ tiene infinitos elementos y que el sistema lineal (1.2.2) tiene más de una solución. Entonces (1.2.2) tiene infinitas soluciones.*

Demostración. Sean c y d dos soluciones y $\lambda \in \mathbb{K}$. Demostremos que $r = \lambda c + (1 - \lambda)d$ es solución del sistema (1.2.2). En efecto,

$$F_i(r) = \lambda F_i(c) + (1 - \lambda)F_i(d) = \lambda b_i + (1 - \lambda)b_i = b_i$$

para todo $i \in \{1, \dots, m\}$. Para ver que hay infinitas soluciones basta observar que si $\lambda, \mu \in \mathbb{K}$ con $\lambda \neq \mu$ entonces $\lambda c + (1 - \lambda)d \neq \mu c + (1 - \mu)d$. $\square$

1.2.9. Ejemplo. En $\mathbb{R}^3$ el sistema

eq:sistema:2x3
$$\begin{cases} x + y + z = 0, \\ \quad\; y + z = 0, \end{cases} \tag{1.2.10}$$

tiene infinitas soluciones. De hecho, el conjunto de soluciones es la recta

$$\{\lambda(0, 1, -1) : \lambda \in \mathbb{R}\}.$$

En cambio, el conjunto de soluciones en $\mathbb{Z}_3^3$ del sistema (1.2.10) es el conjunto finito $\{(0, 0, 0), (0, 1, -1), (0, -1, 1)\}$.

1.2.11. Diremos que dos sistemas lineales son **equivalentes** si tienen exactamente el mismo conjunto de soluciones.

1.2.12. Los sistemas lineales

$$\begin{cases} x + y + z = 0, \\ \quad x + y = 0, \end{cases} \qquad\qquad \begin{cases} x + y + z = 0, \\ \qquad\quad z = 0, \end{cases}$$

son equivalentes. En efecto, el conjunto de soluciones es $\{(\lambda, -\lambda, 0) : \lambda \in \mathbb{K}\}$.

1.2.13. Dado un sistema lineal como en (1.2.2) vamos a definir dos **operaciones elementales** que no alteran el conjunto de soluciones del sistema:

1) Intercambiar el lugar de dos ecuaciones.
2) Reemplazar una ecuación, digamos $F_i(x_1,\dots,x_n)=b_i$, por la ecuación $(F_i+\lambda F_j)(x_1,\dots,x_n)=b_i+\lambda b_j$, donde $\lambda\in\mathbb{K}$.

Teorema. *Al aplicar estas operaciones elementales se obtiene un sistema equivalente al original.*

Demostración. El conjunto de soluciones de un sistema lineal es la intersección del conjunto de soluciones de cada ecuación lineal

$$F_i(x_1,\dots,x_n)=b_i.$$

Como la intersección de conjuntos es conmutativa, intercambiar dos ecuaciones no altera el conjunto de soluciones del sistema lineal.

Para demostrar que al aplicar la segunda operación elemental el conjunto de soluciones del sistema lineal no se altera, observemos que

$$\begin{cases} F_i(x_1,\dots,x_n)=b_i,\\ F_j(x_1,\dots,x_n)=b_j,\end{cases}$$

si y sólo si

$$\begin{cases} (F_i+\lambda F_j)(x_1,\dots,x_n)=b_i+\lambda b_j,\\ F_j(x_1,\dots,x_n)=b_j,\end{cases}$$

para todo $\lambda\in\mathbb{K}$. $\square$

ock:sistemas:gauss

1.2.14 (Método de Gauss). Supongamos que se tiene un sistema lineal como en (1.2.2) tal que $a_{11}\neq 0$. Para cada $i\in\{2,\dots,m\}$ podemos utilizar la segunda operación elemental y reemplazar la i-ésima ecuación $F_i(x_1,\dots,x_n)=b_i$ por

$$F_i(x_1,\dots,x_n)-\frac{a_{i1}}{a_{11}}F_1(x_1,\dots,x_n)=b_i-\frac{a_{i1}}{a_{11}}b_1.$$

De esta forma obtenemos un sistema lineal

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\ \qquad\quad a'_{22}x_2+\cdots+a'_{2n}x_n=b'_2,\\ \qquad\qquad\qquad\vdots\\ \qquad\quad a'_{m2}x_2+\cdots+a'_{mn}x_n=b'_m,\end{cases}$$

equivalente a (1.2.2).

thm:sistemas:n>m

1.2.15. Teorema. *Sean $n,m\in\mathbb{N}$ tales que $n>m$. Un sistema de m ecuaciones con n incógnitas es incompatible o compatible indeterminado.*

Demostración. Procederemos por inducción en m.

Supongamos primero que $m=1$. En este caso tenemos una ecuación tipo $\sum_{i=1}^n a_ix_i=b$. Si todos los a_i son cero, el sistema es incompatible si $b\neq 0$ y compatible indeterminado si $b=0$. En cambio, si existe algún $a_i\neq 0$, se tienen infinitas soluciones: las incógnitas x_j con $j\neq i$ pueden tomar cualquier valor en $\mathbb{K}$ y $x_i=\frac{1}{a_i}\sum_{j\neq i}a_jx_j$.

Supongamos ahora que $m>1$. Si todos los a_{ij} son cero el resultado es trivialmente válido. Si suponemos que existe al menos un $a_{ij}\neq 0$, después de utilizar la operación elemental de intercambiar filas y (si fuera necesario) de cambiar el nombre de las variables, podemos suponer que $a_{11} \neq 0$. El procedimiento visto en 1.2.14 nos da un sistema de la forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ \qquad\qquad a'_{22}x_2 + \cdots + a'_{2n}x_n = b'_2, \\ \qquad\qquad\qquad\qquad\qquad \vdots \\ \qquad\qquad a'_{m2}x_2 + \cdots + a'_{mn}x_n = b'_m, \end{cases}$$

que es equivalente al sistema original. Como $n-1 > m-1$, la hipótesis inductiva aplicada al sistema

$$\begin{cases} a'_{22}x_2 + \cdots + a'_{2n}x_n = b'_2, \\ \qquad\qquad\qquad \vdots \\ a'_{m2}x_2 + \cdots + a'_{mn}x_n = b'_m, \end{cases}$$

de $m-1$ ecuaciones y $n-1$ incógnitas nos dice que ese sistema es compatible indeterminado o incompatible. Esto implica que el sistema original es compatible indeterminado o incompatible. □

cor:homogeneo 1.2.16. COROLARIO. *En un sistema homogéneo que tiene más incógnitas que ecuaciones existe al menos una solución no trivial.*

DEMOSTRACIÓN. Es consecuencia inmediata del teorema 1.2.15 ya que los sistemas homogéneos son siempre compatibles. □

1.2.17. Sean $\mathbb{K}$ un cuerpo y $n, m \in \mathbb{N}$. Una **matriz** $A = (a_{ij})$ sobre $\mathbb{K}$ de tamaño $m \times n$ es un arreglo de elementos de $\mathbb{K}$ de la forma

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

Los a_{ij} son los **elementos** de la matriz A. El conjunto de matrices de tamaño $m \times n$ con elementos en $\mathbb{K}$ se denotará por $\mathbb{K}^{m\times n}$.

1.2.18 (Suma de matrices). Sean las matrices $A = (a_{ij}) \in \mathbb{K}^{m\times n}$ y $B = (b_{ij}) \in \mathbb{K}^{m\times n}$. Se define la **suma** de A y B como la matriz de $\mathbb{K}^{m\times n}$ dada por

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

para todo i, j. Queda como ejercicio demostrar que el neutro para la suma es la matriz $0 \in \mathbb{K}^{m\times n}$ dada por $0_{ij} = 0$ para todo i, j, que el inverso es $(-A)_{ij} = -a_{ij}$ para todo i, j, y que la suma de matrices es asociativa y conmutativa.

1.2.19 (Producto por un escalar). Sean $A = (a_{ij}) \in \mathbb{K}^{m\times n}$ y $\lambda \in \mathbb{K}$. Se define λA como la matriz de $m \times n$ cuyos elementos son

$$(\lambda A)_{ij} = \lambda a_{ij}$$

para todo i, j. Queda como ejercicio demostrar que las siguientes propiedades: $(\lambda\mu)A = \lambda(\mu A)$, $(\lambda + \mu)A = \lambda A + \mu A$ y $\lambda(A + B) = \lambda A + \lambda B$ para todo $\lambda, \mu \in \mathbb{K}$ y $A, B \in \mathbb{K}^{m\times n}$.

1.2.20 (Producto de matrices). Sean las matrices $A = (a_{ij}) \in \mathbb{K}^{m\times n}$ y $B = (b_{ij}) \in \mathbb{K}^{n\times p}$. Se define el **producto** AB como la matriz de $m \times p$ cuyos elementos son

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

para todo i, j.

1.2.21. Ejemplo. El producto de matrices en general no es conmutativo:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

1.2.22. Ejemplo. Sea $A = (A_{ij}) \in \mathbb{K}^{m\times n}$. Para $i \in \{1, \dots, n\}$ consideremos el vector columna $e_j \in \mathbb{K}^{n\times 1}$ dado por

$$(e_i)_j = \delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

Entonces $Ae_i \in \mathbb{K}^{m\times 1}$ y Ae_i es la i-ésima columna de la matriz A.

1.2.23. Se define la **matriz identidad** I de tamaño n como la matriz cuadrada de $n \times n$ dada por $I_{ij} = \delta_{ij}$ para todo i, j. Si $A \in \mathbb{K}^{m\times n}$, se tienen las siguientes dos propiedades:

1) Si I es la identidad de $m \times m$ entonces $IA = A$.
2) Si I es la identidad de $n \times n$ entonces $AI = A$.

1.2.24. Teorema. *El producto de matrices es asociativo.*

Demostración. Supongamos que $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{n\times p}$ y $C \in \mathbb{K}^{p\times q}$. Como el producto en $\mathbb{K}$ es asociativo,

$$\begin{aligned}(A(BC))_{ij} &= \sum_{k=1}^{n} A_{ik}(BC)_{kj} = \sum_{k=1}^{n} A_{ik}\left(\sum_{l=1}^{p} B_{kl}C_{lj}\right) = \sum_{k=1}^{n}\sum_{l=1}^{p} A_{ik}B_{kl}C_{lj} \\ &= \sum_{l=1}^{p}\left(\sum_{k=1}^{n} A_{ik}B_{kl}\right) C_{lj} = \sum_{l=1}^{p}(AB)_{il}C_{lj} = ((AB)C)_{ij}\,,\end{aligned}$$

tal como se quería demostrar. □

1.2.25. Observemos que el producto de matrices nos permite escribir el sistema lineal (1.2.2) como la ecuación matricial $Ax = b$, donde

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

La matriz A se denomina **matriz del sistema**. La **matriz ampliada** es la matriz de tamaño $m \times (n + 1)$ dada por

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

xca:Ax=b_y_Ax=0 1.2.26. Ejercicio. Sea p una solución del sistema $Ax = b$ y S el conjunto de soluciones del sistema homogéneo $As = 0$. Entonces toda solucion de $Ax = b$ es de la forma $s + p$ para algún $s \in S$.

block:traspuesta 1.2.27. Si $A \in \mathbb{K}^{m\times n}$ se define la **traspuesta** de A como la matriz $A^T \in \mathbb{K}^{n\times m}$ dada por

$$(A^T)_{ij} = A_{ji}$$

para todo i, j. De la definición es evidente que $(A^T)^T = A$. Vale además que $(AB)^T = B^TA^T$ pues si $A \in \mathbb{K}^{m\times n}$ y $B \in \mathbb{K}^{n\times p}$ entonces

$$(B^TA^T)_{ij} = \sum_{k=1}^{n}(B^T)_{ik}(A^T)_{kj} = \sum_{k=1}^{n} B_{ki}A_{jk} = \sum_{k=1}^{n} A_{jk}B_{ki} = (AB)_{ji}.$$

xca:I^T=I

1.2.28. Ejercicio. Demuestre que si I es la matriz identidad de $n \times n$ entonces $I^T = I$.

1.2.29. Si $A \in \mathbb{K}^{n\times n}$ se define la **traza** de A como

$$\operatorname{tr}(A) = A_{11} + A_{22} + \cdots + A_{nn}.$$

Valen las siguientes propiedades:

1) $\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$ para todo $A, B \in \mathbb{K}^{n\times n}$.
2) $\operatorname{tr}(\lambda A) = \lambda \operatorname{tr}(A)$ para todo $A \in \mathbb{K}^{n\times n}$ y $\lambda \in \mathbb{K}$.
3) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ para todo $A, B \in \mathbb{K}^{n\times n}$.

Demostremos la tercera afirmación:

$$\begin{aligned}\operatorname{tr}(AB) &= \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}B_{ji} \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} B_{ji}A_{ij} = \sum_{j=1}^{n} (BA)_{jj} = \operatorname{tr}(BA).\end{aligned}$$

pro:Av=Bv=>A=B

1.2.30. Proposición. *Sean A y B dos matrices de $n \times n$ tales que $Av = Bv$ para todo $v \in \mathbb{K}^{n\times 1}$. Entonces $A = B$.*

Demostración. Basta tomar $v = e_i$ para todo $i \in \{1, \dots, n\}$, donde los e_i están dados por $(e_i)_j = \delta_{ij}$. □

1.3. Matrices inversibles

1.3.1. Una matriz cuadrada $A \in \mathbb{K}^{n\times n}$ es **inversible** si existe $B \in \mathbb{K}^{n\times n}$ tal que $AB = BA = I$. El conjunto de matrices inversibles de $n \times n$ con elementos en $\mathbb{K}$ se denota por

$$\mathbf{GL}(n, \mathbb{K}) = \{A \in \mathbb{K}^{n\times n} \mid A \text{ es inversible}\}.$$

rem:inversa

1.3.2. Observaciones.

1) La identidad I es inversible pues $II = I$.
2) Si $AB = I$ y $CA = I$ entonces $B = C$ pues
$$B = IB = (CA)B = C(AB) = CI = C.$$
3) La inversa de una matriz, si existe, es única.
4) Si A es inversible entonces A^{-1} es inversible y $(A^{-1})^{-1} = A$.
5) Si A y B son matrices inversibles entonces AB es inversible y $(AB)^{-1} = B^{-1}A^{-1}$.

1.3.3. Observación. Un **grupo** es un conjunto G junto con una operación binaria $G \times G \to G$, que denotamos por $(x, y) \mapsto xy$, tal que:

1) $x(yz) = (xy)z$ para todo $x, y, z \in G$.
2) Existe $e \in G$ tal que $xe = ex = x$ para todo $x \in G$.
3) Para cada $x \in G$ existe $x^{-1} \in G$ tal que $xx^{-1} = x^{-1}x = e$.

Lo hecho en 1.3.2 nos permite demostrar que $\mathbf{GL}(n, \mathbb{K})$ es un **grupo**. Este grupo se denomina **grupo lineal general** de grado n sobre $\mathbb{K}$. En general $\mathbf{GL}(n, \mathbb{K})$ es no conmutativo.

1.3.4. Ejercicio. Encuentre dos matrices inversibles $A, B \in \mathbb{R}^{2\times 2}$ tales que $AB \neq BA$.

1.3.5. Ejemplo. Vamos a calcular la inversa de la matriz $A = \begin{pmatrix} 2 & 7 \\ 1 & 4 \end{pmatrix}$. Sea $B = (b_{ij})$ tal que $AB = I$. Las ecuaciones correspondientes a $AB = I$ son

$$2b_{11} + 7b_{21} = 1, \quad b_{11} + 4b_{21} = 0, \quad 2b_{12} + 7b_{22} = 0, \quad b_{12} + 4b_{22} = 1.$$

Observemos que para resolver este sistema de 4×4 alcanza con aplicar el método de Gauss en la matriz ampliada

$$\left(\begin{array}{cc|cc} 2 & 7 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{array}\right)$$

El método de Gauss nos da un matriz $B = \begin{pmatrix} 4 & -7 \\ -1 & 2 \end{pmatrix}$ que cumple $AB = I$. Esta matriz B también cumple que $BA = I$. ¿Por qué?

1.3.6. Ejercicio. Calcule la inversa de la matriz $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

1.3.7. Si A es inversible entonces A^T es inversible y $(A^T)^{-1} = (A^{-1})^T$. En efecto, si A es inversible entonces $AA^{-1} = A^{-1}A = I$. Entonces, al aplicar la traspuesta y utilizar las propiedades de la traza vistas en 1.2.27,

$$(A^{-1})^T A^T = I = A^T (A^{-1})^T,$$

es decir: A^T es inversible y $(A^T)^{-1} = (A^{-1})^T$.

1.4. Una aplicación a la criptografía

1.4.1. Vamos a utilizar matrices inversibles para codificar mensajes. Supongamos que queremos codificar el mensaje:

la pelota no se mancha

Si a cada letra del alfabeto le asignamos un número tal como vemos en la tabla

a	b	c	d	e	$\cdots$	z
1	2	3	4	5	$\cdots$	26

y al espacio le asignamos el número 27, el mensaje a encriptar se transforma en la siguiente tira de números:

apelotanosemancha

(1.4.2) 12 1 27 16 5 12 15 20 1 27 14 15 27 19 5 27 13 1 14 3 8 1

Identificaremos a nuestro mensaje con la tira (1.4.2). Para encriptar el mensaje utilizaremos una matriz inversible A de $\mathbb{R}^{2\times 2}$ y es por eso que necesitamos escribir a (1.4.2) como una matriz en $\mathbb{R}^{2\times 11}$:

$$B = \begin{pmatrix} 12 & 27 & 5 & 15 & 1 & 14 & 27 & 5 & 13 & 14 & 8 \\ 1 & 16 & 12 & 20 & 27 & 15 & 19 & 27 & 1 & 3 & 1 \end{pmatrix}.$$

Nuestro método de encriptación consiste en multiplicar a izquierda a la matriz B por A. Si, por ejemplo, elegimos la matriz inversible

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

el mensaje encriptado será entonces

$$AB = \begin{pmatrix} -45 & -60 & 16 & 0 & 77 & -11 & -51 & 61 & -49 & -47 & -29 \\ 34 & 49 & -9 & 5 & -51 & 12 & 43 & -39 & 37 & 36 & 22 \end{pmatrix}.$$

Observemos que al multiplicar al mensaje encriptado por A^{-1} a izquierda se obtiene el mensaje original.

Capítulo 2

Espacios vectoriales

Tres clases. Temas: Espacios vectoriales y subespacios. Suma e intersección. Sistemas de generadores e independencia lineal. Bases y dimensión. Rango. Coordenadas y cambio de base. Aplicaciones.

2.1. Espacios vectoriales

2.1.1. Un **espacio vectorial** sobre un cuerpo $\mathbb{K}$ es un conjunto no vacío V con dos operaciones, una suma $V \times V \to V$ denotada por $(x, y) \mapsto x + y$, y un producto por escalares $\mathbb{K} \times V \to V$, denotado por $(\lambda, x) \mapsto \lambda x$, tal que

1) $(x + y) + z = x + (y + z)$ para todo $x, y, z \in V$.
2) Existe un único $0_V \in V$ tal que $x + 0_V = 0_V + x = x$ para todo $x \in V$.
3) Para cada $x \in V$ existe un único $(-x) \in V$ tal que
$$x + (-x) = (-x) + x = 0_V.$$
4) $x + y = y + x$ para todo $x, y \in V$.
5) $1x = x$ para todo $x \in V$.
6) $(\lambda\mu)x = \lambda(\mu x)$ para todo $\lambda, \mu \in \mathbb{K}$ y $x \in V$.
7) $(\lambda + \mu)x = \lambda x + \mu x$ para todo $\lambda, \mu \in \mathbb{K}$ y $x \in V$.
8) $\lambda(x + y) = \lambda x + \lambda y$ para todo $\lambda \in \mathbb{K}$ y $x, y \in V$.

Los elementos de V se denominan **vectores**. Los elementos de $\mathbb{K}$ se denominan **escalares**. El elemento $0_V \in V$ se denomina el **origen** de V. Cuando no haya peligro de confusión, denotaremos por 0 al cero del cuerpo $\mathbb{K}$ y también al origen de V.

2.1.2. Demostremos algunas propiedades básicas:

1) $0v = 0_V$ para todo $v \in V$. En efecto, $0v = (0 + 0)v = 0v + 0v$ y luego, al sumar en ambos miembros el elemento $-(0v)$, obtenemos $0v = 0_V$.
2) $(-1)v = -v$ para todo $v \in V$. En efecto, como $0v = 0_V$, entonces
$$0_V = 0v = (1 + (-1))v = 1v + (-1)v = v + (-1)v$$
y luego $-v = (-1)v$, por la unicidad del inverso aditivo.
3) $\lambda 0_V = 0_V$ para todo $\lambda \in \mathbb{K}$ pues $\lambda 0_V = \lambda(0_V + 0_V) = \lambda 0_V = \lambda 0_V$ y luego $\lambda 0_V = 0_V$.

2.1.3. Ejercicio. Sea I un conjunto no vacío y sea $\{V_i : i \in I\}$ una colección de espacios vectoriales sobre $\mathbb{K}$. Se define el **producto directo** de los V_i como
$$\prod_{i\in I} V_i = \left\{ f \colon I \to \bigcup_{i\in I} V_i : f(i) \in V_i \text{ para todo } i \in I \right\}.$$
Demuestre que con las operaciones
$$(f + g)(i) = f(i) + g(i),$$
$$(\lambda f)(i) = \lambda f(i),$$
donde la suma es la suma del espacio vectorial V_i y el producto por escalares es el producto por escalares de V_i, es un espacio vectorial sobre $\mathbb{K}$. Un poco de notación: si $I = \{1, \dots, n\}$ es un conjunto finito escribiremos $\prod_{i\in I} V_i = V_1 \times \cdots \times V_n$. Si todos los V_i

son iguales a un espacio vectorial V entonces $\prod_{i\in I} V_i = V^I$. Por último, si $I = \{1, \dots, n\}$ y todos los V_i son iguales a V entonces $\prod_{i\in I} V_i = V^n$.

2.1.4. Ejemplos.

1) Un cuerpo $\mathbb{K}$ con las operaciones usuales es un espacio vectorial sobre $\mathbb{K}$ donde vectores y escalares coinciden.
2) El conjunto $\mathbb{K}^n$ de n-tuplas de elementos de $\mathbb{K}$ con las operaciones
$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n),$$
$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n),$$
es un espacio vectorial sobre $\mathbb{K}$.
3) El conjunto $\mathbb{K}^{m\times n}$ de matrices de $m\times n$ con las operaciones usuales es un espacio vectorial sobre $\mathbb{K}$.
4) El conjunto $\mathbb{K}[X]$ de polinomios con coeficientes en $\mathbb{K}$ es un espacio vectorial sobre $\mathbb{K}$.
5) $V = C(\mathbb{R}) = \{f\colon \mathbb{R} \to \mathbb{R} : f \text{ es continua}\}$
6) $V = C^\infty(\mathbb{R}) = \{f\colon \mathbb{R} \to \mathbb{R} : f \text{ es de clase } C^\infty\}$.
7) Las sucesiones en $\mathbb{K}$ con las operaciones
$$(x_1, x_2, \dots) + (y_1, y_2, \dots) = (x_1 + y_1, x_2 + y_2, \dots),$$
$$\lambda(x_1, x_2, \dots) = (\lambda x_1, \lambda x_2, \dots),$$
forman un espacio vectorial sobre $\mathbb{K}$. Este espacio vectorial será denotado por $\mathbb{K}^\infty$.

2.1.5. Ejercicio. Si V un espacio vectorial sobre $\mathbb{K}$ y $p \in V$. Pruebe que las operaciones $v +_p w = v + w - p$ y $\lambda \cdot_p v = p + \lambda(x - p)$ definen una estructura de espacio vectorial sobre V cuyo origen es p.

2.1.6. Ejercicio. Sea V un espacio vectorial sobre $\mathbb{R}$. Pruebe que las operaciones
$$(x, y) + (x', y') = (x + x', y + y'), \qquad (a + bi)(x, y) = (ax - by, ay + bx),$$
definen sobre $V \times V$ una estructura de espacio vectorial sobre $\mathbb{C}$. Este espacio vectorial se denomina la **complexificación** de V y se denota por $V_{\mathbb{C}}$.

2.2. Subespacios, suma e intersección

2.2.1. Sea V un espacio vectorial sobre $\mathbb{K}$. Un subconjunto no vacío $S \subseteq V$ es un **subespacio** de V si $0 \in S$ y si $x, y \in S$ y $\lambda \in \mathbb{K}$ entonces $x + \lambda y \in S$. Observemos que si S es un subespacio de V entonces S, con las operaciones de V restringidas a S, es un espacio vectorial.

xca:subespacio

2.2.2. Ejercicio. Sea V un espacio vectorial sobre $\mathbb{K}$ y sea $S \subseteq V$ un subconjunto no vacío. Pruebe que S es subespacio si y sólo si $v + \lambda w \in S$ para todo $v, w \in S$ y $\lambda \in \mathbb{K}$.

2.2.3. Ejercicio. Sea I un conjunto no vacío y $\{V_i : i \in I\}$ una colección de espacios vectoriales sobre $\mathbb{K}$. Demuestre que el conjunto
$$\coprod_{i\in I} V_i = \left\{ f \in \prod_{i\in I} V_i : f(i) = 0 \text{ salvo para finitos } i \in I \right\}$$
es un subespacio de $\prod_{i\in I} V_i$. Este subespacio se conoce como el **coproducto directo** de los V_i. Notación: si todos los V_i son iguales a un espacio vectorial V escribimos $\coprod_{i\in I} V_i = V^{(I)}$.

2.2.4. Ejemplos.

1) Si V es un espacio vectorial entonces $\{0\}$ y V son subespacios de V.

2) Si $A = (a_{ij}) \in \mathbb{K}^{m\times n}$ entonces

$$\left\{ (x_1, \dots, x_n) \in \mathbb{K}^n : \sum_{j=1}^{n} a_{ij}x_j = 0 \text{ para todo } i \in \{1, \dots, m\} \right\}$$

es un subespacio de $\mathbb{K}^n$.

3) Para cada $n \in \mathbb{N}$, el conjunto

$$\mathbb{K}_n[X] = \{f \in \mathbb{K}[X] : \deg f \leq n \text{ o } f = 0\}$$

es un subespacio de $\mathbb{K}[X]$.

4) El conjunto $C^\infty(\mathbb{R})$ es un subespacio de $C(\mathbb{R})$.
5) El conjunto $\{f \in C^\infty(\mathbb{R}) : f'' + f = 0\}$ es un subespacio de $C^\infty(\mathbb{R})$.
6) El conjunto de sucesiones en $\mathbb{K}$ con un número finito de elementos no nulos es un subespacio de $\mathbb{K}^\infty$.

2.2.5. Ejercicio. Si $n \in \mathbb{N}$ entonces $\{f \in \mathbb{R}[X] : \deg f \geq n \text{ o } f = 0\}$ no es un subespacio de $\mathbb{R}[X]$. ¿Por qué?

2.2.6. Ejercicio. Sea I un conjunto de índices y sea $(S_i)_{i\in I}$ una colección arbitraria de subespacios de un espacio vectorial V. Demuestre que $\cap_{i\in I} S_i$ es un subespacio de V.

2.2.7. Si S y T son dos subespacios de un espacio vectorial V entonces $S \cap T$ es un subespacio de V. Observemos que $S \cap T$ es el mayor subespacio de V contenido en S y en T.

2.2.8. Ejercicio. Sea I un conjunto no vacío y sea $\{W_i : i \in I\}$ una colección de subespacios de un espacio vectorial V. Demuestre que

$$\sum_{i\in I} W_i = \left\{ v \in V : v = \sum_j w_j \text{ (suma finita), donde } w_j \in W_j \text{ para todo } j \right\}$$

es un subespacio de V y que es el menor subespacio de V que contiene a $\cup_{i\in I} W_i$.

2.2.9. Sean V un espacio vectorial y S y T dos subespacios de V. Entonces

$$S + T = \{s + t \in V : s \in S,\ t \in T\}$$

es un subespacio de V. Demostremos que $S+T$ es el menor subespacio de V que contiene a $S \cup T$. En efecto, si W es un subespacio de V que contiene a $S \cup T$ entonces, como $S \subseteq S \cup T \subseteq W$ y $T \subseteq S \cup T \subseteq W$, se tiene que $S + T \subseteq W$.

2.2.10. Observación. La unión de subespacios no siempre es subespacio. Por ejemplo: si $V = \mathbb{R}^2$, $S = \{(x, y) : x = 0\}$ y $T = \{(x, y) : y = 0\}$ entonces $(1, 1) \notin S \cup T$.

exa:R2x2 2.2.11. Ejemplo. Consideremos el subespacio U de $\mathbb{R}^{2\times 2}$ dado por

$$U = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} : a_{11}, a_{12}, a_{22} \in \mathbb{K} \right\} = \left\{ \begin{pmatrix} \star & \star \\ 0 & \star \end{pmatrix} \right\},$$

y los siguientes subespacios

$$U' = \left\{ \begin{pmatrix} 0 & \star \\ 0 & 0 \end{pmatrix} \right\}, \qquad L = \left\{ \begin{pmatrix} \star & 0 \\ \star & \star \end{pmatrix} \right\}, \qquad L' = \left\{ \begin{pmatrix} 0 & 0 \\ \star & 0 \end{pmatrix} \right\}, \qquad H = \left\{ \begin{pmatrix} \star & 0 \\ 0 & \star \end{pmatrix} \right\},$$

Entonces, por ejemplo, se demuestra fácilmente que:

1) $V = U + L$.
2) $V = U + L'$.

2.2.12. Sean V un espacio vectorial y S y T dos subespacios. Se dice que V es **suma directa** de S y T si $V = S+T$ y $S \cap T = \{0\}$. Si V es suma directa de S y T la notación es: $V = S \oplus T$.

2.2.13. Ejemplo. En el ejemplo 2.2.11 vimos que $V = U + L$. Es fácil ver que V no es suma directa de U y L pues $U \cap L \neq \{0\}$. En cambio, $V = U \oplus L'$ pues $V = U + L'$ y además $U \cap L' = \{0\}$.

2.2.14. Ejemplo. Sea V el espacio vectorial de las funciones $f \colon \mathbb{R} \to \mathbb{R}$. Consideremos los subespacios $V_P = \{f \in V : f(x) = f(-x) \text{ para todo } x\}$ y $V_I = \{f \in V : f(x) = -f(-x) \text{ para todo } x\}$. Como toda f puede escribirse como $f = f_P + f_I$, donde

$$f_P(x) = \frac{f(x)+f(-x)}{2} \in V_P, \quad f_I(x) = \frac{f(x)-f(-x)}{2} \in V_I,$$

entonces $V = V_P + V_I$. Más aún, si $f \in V_P \cap V_I$ entonces $f(x) = 0$ para todo x. Luego $V_P + F_I = \{0\}$ y entonces $V = V_P \oplus V_I$.

2.2.15. Proposición. *Son equivalentes:*

1) $V = S \oplus T$.
2) $V = S + T$ *y si* $s + t = 0$ *con* $s \in S$ *y* $t \in T$ *entonces* $s = 0$ *y* $t = 0$.
3) *Para cada* $v \in V$ *existen únicos* $s \in S$ *y* $t \in T$ *tales que* $v = s + t$.

Demostración. Demostremos que (1) implica (2). Como $V = S \oplus T$, entonces, por definición, $V = S + T$. Si $s + t = 0$ con $s \in S$ y $t \in T$ entonces $s = -t \in S \cap T = \{0\}$ y por lo tanto $s = 0$ y $t = 0$. Demostremos ahora que (2) implica (3). Como la existencia es trivial, demostraremos la unicidad: si $s + t = s' + t'$ con $s, s' \in S$ y $t, t' \in T$ entonces $(s - s') + (t - t') = 0$ con $s - s' \in S$ y $t - t' \in T$. Luego $s - s' = 0$ y $t - t' = 0$ como se quería demostrar. Por último, demostremos (3) $\Rightarrow$ (1). Por hipótesis, $V = S + T$. Si $v \in S \cap T$ entonces $v \in S$. Además $-v \in S \cap T$ y entonces $-v \in T$. Como $0 = v + (-v)$ con $v \in S$ y $-v \in T$, entonces $v = 0$. $\square$

2.2.16. Si V es un espacio vectorial y $S_1, \dots, S_n$ son subespacios de V se dice que V es **suma directa** de los S_i si $V = S_1 + \cdots + S_n$ y la igualdad $s_1 + \cdots + s_n = 0$, donde $s_i \in S_i$ para cada $i \in \{1, \dots, n\}$, implica que $s_i = 0$ para todo $i \in \{1, \dots, n\}$.

2.2.17. Ejercicio. Sea V un espacio vectorial y sean $S_1, \dots, S_n$ subespacios de V. Demuestre que las siguientes afirmaciones son equivalentes:

1) $S_1 + \cdots + S_n = S_1 \oplus \cdots \oplus S_n$.
2) Todo $v \in S_1 + \cdots + S_n$ puede escribirse en forma única como $v = s_1 + \cdots + s_n$ con $s_i \in S_i$ para todo $i \in \{1, \dots, n\}$.
3) Para cada $i \in \{1, \dots, n\}$, $S_i \cap \sum_{j \neq i} S_j = \{0\}$.

2.2.18. Ejercicio. Demuestre que $\mathbb{R}^{2\times 2} = U' \oplus H \oplus L'$.

2.3. Sistemas de generadores y dependencia lineal

2.3.1. Sean V un espacio vectorial y $v_1, \dots, v_n \in V$. Diremos que un vector $v \in V$ es **combinación lineal** de $v_1, \dots, v_n$ si existen $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tales que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. El conjunto de combinaciones lineales de $v_1, \dots, v_n$ será denotado por $\langle v_1, \dots, v_n \rangle$.

2.3.2. Ejemplo. Consideremos los vectores $v_1 = (2, 3, 4)$ y $v_2 = (0, 1, 1)$ de $\mathbb{R}^3$. Como $v = 2v_1 - v_2$, el vector $v = (4, 5, 7)$ es combinación lineal de v_1 y v_2.

2.3.3. Ejemplo. $\mathbb{R}[X] = \{f \in \mathbb{R}[X] : f(0) = 0\} \oplus \langle 1 \rangle$.

2.3.4. Sean V un espacio vectorial y $v_1, \dots, v_n \in V$. Entonces $\langle v_1, \dots, v_n \rangle$ es el menor subespacio de V que contiene a los vectores $v_1, \dots, v_n$ y se denomina el **subespacio generado** por $v_1, \dots, v_n$. Como consecuencia se obtienen fácilmente las siguientes propiedades:

1) $\langle v_1, \dots, v_n \rangle = \langle v_1 \rangle + \cdots + \langle v_n \rangle$.
2) Si $S = \langle v_1, \dots, v_n \rangle$ y $T = \langle w_1, \dots, w_m \rangle$ son subespacios de V entonces $S + T = \langle v_1, \dots, v_n, w_1, \dots, w_m \rangle$.

2.3.5. Ejercicio. Demuestre las siguientes propiedades:

1) $\langle v_1,\dots,v_i,\dots,v_j,\dots,v_n\rangle = \langle v_1,\dots,v_j,\dots,v_i,\dots,v_n\rangle$.
2) $\langle v_1,\dots,v_i,\dots,v_n\rangle = \langle v_1,\dots,\lambda v_i,\dots,v_n\rangle$ para todo $\lambda\in\mathbb{K}\setminus\{0\}$.
3) $\langle v_1,\dots,v_i,\dots,v_j,\dots,v_n\rangle = \langle v_1,\dots,v_i+\lambda v_j,\dots,v_j,\dots,v_n\rangle$ para todo $\lambda\in\mathbb{K}$.

2.3.6. Sea V un espacio vectorial y $X\subseteq V$ un subconjunto no vacío. Diremos que X es un **conjunto de generadores** para V (o que X genera a V) si para cada $v\in V$ existen $x_1,\dots,x_n\in X$ tales que $v\in\langle x_1,\dots,x_n\rangle$.

`exa:generadores` 2.3.7. Ejemplos.

1) El conjunto $\{(2,3),(1,2)\}$ es un conjunto de generadores de $\mathbb{R}^2$ pues
$$(x,y)=(2x-y)(2,3)+(-3x+2y)(1,2).$$
2) $\{e_1,\dots,e_n\}$ es un conjunto de generadores de $\mathbb{K}^n$.
3) $\{E^{ij}:1\leq i\leq m,\ 1\leq j\leq n\}$ es un conjunto de generadores de $\mathbb{K}^{m\times n}$.
4) $\{1,X,X^2,\dots,\}$ es un conjunto de generadores de $\mathbb{K}[X]$.
5) Sean $V=\mathbb{R}^{2\times 2}$ y $S=\{A\in V:A^T=A,\ \operatorname{tr}(A)=0\}$. Entonces
$$\left\{\begin{pmatrix}1&0\\0&-1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix}\right\}$$
es un conjunto de generadores de S. ¿Por qué?

`xca:f(1)=0` 2.3.8. Ejercicio. Sean $V=\mathbb{R}[X]$ y $S=\{f\in\mathbb{R}[X]:f(1)=0\}$. Demuestre que
$$\{X^{m+1}-X^m:m\in\mathbb{N}\}$$
es un conjunto de generadores de S.

`xca:X^2-3X+2|f` 2.3.9. Ejercicio. Sea $S=\{f\in\mathbb{R}[X]:f(1)=f(2)=0\}$. Pruebe que
$$\{(X^2-3X+2)X^m:m\geq 0\}$$
es un conjunto de generadores para S.

2.3.10. Ejercicio. Sea V un espacio vectorial y $X\subseteq V$ un subconjunto no vacío. Pruebe que
$$\bigcap_{S:X\subseteq S} S,$$
donde la intersección se toma sobre todo los subespacios de V que contienen a X es el menor subespacio de V que contiene a X. Concluya que este subespacio de V es igual al subespacio generado por X.

2.3.11. Ejemplo. $\mathbb{R}[X]=\mathbb{R}_n[X]\oplus\langle X^{n+1},X^{n+2},\dots\rangle$.

2.3.12. Un espacio vectorial V es **finitamente generado** si admite un conjunto finito de generadores.

2.3.13. Ejemplos. Vimos en el ejemplo 2.3.7 que los espacios vectoriales $\mathbb{K}^n$ y $\mathbb{K}^{m\times n}$ son finitamente generados.

2.3.14. El conjunto $\mathbb{R}$ de los números reales, visto como espacio vectorial sobre $\mathbb{R}$, es finitamente generado. Con un argumento de cardinalidad puede demostrarse que $\mathbb{R}$, como espacio vectorial sobre $\mathbb{Q}$, no es finitamente generado.

`xca:K[X]_no_fg` 2.3.15. Ejercicio. Demuestre que $\mathbb{K}[X]$ no es finitamente generado.

2.3.16. Sea V un espacio vectorial sobre un cuerpo $\mathbb{K}$. Un subconjunto finito y no vacío $\{v_1,\dots,v_n\}\subseteq V$ se dice **linealmente dependiente** si existen $\alpha_1,\dots,\alpha_n\in\mathbb{K}$, no todos cero, tales que $\alpha_1v_1+\cdots+\alpha_nv_n=0$. Observemos que $\{v_1,\dots,v_n\}$ es linealmente dependiente si y sólo si existe $j\in\{1,\dots,n\}$ tal que $v_j\in\langle v_1,\dots,\widehat{v_j},\dots,v_n\rangle$, donde
$$\langle v_1,\dots,\widehat{v_j},\dots,v_n\rangle=\langle v_1,\dots,v_{j-1},v_{j+1},\dots,v_n\rangle$$

2.3.17. Ejemplos.

1) El conjunto $\{(2,3),(1,2),(3,5)\} \subseteq \mathbb{R}^2$ es linealmente dependiente. ¿Por qué?
2) El conjunto $\{(1,1,1),(0,1,1),(0,0,1),(1,2,3)\} \subseteq \mathbb{R}^3$ es linealmente dependiente. ¿Por qué?

2.3.18. Ejercicio. Demuestre las siguientes propiedades:

1) Sea $\lambda \in \mathbb{K} \setminus \{0\}$. Entonces $\{v_1,\dots,v_i,\dots,v_n\}$ es linealmente dependiente si y sólo si $\{v_1,\dots,\lambda v_i,\dots,v_n\}$ es linealmente dependiente.
2) Sea $\lambda \in \mathbb{K}$. Entonces $\{v_1,\dots,v_i,\dots,v_j,\dots,v_n\}$ es linealmente dependiente si y sólo si $\{v_1,\dots,v_i+\lambda v_j,\dots,v_j,\dots,v_n\}$ es linealmente dependiente.

2.3.19. Sea V un espacio vectorial sobre $\mathbb{K}$. Un subconjunto finito se dice **linealmente independiente** si no es linealmente dependiente. Por definición, el conjunto vacío $\emptyset$ es linealmente independiente. Un subconjunto finito y no vacío $\{v_1,\dots,v_n\} \subseteq V$ es linealmente independiente si y sólo si $\alpha_1 v_1+\cdots+\alpha_n v_n = 0$ implica que $\alpha_1 = \cdots = \alpha_n = 0$.

2.3.20. Ejemplo. El conjunto $\{(1,2,1),(2,1,2),(1,0,3)\} \subseteq \mathbb{R}^3$ es linealmente independiente. ¿Por qué?

2.3.21. Un conjunto infinito de vectores es linealmente dependiente si tiene un subconjunto finito y no vacío linealmente dependiente. Un conjunto infinito de vectores es linealmente independiente si todo subconjunto finito es linealmente independiente.

2.3.22. Ejemplo. Sea V el espacio vectorial de funciones $\mathbb{R} \to \mathbb{R}$. Demostremos que el conjunto $\{f_a\colon \mathbb{R} \to \mathbb{R} : a \in \mathbb{R}\}$, donde

$$f_a\colon \mathbb{R} \to \mathbb{R}, \quad f_a(x) = |x-a|,$$

es linealmente independiente. En caso contrario, existirían $\lambda_2,\dots,\lambda_n \in \mathbb{R}$ y $a_1,\dots,a_n \in \mathbb{R}$, con $a_i \neq a_j$ para $i \neq j$, tales que

$$f_{a_1} = \lambda_2 f_{a_2} + \cdots + \lambda_n f_{a_n}.$$

Observemos que la función f_{a_1} no es derivable en $x = a_1$, y, en cambio, la función $\lambda_2 f_{a_2} + \cdots + \lambda_n f_{a_n}$ sí lo es. Como esto es una contradicción, las f_a son linealmente independientes.

2.3.23. Ejercicio. Sea V el espacio vectorial de funciones $\mathbb{R} \to \mathbb{R}$. Demuestre las siguientes afirmaciones:

1) El conjunto $\{e^{ax} : a \in \mathbb{R}\}$ es linealmente independiente.
2) El conjunto $\{2^{x-1}x^{i-1} : i \in \mathbb{N}\}$ es linealmente independiente.

2.3.24. Ejemplo. Consideremos a $\mathbb{R}$ como espacio vectorial sobre $\mathbb{Q}$ y demostremos que el conjunto $\{\log p : p \text{ primo positivo}\}$ es linealmente independiente. Si no lo fuera, existirían $n \in \mathbb{N}$, $\alpha_1,\dots,\alpha_n \in \mathbb{Q}$, no todos cero, y primos positivos $p_1,\dots,p_n$ tales que $\alpha_1 \log p_1 + \cdots + \alpha_n \log p_n = 0$. Sin pérdida de generalidad podemos suponer que $\alpha_i \in \mathbb{Z}$ para todo i. Luego

$$1 = e^0 = e^{\sum_{j=1}^n \alpha_j \log p_j} = \prod_{j=1}^n e^{\alpha_j \log p_j} = \prod_{j=1}^n p_j^{\alpha_j},$$

una contradicción.

2.3.25. Ejercicio. Pruebe que si $\{v_1,\dots,v_n\} \subseteq V$ es linealmente independiente entonces todo subconjunto de $\{v_1,\dots,v_n\}$ es también linealmente independiente.

2.3.26. Ejercicio. Pruebe que si $\{v_1,\dots,v_n\}$ es linealmente independiente entonces $\langle v_1,\dots,v_n\rangle = \langle v_1\rangle \oplus \cdots \oplus \langle v_n\rangle$.

2.4. Bases y dimensión

2.4.1. Una **base** de V es un conjunto de generadores de V linealmente independiente. Si V es finitamente generado, una **base ordenada** de V es una sucesión de vectores $v_1, \dots, v_n$ de V tal que $\{v_1, \dots, v_n\}$ es linealmente independiente y genera V.

2.4.2. Ejemplos.

1) $\{(2,3),(1,2)\}$ y $\{(1,2),(2,3)\}$ son dos bases ordenadas distintas de $\mathbb{R}^2$.
2) $\{e_1, \dots, e_n\}$ es base de $\mathbb{K}^n$ y se llama base canónica de $\mathbb{K}^n$.
3) $\{E^{ij} : 1 \leq i \leq m,\ 1 \leq j \leq n\}$, donde
$$(E^{ij})_{kl} = \delta_{ik}\delta_{jl}$$
para todo i, j, k, l, es base de $\mathbb{K}^{m\times n}$ y se llama **base canónica** de $\mathbb{K}^{m\times n}$.
4) $\{1, X, X^2, \dots\}$ es base de $\mathbb{K}[X]$.

2.4.3. Proposición. *Sea V un espacio vectorial. El conjunto $\{v_1, \dots, v_n\}$ es una base de V si y sólo si para cada $v \in V$ existen únicos $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tales que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$.*

Demostración. Supongamos que $\{v_1, \dots, v_n\}$ es una base. Entonces para cada $v \in V$ existen $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tales que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Veamos que los α_i son únicos: si suponemos que
$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n = \beta_1 v_1 + \cdots + \beta_n v_n,$$
donde los $\alpha_i, \beta_i \in \mathbb{K}$, entonces
$$(\alpha_1 - \beta_1)v_1 + \cdots + (\alpha_n - \beta_n)v_n = 0.$$
Luego $\alpha_i = \beta_i$ para todo $i \in \{1, \dots, n\}$ por la independencia lineal del conjunto $\{v_1, \dots, v_n\}$.

Recíprocamente, si todo $v \in V$ se escribe como combinación lineal de los v_i entonces $\{v_1, \dots, v_n\}$ es un conjunto de generadores de V. Veamos que es linealmente independiente: si $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$ entonces, como también $0 = 0v_1 + \cdots + 0v_n$, la unicidad de la escritura implica que $\alpha_i = 0$ para todo $i \in \{1, \dots, n\}$. □

:extraer_una_base

2.4.4. Proposición. *Sea V un espacio vectorial finitamente generado y sea $\{v_1, \dots, v_n\}$ un conjunto de generadores de V. Entonces existe un subconjunto de $\{v_1, \dots, v_n\}$ que es base de V.*

Demostración. Sean $X_0 = \{v_1, \dots, v_n\}$ y
$$X_1 = \begin{cases} X_0 \setminus \{v_1\} & \text{si } v_1 = 0, \\ X_0 & \text{si } v_1 \neq 0. \end{cases}$$
Para cada $j \geq 2$ definimos inductivamente
$$X_j = \begin{cases} X_{j-1} \setminus \{v_j\} & \text{si } v_j \in \langle X_{j-1} \rangle, \\ X_{j-1} & \text{si } v_j \notin \langle X_{j-1} \rangle, \end{cases}$$
Como V está generado por X_0, el proceso termina con X_n después de n pasos. Por construcción, V está generado por X_n, pues V está generado por X_0 y si V está generado por X_j, también X_{j+1} es un conjunto de generadores de V, ya que si $v_{j+1} \notin \langle X_j \rangle$, entonces $X_j = X_{j+1}$ y si $v_{j+1} \in \langle X_j$, entonces $\langle X_j \rangle = \langle X_{j+1} \rangle$. Además el conjunto X_n es linealmente independiente.. $\boxed{\text{FIXME}}$ □

2.4.5. Corolario. *Todo espacio vectorial no nulo finitamente generado admite una base.*

Demostración. El corolario es consecuencia de la proposición 2.4.4, que afirma que de todo conjunto de generadores puede extraerse una base. □

thm:>n_es_LD

2.4.6. Teorema. *Sea V un espacio vectorial finitamente generado por los vectores $v_1, \dots, v_n$. Entonces todo conjunto con más de n elementos es linealmente dependiente.*

Demostración. Sea $\{w_1, \dots, w_m\} \subseteq V$ con $m > n$. Como por hipótesis $V = \langle v_1, \dots, v_n \rangle$, para cada $j \in \{1, \dots, m\}$ existen escalares a_{ij} tales que $w_j = \sum_{i=1}^{n} a_{ij} v_i$. Como $m > n$, el sistema lineal

$$\sum_{j=1}^{m} a_{ij} v_j = 0, \quad i \in \{1, \dots, n\},$$

admite una solución no trivial, digamos $(\lambda_1, \dots, \lambda_m)$. Entonces

$$\sum_{j=1}^{m} \lambda_j w_j = \sum_{j=1}^{m} \lambda_j \left(\sum_{i=1}^{n} a_{ij} v_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} \lambda_j \right) v_i = 0,$$

y luego $\{w_1, \dots, w_m\}$ es un conjunto linealmente dependiente. $\square$

2.4.7. Corolario. *Sea V un espacio vectorial finitamente generado. Entonces dos bases cualesquiera tienen la misma cantidad de elementos.*

Demostración. Es consecuencia inmediata del teorema 2.4.6. $\square$

2.4.8. La **dimensión** de un espacio vectorial finitamente generado V es el cardinal de una base cualquiera de V y se denotará con $\dim V$. Por convención: $\dim\{0\} = 0$.

2.4.9. Ejemplos.

1) $\dim \mathbb{K}^n = n$.
2) $\dim \mathbb{K}^{m \times n} = mn$.
3) $\dim \mathbb{K}_n[X] = n + 1$.

block:agregar_v

2.4.10. Sean V un espacio vectorial y $v, v_1, \dots, v_n \in V$. Si $\{v_1, \dots, v_n\}$ es linealmente independiente y $v \notin \langle v_1, \dots, v_n \rangle$ entonces $\{v, v_1, \dots, v_n\}$ es linealmente independiente. En efecto, si $\alpha v + \sum \alpha_i v_i = 0$ entonces, como $v \notin \langle v_1, \dots, v_n \rangle$, se tiene que $\alpha = 0$. Luego $\alpha_i = 0$ para todo $i \in \{1, \dots, n\}$ pues $\{v_1, \dots, v_n\}$ es linealmente independiente.

2.4.11. Corolario. *Sea V un espacio vectorial sobre $\mathbb{K}$ de dimensión n y sean $v_1, \dots, v_n \in V$. Son equivalentes:*

1) *$\{v_1, \dots, v_n\}$ es base de V.*
2) *$\{v_1, \dots, v_n\}$ es linealmente independiente.*
3) *$\{v_1, \dots, v_n\}$ genera a V.*

Demostración. La implicación (1) $\Rightarrow$ (2) es trivial. Para demostrar que (2) implica (3) observemos que si existiera $v \in V \setminus \langle v_1, \dots, v_n \rangle$ entonces $\{v, v_1, \dots, v_n\}$ sería linealmente independiente por 2.4.10, y esto contradice el teorema 2.4.6. Finalmente, para demostrar que (3) implica (1) basta ver que $\{v_1, \dots, v_n\}$ es linealmente independiente. Para eso, extraemos un subconjunto $\{v_{i_1}, \dots, v_{i_k}\} \subseteq \{v_1, \dots, v_n\}$ linealmente independiente tal que $V = \langle v_{i_1}, \dots, v_{i_k} \rangle$. Luego $k = n$ y entonces $\{v_1, \dots, v_n\}$ es base de V. $\square$

tender_a_una_base

2.4.12. Proposición. *Sea V un espacio vectorial sobre $\mathbb{K}$ de dimensión n y para $m < n$ sea $\{v_1, \dots, v_m\} \subseteq V$ un conjunto linealmente independiente. Entonces existen $v_{m+1}, \dots, v_n \in V$ tales que $\{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$ es una base de V.*

Demostración. Sea $\{w_1, \dots, w_n\}$ un conjunto de generadores de V. Procederemos inductivamente. Para eso, sea $X_0 = \{v_1, \dots, v_m\}$. En el paso j-ésimo, donde $j \geq 1$, se define

$$X_j = \begin{cases} X_{j-1} & \text{si } w_j \in \langle X_{j-1} \rangle, \\ X_{j-1} \cup \{w_j\} & \text{si } w_j \notin \langle X_{j-1} \rangle. \end{cases}$$

El conjunto X_j es linealmente independiente para todo j. Además, por construcción, todos los w_j están en el subespacio de V generado por X_n. Hemos probado entonces que $X_n = \{v_1, \dots, v_n, w_{i_1}, \dots, w_{i_k}\}$ es una base de V. $\square$

2.4.13. Proposición. *Sea V un espacio vectorial de dimensión finita y sea $S \subseteq V$ un subespacio. Entonces S es de dimensión finita y $\dim S \leq \dim V$.*

Demostración. Sea $n = \dim V$. Si $S = \{0\}$ entonces S es de dimensión finita. Supongamos entonces que $S \neq \{0\}$ y sea $s_1 \in S \setminus \{0\}$. Si $S = \langle s_1 \rangle$ entonces S es de dimensión finita. En caso contrario, existe $s_2 \in S \setminus \{s_1\}$. Si procedemos inductivamente, en el paso j-ésimo, donde $j \geq 1$, se tiene un conjunto $\{s_1, \dots, s_{j-1}\}$. Si $S = \langle s_1, \dots, s_{j-1} \rangle$ entonces S es de dimensión finita. En caso contrario, existe $s_j \in S \setminus \langle s_1, \dots, s_{j-1} \rangle$. Como, por construcción, el conjunto de los s_j es linealmente independiente, el teorema 2.4.6 implica que este proceso tiene que terminar en a lo sumo n pasos. $\square$

2.4.14. Corolario. *Sean V un espacio vectorial de dimensión finita y $S \subseteq V$ un subespacio. Entonces existe un **complemento** de S, es decir: existe un subespacio $T \subseteq V$ tal que $V = S \oplus T$.*

Demostración. Como V es de dimensión finita, S es de dimensión finita. Sea $\{s_1, \dots, s_n\}$ una base de S. Por la proposición 2.4.12, podemos extender la base de S a una base de V, es decir: existen $t_1, \dots, t_m \in V$ tales que $\{s_1, \dots, s_n, t_1, \dots, t_m\}$ es base de V. Sea $T = \langle t_1, \dots, t_m \rangle$. Es evidente que $V = S + T$. Demostremos entonces que $S \cap T = \{0\}$. Si $v \in S \cap T$ entonces $v = \sum_{i=1}^n \alpha_i s_i = \sum_{j=1}^m \beta_j t_j$. Como

$$\alpha_1 s_1 + \cdots + \alpha_n s_n + (-\beta_1) t_1 + \cdots + (-\beta_m) t_m = 0$$

y $\{s_1, \dots, s_n, t_1, \dots, t_m\}$ es linealmente independiente, los α_i y los β_j son cero y por lo tanto $v = 0$. $\square$

2.4.15. Ejercicio. Sea V un espacio vectorial de dimensión n y sean S y T dos subespacios. Pruebe las siguientes afirmaciones:

1) Si $\dim S = n$ entonces $S = V$.
2) Si $S \subseteq T$ entonces $\dim S \leq \dim T$.
3) Si $S \subseteq T$ y $\dim S = \dim T$ entonces $S = T$.

m:de_la_dimension

2.4.16. Teorema (teorema de la dimensión). *Sea V un espacio vectorial de dimensión finita y sean S y T dos subespacios de V. Entonces*

$$\dim(S + T) = \dim S + \dim T - \dim(S \cap T).$$

Demostración. Supongamos que $\dim S = s$ y que $\dim T = t$ y sea entonces $\{u_1, \dots, u_m\}$ una base de $S \cap T$. Completamos a una base

$$\{u_1, \dots, u_m, v_{m+1}, \dots, v_s\}$$

de S y a una base

$$\{u_1, \dots, u_m, w_{m+1}, \dots, w_t\}$$

de T. Entonces $S + T$ está generado por

$$\{u_1, \dots, u_m, w_{m+1}, \dots, w_t, v_{m+1}, \dots, v_s\}.$$

Para demostrar el teorema basta probar que este conjunto es linealmente independiente. Si

$$0 = \sum_{i=1}^m \alpha_i u_i + \sum_{j=m+1}^s \beta_j v_j + \sum_{k=m+1}^t \gamma_k w_k$$

entonces, como $\sum_{k=m+1}^t \gamma_j w_j = -\sum_{i=1}^m \alpha_i u_i - \sum_{j=m+1}^s \beta_j v_j \in S \cap T$, existen $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ tales que

$$\sum_{k=1}^m \gamma_k w_k = \sum_{l=1}^m \lambda_l u_l.$$

Luego, como $\{u_1, \dots, u_m, w_{m+1}, \dots, w_t\}$ es una base de T, concluimos que $\gamma_k = 0$ para todo $k \in \{m+1, \dots, t\}$ y entonces $\alpha_i = 0$ para todo $i \in \{1, \dots, m\}$ y $\beta_j = 0$ para todo $j \in \{m+1, \dots, s\}$. Luego

$$\begin{aligned}\dim(S+T) &= m + (t-m) + (s-m) = s + t - m \\ &= \dim S + \dim T - \dim(S \cap T),\end{aligned}$$

que es lo que se quería demostrar. □

2.5. Rango de matrices

2.5.1. Sea $A \in \mathbb{K}^{m \times n}$. El **espacio fila** de A es el subespacio de $\mathbb{K}^n$ generado por las filas de A,

$$E_F(A) = \langle (a_{i1}, a_{i2}, \dots, a_{in}) : i \in \{1, \dots, m\} \rangle,$$

y se denota por $E_F(A)$. El **espacio columna** de A es el subespacio de $\mathbb{K}^n$ generado por las columnas de A

$$E_C(A) = \langle (a_{1j}, a_{2j}, \dots, a_{mj}) : j \in \{1, \dots, n\} \rangle.$$

y se denota por $E_C(A)$.

2.5.2. Ejemplo. Sea $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 5 \end{pmatrix}$. Entonces $E_F(A) = \langle (1,2,1), (0,1,5) \rangle$ y $E_C(A) = \langle (1,0), (2,1), (1,5) \rangle$. Observemos que $\dim E_F(A) = \dim E_C(A) = 2$.

2.5.3. Sea $A \in \mathbb{K}^{m \times n}$. La dimensión de $E_F(A)$ se denomina el **rango fila** de A y se denota por $\mathrm{rg}_F(A)$. La dimensión de $E_C(A)$ se denomina el **rango columna** de A y se denota por $\mathrm{rg}_C(A)$.

rem:producto

2.5.4. Observación. Sean $A \in \mathbb{K}^{m \times n}$, $B \in \mathbb{K}^{m \times r}$ y $C = \mathbb{K}^{r \times n}$ tales que $A = BC$. Entonces cada columna de A puede escribirse como combinación lineal de las columnas de B, es decir:

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = c_{1j} \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{pmatrix} + c_{2j} \begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{m2} \end{pmatrix} + \cdots + c_{rj} \begin{pmatrix} b_{1r} \\ b_{2r} \\ \vdots \\ b_{mr} \end{pmatrix}$$

para todo $j \in \{1, \dots, n\}$. Similarmente, toda fila de A puede escribirse como combinación lineal de las filas de C:

$$(a_{i1}\ a_{i2}\ \cdots\ a_{in}) = b_{i1}(c_{11}\ c_{12}\ \cdots\ c_{in}) + \cdots + b_{ir}(c_{r1}\ \cdots\ c_{rn})$$

para todo $i \in \{1, \dots, m\}$.

thm:rgC=rgF

2.5.5. Teorema. *Sea* $A \in \mathbb{K}^{m \times n}$. *Entonces* $\mathrm{rg}_F(A) = \mathrm{rg}_C(A)$.

Demostración. $\boxed{\text{FIXME}}$ Si $A = 0$ el resultado es trivial. Supongamos entonces que $A \neq 0$. Sea r el menor entero positivo tal que existen $B \in \mathbb{K}^{m \times r}$ y $C \in \mathbb{K}^{r \times n}$ tales que $A = BC$. (Observemos que si I es la identidad de $n \times n$ entonces $A = AI$ y entonces la existencia de un tal r está garantizada.) Por lo visto en la observación 2.5.4, las filas de C forman un conjunto minimal de generadores de $E_F(A)$ y las r columnas de B forman un conjunto minimal de generadores de $E_C(A)$. Luego, $\dim E_F(A) = \dim E_C(A) = r$. □

2.5.6. El **rango** de una matriz $A \in \mathbb{K}^{m \times n}$ se define como el rango fila (o columna) de A y se denota por $\mathrm{rg}(A)$.

2.5.7. Corolario. *Si* $A \in \mathbb{K}^{m \times n}$ *entonces* $\mathrm{rg}(A) = \mathrm{rg}(A^T)$.

Demostración. Es consecuencia inmediata del teorema 2.5.5. □

2.5.8. Proposición. *Sean* $A \in \mathbb{K}^{m\times n}$ *y* $B \in \mathbb{K}^{n\times p}$. *Entonces*

$$\operatorname{rg}(AB) \leq \text{mín}\{\operatorname{rg} A, \operatorname{rg} B\}.$$

Demostración. La proposición es consecuencia de la observación 2.5.4. Como el espacio generado por las columnas de AB es un subespacio del espacio generado por las columnas de A, $\operatorname{rg}(AB) \leq \operatorname{rg}(A)$. Similarmente, como el espacio generado por las filas de AB es un subespacio del espacio generado por las filas de B, se tiene que $\operatorname{rg}(AB) \leq \operatorname{rg}(B)$. □

2.5.9. Ejemplo. Sean $x_1, \dots, x_n$ números reales distintos. Veamos que el conjunto

$$\left\{ \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_1^{n-1} \end{pmatrix}, \dots, \begin{pmatrix} 1 \\ x_n \\ \vdots \\ x_n^{n-1} \end{pmatrix} \right\}$$

es una base de $\mathbb{R}^{n\times 1}$. Para esto veamos que la matriz

$$\begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \\ \vdots & \ddots & \vdots \\ x_1^{n-1} & \cdots & x_n^{n-1} \end{pmatrix},$$

de tamaño $n \times n$, tiene rango fila igual a n. En efecto, si las filas fueran linealmente dependientes, existirían $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{K}$, no todos cero, tales que

$$0 = \alpha_0(1, 1, \dots, 1) + \alpha_1(x_1, x_2, \dots, x_n) + \cdots + \alpha_{n-1}(x_1^{n-1}, x_2^{n-1}, \dots, x_n^{n-1}).$$

Esto implicaría que los $x_1, \dots, x_n$ son raíces del polinomio $f = \sum_{i=0}^{n-1} \alpha_i X^i$. Esto es una contradicción pues f tiene grado $\leq n - 1$.

2.6. Coordenadas y matriz de cambio de base

2.6.1. Sea V un espacio vectorial de dimensión n y sea $\mathscr{B} = \{v_1, \dots, v_n\}$ una base (ordenada) de V. Todo elemento $v \in V$ puede escribirse como combinación lineal de los elementos de $\mathscr{B}$, digamos $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$, donde los $\alpha_i \in \mathbb{K}$. La n-tupla $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$ es la tupla de **coordenadas** del vector v en la base $\mathscr{B}$. El vector de coordenadas de v en la base $\mathscr{B}$ se denota por $(v)_{\mathscr{B}}$.

2.6.2. Observación. Se tienen las siguientes propiedades:

1) $(v)_{\mathscr{B}} + (w)_{\mathscr{B}} = (v + w)_{\mathscr{B}}$.
2) $(\lambda v)_{\mathscr{B}} = \lambda(v)_{\mathscr{B}}$.

2.6.3. Ejemplos.

1) Sea $V = \mathbb{C}$ como espacio vectorial sobre $\mathbb{R}$. El conjunto $\{1, i\}$ es base de V. Las coordenadas de $a + bi$ en la base $\{1, i\}$ son (a, b).
2) Sean $V = \mathbb{R}^2$. Las coordenadas de un vector $(x, y) \in V$ en la base $\{(2, 3), (1, 2)\}$ son $(2x - y, -3x + 2y)$.
3) Sea $V = \{f \in \mathbb{R}[X] : \deg f \leq 2\}$. El vector de coordenadas del elemento $f = aX^2 + bX + c$ en la base $\{X^2, X, 1\}$ es (a, b, c).

2.6.4. Sea V un espacio vectorial de dimensión n y sean $\mathscr{B} = \{v_1, \dots, v_n\}$ y $\mathscr{B}' = \{v_1', \dots, v_n'\}$ dos bases de V. Escribamos a cada elemento de $\mathscr{B}$ en la base $\mathscr{B}'$:

$$v_j = \sum_{i=1}^{n} a_{ij} v_i' = a_{1j} v_1' + \cdots + a_{nj} v_n', \quad j \in \{1, \dots, n\}.$$

La matriz $A = (a_{ij})$ es la matriz de **cambio de base** de $\mathscr{B}$ a $\mathscr{B}'$ y se denota por $C(\mathscr{B}, \mathscr{B}')$. Observemos que las columnas de la matriz $C(\mathscr{B}, \mathscr{B}')$ son las coordenadas de los v_i en la base $\mathscr{B}'$, es decir:

$$C(\mathscr{B}, \mathscr{B}') = ((v_1)_{\mathscr{B}'} \mid \cdots \mid (v_n)_{\mathscr{B}'}).$$

Es evidente que $C(\mathscr{B}, \mathscr{B}) = I$.

2.6.5. Ejemplo. Sean $\{e_1, e_2\}$ la base canónica de $\mathbb{R}^2$ y $\mathscr{B} = \{(1,1), (1,-1)\}$. Un cálculo sencillo muestra que

$$C(\mathscr{B}, \{e_1, e_2\}) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad C(\{e_1, e_2\}, \mathscr{B}) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}.$$

pro:coordenadas

2.6.6. Proposición. *Sea V un espacio vectorial de dimensión n y supongamos que $\mathscr{B} = \{v_1, \dots, v_n\}$ y $\mathscr{B}' = \{v_1', \dots, v_n'\}$ son dos bases de V. Si los α_i son las coordenadas de v en la base $\mathscr{B}$ y los β_j son las coordenadas de v en la base $\mathscr{B}'$ entonces*

io_de_coordenadas

$$C(\mathscr{B}, \mathscr{B}') \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}. \tag{2.6.7}$$

Demostración. Un cálculo directo muestra que

$$v = \sum_{j=1}^n \alpha_j v_j = \sum_{j=1}^n \alpha_j \left(\sum_{i=1}^n a_{ij} v_i' \right) = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} \alpha_j \right) v_i',$$

tal como se quería demostrar. □

hm:cambio_de_base

2.6.8. Teorema. *Sea V un espacio vectorial de dimensión n y supongamos que $\mathscr{B} = \{v_1, \dots, v_n\}$ y $\mathscr{B}' = \{v_1', \dots, v_n'\}$ y $\mathscr{B}'' = \{v_1'', \dots, v_n''\}$ son tres bases de V. Entonces*

$$C(\mathscr{B}, \mathscr{B}'') = C(\mathscr{B}', \mathscr{B}'') C(\mathscr{B}, \mathscr{B}').$$

Demostración. Supongamos que $C(\mathscr{B}, \mathscr{B}'') = (c_{ij})$, $C(\mathscr{B}', \mathscr{B}'') = (b_{ij})$ y que $C(\mathscr{B}, \mathscr{B}') = (a_{ij})$, es decir:

$$v_j = \sum_{i=1}^n a_{ij} v_i', \qquad v_j' = \sum_{i=1}^n b_{ij} v_i'', \qquad v_j = \sum_{i=1}^n c_{ij} v_i''.$$

Como $v_j = \sum_{i=1}^n c_{ij} v_i''$,

$$v_j = \sum_{k=1}^n a_{kj} v_k' = \sum_{k=1}^n a_{kj} \left(\sum_{i=1}^n b_{ik} v_i'' \right) = \sum_{i=1}^n \left(\sum_{k=1}^n b_{ik} a_{kj} \right) v_i''$$

y los v_i'' forman una base de V, se tiene que $c_{ij} = \sum_{k=1}^n b_{ik} a_{kj}$ que es lo que se quería demostrar. □

cor:C_inversible

2.6.9. Corolario. *La matriz $C(\mathscr{B}, \mathscr{B}')$ es inversible y $C(\mathscr{B}, \mathscr{B}')^{-1} = C(\mathscr{B}', \mathscr{B})$.*

Demostración. Por el teorema 2.6.8 sabemos que $C(\mathscr{B}, \mathscr{B}) = C(\mathscr{B}', \mathscr{B}) C(\mathscr{B}, \mathscr{B}')$ y $C(\mathscr{B}', \mathscr{B}') = C(\mathscr{B}, \mathscr{B}') C(\mathscr{B}', \mathscr{B})$. El corolario queda demostrado pues $C(\mathscr{B}, \mathscr{B}) = I$ y $C(\mathscr{B}', \mathscr{B}') = I$. □

pro:C(-,B)

2.6.10. Proposición. *Sea $A \in \mathbb{K}^{n \times n}$ una matriz inversible y $\mathscr{B}$ una base de $\mathbb{K}^n$. Entonces existe una base $\mathscr{B}'$ de $\mathbb{K}^n$ tal que $A = C(\mathscr{B}', \mathscr{B})$.*

Demostración. Supongamos que $A = (a_{ij})$. Para cada $j \in \{1, \dots, n\}$ sea $v_j' = \sum_{i=1}^n a_{ij} v_i$. Veamos que el conjunto $\{v_1', \dots, v_n'\}$ es linealmente independiente. Como

$$0 = \sum_{j=1}^n \alpha_j v_j' = \sum_{j=1}^n \alpha_j \left(\sum_{i=1}^n a_{ij} v_i \right) = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} \alpha_j \right) v_i$$

y los v_i son linealmente independientes, $\sum_{j=1}^{n} a_{ij}\alpha_j = 0$ para todo i. Entonces

$$A\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix} = 0$$

y luego, como A es inversible, $\alpha_j = 0$ para todo $j \in \{1,\dots,n\}$. Como tenemos n de los v'_j, los v'_j son entonces una base de $\mathbb{K}^n$. Luego basta tomar $\mathscr{B} = \{v'_1,\dots,v'_n\}$ y la proposición queda demostrada. $\square$

cor:C(B,-)

2.6.11. Corolario. *Sea $A \in \mathbb{K}^{n\times n}$ una matriz inversible y $\mathscr{B}$ una base de $\mathbb{K}^n$. Entonces existe una base $\mathscr{B}'$ de $\mathbb{K}^n$ tal que $A = C(\mathscr{B},\mathscr{B}')$.*

Demostración. Basta aplicar la proposición 2.6.10 a la matriz A^{-1} y utilizar el corolario 2.6.9 que afirma que $C(\mathscr{B}',\mathscr{B})^{-1} = C(\mathscr{B},\mathscr{B}')$. $\square$

2.6.12. Ejemplo. Consideremos en el espacio vectorial $\mathbb{R}[X]$ los primeros cuatro **polinomios de Hermite**:

eq:Hermite

$$H_0 = 1, \qquad H_1 = 2X, \qquad H_2 = 4X^2 - 2, \qquad H_3 = 8X^3 - 12X. \tag{2.6.13}$$

Dejamos como ejercicio demostrar que $\mathscr{B} = \{H_0, H_1, H_2, H_3\}$ es un conjunto linealmente independiente. Escribamos cada polinomio de Hermite en la base canónica $\mathscr{E} = \{1, X, X^2, X^3\}$ de $\mathbb{R}_3[X]$ y formemos las matrices de cambio de base:

$$C(\mathscr{B},\mathscr{E}) = \begin{pmatrix}1 & 0 & -2 & 0\\ 0 & 2 & 0 & -12\\ 0 & 0 & 4 & 0\\ 0 & 0 & 0 & 8\end{pmatrix}, \qquad C(\mathscr{B},\mathscr{E})^{-1} = C(\mathscr{E},\mathscr{B}) = \frac{1}{8}\begin{pmatrix}8 & 0 & 4 & 0\\ 0 & 4 & 0 & 6\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}.$$

Luego, para escribir a $f = a_0 + a_1X + a_2X^2 + a_3X^3$ en la base de los polinomios de Hermite, hacemos

$$C(\mathscr{E},\mathscr{B})\begin{pmatrix}a_0\\ a_1\\ a_2\\ a_3\end{pmatrix} = \frac{1}{8}\begin{pmatrix}8a_0 + 4a_2\\ 4a_1 + 6a_3\\ 2a_2\\ a_3\end{pmatrix},$$

es decir:

$$f = \frac{1}{8}(8a_0 + 4a_2)H_0 + \frac{1}{8}(4a_1 + 6a_3)H_1 + \frac{1}{4}a_2H_2 + \frac{1}{8}a_3H_3.$$

2.7. Una aplicación del rango

2.7.1. Problema. *Una ciudad tiene dos reglas que controlan la creación de clubes: 1) Cada club debe tener un número impar de miembros, y 2) Dos clubes cualesquiera tienen a lo sumo un número par de miembros en común. Demuestre que la cantidad de clubes que pueden crearse con estas reglas es menor o igual que la cantidad de habitantes de la ciudad.*

Solución. Supongamos que la ciudad tiene n habitantes y sean $C_1,\dots,C_m$ los clubes. Queremos demostrar que $m \leq n$. Sea $A = (a_{ij}) \in \mathbb{Z}_2^{m\times n}$ la matriz dada por

$$a_{ij} = \begin{cases}1 & \text{si } j \in C_i,\\ 0 & \text{si } j \notin C_i.\end{cases}$$

Entonces $AA^T \in \mathbb{Z}_2^{m\times m}$ y $(AA^T)_{ij} = \sum_{k=1}^{n} a_{ik}a_{jk}$ es la cantidad de personas que hay en $C_i \cap C_j$. En particular,

$$\begin{cases}1 & \text{si } |C_i \cap C_j| \text{ es impar},\\ 0 & \text{si } |C_i \cap C_j| \text{ es par}.\end{cases}$$

La condiciones para la creación de clubes nos dicen que $AA^T = I$. Luego

$$m = \operatorname{rg}(I) = \operatorname{rg}(AA^T) \leq \operatorname{rg}(A) \leq \min\{n, m\}$$

y entonces $m \leq n$.

2.8. Una aplicación a las fracciones simples

2.8.1. Sean $r_1, \dots, r_n \in \mathbb{R}$ y sea $g = (X - r_1) \cdots (X - r_n) \in \mathbb{R}[X]$. Como aplicación de la teoría de espacios vectoriales vamos a demostrar que si $f \in \mathbb{R}[X]$ con $\deg f < \deg g$ entonces existen $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tales que

$$\frac{f}{g} = \sum_{i=1}^{n} \alpha_i \frac{1}{X - r_i}.$$

En efecto, primero observamos que el conjunto

$$S = \left\{ \frac{f}{g} : \deg f < \deg g \right\}$$

con las operaciones

$$\frac{f_1}{g} + \frac{f_2}{g} = \frac{f_1 + f_2}{g}, \quad \lambda \frac{f}{g} = \frac{\lambda f}{g},$$

es un espacio vectorial sobre $\mathbb{R}$. Además tiene dimensión finita pues el conjunto $\left\{ \frac{1}{g}, \frac{X}{g}, \dots, \frac{X^{n-1}}{g} \right\} \subseteq S$ es base de S. Por otro lado, como

$$\left\{ \frac{1}{X - r_1}, \dots, \frac{1}{X - r_n} \right\} \subseteq S$$

es un conjunto linealmente independiente, es también base de S y entonces existen escalares $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tales que

$$\frac{f}{g} = \sum_{i=1}^{n} \alpha_i \frac{1}{X - r_i},$$

tal como queríamos demostrar.

Capítulo 3

Transformaciones lineales

3.1. Transformaciones lineales

3.1.1. Sean V y W dos espacios vectoriales sobre $\mathbb{K}$. Una función $f\colon V \to W$ es una **transformación lineal** si

$$f(x+y) = f(x) + f(y), \quad f(\lambda x) = \lambda f(x),$$

para todo $x, y \in V$ y $\lambda \in \mathbb{K}$. Denotaremos por $\hom(V, W)$ al conjunto de transformaciones lineales $V \to W$. En $\hom(V, W)$ se tiene una multiplicación y está dada por la composición: si $f \in \hom(U, V)$ y $g \in \hom(V, W)$ entonces la composición $g \circ f\colon U \to W$, dada por $x \mapsto g(f(x))$, es un elemento de $\hom(U, W)$.

Utilizaremos la notación $gf = g \circ f$.

La composición de transformaciones lineales tiene las siguientes propiedades:

1) Es asociativa.
2) La identidad $\mathrm{id}_V\colon V \to V$, dada por $\mathrm{id}_V(x) = x$, cumple $f \,\mathrm{id}_V = \mathrm{id}_V$ para toda $f \in \hom(V, W)$.
3) La identidad $\mathrm{id}_W\colon W \to W$ cumple $\mathrm{id}_W f = \mathrm{id}_W$ para toda $f \in \hom(V, W)$.
4) Valen las propiedades distributivas:

$$(f_1 + f_2)g = f_1 g + f_2 g, \quad f(g_1 + g_2) = f g_1 + f g_2,$$

para toda $f_1, f_2, f \in \hom(U, V)$ y $g_1, g_2, g \in \hom(V, W)$.

Tal como sucede con las matrices, la composición de aplicaciones lineales en general no es conmutativa.

3.1.2. Ejercicio. Una función $f\colon V \to W$ es una transformación lineal si y sólo si $f(v + \lambda w) = f(v) + \lambda f(w)$ para todo $v, w \in V$ y $\lambda \in \mathbb{K}$.

3.1.3. Si V es un espacio vectorial, la mayoría de las funciones, digamos $V \to V$, no son transformaciones lineales. De hecho, supongamos por ejemplo que $V = \mathbb{Z}_2^4$ como espacio vectorial sobre el cuerpo $\mathbb{Z}_2$. Entonces V tiene $2^4 = 16$ elementos y hay 16^{16} funciones $V \to V$. Por otro lado, cada transformación lineal queda determinada por su valor en una base, y como cada base de V tiene cuatro elementos, hay, como mucho, 16^4 transformaciones lineales $V \to V$. Luego la probabilidad de que una función $V \to V$ elegida al azar sea una transformación lineal es menor o igual a $1/16^{12}$.

3.1.4. Ejemplos.

1) Sea $A \in \mathbb{K}^{m \times n}$. La función $f\colon \mathbb{K}^{n \times 1} \to \mathbb{K}^m$ dada por $x \mapsto Ax$ es una transformación lineal pues

$$\begin{aligned} f(x+y) &= A(x+y) = Ax + Ay = f(x) + f(y), \\ f(\lambda x) &= A(\lambda x) = \lambda A x = \lambda f(x). \end{aligned}$$

2) Sea $A \in \mathbb{K}^{m \times n}$. La función $f\colon \mathbb{K}^m \to \mathbb{K}^n$ dada por

$$(x_1, \dots, x_m) \mapsto \left(\sum_{i=1}^m a_{i1} x_i, \dots, \sum_{i=1}^n a_{in} x_i \right)$$

es una transformación lineal.
3) La traza $\mathrm{tr}\colon \mathbb{K}^{n \times n} \to \mathbb{K}$ es una transformación lineal.

4) La derivada
$$K[X] \to \mathbb{K}[X], \quad f \mapsto f'$$
donde f' denota el polinomio derivado de f, es una transformación lineal.

5) La derivada
$$C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), \quad f \mapsto f',$$
donde f' denota la función derivada de f, es una transformación lineal.

6) Sean $a, b \in \mathbb{R}$ con $a < b$. La aplicación
$$C(\mathbb{R}) \to C(\mathbb{R}), \quad f \mapsto \int_a^b f(x)dx$$
es una transformación lineal.

7) Las funciones $f\colon \mathbb{K}^\infty \to \mathbb{K}^\infty$ y $g\colon \mathbb{K}^\infty \to \mathbb{K}^\infty$,
$$f(x_1, x_2, x_3 \dots) = (x_2, x_3, \dots),$$
$$g(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots),$$
son transformaciones lineales.

3.1.5. Teorema (Propiedad fundamental de las transformaciones lineales). *Sea V un espacio vectorial sobre $\mathbb{K}$ de dimensión finita y sea $\{v_1, \dots, v_n\}$ una base de V. Sea W un espacio vectorial sobre $\mathbb{K}$ y sean $w_1, \dots, w_n \in W$. Entonces existe una única $f \in \hom(V, W)$ tal que $f(v_i) = w_i$ para todo $i \in \{1, \dots, n\}$.*

Demostración. Demostremos la existencia: todo $v \in V$ se escribe unívocamente como $v = \sum_{i=1}^n \alpha_i v_i$. Definimos entonces a f como $f(v) = \sum_{i=1}^n \alpha_i w_i$. Es evidente que f es una transformación lineal.

Demostremos ahora la unicidad: si $g \in \hom(V, W)$ y $g(v_i) = w_i = f(v_i)$ para todo $i \in \{1, \dots, n\}$ entonces
$$g(v) = g\left(\sum_{i=1}^n \alpha_i v_i\right) = \sum_{i=1}^n \alpha_i g(v_i) = \sum_{i=1}^n \alpha_i f(v_i) = f\left(\sum_{i=1}^n \alpha_i v_i\right) = f(v),$$
y luego $f = g$. $\square$

3.1.6. Ejemplo. Encontraremos explícitamente las transformaciones lineales $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ tales que $f(1, 2) = (1, 0, 2)$ y $f(-1, 3) = (0, 0, 1)$. Como $\{(1, 2), (-1, 3)\}$ es base de $\mathbb{R}^2$, el teorema anterior nos dice que existe una única f. Si $(x, y) \in \mathbb{R}^2$ entonces
$$(x, y) = \frac{3x + y}{5}(1, 2) + \frac{-2x + y}{5}(-1, 3)$$
y luego
$$f(x, y) = \frac{3x + y}{5}(1, 0, 2) + \frac{-2x + y}{5}(0, 0, 1).$$

3.2. Subespacios asociados a transformaciones lineales

magen_y_preimagen

3.2.1. Ejercicio. Sean V y W dos espacios vectoriales sobre $\mathbb{K}$ y $f \in \hom(V, W)$. Demuestre las siguientes afirmaciones:

1) Si $S \subseteq V$ es subespacio entonces $f(S) \subseteq W$ es subespacio.
2) Si $T \subseteq W$ es subespacio entonces $f^{-1}(T) \subseteq V$ es subespacio.

3.2.2. Sean V y W dos espacios vectoriales sobre $\mathbb{K}$ y $f \in \hom(V, W)$. Se define el **núcleo** de f como el conjunto
$$\ker f = \{v \in V : f(v) = 0\}$$
y la **imagen** de f como
$$\operatorname{im} f = f(V) = \{f(v) : v \in V\}.$$
Como consecuencia del ejercicio 3.2.1 se tiene que $\ker f \subseteq V$ e $\operatorname{im} f \subseteq W$ son subespacios.

3.2.3. Ejemplo. Sea $A \in \mathbb{K}^{m\times n}$ y sea f la transformación $x \mapsto Ax$. Más precisamente:

$$f\colon \mathbb{K}^{n\times 1} \to \mathbb{K}^{m\times 1}, \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}.$$

Entonces $\ker f$ es el conjunto de soluciones del sistema homogéneo $Ax = 0$ y la imagen $\operatorname{im} f$ es el conjunto de matrices b de $m \times 1$ tales que existe x de tamaño $n \times 1$ con $Ax = b$.

3.2.4. Sea $f \in \hom(V, W)$. Entonces f es inyectiva si y sólo si $\ker f = 0$. En efecto, si f es inyectiva y $v \in \ker f$ entonces $f(0) = 0 = f(v)$ y luego $v = 0$. Recíprocamente, si $\ker f = 0$ y $f(v) = f(v')$ entonces $f(v - v') = 0$ y por lo tanto $v - v' = 0$.

3.2.5. Sea $f \in \hom(V, W)$. Diremos que f es **monomorfismo** si $\ker f = 0$, f es **epimorfismo** si f es sobreyectiva, f es **isomorfismo** si f es monomorfismo y epimorfismo. Una transformación lineal $f \in \hom(V, V)$ se denomina **endomorfismo**. Un endomorfismo biyectivo se denomina **automorfismo**.

xca:fg=id

3.2.6. Ejercicio. Sean $f \in \hom(V, W)$ y $g \in \hom(W, V)$ tales que $fg = \mathrm{id}$. Pruebe que f es epimorfismo y g es monomorfismo.

Solución. Si $w \in W$ entonces $g(w) \in V$ y $w = f(g(w))$. Si $w \in \ker g$ entonces $g(w) = 0$ y luego $w = (fg)(w) = f(0) = 0$.

3.2.7. Ejemplos. Veamos algunos ejemplos de transformaciones lineales:

1) $\mathbb{R}^2 \to \mathbb{R}^3$, $(x, y) \mapsto (x, y, x)$, es monomorfismo y no es epimorfismo.
2) $\mathbb{R}^3 \to \mathbb{R}$, $(x, y, y) \mapsto x$, es epimorfismo y no es monomorfismo.
3) $\mathbb{K}^{n+1} \to \mathbb{K}_n[X]$, $(a_0, \dots, a_n) \mapsto \sum_{i=0}^n a_0 X^i$, es isomorfismo.
4) Si $A \in \mathbb{K}^{2\times 2}$ es inversible entonces

$$\mathbb{K}^{2\times 1} \to \mathbb{K}^{2\times 1}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto A \begin{pmatrix} x \\ y \end{pmatrix},$$

es un automorfismo de $\mathbb{R}^{2\times 1}$.

3.2.8. Si $f \in \hom(V, W)$ es un isomorfismo y g es la inversa de f entonces $g \in \hom(W, V)$. Sean $w, w' \in W$. Entonces existen $v, v' \in V$ tales que $f(v) = w$ y $f(v') = w'$. Luego $g(w + w') = g(f(v) + f(v')) = g(f(v + v')) = v + v' = g(w) + g(w')$. Además, si $\lambda \in \mathbb{K}$, entonces $g(\lambda w) = g(\lambda f(v)) = g(f(\lambda v)) = \lambda v = \lambda g(w)$.

3.2.9. Sean V y W dos espacios vectoriales sobre $\mathbb{K}$. Si existe un isomorfismo $V \to W$ entonces V y W se dicen **isomorfos**. La notación es: $V \simeq W$.

lem:mono

3.2.10. Lema. *Sean V y W dos espacios vectoriales sobre $\mathbb{K}$, y $f \in \hom(V, W)$ un monomorfismo. Si $\{v_1, \dots, v_n\} \subseteq V$ es linealmente independiente entonces $\{f(v_1), \dots, f(v_n)\}$ es linealmente independiente.*

Demostración. Si $\sum_{i=1}^n \alpha_i f(v_i) = 0$ entonces $f\left(\sum_{i=1}^n \alpha_i v_i\right) = 0$ y luego, como f es monomorfismo, $\sum_{i=1}^n \alpha_i v_i = 0$. Como los v_i son linealmente independientes, $\alpha_i = 0$ para todo i. □

lem:epi

3.2.11. Lema. *Sean V y W dos espacios vectoriales sobre $\mathbb{K}$, y $f \in \hom(V, W)$ un epimorfismo. Si $\{v_1, \dots, v_n\} \subseteq V$ es un conjunto de generadores de V entonces $\{f(v_1), \dots, f(v_n)\}$ genera W.*

Demostración. Sea $w \in W$. Como f es epimorfismo, existe $v \in V$ tal que $w = f(v)$. Si escribimos $v = \sum_{i=1}^n \alpha_i v_i$ entonces $w = f(v) = \sum_{i=1}^n \alpha_i f(v_i)$. □

pro:iso

3.2.12. Proposición. *Sean V y W dos espacios vectoriales sobre $\mathbb{K}$ y supongamos que V tiene dimensión finita y que $\{v_1, \dots, v_n\}$ es una base de V. Sea $f \in \hom(V, W)$. Entonces f es un isomorfismo si y sólo si $\{f(v_1), \dots, f(v_n)\}$ es base de W.*

Demostración. Si f es un isomorfismo, es monomorfismo y entonces, por el lema 3.2.10, el conjunto $\{f(v_1),\dots,f(v_n)\}$ es linealmente independiente. Como f es también un epimorfismo, $\{f(v_1),\dots,f(v_n)\}$ genera a W por el lema 3.2.11.

Supongamos ahora que los $f(v_i)$ son base de W. Probemos que f es epimorfismo: si $w\in W$ entonces $w\in\operatorname{im} f$ pues $w=\sum_{i=1}^n \alpha_i f(v_i)=f(\sum_{i=1}^n\alpha_i v_i)$. Probemos ahora que f es monomorfismo: sea v tal que $f(v)=0$. Entonces $v=\sum_{i=1}^n\alpha_i v_i$ y luego $0=f(v)=f(\sum_{i=1}^n\alpha_i v_i)=\sum_{i=1}^n\alpha_i f(v_i)$. Como los $f(v_i)$ son base, $\alpha_i=0$ para todo i. Luego $v=0$. □

3.2.13. Teorema. *Sean V y W dos espacios vectoriales sobre $\mathbb{K}$ de dimensión finita. Entonces $V\simeq W$ si y sólo si $\dim V=\dim W$.*

Demostración. Supongamos que $V\simeq W$ y sea $f\colon V\to W$ un isomorfismo. Si $\{v_1,\dots,v_n\}$ es una base de V entonces $\{f(v_1),\dots,f(v_n)\}$ es una base de W por la proposición 3.2.12. Luego $\dim V=\dim W$.

Recíprocamente, supongamos que $\dim V=\dim W$ y sean $\{v_1,\dots,v_n\}$ un base de V y $\{w_1,\dots,w_n\}$ una base de W. Entonces la función $f\colon V\to W$ que cumple $f(v_i)=w_i$ para todo i, es un isomorfismo por la proposición 3.2.12. □

thm:dimension

3.2.14. Teorema (teorema de la dimensión). *Sean V y W dos espacios vectoriales sobre $\mathbb{K}$ y $f\in\hom(V,W)$. Supongamos que V es de dimensión finita. Entonces*

$$\dim V=\dim\ker f+\dim\operatorname{im} f.$$

Demostración. Sea $\{v_1,\dots,v_k\}$ una base de $\ker f$. Consideremos la extensión a una base $\{v_1,\dots,v_k,v_{k+1},\dots,v_n\}$ de V. Afirmamos que $\{f(v_{k+1}),\dots,f(v_n)\}$ es una base de $\operatorname{im} f$. Veamos que es un conjunto linealmente independiente: si

$$0=\sum_{i=k+1}^n\alpha_i f(v_i)=f\left(\sum_{i=k+1}^n\alpha_i v_i\right)$$

entonces, como $\{v_1,\dots,v_k\}$ es base de $\ker f$, existen $\beta_1,\dots,\beta_k\in\mathbb{K}$ tales que

$$\sum_{i=k+1}^n\alpha_i v_i-\sum_{j=1}^k\beta_j v_j=0.$$

Como los v_i son linealmente independientes, $\alpha_i=0$ para todo $i\in\{k+1,\dots,n\}$ y $\beta_j=0$ para todo $j\in\{1,\dots,k\}$. Veamos ahora que $\operatorname{im} f$ está generado por $\{f(v_{k+1}),\dots,f(v_n)\}$: si $v\in V$ entonces $v=\sum_{i=1}^n\alpha_i v_i$. Luego

$$f(v)=\sum_{i=1}^n\alpha_i f(v_i)=\sum_{i=k+1}^n\alpha_i f(v_i)$$

pues $f(v_j)=0$ para todo $j\in\{1,\dots,k\}$. Esto demuestra el teorema. □

or:no_monomorfismo

3.2.15. Corolario. *Sean V y W dos espacios vectoriales sobre $\mathbb{K}$ de dimensión finita. Sea f un monomorfismo. Entonces $\dim V\leq\dim W$.*

Demostración. Si f es monomorfismo entonces $\ker f=\{0\}$. Luego, por el teorema de la dimensión,

$$0=\dim\ker f=\dim V-\dim\operatorname{im} f\geq\dim V-\dim W,$$

como se quería demostrar. □

or:no_epimorfismo

3.2.16. Corolario. *Sean V y W dos espacios vectoriales sobre $\mathbb{K}$ de dimensión finita. Sea f un epimorfismo. Entonces $\dim V\geq\dim W$.*

Demostración. Si f es un epimorfismo entonces $\operatorname{im} f = W$. Luego, por el teorema de la dimensión,

$$\dim W = \dim \operatorname{im} f = \dim V - \dim \ker f \leq \dim V,$$

tal como se quería demostrar. □

:mono<=>epi<=>iso

3.2.17. Corolario. *Sean V y W dos espacios vectoriales sobre $\mathbb{K}$ de dimensión finita tales que $\dim V = \dim W$ y sea $f \in \hom(V, W)$. Son equivalentes:*

1) *f es isomorfismo.*
2) *f es monomorfismo.*
3) *f es epimorfismo.*

Demostración. La implicación (1) $\Rightarrow$ (2) es trivial. Para demostrar (2) $\Rightarrow$ (3) observemos que, si f es monomorfismo, $\ker f = \{0\}$ y luego $\operatorname{im} f = W$, ya que por el teorema de la dimensión se tiene $\dim V = \dim \operatorname{im} f$. Queda demostrar que (3) $\Rightarrow$ (1). Si f es epimorfismo entonces $\operatorname{im} f = W$ y luego, por el teorema de la dimensión, $\dim V = \dim \ker f + \dim W$. Como $\dim V = \dim W$, $\ker f = \{0\}$ y entonces f es isomorfismo. □

3.2.18. Ejemplo. El corolario anterior no vale en dimensión infinita. Por ejemplo, si $f\colon \mathbb{K}^\infty \to \mathbb{K}^\infty$ es la transformación dada por $f(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots)$, $g\colon \mathbb{K}^\infty \to \mathbb{K}^\infty$ está dada por $g(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$ entonces $fg = \mathrm{id}$ pues

$$(fg)(x_1, x_2, \dots) = f(0, x_1, x_2, \dots) = (x_1, x_2, \dots).$$

Por el ejercicio 3.2.6 sabemos que entonces f es epimorfismo y g es monomorfismo. Sin embargo, f no es monomorfismo y g no es epimorfismo.

xca:Silvester

3.2.19. Ejercicio. Sean U, V, W espacios vectoriales de dimensión finita y sean $f \in \hom(U, V)$ y $g \in \hom(V, W)$. Demuestre las siguientes afirmaciones:

1) $\dim \ker(gf) \leq \dim \ker f + \dim \ker g$.
2) $\dim \operatorname{im}(gf) \leq \min\{\dim \operatorname{im} f, \dim \operatorname{im} g\}$.
3) $\dim \operatorname{im} f + \dim \operatorname{im} g - \dim V \leq \dim \operatorname{im}(gf)$.

3.3. Proyectores

3.3.1. Una transformación lineal $f\colon V \to V$ es un **proyector** si $f^2 = f$.

3.3.2. Ejemplo. La transformación lineal $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ dada por

$$f(x, y, z) = (3x - 2z, -x + y + z, 3x - 2z)$$

es un proyector.

xca:proyector

3.3.3. Ejercicio. Sea $f \in \hom(V, V)$. Pruebe que f es un proyector si y sólo si $f(w) = w$ para todo $w \in \operatorname{im} f$.

proyector(2f-1)^2

3.3.4. Ejercicio. Sea V un espacio vectorial sobre $\mathbb{R}$. Demuestre que $f \in \hom(V, V)$ es un proyector si y sólo si $(2f - \mathrm{id}_V)^2 = \mathrm{id}_V$.

3.3.5. Proposición. *Sea $f \in \hom(V, V)$ un proyector. Entonces*

$$V = \ker f \oplus \operatorname{im} f.$$

Demostración. Veamos primero que $V = \ker f + \operatorname{im} f$. En efecto, si $v \in V$ entonces $v = v - f(v) + f(v) \in \ker f + \operatorname{im} f$ pues $f(v - f(v)) = f(v) - f^2(v) = 0$. Si $v \in \ker f \cap \operatorname{im} f$ entonces, como $v \in \operatorname{im} f$, existe $x \in V$ tal que $v = f(x)$. Luego, como $v \in \ker f$, $0 = f(v) = f(f(x)) = f(x)$ y entonces $v = 0$. □

xca:pfp=fp

3.3.6. Ejercicio. Sean V un espacio vectorial y $f \in \hom(V, V)$. Pruebe que $S \subseteq V$ es un **subespacio invariante** por f (es decir: $f(S) \subseteq S$) si y sólo si $pfp = fp$ para todo proyector $p\colon V \to V$ con $\operatorname{im} p = S$.

3.3.7. Proposición. *Sea V un espacio vectorial y sean S y T dos subespacios de V tales que $V = S \oplus T$. Entonces existe un único proyector $f\colon V \to V$ tal que* $\operatorname{im} f = T$ *y* $\ker f = S$.

Demostración. Demostremos primero la existencia: sea $f\colon V \to V$ dado por $f(s+t) = t$ para todo $s \in S$ y $t \in T$. Como $V = S \oplus T$, f está bien definida. Es claro que $\ker f = S$ y que $\operatorname{im} f = T$. Además f es un proyector pues si $v \in V$ entonces v se escribe unívocamente como $v = s + t \in S + T$ y entonces $f(f(v)) = t = f(v)$.

Demostremos la unicidad: si $g \in \hom(V, V)$ es un proyector tal que $\operatorname{im} g = T$ y $\ker f = S$ entonces, como todo $v \in V$ se escribe unívocamente como $v = s + t$ con $s \in S$ y $t \in T$, se tiene que $g(v) = g(s) + g(t) = g(t)$. Como g es proyector, $g(y) = y$ para todo $y \in \operatorname{im} g$ por el ejercicio 3.3.3. Luego $g(v) = t = f(v)$. $\square$

oplus_proyectores

3.3.8. Ejercicio. Sea V un espacio vectorial y sean $S_1, \dots, S_n \subseteq V$ subespacios. Pruebe que las siguientes afirmaciones son equivalentes:

1) $V = S_1 \oplus \cdots \oplus S_n$.
2) Existen proyectores $p_1, \dots, p_n \in \hom(V, V)$ con $S_i = \operatorname{im} p_i$ para todo i tales que $p_i p_j = 0$ si $i \neq j$ y $p_1 + \cdots + p_n = \mathrm{id}_V$.

:proyector_matriz

3.3.9. Ejercicio. Sean V un espacio vectorial de dimensión finita sobre $\mathbb{K}$ y $f \in \hom(V, V)$. Pruebe que f es un proyector si y sólo si existe una base $\mathscr{B}$ de V tal que

$$\|f\|_{\mathscr{B}} = \left(\begin{array}{c|c} \mathrm{id}_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

para algún $r \leq \dim V$.

3.4. Matriz de una transformación lineal

3.4.1. Sean V y W dos espacios vectoriales sobre $\mathbb{K}$ de dimensión finita y sean $\mathscr{B}_V = \{v_1, \dots, v_n\}$ y $\mathscr{B}_W = \{w_1, \dots, w_m\}$ bases ordenadas de V y W respectivamente. Sea $f \in \hom(V, W)$. Si escribimos cada $f(v_j)$ en la base $\mathscr{B}_W$

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

se define la **matriz de la transformación lineal** f con respecto a las bases $\mathscr{B}_V$ y $\mathscr{B}_W$ como la matriz $\|f\|_{\mathscr{B}_V, \mathscr{B}_W} \in \mathbb{K}^{m \times n}$ dada por

$$\|f\|_{\mathscr{B}_V, \mathscr{B}_W} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

3.4.2. Ejemplos.

1) Si $V = W$ y $\mathscr{B}$ y $\mathscr{B}'$ son dos bases de V entonces la matriz de la identidad $\mathrm{id}_V\colon V \to V$, $v \mapsto v$, es $\|\mathrm{id}_V\|_{\mathscr{B}, \mathscr{B}'} = C(\mathscr{B}, \mathscr{B}')$.
2) Si $A \in \mathbb{K}^{m \times n}$ y $f\colon \mathbb{K}^{n \times 1} \to \mathbb{K}^{m \times 1}$ es la transformación lineal dada por

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

entonces la matriz de f con respecto a las bases canónicas de $\mathbb{K}^{n \times 1}$ y $\mathbb{K}^{m \times 1}$ es $\|f\| = A$.

3.4.3. Ejemplo. Sea $v = (a,b,c) \in \mathbb{R}^3$ y sea $f_v \colon \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal dada por $w \mapsto v \times w$, donde $v \times w$ denota el producto vectorial entre v y w. Como $f_v(1,0,0) = (0,c,-b)$, $f_v(0,1,0) = (-c,0,a)$ y $f_v(0,0,1) = (b,-a,0)$, la matriz de f_v con respecto a las bases canónicas es

$$\|f_v\| = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}.$$

3.4.4. Ejemplo. Consideremos la transformación lineal $\partial \colon \mathbb{K}_n[X] \to \mathbb{K}_n[X]$ dada por

$$\partial(a_0 + a_1X + \cdots + a_nX^n) = a_1 + a_2X + 2a_3X^2 + \cdots + na_nX^{n-1}.$$

Como $\partial(1) = 0$ y $\partial(X^j) = jX^{j-1}$ para todo $j \in \{1,\dots,n\}$, la matriz de ∂ con respecto a las bases canónicas es

$$\|\partial\| = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{K}^{(n+1)\times(n+1)}.$$

3.4.5. Como toda transformación lineal queda unívocamente determinada por su valor en una base, se obtiene el siguiente resultado: Sean $f, g \in \hom(V,W)$ y sean $\mathscr{B}_V$ y $\mathscr{B}_W$ bases de V y W respectivamente. Entonces $f = g$ si y sólo si $\|f\|_{\mathscr{B}_V,\mathscr{B}_W} = \|g\|_{\mathscr{B}_V,\mathscr{B}_W}$.

pro:|f|a=b

3.4.6. Proposición. *Sean $\mathscr{B}_V = \{v_1,\dots,v_n\}$ y $\mathscr{B}_W = \{w_1,\dots,w_m\}$ bases ordenadas de V y W respectivamente. Sea $f \in \hom(V,W)$, sea $v \in V$ y supongamos que*

$$v = \sum_{i=j}^{n} \alpha_j v_j, \quad f(v) = \sum_{i=1}^{m} \beta_i w_i,$$

es decir: $(v)_{\mathscr{B}_V} = (\alpha_1,\dots,\alpha_n)$ y $(f(v))_{\mathscr{B}_W} = (\beta_1,\dots,\beta_m)$. Entonces

$$\|f\|_{\mathscr{B}_V,\mathscr{B}_W} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}.$$

Demostración. La matriz de f con respecto a $\mathscr{B}_V$ y $\mathscr{B}_W$ es $(a_{ij}) \in \mathbb{K}^{m\times n}$ donde $f(v_j) = \sum_{i=1}^m a_{ij}w_i$ para todo j. Como $v = \sum_{j=1}^n \alpha_j v_j$ entonces

$$f(v) = \sum_{j=1}^{n} \alpha_j f(v_j) = \sum_{j=1}^{n} \alpha_j \left(\sum_{i=1}^{m} a_{ij} w_i \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}\alpha_j \right) w_i.$$

Como los w_i son base de W entonces $\sum_{j=1}^n a_{ij}\alpha_j = \beta_i$ para todo i, que es lo que queríamos demostrar. $\square$

thm:|gf|=|g||f|

3.4.7. Teorema. *Sean U, V y W espacios vectoriales sobre $\mathbb{K}$ de dimensión finita, y sean $\mathscr{B}_U = \{u_1,\dots,u_r\}$, $\mathscr{B}_V = \{v_1,\dots,v_n\}$ y $\mathscr{B}_W = \{w_1,\dots,w_m\}$ son bases ordenadas de U, V y W, respectivamente. Si $f \in \hom(U,V)$ y $g \in \hom(V,W)$ entonces*

$$\|gf\|_{\mathscr{B}_U,\mathscr{B}_W} = \|g\|_{\mathscr{B}_V,\mathscr{B}_W} \|f\|_{\mathscr{B}_U,\mathscr{B}_V}.$$

Demostración. Supongamos que

$$f(u_j) = \sum_{k=1}^{n} b_{kj} v_k, \qquad g(v_k) = \sum_{i=1}^{m} a_{ik} w_i, \qquad (gf)(u_j) = \sum_{i=1}^{m} c_{ij} w_i,$$

para todo $j \in \{1,\dots,r\}$ y $k \in \{1,\dots,n\}$, es decir:

$$\|f\|_{\mathscr{B}_U,\mathscr{B}_V} = (b_{ij}), \qquad \|g\|_{\mathscr{B}_V,\mathscr{B}_W} = (a_{ij}), \qquad \|gf\|_{\mathscr{B}_U,\mathscr{B}_W} = (c_{ij}).$$

Para cada $j \in \{1, \dots, r\}$ se tiene entonces

$$\begin{aligned}(gf)(u_j) = g(f(u_j)) = g\left(\sum_{k=1}^{n} b_{kj} v_k\right) &= \sum_{k=1}^{n} b_{kj} g(v_k)\\ = \sum_{k=1}^{n} b_{kj}\left(\sum_{i=1}^{m} a_{ik} w_i\right) &= \sum_{i=1}^{m}\left(\sum_{k=1}^{n} a_{ik} b_{kj}\right) w_i.\end{aligned}$$

Como los w_k son base de W y $(gf)(u_j) = \sum_{i=1}^{m} c_{ij} w_i$ para todo j, se tiene que $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$, tal como se quería demostrar. □

o<=>|f|inversible

3.4.8. Corolario. *Sean V y W dos espacios vectoriales sobre $\mathbb{K}$ de dimensión n, y sean $\mathscr{B}_V$ y $\mathscr{B}_W$ bases ordenadas de V y W respectivamente. Si $f \in \hom(V, W)$ entonces f es un isomorfismo si y sólo si $\|f\|_{\mathscr{B}_V,\mathscr{B}_W}$ es inversible.*

Demostración. Supongamos que $\mathscr{B}_V = \{v_1, \dots, v_n\}$, $\mathscr{B}_W = \{w_1, \dots, w_n\}$. Si $\|f\|_{\mathscr{B}_V,\mathscr{B}_W}$ es inversible, sean $v = \sum_{i=1}^{n} \alpha_i v_i \in \ker f$ y $f(v) = \sum_{j=1}^{n} \beta_j w_j$. Entonces $\beta_j = 0$ para todo j. Por la proposición 3.4.6,

$$\|f\|_{\mathscr{B}_V,\mathscr{B}_W} \begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix} = \begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}$$

y entonces $\alpha_i = 0$ para todo i pues $\|f\|_{\mathscr{B}_V,\mathscr{B}_W}$ es inversible. Luego f es monomorfismo y entonces f es isomorfismo por el corolario 3.2.17.

Recíprocamente, supongamos que f es un isomorfismo y sea $g \in \hom(W, V)$ la inversa de f. Entonces $\mathrm{id}_V = gf$ y $\mathrm{id}_W = fg$. Por el teorema 3.4.7,

$$I = \|\,\mathrm{id}_V\,\|_{\mathscr{B}_V,\mathscr{B}_W} = \|gf\|_{\mathscr{B}_V,\mathscr{B}_W} = \|g\|_{\mathscr{B}_W,\mathscr{B}_V} \|f\|_{\mathscr{B}_V,\mathscr{B}_W}.$$

Análogamente se obtiene que $\|f\|_{\mathscr{B}_V,\mathscr{B}_W} \|g\|_{\mathscr{B}_W,\mathscr{B}_V} = I$ y entonces la matriz $\|f\|_{\mathscr{B}_V,\mathscr{B}_W}$ es inversible. □

cor:semejanza:tls

3.4.9. Corolario. *Sea V un espacio vectorial de dimensión n y sean $\mathscr{B}_V$ y $\mathscr{B}'_V$ bases ordenadas de V. Sea W un espacio vectorial de dimensión m y sean $\mathscr{B}_W$ y $\mathscr{B}'_W$ bases ordenadas de W. Si $f \in \hom(V, W)$ entonces*

$$\|f\|_{\mathscr{B}'_V,\mathscr{B}'_W} = C(\mathscr{B}_W, \mathscr{B}'_W) \|f\|_{\mathscr{B}_V,\mathscr{B}_W} C(\mathscr{B}_V, \mathscr{B}'_V)^{-1}.$$

Demostración. El teorema 3.4.7 implica que

$$\begin{aligned}\|f\|_{\mathscr{B}'_V,\mathscr{B}'_W} &= \|\,\mathrm{id}_W \circ f\|_{\mathscr{B}'_V,\mathscr{B}'_W}\\ &= \|\,\mathrm{id}_W\,\|_{\mathscr{B}_W,\mathscr{B}'_W} \|f\|_{\mathscr{B}'_V,\mathscr{B}_W}\\ &= C(\mathscr{B}_W, \mathscr{B}'_W) \|f \circ \mathrm{id}_V\,\|_{\mathscr{B}'_V,\mathscr{B}_W}\\ &= C(\mathscr{B}_W, \mathscr{B}'_W) \|f\|_{\mathscr{B}_V,\mathscr{B}_W} \|\,\mathrm{id}_V\,\|_{\mathscr{B}'_V,\mathscr{B}_V}\\ &= C(\mathscr{B}_W, \mathscr{B}'_W) \|f\|_{\mathscr{B}_V,\mathscr{B}_W} C(\mathscr{B}'_V, \mathscr{B}_V),\end{aligned}$$

de donde se deduce el corolario. □

3.4.10. Diremos que dos matrices cuadradas $A \in \mathbb{K}^{n\times n}$ y $B \in \mathbb{K}^{n\times n}$ son **semejantes** si existe una matriz inversible $C \in \mathbb{K}^{n\times n}$ tal que $B = CAC^{-1}$.

cor:semejanza

3.4.11. Corolario. *Dos matrices $A, B \in \mathbb{K}^{n\times n}$ son semejantes si y sólo si existe $f \in \hom(\mathbb{K}^{n\times 1}, \mathbb{K}^{n\times 1})$ y existen $\mathscr{B}$ y $\mathscr{B}'$ bases ordenadas de $\mathbb{K}^{n\times 1}$ tales que $\|f\|_{\mathscr{B},\mathscr{B}} = A$ y $\|f\|_{\mathscr{B}',\mathscr{B}'} = B$.*

Demostración. Supongamos que $\|f\|_{\mathscr{B},\mathscr{B}} = A$ y que $\|f\|_{\mathscr{B}',\mathscr{B}'} = B$. Para demostrar que las matrices A y B son semejantes hay que aplicar el corolario 3.4.9 con $V = W = \mathbb{K}^{n\times 1}$, $\mathscr{B}_V = \mathscr{B}_W = \mathscr{B}$ y $\mathscr{B}'_V = \mathscr{B}'_W = \mathscr{B}'$.

Recíprocamente, supongamos que $B = CAC^{-1}$. Sea $f\colon \mathbb{K}^{n\times 1} \to \mathbb{K}^{n\times 1}$ la transformación lineal definida por $f(x) = Ax$. Entonces $\|f\| = A$. Como C es una matriz inversible, la proposición 2.6.10 nos dice que existe una base $\mathscr{B}$ de $\mathbb{K}^{n\times 1}$ tal que C es la matriz de cambio de base entre $\mathscr{B}$ y la base canónica $\{e_1, \dots, e_n\}$, es decir $C = C(\mathscr{B}, \{e_1, \dots, e_n\})$. Entonces, por el corolario 3.4.9,

$$\|f\|_{\mathscr{B},\mathscr{B}} = C\|f\|C^{-1} = CAC^{-1} = B,$$

tal como queríamos demostrar. $\square$

3.4.12. Observación. Sean $A \in \mathbb{K}^{m\times n}$ y

$$f\colon \mathbb{K}^{n\times 1} \to \mathbb{K}^{m\times 1}, \quad x \mapsto Ax.$$

Si $\{e_1, \dots, e_n\}$ es la base canónica de $\mathbb{K}^{n\times 1}$. Entonces $\operatorname{im} f = \langle Ae_1, \dots, Ae_n\rangle$ y luego, como los Ae_j son las columnas de A, obtenemos $\dim \operatorname{im} f = \operatorname{rg}(A)$.

3.5. El espacio vectorial $\hom(V, W)$

3.5.1. Sean V y W dos espacios vectoriales. El conjunto $\hom(V, W)$ de transformaciones lineales $V \to W$ con las operaciones

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

es un espacio vectorial sobre $\mathbb{K}$.

pro:hom(V,W)

3.5.2. Proposición. *Sean V y W dos espacios vectoriales de dimensión finita y supongamos que* $\dim V = n$ *y* $\dim W = m$*. Entonces* $\hom(V, W) \simeq \mathbb{K}^{m\times n}$*. En particular,* $\dim \hom(V, W) = (\dim V)(\dim W)$.

Demostración. Sean $\mathscr{B}_V = \{v_1, \dots, v_n\}$ y $\mathscr{B}_W = \{w_1, \dots, w_m\}$ bases de V y W respectivamente. La función

$$T\colon \hom(V, W) \to \mathbb{K}^{m\times n}, \quad f \mapsto \|f\|_{\mathscr{B}_V,\mathscr{B}_W}.$$

es una transformación lineal. Primero observemos que T es monomorfismo pues si $Tf = 0$ entonces $f(v_j) = 0$ para todo j y luego $f = 0$. Además T es epimorfismo pues si $(a_{ij}) \in \mathbb{K}^{m\times n}$ entonces la función $f\colon V \to W$, $f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$ para todo j, es una función lineal que cumple $\|f\|_{\mathscr{B}_V,\mathscr{B}_W} = (a_{ij})$. Luego $\hom(V, W) \simeq \mathbb{K}^{m\times n}$ y entonces $\dim \hom(V, W) = mn$. $\square$

3.6. Aplicaciones a sistemas lineales

cor:Ax=0<=>Ax=b

3.6.1. El corolario 3.2.17 nos permite demostrar el siguiente resultado.

Corolario. *Un sistema lineal no homogéneo de $n \times n$ tiene solución única si y sólo si el sistema lineal homogéneo asociado tiene solución única.*

Demostración. Sean $A = (a_{ij}) \in \mathbb{K}^{n\times n}$ y $f \in \hom(\mathbb{K}^{n\times 1}, \mathbb{K}^{n\times 1})$ dada por $x \mapsto Ax$. El resultado se deduce de la proposición 3.2.17 que afirma que f es monomorfismo si y sólo si f es epimorfismo. $\square$

3.6.2. Proposición. *Sea $A = (a_{ij})$ una matriz de tamaño $m \times n$. Entonces*

$$\dim\{x \in \mathbb{K}^{n\times 1} : Ax = 0\} = n - \operatorname{rg}(A).$$

Demostración. Sea $f\colon \mathbb{K}^{n\times 1} \to \mathbb{K}^{m\times 1}$ la transformación lineal definida por $x \mapsto Ax$. Entonces $\ker f = \{x \in \mathbb{K}^{n\times 1} : Ax = 0\}$ y luego

$$\dim \ker f = n - \dim \operatorname{im} f = n - \operatorname{rg}(A),$$

tal como se quería demostrar. $\square$

3.6.3. Proposición. *Sean $A = (a_{ij}) \in \mathbb{K}^{m\times n}$ y $b \in \mathbb{K}^{m\times 1}$. Entonces el sistema lineal $Ax = b$ tiene solución si y sólo si* $\operatorname{rg}(A|b) = \operatorname{rg}(A)$*, donde $(A|b)$ es la matriz ampliada del sistema $Ax = b$.*

Demostración. Primero observemos que el sistema lineal $Ax = b$ puede rescribirse como

eq:rg(A|b)=rg(A)

$$\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = \sum_{i=1}^{m} x_i \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}. \tag{3.6.4}$$

Ahora, $x \in \mathbb{K}^{n\times 1}$ es solución de $Ax = b$ si y sólo si x satisface la ecuación (3.6.4), y esto es equivalente a decir que b es combinación lineal de las columnas de A, que es equivalente a $\operatorname{rg}(A|b) = \operatorname{rg}(A)$. $\square$

3.7. Aplicaciones a matrices inversibles

pro:inversa

3.7.1. Proposición. *Sea $A \in \mathbb{K}^{n\times n}$. Entonces A es inversible si y sólo si existe $B \in \mathbb{K}^{n\times n}$ tal que $AB = I$.*

Demostración. Sea $B \in \mathbb{K}^{n\times n}$ tal que $AB = I$. Sea $f \in \mathbb{K}^{n\times n} \to \mathbb{K}^{n\times n}$ la transformación lineal dada por $X \mapsto AX$. Veamos que f es epimorfismo: si $C \in \mathbb{K}^{n\times n}$ entonces $f(BC) = A(BC) = (AB)C = IC = C$. Como

$$f(BA - I) = A(BA - I) = A(BA) - AI = (AB)A - A = IA - A = A - A = 0,$$

y f es monomorfismo por el corolario 3.2.17, se tiene que $BA - I = 0$. Luego A es inversible y $A^{-1} = B$. $\square$

cor:inversa

3.7.2. Corolario. *Sea $A \in \mathbb{K}^{n\times n}$. Entonces A es inversible si y sólo si existe $C \in \mathbb{K}^{n\times n}$ tal que $CA = I$.*

Demostración. Si existe $C \in \mathbb{K}^{n\times n}$ tal que $CA = I$ entonces C es inversible por la proposición anterior. Luego $C^{-1} = A$ y entonces A es inversible. $\square$

3.7.3. Corolario. *Sean $A, B \in \mathbb{K}^{n\times n}$. Entonces AB es inversible si y sólo si A y B son inversibles.*

Demostración. Si A y B son inversibles entonces AB es inversible con inversa $B^{-1}A^{-1}$. Supongamos que AB es inversible. Entonces existe $C \in \mathbb{K}^{n\times n}$ tal que $(AB)C = C(AB) = I$. En particular, como $A(BC) = I$ y $(CA)B = I$, la matrices A y B son inversibles por la proposición 3.7.1. $\square$

3.7.4. Proposición. *Sea $A \in \mathbb{K}^{n\times n}$. Son equivalentes:*

1) *A es inversible.*
2) *El sistema lineal $Ax = 0$ tiene una única solución.*
3) *Para cada $b \in \mathbb{K}^{n\times 1}$ el sistema $Ax = b$ tiene una única solución.*
4) *Para cada $b \in \mathbb{K}^{n\times 1}$ el sistema $Ax = b$ tiene al menos una solución.*
5) $\operatorname{rg}(A) = n$.

Demostración. Demostremos primero (1) $\Rightarrow$ (2). Si $Ax = 0$ entonces, como A es inversible, $x = Ix = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}0 = 0$ y $Ax = 0$ tiene una única solución. La implicación (2) $\Rightarrow$ (3) es el corolario 3.6.1. La implicación (3) $\Rightarrow$ (4) es trivial. Para demostrar (4) $\Rightarrow$ (5) sea $f\colon \mathbb{K}^{n\times 1} \to \mathbb{K}^{n\times 1}$ dada por $x \mapsto Ax$. Como f es epimorfismo por hipótesis, $\dim \operatorname{im} f = \operatorname{rg}(A) = n$. Finalmente, para demostrar que (5) $\Rightarrow$ (1) basta observar que $f\colon \mathbb{K}^{n\times 1} \to \mathbb{K}^{n\times 1}$ dada por $x \mapsto Ax$ es un epimorfismo y entonces es un isomorfismo. $\square$

Capítulo 4

Espacio dual

4.1. Espacio dual y base dual

4.1.1. Sea V un espacio vectorial sobre $\mathbb{K}$. Se define el **espacio dual** V^* como el espacio vectorial $\hom(V, \mathbb{K}^*)$. Los elementos de V^* se denominal **funcionales lineales** de V. Observemos que si V es un espacio vectorial de dimensión finita n entonces, por lo visto en la proposición 3.5.2, se tiene que

$$\dim V^* = \dim \hom(V, \mathbb{K}) = \dim V.$$

4.1.2. Ejemplos. En el espacio de matrices de $n \times n$ la traza $\operatorname{tr}\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ es una funcional lineal. En el espacio de polinomios, la evaluación en $\lambda \in \mathbb{K}$,

$$\mathbb{K}[X] \to \mathbb{K}, \quad \sum_{i=1}^{n} a_i X^i \mapsto \sum_{i=1}^{n} a_i \lambda^i,$$

es una funcional lineal. En el espacio vectorial real $C[a,b]$ de funciones continuas $[a,b] \to \mathbb{R}$ la aplicación $g \mapsto \int_a^b g(x)dx$ es una funcional lineal.

4.1.3. Ejemplo. Sea V el subespacio de $C^\infty(\mathbb{R})$ formado por las funciones f tales que $f(x) = 0$ para todo x fuera de algún intervalo cerrado y acotado. La aplicación $f \mapsto \int_{-\infty}^{+\infty} f(x)dx$ es una funcional lineal de V.

:funcional_lineal 4.1.4. Ejemplo. La función $f\colon \mathbb{K}^n \to \mathbb{K}$ dada por

$$f(x_1, \dots, x_n) = \alpha_1 x_1 + \cdots + \alpha_n x_n$$

es una funcional lineal. Más aún, la matriz de f con respecto a las bases canónicas de $\mathbb{K}^n$ y $\mathbb{K}$ es $\|f\| = (\alpha_1 \cdots \alpha_n)$. Veamos que toda funcional lineal de $\mathbb{K}^n$ es de esta forma. En efecto, si $\{e_1, \dots, e_n\}$ es la base canónica de $\mathbb{K}^n$ basta con definir $f(e_j) = \alpha_j$ para todo j pues entonces

$$\begin{aligned} f(x_1, \dots, x_n) &= f(x_1 e_1 + \cdots + x_n e_n) \\ &= x_1 f(e_1) + \cdots + x_n f(e_n) \\ &= \alpha_1 x_1 + \cdots + \alpha_n x_n. \end{aligned}$$

4.1.5. Ejemplo. Para $i \in \{1,2,3\}$ sean $\delta_i\colon \mathbb{R}^3 \to \mathbb{R}$ dada por $(x_1, x_2, x_3) \mapsto x_i$. Entonces por lo visto en el ejemplo 4.1.4 se tiene que $(\mathbb{R}^3)^* = \langle \delta_1, \delta_2, \delta_3 \rangle$.

xca:traza 4.1.6. Ejercicio. Sea $V = \mathbb{K}^{n\times n}$ y sea $\delta \in V^*$ tal que $\delta(AB) = \delta(BA)$ para todo $A, B \in V$. Pruebe que existe $\lambda \in \mathbb{K}$ tal que $\delta(A) = \lambda \operatorname{tr}(A)$ para todo $A \in V$.

4.1.7. Proposición. *Sea V un espacio vectorial de dimensión finita y sea $\{v_1, \dots, v_n\}$ una base de V. Entonces existe una única base $\{f_1, \dots, f_n\}$ de V^* tal que $f_i(v_j) = \delta_{ij}$ para todo i, j. La base $\{f_1, \dots, f_n\}$ se denomina **base dual** a la base $\{v_1, \dots, v_n\}$.*

Demostración. Como los v_j son una base de V, para cada i las funcionales lineales f_i quedan bien definidas por $f_i(v_j) = \delta_{ij}$. Veamos que el conjunto $\{f_1, \dots, f_n\}$ es linealmente independientes: si $\sum_{i=1}^n f_i = 0$ entonces para todo v_j se tiene

$$0 = \left(\sum_{i=1}^{n} \alpha_i f_i \right)(v_j) = \sum_{i=1}^{n} \alpha_i f_i(v_j) = \alpha_j$$

y luego $\alpha_j = 0$ para todo j. Como $\dim V^* = n$, esto demuestra además que los f_j forman una base de V^*.

Veamos la unicidad: si las g_i son funcionales lineales tales que $g_i(v_j) = f_i(v_j)$ para todo i, j entonces

$$g_i(v) = g_i\left(\sum_{j=1}^{n} \alpha_j v_j\right) = \sum_{j=1}^{n} \alpha_j g_i(v_j) = \sum_{j=1}^{n} \alpha_j f_i(v_j) = f_i(v)$$

y luego $g_i = f_i$ para todo i. □

`exa:base_dual` 4.1.8. Ejemplos.

1) La base dual de $\{(1,0),(0,1)\} \subseteq \mathbb{R}^2$ es $\{f_1, f_2\}$ donde $f_1(x,y) = x$ y $f_2(x,y) = y$.
2) La base dual de $\{(1,1),(1,-1)\} \subseteq \mathbb{R}^2$ es $\{f_1, f_2\}$ donde $f_1(x,y) = \frac{x+y}{2}$ y $f_2(x,y) = \frac{x-y}{2}$.

4.1.9. Ejemplo. Sea $V = \mathbb{K}_n[X]$ y consideremos la base

$$\mathscr{B} = \{1, X-a, (X-a)^2, \dots, (X-a)^n\}.$$

La base dual a $\mathscr{B}$ es $\{f_0, f_1, \dots, f_n\}$, donde

$$f_k(p) = \frac{p^{(k)}(a)}{k!}$$

para todo $k \in \{0, \dots, n\}$ y $p \in \mathbb{K}_n[X]$.

4.1.10. Proposición. *Sean $\mathscr{B} = \{v_1, \dots, v_n\}$ una base de V y $\mathscr{B}^* = \{f_1, \dots, f_n\}$ su base dual. Si $v \in V$ y $f \in V^*$ entonces*

$$(v)_{\mathscr{B}} = (f_1(v), \dots, f_n(v)), \qquad (f)_{\mathscr{B}^*} = (f(v_1), \dots, f_n(v)). \tag{4.1.11}$$

Demostración. Si $v = \sum_{i=1}^{n} \alpha_i v_i$ entonces $f_j(v) = \sum_{i=1}^{n} \alpha_i f_j(v_i) = \alpha_j$ y luego $v = \sum_{i=1}^{n} f_i(v) v_i$. Por otro lado, si escribimos a f como $f = \sum_{i=1}^{n} \alpha_i f_i$ entonces $f(v_j) = \alpha_j$ y luego $f = \sum_{i=1}^{n} f(v_i) f_i$. □

`xca:v=0<=>f(v)=0` 4.1.12. Ejercicio. Sea V un espacio vectorial de dimensión finita. Sean $f \in V^*$ y $v \in V$. Pruebe que $v = 0$ si y sólo si $f(v) = 0$ para todo $f \in V^*$.

`xca:V=<v>+kerf` 4.1.13. Ejercicio. Sea V un espacio vectorial de dimensión finita. Sean $f \in V^*$ y $v \in V$ con $f(v) \neq 0$. Pruebe que $V = \langle v \rangle \oplus \ker f$.

`xca:kerf=kerg` 4.1.14. Ejercicio. Sea V un espacio vectorial de dimensión finita. Si $f, g \in V^*$ pruebe que $\ker f = \ker g$ si y sólo si existe $\lambda \in \mathbb{K} \setminus \{0\}$ tal que $f = \lambda g$.

4.1.15. Ejemplo. Sea $V = \mathbb{R}_2[X]$ y sean $\varphi_0, \varphi_1, \varphi_2 \in V^*$ dadas por $\varphi_i(p) = p(i)$ para todo $p \in V$. Vamos a demostrar que $\{\varphi_0, \varphi_1, \varphi_2\}$ es base de V^*. Veamos primero que es un conjunto linealmente independiente. Si

$$\alpha_0 \varphi_0 + \alpha_1 \varphi_1 + \alpha_2 \varphi_2 = 0,$$

al evaluar en el polinomio $(X-1)(X-2)$ obtenemos $\alpha_0 = 0$. Al evaluar en $X(X-1)$ obtenemos $\alpha_2 = 0$. Finalmente, al evaluar en $X(X-2)$ obtenemos $\alpha_1 = 0$. Luego $\{\varphi_0, \varphi_1, \varphi_2\}$ es una base de V^* porque es un conjunto linealmente independiente y $\dim V = \dim V^* = 3$.

4.2. Matriz de cambio de base

`al:cambio_de_base` 4.2.1. Proposición. *Sean $\mathscr{B} = \{v_1, \dots, v_n\}$ y $\mathscr{B}' = \{v_1', \dots, v_n'\}$ bases de V. Entonces*

$$C(\mathscr{B}^*, \mathscr{B}'^*) = C(\mathscr{B}', \mathscr{B})^T.$$

Demostración. Supongamos que $C(\mathscr{B}^*, \mathscr{B}'^*) = (a_{ij})$ y que $C(\mathscr{B}', \mathscr{B}) = (b_{ij})$. Para cada i, j escribimos

$$v_i' = \sum_{k=1}^n b_{ki} v_k, \qquad f_j = \sum_{k=1}^n a_{kj} f_k'.$$

Luego, si calculamos $f_j(v_i')$ de dos formas, obtenemos

$$b_{ji} = \sum_{k=1}^n b_{ki}\delta_{jk} = \sum_{k=1}^n b_{ki} f_j(v_k) = f_j(v_i') = \sum_{k=1}^n a_{kj} f_k'(v_i') = \sum_{k=1}^n a_{kj}\delta_{ki} = a_{ij},$$

que es lo que queríamos demostrar. □

4.2.2. Ejemplo. Vamos a utilizar la proposición 4.2.1 para calcular la base dual a la base $\mathscr{B} = \{(1,1),(1,-1)\}$ de $\mathbb{R}^2$ vista en el ejemplo 4.1.8. Si $\{e_1, e_2\}$ es la base canónica de $\mathbb{R}^2$ entonces

$$C(\mathscr{B}, \{e_1, e_2\}) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad C(\{e_1, e_2\}, \mathscr{B}) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

Sea $\{f_1, f_2\}$ la base dual a $\{e_1, e_2\}$ y sea $\mathscr{B}^* = \{g_1, g_2\}$ la base dual a $\mathscr{B}$. Por la proposición 4.2.1 sabemos que

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = C(\mathscr{B}, \{e_1, e_2\})^T = C(\{f_1, f_2\}, \mathscr{B}^*).$$

Entonces, como

$$C(\mathscr{B}^*, \{f_1, f_2\}) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

podemos calcular las coordenadas de las g_i en la base $\{f_1, f_2\}$:

$$\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Luego $g_1(x,y) = \frac{1}{2}(x+y)$ y $g_2(x,y) = \frac{1}{2}(x-y)$.

4.3. El anulador de un subespacio

4.3.1. Sean V un espacio vectorial y $S \subseteq V$ un subespacio. Se define el **anulador** de S en V como el subespacio

$$\operatorname{Ann} S = \{f \in V^* : f(s) = 0 \text{ para todo } s \in S\} = \{f \in V^* : S \subseteq \ker f\}$$

Observemos que el anulador $\operatorname{Ann} S$ puede definirse si S es un subconjunto no vacío de V. Por ejemplo, el anulador en $\mathbb{R}^2$ de $\{(1,1)\}$ es el subespacio de $\hom(\mathbb{R}^2, \mathbb{R})$ generado por la función $(x,y) \mapsto x - y$. De la definición es evidente que $\operatorname{Ann} S$ es un subespacio de V^* y que $\operatorname{Ann} V = \{0\}$ y $\operatorname{Ann}\{0\} = V^*$.

xca:annX=<X> 4.3.2. Ejercicio. Sea V un espacio vectorial y X un subconjunto de V. Pruebe que $\operatorname{Ann} X = \operatorname{Ann}\langle X \rangle$.

4.3.3. Ejemplo. Sea $S = \langle (1,1,1,1),(1,1,0,0),(1,0,1,0)\rangle \subseteq \mathbb{R}^4$. Entonces el anulador de S está generado por la funcional lineal

$$(x_1, x_2, x_3, x_4) \mapsto -x_1 + x_2 + x_3 - x_4.$$

:dual:fundamental 4.3.4. Lema. *Sean V un espacio vectorial de dimensión finita, $S \subseteq V$ un subespacio y $\{v_1, \dots, v_k\}$ una base de S. Supongamos que $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ es una base de V y sea $\{f_1, \dots, f_n\}$ su base dual. Entonces $\{f_{k+1}, \dots, f_n\}$ es una base de* $\operatorname{Ann} S$.

Demostración. Sea $f \in \operatorname{Ann} S$ y escribamos a f en la base $\{f_1, \dots, f_n\}$. Entonces, como $f(v_j) = 0$ para todo $j \in \{1, \dots, k\}$,

$$f = \sum_{i=1}^{n} \alpha_i f_i = \sum_{i=1}^{n} f(v_i) f_i = \sum_{i=k+1}^{n} \alpha_i f_i.$$

Luego $\{f_{k+1}, \dots, f_n\}$ es un conjunto de generadores de $\operatorname{Ann} S$. Como es un conjunto linealmente independiente, es también una base de $\operatorname{Ann} S$. $\square$

4.3.5. Proposición. *Sean V un espacio vectorial de dimensión finita y $f \in V^*$. Entonces* $\operatorname{Ann} \ker f = \langle f \rangle$.

Demostración. Vamos a demostrar que $\operatorname{Ann} \ker f \subseteq \langle f \rangle$ ya que la otra inclusión es trivial. Como el resultado es trivialmente válido para $f = 0$ vamos a suponer que $f \neq 0$. Si $g \in \operatorname{Ann} \ker f$ entonces $\ker f \subseteq \ker g$. Si $g = 0$ el resultado es trivial. Si $g \neq 0$ entonces $\ker f = \ker g$ pues ambos tienen la misma dimensión. Luego $g \in \langle f \rangle$ por el ejercicio 4.1.12. $\square$

imension_anulador 4.3.6. Teorema (teorema de la dimensión del anulador). *Si V es un espacio vectorial de dimensión n y $S \subseteq V$ es un subespacio entonces*

$$\dim \operatorname{Ann} S = n - \dim S.$$

Demostración. Supongamos que $\dim S = k$ y sea $\{v_1, \dots, v_k\}$ una base de S. Sean $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ una base de V y $\{f_1, \dots, f_n\}$ su base dual. El lema 4.3.4 nos dice que $\{f_{k+1}, \dots, f_n\}$ es base de $\operatorname{Ann} S$. Luego $\dim \operatorname{Ann} S = n - k$ y el teorema queda demostrado. $\square$

4.3.7. Ejemplo. Consideremos los siguientes vectores de $\mathbb{R}^5$:

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

y sea $S = \langle v_1, v_2, v_3, v_4 \rangle$. Vamos a calcular $\operatorname{Ann} S$. Como $\dim S = 3$, por el teorema de la dimensión del anulador sabemos que $\dim \operatorname{Ann} S = 2$. Si $f \in (\mathbb{R}^5)^*$ entonces

$$f(x_1, \dots, x_5) = \sum_{i=1}^{5} \alpha_i x_i.$$

Luego $f \in \operatorname{Ann} S$ si y sólo si $f(v_i) = 0$ para todo i, es decir si y sólo si

$$\begin{cases} 2\alpha_1 - 2\alpha_2 + 3\alpha_3 + 4\alpha_4 - \alpha_5 = 0, \\ -\alpha_1 + \alpha_2 + 2\alpha_3 + 5\alpha_4 + 2\alpha_5 = 0, \\ -\alpha_3 - 2\alpha_4 + 3\alpha_5 = 0, \\ \alpha_1 - \alpha_2 + 2\alpha_3 + 3\alpha_4 = 0. \end{cases}$$

Como este sistema es equivalente a

$$\begin{cases} \alpha_1 - \alpha_2 - \alpha_4 = 0, \\ \alpha_3 + 2\alpha_4 = 0, \\ \alpha_5 = 0, \end{cases}$$

se deduce que $f(x_1, \dots, x_5) = (\alpha_2 + \alpha_4) x_1 + \alpha_2 x_2 - 2\alpha_4 x_3 + \alpha_4 x_4$. Luego una base de $\operatorname{Ann} S$ es $\{f_1, f_2\}$, donde

$$f_1(x_1, \dots, x_5) = x_1 + x_2, \quad f_2(x_1, \dots, x_5) = x_1 - 2x_3 + x_4.$$

4.3.8. Corolario. *Sea V un espacio vectorial de dimensión finita n y sea $S \subseteq V$ un subespacio de dimensión $m < n$. Entonces existen **hiperplanos** (es decir: subespacios de dimensión $n-1$) $H_1, \dots, H_{n-m} \subseteq V$ tales que $S = H_1 \cap \dots \cap H_{n-m}$.*

Demostración. Sea $\{v_1,\dots,v_m\}$ una base de S. Extendemos esta base de S a una base $\{v_1,\dots,v_m,v_{m+1},\dots,v_n\}$ de V y sea $\{f_1,\dots,f_m,f_{m+1},\dots,f_n\}$ su base dual. Para cada $i\in\{1,\dots,n-m\}$ sea $H_i=\ker f_{m+i}$. Si $v\in V$ entonces $v=\sum_{i=1}^n f_i(v)v_i$. Tenemos entonces que $v\in S$ si y sólo si $v\in\cap_{i=1}^{n-m}\ker f_i$ y por lo tanto $S=\cap_{i=1}^{n-m}H_i$. □

cor:annS=annT 4.3.9. Corolario. *Sea V un espacio vectorial de dimensión finita y sean S y T dos subespacios de V. Entonces $S=T$ si y sólo si $\operatorname{Ann}S=\operatorname{Ann}T$.*

Demostración. Si $S=T$ entonces trivialmente $\operatorname{Ann}S=\operatorname{Ann}T$. Recíprocamente, supongamos que $S\neq T$. Sin pérdida de generalidad podemos suponer entonces que existe $v\in T\setminus S$. Vamos a demostrar que existe $f\in V^*$ tal que $f(s)=0$ para todo $s\in S$ y $f(v)\neq 0$, lo que significa que $\operatorname{Ann}S\neq\operatorname{Ann}T$ pues $f\notin\operatorname{Ann}T$ y $f\in\operatorname{Ann}S$. En efecto, sea $\{v_1,\dots,v_k\}$ una base de S. Como $v\notin S$, el conjunto $\{v,v_1,\dots,v_k\}$ es linealmente independiente y entonces puede extenderse a una base $\{v,v_1,\dots,v_n\}$ de V. Sea $\{f,f_1,\dots,f_n\}$ la base dual a $\{v,v_1,\dots,v_n\}$. Luego $f(v)=1$ y $f(v_j)=0$ para todo j. En particular $f(s)=0$ para todo $s\in S$. □

4.3.10. Observemos que en el corolario 4.3.9 es necesario asumir que S y T son subespacios de V. En efecto, si $V=\mathbb{R}$, $S=\{1\}$ y $T=\mathbb{R}$ entonces $S\subsetneq T$ pero $\operatorname{Ann}S=\operatorname{Ann}T=\{0\}$.

cor:anuladores 4.3.11. Corolario. *Sean V un espacio vectorial de dimensión finita y $S\subseteq V$ un subespacio. Entonces*

$$\{v\in V: f(v)=0 \textit{ para todo } f\in\operatorname{Ann}S\}=S.$$

Demostración. Si $v\in S$ entonces trivialmente $f(v)=0$ para todo $f\in\operatorname{Ann}S$. Recíprocamente, sea $v\in V$ tal que $f(v)=0$ para todo $f\in\operatorname{Ann}S$. Sea $\{v_1,\dots,v_k\}$ una base de S. Si $v\notin S$ entonces $\{v,v_1,\dots,v_k,\}$ es linealmente independiente. Extendamos este conjunto a una base $\{v,v_1,\dots,v_n\}$ de V y sea $\{f,f_1,\dots,f_n\}$ su base dual. Por construcción, $f(v_1)=\cdots=f(v_n)=0$ y entonces $S\subseteq\ker(f)$, es decir: $f\in\operatorname{Ann}S$. Sin embargo, por construcción, $f(v)=1$, una contradicción. □

cor:base_annS 4.3.12. Corolario. *Sean V un espacio vectorial de dimensión finita, $S\subseteq V$ un subespacio, y $\{f_1,\dots,f_k\}$ una base de $\operatorname{Ann}S$. Entonces*

$$S=\bigcap_{i=1}^k\ker(f_i).$$

Demostración. Si $\{f_1,\dots,f_k\}$ es una base de $\operatorname{Ann}S$ entonces, por el corolario 4.3.11,

$$\begin{aligned}S&=\{v\in V: f(v)=0 \text{ para todo } f\in\operatorname{Ann}S\}\\&=\{v\in V: f_i(v)=0 \text{ para todo } i\in\{1,\dots,k\}\}=\bigcap_{i=1}^k\ker f_i,\end{aligned}$$

tal como queríamos demostrar. □

pro:ann(S+T) 4.3.13. Proposición. *Sean S y T dos subespacios de un espacio vectorial V. Entonces*

eq:ann(S+T) (4.3.14) $$\operatorname{Ann}(S+T)=\operatorname{Ann}S\cap\operatorname{Ann}T.$$

Demostración. Si $f\in\operatorname{Ann}S\cap\operatorname{Ann}T$ entonces $S\subset\ker f$ y $T\subseteq\ker f$ y luego $S+T\subseteq\ker f$, es decir: $f\in\operatorname{Ann}(S+T)$. Recíprocamente, si $S+T\subseteq\ker f$ entonces S y T están contenidos en $\ker f$, es decir: $f\in\operatorname{Ann}S\cap\operatorname{Ann}T$. □

4.3.15. Corolario. *Si $S_1,\dots,S_k$ son subespacios de V entonces*

eq:ann(S1+...+S_k) (4.3.16) $$\operatorname{Ann}(S_1+\cdots+S_k)=\operatorname{Ann}(S_1)\cap\cdots\cap\operatorname{Ann}(S_k).$$

DEMOSTRACIÓN. Es consecuencia de la proposición 4.3.13 y del principio de inducción. □

(ScapT)=annS+annT

4.3.17. PROPOSICIÓN. *Sea V un espacio vectorial de dimensión finita y sean S y T subespacios de V. Entonces:*

$$\operatorname{Ann}(S \cap T) = \operatorname{Ann} S + \operatorname{Ann} T.$$

DEMOSTRACIÓN. Sea $f \in \operatorname{Ann} S + \operatorname{Ann} T$ y escribamos $f = g + h$, donde $S \subseteq \ker g$ y $T \subseteq \ker h$. Si $v \in S \cap T$ entonces $v \in \ker f$ pues $f(v) = g(v) + h(v) = 0$, es decir: $f \in \operatorname{Ann}(S \cap T)$. Para ver que $\operatorname{Ann}(S \cap T) = \operatorname{Ann} S + \operatorname{Ann} T$ calculemos la dimensión de $\operatorname{Ann} S + \operatorname{Ann} T$ con la proposición 4.3.13:

$$\begin{aligned}
\dim(\operatorname{Ann} S + \operatorname{Ann} T) &= \dim \operatorname{Ann} S + \dim \operatorname{Ann} T - \dim(\operatorname{Ann} S \cap \operatorname{Ann} T) \\
&= n - \dim S + n - \dim T - \dim \operatorname{Ann}(S + T) \\
&= n - \dim S + n - \dim T - (n - \dim(S + T)) \\
&= n - \dim S - \dim T + \dim(S + T) \\
&= n - \dim(S \cap T) \\
&= \dim \operatorname{Ann}(S \cap T),
\end{aligned}$$

y entonces la proposición queda demostrada. □

4.3.18. COROLARIO. *Sea V un espacio vectorial de dimensión finita y sean $S_1, \dots, S_k$ subespacios de V. Entonces:*

$$\operatorname{Ann}(S_1 \cap \cdots \cap S_k) = \operatorname{Ann} S_1 + \cdots + \operatorname{Ann} S_k.$$

DEMOSTRACIÓN. Es consecuencia de la proposición 4.3.17 y del principio de inducción. □

cor:basis_of_V*

4.3.19. COROLARIO. *Sea V un espacio vectorial de dimensión finita n y sea $\{f_1, \dots, f_n\}$ un subconjunto de V^*. Entonces $\{f_1, \dots, f_n\}$ es base de V^* si y sólo si*

eq:capker(f_i)

$$\bigcap_{i=1}^{n} \ker f_i = \{0\}. \tag{4.3.20}$$

DEMOSTRACIÓN. Si $\{f_1, \dots, f_n\}$ es base, entonces el resultado se sigue del corolario 4.3.12 con $S = \{0\}$. Recíprocamente, si se asume (4.3.20) entonces

$$\begin{aligned}
V^* &= \operatorname{Ann}\left(\bigcap_{i=1}^{n} \ker(f_i)\right) \\
&= \sum_{i=1}^{n} \operatorname{Ann}(\ker(f_i)) = \langle f_1 \rangle + \cdots + \langle f_n \rangle = \langle f_1, \dots, f_n \rangle.
\end{aligned}$$

Como $\dim V^* = n$, entonces $\{f_1, \dots, f_n\}$ es una base de V^*. □

4.3.21. EJEMPLO. Vamos a usar el corolario 4.3.11 para calcular las ecuaciones de un subespacio. Sea $S = \langle (1,1,1), (1,2,1) \rangle \subseteq \mathbb{R}^3$. Como $\dim S = 2$ e teorema de la dimensión del anulador nos dice que $\dim \operatorname{Ann} S = 1$. Sea $f \in \operatorname{Ann} S$ y supongamos que $f(x,y,z) = \alpha x + \beta y + \gamma z$. Entonces f puede escribirse como $f(x,y,z) = \alpha(x - z)$ y luego $\operatorname{Ann} S$ está generado por la funcional $(x,y,z) \mapsto x - z$. El corolario 4.3.11 nos dice que

$$S = \{v \in V : f(v) = 0 \text{ para todo } f \in \operatorname{Ann} S\} = \{(x,y,z) \in \mathbb{R}^3 : x - z = 0\}.$$

4.4. El doble dual

4.4.1. Como V^* es un espacio vectorial es posible considerar el **doble dual** $(V^*)^* = V^{**}$. Si V es de dimensión finita entonces

$$\dim V = \dim V^* = \dim V^{**}.$$

4.4.2. Todo $v \in V$ induce una funcional lineal en V^*. En efecto, la función $L_v \colon V^* \to \mathbb{K}$ dada por $f \mapsto \langle L_v | f\rangle = f(v)$, es lineal pues

$$\langle L_v | f + \lambda g\rangle = (f+\lambda g)(v) = f(v) + (\lambda g)(v) = f(v) + \lambda g(v) = \langle L_v|f\rangle + \lambda\langle L_v|f\rangle$$

para todo $f, g \in V^*$ y $\lambda \in \mathbb{K}$.

thm:doble_dual

4.4.3. Teorema. *Sea V un espacio vectorial de dimensión finita. Entonces $L\colon V \to V^{**}$, $v \mapsto L_v$, es un isomorfismo.*

Demostración. Veamos que L es lineal: si $v, w \in V$, $\lambda \in \mathbb{K}$ y $\varphi \in V^*$ entonces

$$\langle L_{v+\lambda w} | f\rangle = f(v+\lambda w) = f(v) + \lambda f(w) = \langle L_v|f\rangle + \lambda \langle L_w|f\rangle.$$

Veamos que L es monomorfismo: si $v \in V$ tal que $L_v = 0$ entonces $L_v(\varphi) = \varphi(v) = 0$ para todo $\varphi \in V^*$. Luego $v = 0$ por el ejercicio 4.1.12.

Como L es monomorfismo y $\dim V^{**} = \dim V$ por ser V de dimensión finita, L es un isomorfismo por el corolario 3.2.17. ☐

or:representacion

4.4.4. Corolario (teorema de representación). *Sea V un espacio vectorial de dimensión finita sobre $\mathbb{K}$. Si $T \in \hom(V^*, \mathbb{K})$ entonces existe un único vector $v \in V$ tal que $T(f) = f(v)$ para todo $f \in V^*$*

Demostración. Como $T \in V^{**}$ y L es biyectiva por el teorema 4.4.3, existe un único $v \in V$ tal que $T = L_v$. ☐

4.4.5. Corolario. *Sea V un espacio vectorial de dimensión finita sobre $\mathbb{K}$. Entonces toda base de V^* es dual de una única base de V.*

Demostración. Demostremos la existencia. Sea $\{f_1, \dots, f_n\}$ una base de V^* y sea $\{T_1, \dots, T_n\} \subseteq V^{**}$ su base dual. Por el corolario anterior, 4.4.4, existen $v_1, \dots, v_n \in V$ tales que $T_i(f) = f(v_i)$ para todo $f \in V^*$. Luego, como L es un isomorfismo, $\{v_1, \dots, v_n\}$ es base de V y $\{f_1, \dots, f_n\}$ es su base dual pues

$$f_j(v_i) = \langle L_{v_i} | f_j\rangle = \langle T_i | f_j\rangle = \delta_{ij}.$$

Demostremos ahora la unicidad. Supongamos que existen bases $\{v_1, \dots, v_n\}$ y $\{v_1', \dots, v_n'\}$ de V tales que ambas tienen a $\{f_1, \dots, f_n\}$ como base dual. Para cada j escribimos

$$v_j' = \sum_{k=1}^n a_{kj} v_k$$

y entonces $\delta_{ij} = f_i(v_j') = \sum_{k=1}^n a_{kj} f_i(v_k) = a_{ij}$ para todo $i, j \in \{1, \dots, n\}$. Luego $v_j' = v_j$ para todo $j \in \{1, \dots, n\}$. ☐

xca:dual:LI

4.4.6. Ejercicio. Sea V un espacio vectorial de dimensión n y sea $\{f_1, \dots, f_n\}$ un subconjunto de V^*. Pruebe que si existe $v \in V \setminus \{0\}$ tal que $f_i(v) = 0$ para todo $i \in \{1, \dots, n\}$ entonces el conjunto $\{f_1, \dots, f_n\}$ es linealmente dependiente.

4.5. La traspuesta de una transformación lineal

4.5.1. Sean V y W dos espacios vectoriales sobre $\mathbb{K}$ y sea $f \in \hom(V,W)$. Se define la **traspuesta** de la transformación f como la función $f^T\colon W^* \to V^*$ dada por $(f^T\varphi)(v) = \varphi(f(v))$, o equivalentemente

$$\langle f^T\varphi|v\rangle = \langle \varphi|f(v)\rangle,$$

para todo $\varphi \in W^*$ y $v \in V$. Veamos que $f^T \in \hom(W^*,V^*)$: si $\varphi,\psi \in W^*$, $\lambda \in \mathbb{K}$ y $v \in V$ entonces

$$\begin{aligned}\langle f^T(\varphi+\lambda\psi)|v\rangle &= \langle \varphi+\lambda\psi|f(v)\rangle = (\varphi+\lambda\psi)\left(f(v)\right)\\ &= \varphi(f(v)) + \lambda\psi(f(v)) = \langle f^T(\varphi)|v\rangle + \langle f^T(\psi)|v\rangle.\end{aligned}$$

4.5.2. Proposición. *Sean V y W espacios vectoriales sobre $\mathbb{K}$ y supongamos que $\dim V = n$ y $\dim W = m$. Sea $f \in \hom(V,W)$. Entonces $L_W \circ f = (f^T)^T \circ L_V$, es decir: el siguiente diagrama es conmutativo:*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "L_V"'] & W \arrow[d, "L_W"] \\ V^{**} \arrow[r, "(f^T)^T"] & W^{**} \end{tikzcd}$$

Demostración. Si $v \in V$ y $\varphi \in W^*$ entonces

$$\langle (f^T)^T L_v|\varphi\rangle = \langle L_v|f^T\varphi\rangle = \langle L_v|\varphi\circ f\rangle = (\varphi\circ f)(v) = \varphi(f(v)) = \langle L_{f(v)}|\varphi\rangle,$$

tal como queríamos demostrar. $\square$

:ker(fT)=ann(imf)

4.5.3. Proposición. *Sean V y W dos espacios vectoriales. Sea $f \in \hom(V,W)$. Entonces*

$$\ker\left(f^T\right) = \operatorname{Ann}\operatorname{im} f.$$

Demostración. Tenemos:

$$\begin{aligned}\varphi \in \ker\left(f^T\right) &\Leftrightarrow \varphi\circ f = f^T\circ\varphi = 0\\ &\Leftrightarrow \varphi(f(v)) = 0 \text{ para todo } v \in V\\ &\Leftrightarrow \varphi(w) = 0 \text{ para todo } w \in \operatorname{im} f.\end{aligned}$$

Esto demuestra la proposición. $\square$

ro:dimimf=dimimfT

4.5.4. Proposición. *Sean V y W espacios vectoriales sobre $\mathbb{K}$ de dimensión finita y sea $f \in \hom(V,W)$. Entonces*

$$\dim\operatorname{im} f = \dim\operatorname{im}\left(f^T\right).$$

Demostración. Supongamos que $\dim W^* = m$. Entonces, por el teorema de la dimensión y el teorema de la dimensión del anulador:

$$m = \dim\ker\left(f^T\right) + \dim\operatorname{im}\left(f^T\right) = \dim\operatorname{Ann}\operatorname{im} f + \dim\operatorname{im} f.$$

Por la proposición 4.5.3 sabemos que $\dim\ker\left(f^T\right) = \dim\operatorname{Ann}\operatorname{im} f$. Entonces

$$\dim\operatorname{im}\left(f^T\right) = \dim\operatorname{im} f,$$

que es lo que queríamos demostrar. $\square$

pro:imf^T=annkerf

4.5.5. Proposición. *Sean V y W espacios vectoriales sobre $\mathbb{K}$ de dimensión finita y sea $f \in \hom(V,W)$. Entonces*

$$\operatorname{im}\left(f^T\right) = \operatorname{Ann}\ker f.$$

Demostración. Demostremos que $\operatorname{im}\left(f^T\right) \subseteq \operatorname{Ann}\ker f$. Sea $\varphi \in \operatorname{im}\left(f^T\right)$. Para ver que $\ker f \subseteq \ker\varphi$ escribamos $f^T\psi = \varphi$ con $\psi \in W^*$. Entonces, si $v \in \ker f$, $\varphi(v) = (\psi \circ f)(v) = 0$ y luego $\varphi \in \operatorname{Ann}\ker f$.

Por otro lado, como

$$\dim V = \dim\ker f + \dim\operatorname{Ann}\ker f = \dim\ker f + \dim\operatorname{im} f,$$

entonces $\dim\operatorname{Ann}\ker f = \dim\operatorname{im} f$. Además $\dim\operatorname{im}\left(f^T\right) = \dim\operatorname{im} f$ por la proposición 4.5.4. Luego

$$\dim\operatorname{Ann}\ker f = \dim\operatorname{im}\left(f^T\right),$$

que implica lo que queríamos demostrar. $\square$

pro:|f^T|=|f|^T

4.5.6. Proposición. *Sean V y W espacios vectoriales sobre $\mathbb{K}$ de dimensión finita y sea $f \in \hom(V,W)$. Si $\mathscr{B}_V = \{v_1,\dots,v_n\}$ es una base ordenada de V y $\mathscr{B}_V^* = \{f_1,\dots,f_n\}$ es su base dual, $\mathscr{B}_W = \{w_1,\dots,w_m\}$ es una base ordenada de W y $\mathscr{B}_W^* = \{g_1,\dots,g_m\}$ es su base dual, entonces*

$$\left\|f^T\right\|_{\mathscr{B}_W^*,\mathscr{B}_V^*} = \|f\|_{\mathscr{B}_V,\mathscr{B}_W}^T.$$

Demostración. Supongamos que

$$\|f\|_{\mathscr{B}_V,\mathscr{B}_W} = (a_{ij}) \in \mathbb{K}^{m\times n}, \qquad \left\|f^T\right\|_{\mathscr{B}_W^*,\mathscr{B}_V^*} = (b_{ij}) \in \mathbb{K}^{n\times m}.$$

Entonces $f(v_j) = \sum_{i=1}^m a_{ij}w_i$ para todo $j \in \{1,\dots,n\}$ y $\left(f^T\right)(g_j) = \sum_{i=1}^n b_{ij}f_i$ para todo $j \in \{1,\dots,m\}$. Para cada $k \in \{1,\dots,n\}$ por un lado tenemos

$$f^T(g_j)(v_k) = \sum_{i=1}^n b_{ij}f_i(v_k) = \sum_{i=1}^n b_{ij}\delta_{ik} = b_{kj}$$

y por otro lado

$$f^T(g_j)(v_k) = g_j(f(v_k)) = g_j\left(\sum_{i=1}^m a_{ik}w_i\right) = \sum_{i=1}^m a_{ik}g_j(w_i) = \sum_{i=1}^m a_{ik}\delta_{ji} = a_{jk}.$$

Luego $a_{ij} = b_{ji}$ para todo i,j. $\square$

ro:f_iso<=>f^T_iso

4.5.7. Proposición. *Sean V y W espacios vectoriales sobre $\mathbb{K}$ de dimensión finita y supongamos que* $\dim V = \dim W$. *Sea $f \in \hom(V,W)$. Entonces f es un isomorfismo si y sólo si f^T es un isomorfismo.*

Demostración. Sea $\mathscr{B}_V$ una base ordenada de V y $\mathscr{B}_V^*$ su base dual. Sea $\mathscr{B}_W$ una base ordenada de W y $\mathscr{B}_W^*$ su base dual. Entonces

$$f \text{ es isomorfismo} \Leftrightarrow \|f\|_{\mathscr{B}_V,\mathscr{B}_W} \text{ es inversible} \Leftrightarrow \|f\|_{\mathscr{B}_V,\mathscr{B}_W}^T \text{ es inversible.}$$

Como $\left\|f^T\right\|_{\mathscr{B}_W^*,\mathscr{B}_V^*} = \|f\|_{\mathscr{B}_V,\mathscr{B}_W}^T$ por la proposición 4.5.6, concluimos que f es isomorfismo si y sólo si f^T es un isomorfismo. $\square$

4.6. Una aplicación al rango de matrices

4.6.1. Vamos a demostrar que el rango fila y el rango columna son iguales.
Sea

$$f\colon \mathbb{K}^{n\times 1} \to \mathbb{K}^{m\times 1}, \quad x \mapsto Ax$$

Con respecto a las bases canónicas de $\mathbb{K}^{n\times 1}$ y $\mathbb{K}^{m\times 1}$ la matriz de f es $\|f\| = A$. En la proposición 4.5.4 vimos que $\dim \operatorname{im} f = \dim \operatorname{im} f^T$. Además $\|f\|^T = \|f^T\|$ por la proposición 4.5.6. Entonces

$$\begin{aligned} \operatorname{rg}_C(A) &= \dim \operatorname{im} f = \dim \operatorname{im} f^T \\ &= \operatorname{rg}_C\left(\left\|f^T\right\|\right) = \operatorname{rg}_C\left(\|f\|^T\right) = \operatorname{rg}_F(\|f\|) = \operatorname{rg}_F(A). \end{aligned}$$

4.7. Una aplicación a los cuadrados mágicos

ock:magic_squares 4.7.1. Vamos a utilizar el teorema de la dimensión del anulador para estudiar cuadrados mágicos. Un **cuadrado mágico** es, por definición, una matriz racional A de $n\times n$ tal que cada una de las sumas de sus filas es igual a la traza $\operatorname{tr}(A)$ de A, cada una de la suma de columnas es igual a $\operatorname{tr}(A)$, y la suma de su antidiagonal es igual a $\operatorname{tr}(A)$, es decir: A es un cuadrado mágico si y sólo si para todo k se tiene que

$$\sum_{i=1}^{n} a_{ik} = \operatorname{tr}(A), \qquad \sum_{j=1}^{n} a_{kj} = \operatorname{tr}(A), \qquad \sum_{i+j=n+1} a_{ij} = \operatorname{tr}(A).$$

Sea $M(n)$ el conjunto de cuadrados mágicos.

Por ejemplo, la matrices

eq:magic
$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 8 & 3 & 4 \\ 1 & 5 & 9 \\ 6 & 7 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{pmatrix}, \tag{4.7.2}$$

son cuadrados mágicos de 3×3.

Ejercicio. Demuestre que el conjunto $M(n)$ de cuadrados mágicos de $n\times n$ es un subespacio vectorial de $\mathbb{Q}^{n\times n}$.

Ahora que sabemos que $M(n)$ es un subespacio vectorial, calculemos su dimensión. Por el teorema de la dimensión del anulador, teorema 4.3.6, sabemos que

$$\dim M(n) + \dim \operatorname{Ann} M(n) = n^2.$$

Para cada $i,j \in \{1,\dots,n\}$ consideremos las funcionales lineales $c_i\colon \mathbb{Q}^{n\times n} \to \mathbb{Q}$, $r_i\colon \mathbb{Q}^{n\times n} \to \mathbb{Q}$ y $\operatorname{antitr}\colon \mathbb{Q}^{n\times n} \to \mathbb{Q}$ dadas por

$$c_j(A) = \sum_{k=1}^{n} a_{kj}, \quad r_i(A) = \sum_{k=1}^{n} a_{ik}, \quad \operatorname{antitr}(A) = \sum_{i+j=n+1} a_{ij}.$$

El conjunto $\{r_1 - \operatorname{tr}, \dots, r_n - \operatorname{tr}, c_1 - \operatorname{tr}, \dots, c_{n-1} - \operatorname{tr}, \operatorname{antitr} - \operatorname{tr}\}$ es un conjunto de generadores para $\operatorname{Ann} M(n)$. Puede demostrarse además que es un conjunto linealmente independiente. Luego, por el teorema de la dimensión del anulador, $\dim M(n) + 2n = n^2$ y entonces

$$\dim M(n) = n(n-2).$$

Demostremos la independencia lineal en el caso $n=3$. Sean los escalares $\alpha_1,\alpha_2,\alpha_3\,\beta_1,\beta_2 \in \mathbb{Q}$ tales que

$$\sum_{i=1}^{3} \alpha_i(r_i - \operatorname{tr}) + \sum_{j=1}^{2} \beta_j(c_j - \operatorname{tr}) + \gamma(\operatorname{antitr} - \operatorname{tr}) = 0.$$

Veamos que $\alpha_1 = \alpha_2 = \alpha_3 = \beta_1 = \beta_2 = \gamma = 0$. Al evaluar esta expresión en la matriz $E_{13} + E_{22} + E_{31}$ se obtiene que $\gamma = 0$. Después, al evaluar en la matrices canónicas E_{13} y E_{23} se obtiene $\alpha_1 = \alpha_2 = 0$. Evaluar en E_{21} y E_{12} nos da $\beta_1 = \beta_2 = 0$. Por último, al evaluar en E_{31} obtenemos que $\alpha_3 = 0$.

4.7.3. Ejercicio. Utilice 4.7.1 para demostrar que la dimensión del espacio de cuadrados mágicos de $n\times n$ es $n(n-2)$.

Capítulo 5

Determinantes

5.1. Permutaciones

5.1.1. Sea $n \in \mathbb{N}$. Denotaremos por $\mathbb{S}_n$ al conjunto de **permutaciones** del conjunto $\{1, \dots, n\}$, es decir:

$$\mathbb{S}_n = \{\sigma \colon \{1, \dots, n\} \to \{1, \dots, n\} : \sigma \text{ es biyectiva}\}.$$

Es fácil demostrar que el conjunto $\mathbb{S}_n$ tiene $n! = n(n-1)\cdots 2$ elementos. Denotaremos a una permutación $\sigma \in \mathbb{S}_n$ de la siguiente forma:

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

Por ejemplo, las permutaciones del conjunto $\{1, 2\}$ son:

$$\mathbb{S}_2 = \left\{ \begin{pmatrix} 12 \\ 12 \end{pmatrix}, \begin{pmatrix} 12 \\ 21 \end{pmatrix} \right\}$$

Las permutaciones del conjunto $\{1, 2, 3\}$ son:

$$\mathbb{S}_3 = \left\{ \begin{pmatrix} 123 \\ 123 \end{pmatrix}, \begin{pmatrix} 123 \\ 213 \end{pmatrix}, \begin{pmatrix} 123 \\ 321 \end{pmatrix}, \begin{pmatrix} 123 \\ 132 \end{pmatrix}, \begin{pmatrix} 123 \\ 231 \end{pmatrix}, \begin{pmatrix} 123 \\ 312 \end{pmatrix} \right\}.$$

5.1.2. Las permutaciones son funciones biyectivas, y la composición de funciones biyectivas es una función biyectiva. Luego la composición de dos permutaciones es una permutación. Por ejemplo:

$$\sigma = \begin{pmatrix} 123 \\ 132 \end{pmatrix}, \quad \tau = \begin{pmatrix} 123 \\ 312 \end{pmatrix}, \quad \sigma\tau = \begin{pmatrix} 123 \\ 213 \end{pmatrix}.$$

Como sabemos, la composición de funciones es asociativa. La identidad id es el neutro para la composición, y cada permutación, por ser una función biyectiva, es inversible.

5.1.3. Una permutación $\sigma \in \mathbb{S}_n$ es un r-**ciclo** si existen $a_1, \dots, a_r \in \{1, \dots, n\}$ todos distintos tales que $\sigma(j) = j$ para todo $j \notin \{a_1, \dots, a_r\}$ y además

$$\sigma(a_i) = \begin{cases} a_{i+1} & \text{si } i < r, \\ a_1 & \text{si } i = r. \end{cases}$$

Si $a_1, \dots, a_r \in \{1, \dots, n\}$ son todos distintos, denotaremos por $(a_1\, a_2\, \dots, a_r)$ al ciclo σ que mueve únicamente a los a_i y tal que

$$\sigma(a_1) = a_2, \qquad \sigma(a_2) = a_3, \qquad \cdots \qquad \sigma(a_{n-1}) = a_n, \qquad \sigma(a_n) = a_1.$$

5.1.4. Ejemplo. En general omitiremos los 1-ciclos al escribir una permutación como producto de ciclos disjuntos. Escribamos algunas permutaciones de $\mathbb{S}_3$:

$$\begin{pmatrix} 123 \\ 213 \end{pmatrix} = (12)(3) = (12), \qquad \begin{pmatrix} 123 \\ 321 \end{pmatrix} = (13)(2) = (13), \qquad \begin{pmatrix} 123 \\ 231 \end{pmatrix} = (123).$$

Veamos como ejemplo algunas permutaciones de $\mathbb{S}_4$:

$$\begin{pmatrix} 1234 \\ 2314 \end{pmatrix} = (123)(4) = (123). \qquad \begin{pmatrix} 1234 \\ 2143 \end{pmatrix} = (12)(34).$$

5.1.5. Las permutaciones σ y τ son **disjuntas** si para cada $j \in \{1, \dots, n\}$ se tiene que $\sigma(j) = j$ o bien $\tau(j) = j$.

isjuntas_conmutan

5.1.6. Ejercicio. Pruebe que las permutaciones disjuntas conmutan.

5.1.7. Ejemplo. Las permutaciones $(123) \in \mathbb{S}_6$ y $(56) \in \mathbb{S}_6$ son disjuntas. Las permutaciones $(123) \in \mathbb{S}_6$ y $(25) \in \mathbb{S}_6$ no son disjuntas.

xca:permutaciones

5.1.8. Ejercicio. Sea $\sigma = \alpha\beta \in \mathbb{S}_n$, donde α y β son disjuntas. Si $\alpha(i) \neq i$ entonces $\sigma^k(i) = \alpha^k(i)$ para todo $k \geq 0$.

5.1.9. Proposición. *Sea $\sigma \neq \mathrm{id}$ una permutacion de $\mathbb{S}_n$. Entonces σ es producto de ciclos disjuntos de longitud ≥ 2. La descomposición es única salvo el orden de los factores.*

Demostración. Hacemos inducción en la cantidad k de elementos del conjunto $\{1, \dots, n\}$ que mueve σ. El caso base es $k = 2$, que es trivial pues σ es una trasposición. Supongamos que el resultado es válido para permutaciones que mueven menos de k elementos, donde $k > 0$. Sea $i_1 \in \{1, \dots, n\}$ tal que $\sigma(i_1) \neq i_1$. Definimos $i_2 = \sigma(i_1)$, $i_3 = \sigma(i_2)$... y sabemos que existe $r \in \mathbb{N}$ tal que $\sigma(i_r) = i_1$ (si existe $j \in \{2, \dots, n\}$ tal que $\sigma(i_r) = i_j$ entonces $\sigma(i_{j-1}) = i_j$, una contradicción). Tomemos $\sigma_1 = (i_1 i_2 \cdots i_r)$. Por hipótesis inductiva, como $\sigma_1^{-1}\sigma$ es una permutación que mueve menos de k elementos, $\sigma_1^{-1}\sigma = \sigma_2 \cdots \sigma_s$, donde $\sigma_2, \dots, \sigma_s$ son ciclos disjuntos.

Probemos ahora la unicidad. Para eso, supongamos que $\sigma = \sigma_1 \dots \sigma_s = \tau_1 \cdots \tau_t$ con $s > 0$. Sea $i_1 \in \{1, \dots, n\}$ tal que $\sigma_1(i_1) \neq i_1$. Por lo que hicimos en el ejercicio 5.1.8, $\sigma^k(i_1) = \sigma_1^k(i_1)$ para todo $k \geq 0$. Existe $j \in \{1, \dots, t\}$ tal que τ_j mueve a i_1. Sin pérdida de generalidad (pues los τ_k conmutan) podemos suponer que $j = 1$, y luego $\sigma_1^k(i_1) = \tau_1^k(i_1)$ para todo $k \geq 0$. De acá se deduce que $\sigma_1 = \tau_1$ y por lo tanto $\sigma_2 \cdots \sigma_s = \tau_2 \cdots \tau_t$. Por inducción en $\max\{s, t\}$ obtenemos que $s = t$ y por lo tanto $\sigma_k = \tau_k$ para todo k. $\square$

or:trasposiciones

5.1.10. Corolario. *Toda permutación es producto de trasposiciones.*

Demostración. Como toda permutación $\neq \mathrm{id}$ es producto de ciclos disjuntos, alcanza con demostrar que cada ciclo es producto de trasposiciones. Luego

$$(a_1 \dots a_r) = (a_1 a_r)(a_1 a_{r-1}) \cdots (a_1 a_3)(a_1 a_2)$$

demuestra el corolario. $\square$

rasposiciones_(1k)

5.1.11. Corolario. *Toda permutación es producto de trasposiciones de la forma $(1k)$.*

Demostración. Por el corolario 5.1.10 sabemos que toda permutación es producto de trasposiciones. Como cada trasposición puede escribirse como

$$(ij) = (1i)(1j)(1i),$$

entonces toda permutación puede escribirse como producto de trasposiciones de la forma $(1k)$. $\square$

rasposiciones_ady

5.1.12. Corolario. *Toda permutación es producto de* ***trasposiciones adyacentes****, es decir: trasposiciones de la forma $(k\,k+1)$.*

Demostración. Por el corolario 5.1.11 sabemos que toda permutación es producto de trasposiciones de la forma $(1k)$. Probemos por inducción que toda permutación de la forma $(1k)$ puede escribirse como producto de trasposiciones adyacentes. El primer caso es $(13) = (12)(23)(12)$. Si suponemos que la afirmación es válida para la permutación $(1k)$ entonces el corolario queda demostrado al utilizar la hipótesis inductiva en la descomposición $(1\,k+1) = (k\,k+1)(1k)(k\,k+1)$. $\square$

5.1.13. El corolario 5.1.10 nos dice que toda $\sigma \in \mathbb{S}_n$ puede escribirse como producto de trasposiciones. Esta escritura no es única, por ejemplo:

$$(123) = (13)(12) = (23)(12)(23)(12).$$

Sin embargo, puede probarse que si $\sigma = \sigma_1 \cdots \sigma_k = \tau_1 \cdots \tau_l$ como producto de trasposiciones entonces k y l tienen la misma paridad. Probaremos esta afirmación como aplicación de la teoría de determinantes.

5.2. Funciones multilineales alternadas

5.2.1. Sea $n \in \mathbb{N}$. Una función $d\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ es **n-lineal** (por filas) si para cada $i \in \{1,\dots,n\}$ la función d es lineal en la i-ésima fila cuando las otras $n-1$ se dejan fijas, es decir:

$$\begin{aligned}&d(A_1,\dots,A_{i-1},x+\lambda y,A_{i+1},\dots,A_n)\\&\qquad= d(A_1,\dots,A_{i-1},x,A_{i+1},\dots,A_n)+\lambda d(A_1,\dots,A_{i-1},y,A_{i+1},\dots,A_n).\end{aligned}$$

ombinacion_nlineal 5.2.2. Ejercicio. Toda combinación lineal de funciones n-lineales es n-lineal.

5.2.3. Una función n-lineal d es **alternada** si $d(A) = 0$ para toda matriz A que tiene dos filas iguales. Una función d es un **determinante** si es n-lineal, alternada y $d(I) = 1$.

xca:2x2 5.2.4. Ejemplos. La función $d\colon \mathbb{K} \to \mathbb{K}$ dada por $d(a) = a$ es una función determinante. Similarmente la función $d\colon \mathbb{K}^{2\times 2} \to \mathbb{K}$ definida por

$$d\begin{pmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

es una función determinante.

5.2.5. Ejemplo. Supongamos que $d\colon \mathbb{K}^{2\times 2} \to \mathbb{K}$ es una función determinante. Si $e_1 = (1,0)$ y $e_2 = (0,1)$ entonces la primera fila de la matriz $A = \begin{pmatrix} a & b\\ c & d\end{pmatrix}$ puede escribirse como $ae_1 + be_2$ y la segunda fila de A puede escribirse como $ce_1 + de_2$. Si usamos la 2-linealidad de d,

$$\begin{aligned}d(A) &= d(ae_1+be_2, ce_1+de_2)\\ &= acd(e_1,e_1)+add(e_1,e_2)+bcd(e_2,e_1)+bdd(e_2,e_2).\end{aligned}$$

Como d es alternada,

$$d(e_1,e_1) = d(e_2,e_2) = 0, \quad d(e_2,e_1) = -d(e_1,e_2).$$

Entonces

$$d(A) = (ad-bc)d(e_1,e_2) = (ad-bc)d(I).$$

Como d es una función determinante, $d(I) = 1$ y luego $d(A) = ad - bc$.

lem:alternada 5.2.6. Lema. *Si d es alternada entonces*

$$d(A_1,\dots,A_i,\dots,A_j,\dots,A_n) = -d(A_1,\dots,A_j,\dots,A_i,\dots,A_n).$$

Demostración. Como d es alternada,

$$d(A_1,\dots,A_{i-1},A_i+A_j,A_{i+1},\dots,A_{j-1},A_i+A_j,A_{j+1},\dots,A_n) = 0.$$

Al usar la n-linealidad obtenemos

$$\begin{aligned}0 &= d(A_1,\dots,A_i,\dots,A_i,\dots,A_n)+d(A_1,\dots,A_i,\dots,A_j,\dots,A_n)\\&\qquad+d(A_1,\dots,A_j,\dots,A_i,\dots,A_n)+d(A_1,\dots,A_j,\dots,A_j,\dots,A_n).\end{aligned}$$

Como d es alternada,

$$d(A_1,\dots,A_i,\dots,A_i,\dots,A_n) = d(A_1,\dots,A_j,\dots,A_j,\dots,A_n) = 0,$$

y entonces el lema queda demostrado. □

5.2.7. Sea $A \in \mathbb{K}^{n\times n}$ y sean $i,j \in \{1,\dots,n\}$. Si d es una función n-lineal y $A \in \mathbb{K}^{n\times n}$ entonces

$$d_{ij}(A) = d(A[i|j]),$$

donde $A[i|j]$ es la matriz de tamaño $(n-1)\times(n-1)$ que se obtiene al eliminar la i-ésima fila y la j-ésima columna de A. Por ejemplo:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \qquad A[1|1] = \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}, \qquad A[1|2] = \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix}, \qquad A[3|2] = \begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix}.$$

m:desarrollo_cols 5.2.8. Teorema. *Si d es una función $(n-1)$-lineal, alternada y $d(I) = 1$ entonces para cada $j \in \{1,\dots,n\}$ la función*

$$E_j(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} d_{ij}(A)$$

es n-lineal, alternada y cumple que $E_j(I) = 1$.

Demostración. Para demostrar que E_j es n-lineal basta observar que la función $A \mapsto a_{ij}d_{ij}(A)$ es n-lineal y utilizar el ejercicio 5.2.2 que afirma que toda combinaciones lineal de funciones n-lineales es n-lineal.

Demostremos que E_j es alternada. Para eso, supongamos que A tiene dos filas iguales, digamos que son las filas k y l, donde $k < l$. Si $i \notin \{k,l\}$ entonces $d_{ij}(A) = 0$ pues la matriz $A[i|j]$ tiene dos filas iguales. Entonces

$$E_j(A) = (-1)^{k+j} a_{kj} d_{kj}(A) + (-1)^{l+j} a_{lj} d_{lj}(A).$$

La $(l-1)$-ésima fila de $A[k|j]$ y la k-ésima fila de $A[l|j]$ son iguales a A_k y entonces $A[k|j]$ y $A[l|j]$ difieren en $l-k-1$ trasposiciones. Luego

$$d_{kj}(A) = (-1)^{l-k-1} d_{lj}(A).$$

Como $a_{kj} = a_{lj}$,

$$E_j(A) = (-1)^{k+j+l-k-1} a_{lj} d_{lj}(A) + (-1)^{l+j} a_{lj} d_{lj}(A) = 0.$$

Para ver que $E_j(I) = 1$ basta observar que

$$E_j(I) = \sum_{i=1}^{n} (-1)^{i+j} \delta_{ij} d_{ij}(I) = d_{jj}(I) = 1.$$

Esto completa la demostración. □

inante:existencia 5.2.9. Corolario. *Existe al menos una función determinante.*

Demostración. Vimos en el ejemplo 5.2.4 que existen determinantes cuando $n \in \{1,2\}$. El caso general es consecuencia directa del teorema anterior y el principio de inducción. □

5.3. Aplicación a la teoría de permutaciones

5.3.1. Sea $\sigma \in \mathbb{S}_n$. El corolario 5.1.10 nos dice que σ puede escribirse como producto de traposiciones. Esta escritura no es única, por ejemplo:

$$(123) = (13)(12) = (23)(12)(23)(12).$$

Sin embargo, puede probarse que si $\sigma = \sigma_1 \cdots \sigma_k = \tau_1 \cdots \tau_l$ como producto de trasposiciones entonces k y l tienen la misma paridad.

cor:signo 5.3.2. Corolario. *Sea $\sigma \in \mathbb{S}_n$ tal que $\sigma = \sigma_1 \cdots \sigma_k = \tau_1 \cdots \tau_l$ como producto de trasposiciones. Entonces $(-1)^k = (-1)^l$.*

Demostración. Por el corolario 5.2.9 sabemos que existe una función determinante $d\colon \mathbb{K}^{n\times n} \to \mathbb{K}$. Si $\{e_1, e_2, \ldots, e_n\}$ es la base canónica de $\mathbb{K}^{n\times 1}$ entonces $d(e_1, \ldots, e_n) = d(I) = 1$ pues d es una función determinante. Como, por hipótesis, $\sigma = \sigma_1 \cdots \sigma_k = \tau_1 \cdots \tau_l$ y d es una función alternada, entonces

$$\begin{aligned}(-1)^k &= (-1)^k d(e_1, \ldots, e_n)\\ &= d(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}) = (-1)^l d(e_1, \ldots, e_n) = (-1)^l,\end{aligned}$$

que es lo que queríamos demostrar. □

5.3.3. El corolario 5.3.2 nos permite definir el signo de una permutación. Sea $\sigma \in \mathbb{S}_n$. Se define el **signo** de σ como el número $\operatorname{sgn}(\sigma) = (-1)^k$ si $\sigma = \sigma_1 \cdots \sigma_k$ como producto de trasposiciones. Una permutación σ es **par** si $\operatorname{sgn}(\sigma) = 1$ y es **impar** si $\operatorname{sgn}(\sigma) = -1$.

5.3.4. Ejemplos. La identidad es una permutación par. Las trasposiciones son permutaciones impares. Un r-ciclo tiene signo $(-1)^{r-1}$. Otros ejemplos:

$$\operatorname{sgn}((123)) = \operatorname{sgn}((12)(34)) = 1, \qquad \operatorname{sgn}((1234)) = \operatorname{sgn}((123)(45)) = -1.$$

r:sgn_es_morfismo

5.3.5. Corolario. *Sean $\sigma, \tau \in \mathbb{S}_n$. Entonces* $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$.

Demostración. Escribamos $\sigma = \sigma_1 \cdots \sigma_k$ y $\tau = \tau_1 \cdots \tau_l$, ambas como producto de trasposiciones. Entonces, como $\sigma\tau$ puede escribirse como producto de $k+l$ trasposiciones,

$$(\operatorname{sgn}\sigma)(\operatorname{sgn}\tau) = (-1)^k(-1)^l = (-1)^{k+l} = \operatorname{sgn}(\sigma\tau),$$

tal como queríamos demostrar. □

5.3.6. Corolario. *Sea $\sigma \in \mathbb{S}_n$. Entonces* $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.

Demostración. Si escribimos a σ como producto de trasposiciones, digamos $\sigma = \sigma_1 \cdots \sigma_k$, entonces, como $\sigma_j^2 = \mathrm{id}$ para todo j, $\sigma^{-1} = \sigma_k \cdots \sigma_1$. Luego $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$. □

5.3.7. Corolario. *Supongamos que $\sigma \in \mathbb{S}_n$ se descompone como producto de m ciclos disjuntos de longitud $l_1, \ldots, l_m$. Entonces*

$$\operatorname{sgn}(\sigma) = (-1)^{\sum_{k=1}^m (l_k - 1)}.$$

Demostración. Supongamos que σ se descompone como producto de ciclos disjuntos $\sigma = c_1 \ldots c_m$ y que cada ciclo c_k tiene longitud l_k. Como cada c_k puede escribirse como producto de $l_k - 1$ trasposiciones, $\operatorname{sgn}(c_k) = (-1)^{l_k - 1}$. Por el corolario 5.3.5,

$$\operatorname{sgn}(\sigma) = \prod_{k=1}^m \operatorname{sgn}(c_k) = \prod_{k=1}^m (-1)^{l_k-1} = (-1)^{\sum_{k=1}^m (l_k-1)},$$

tal como queríamos demostrar. □

5.4. Unicidad y propiedades básicas

lem:sigma

5.4.1. Lema. *Sean $d\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ una función alternada y $\sigma \in \mathbb{S}_n$. Entonces, para todo $A \in \mathbb{K}^{n\times n}$,*

$$d(A_{\sigma(1)}, \ldots, A_{\sigma(n)}) = (\operatorname{sgn}\sigma) d(A).$$

Demostración. Supongamos que σ se escribe como producto de k trasposiciones. Entonces

$$d(A_{\sigma(1)}, \ldots, A_{\sigma(n)}) = (-1)^k \det A = (\operatorname{sgn}\sigma) d(A),$$

tal como queríamos demostrar. □

erminante:unicidad 5.4.2. Teorema. *Existe una única función determinante* $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$. *Más aún,*

det:permutaciones (5.4.3)
$$\det(A) = \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)}$$

para todo $A \in \mathbb{K}^{n\times n}$.

Demostración. Supongamos que d es una función determinante y sea $A \in \mathbb{K}^{n\times n}$. Escribamos a cada fila de A como $A_i = \sum_{j=1}^n a_{ij}e_j$ donde $\{e_1,\dots,e_n\}$ es la base canónica de $\mathbb{K}^{1\times n}$. Entonces

$$d(A) = d(A_1,\dots,A_n) = \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n a_{1j_1}\cdots a_{nj_n} d(e_{j_1},\dots,e_{j_n}).$$

Como d es alternada, la suma es no nula únicamente cuando todos los j_k son distintos, es decir cuando $|\{j_1,\dots,j_n\}| = n$. Entonces, la suma anterior se hace sobre todas las n-uplas $(j_1,\dots,j_n)$ de elementos distintos del conjunto $\{1,\dots,n\}$. Luego, por el lema 5.4.1,

$$\begin{aligned} d(A) &= \sum_{\sigma\in\mathbb{S}_n} a_{1\sigma(1)}\cdots a_{n\sigma(n)} d(e_{\sigma(i)},\dots,e_{\sigma(n)}) \\ &= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)} d(e_1,\dots,e_n). \end{aligned}$$

Como d es un determinante, $d(e_1,\dots,e_n) = d(I) = 1$ y entonces

$$d(A) = \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)}.$$

En particular, existe una única función determinante. □

xca:dA=(detA)dI 5.4.4. Ejercicio. Utilice lo hecho en la demostración del teorema 5.4.2 y pruebe que si d es n-lineal y alternada entonces

$$d(A) = (\det A)d(I)$$

para todo $A \in \mathbb{K}^{n\times n}$.

5.4.5. Corolario (Desarrollo por columnas). *Sea* $A \in \mathbb{K}^{n\times n}$. *Entonces, para cada* $j \in \{1,\dots,n\}$,

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} M_{ij}(A),$$

donde $M_{ij}(A) = \det A[i|j]$.

Demostración. Es consecuencia directa del teorema 5.4.2. □

5.4.6. Ejemplo. Los números 917, 854 y 994 son divisibles por 7. Vamos a utilizar esta información para demostrar que

$$\det\begin{pmatrix} 9 & 1 & 7 \\ 8 & 5 & 4 \\ 9 & 9 & 4 \end{pmatrix}$$

es divisible por 7. Observemos que al efectuar la operación de columnas $100C_1 + 10C_2 + C_3 \to C_3$ se tiene que

$$\det\begin{pmatrix} 9 & 1 & 7 \\ 8 & 5 & 4 \\ 9 & 9 & 4 \end{pmatrix} = \det\begin{pmatrix} 9 & 1 & 917 \\ 8 & 5 & 854 \\ 9 & 9 & 994 \end{pmatrix}.$$

Si desarrollamos este último determinante por la última columna, el valor del determinante es divisible por siete.

(AB)=(detA)(detB) 5.4.7. Teorema. *Sean* $A, B \in \mathbb{K}^{n\times n}$. *Entonces*

$$\det(AB) = (\det A)(\det B).$$

Demostración. Sea $d\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ dada por

$$d(X) = \det(XB)$$

Si las filas de X son los vectores $X_1, \dots, X_n$ entonces las filas de la matriz XB son $X_1B, \dots, X_nB$. Demostremos que d es alternada: si $X_i = X_j$ entonces $X_iB = X_jB$ y usamos que det es alternada. Demostremos que d es n-lineal: si la fila i-ésima de X es $X_i + \lambda Y_i$ entonces la i-ésima fila de XB es $(X_i + \lambda Y_i)B = X_iB + \lambda Y_iB$ y usamos que det es n-lineal. Por el ejercicio 5.4.4,

$$\det(AB) = d(A) = (\det A)d(I) = (\det A)(\det B),$$

tal como queríamos demostrar. $\square$

rminante:semejanza 5.4.8. El teorema anterior implica que si $A, B \in \mathbb{K}^{n\times n}$ son semejantes entonces, como $B = CAC^{-1}$ para alguna matriz inversible $C \in \mathbb{K}^{n\times n}$, se tiene que

$$\begin{aligned} \det B &= \det(CAC^{-1}) = (\det C)(\det A)(\det C^{-1}) \\ &= (\det C)(\det A)(\det C)^{-1} = \det A. \end{aligned}$$

5.4.9. Sean V un espacio vectorial de dimensión finita y $f \in \hom(V, V)$. Como las matrices de f con respecto a bases distintas son semejantes, la observación hecha en 5.4.8 implica que tiene sentido definir el **determinante** de f como

$$\det f = \det[f]_{\mathscr{B},\mathscr{B}},$$

donde $\mathscr{B}$ es alguna base de V.

exa:det:fibonacci 5.4.10. Ejemplo. La **sucesión de Fibonacci** F_n es la sucesión

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, \dots$$

definida recursivamente por $F_0 = F_1 = 1$ y $F_{n+1} = F_n + F_{n-1}$ para $n \geq 1$. Por inducción puede demostrarse que

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n = \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix}$$

para todo $n \geq 1$. Si aplicamos la función determinante a la fórmula anterior y usamos que el determinante es una función multiplicativa, obtenemos

$$F_{n+1}F_{n-1} - F_n^2 = (-1)^n$$

para todo $n \geq 1$.

5.4.11. Proposición. *Si $A = (a_{ij})$ es triangular superior (es decir $a_{ij} = 0$ si $i > j$) entonces* $\det A = a_{11} \cdots a_{nn}$.

Demostración. Procedemos por inducción en n. Los casos $n \in \{1, 2\}$ son triviales. Supongamos entonces que el resultado es válido para matrices de tamaño $(n-1)\times(n-1)$. Entonces, al desarrollar el determinante por la primera columna y utilizar la hipótesis inductiva en la matriz $A[1|1]$,

$$\det A = a_{11} \det A[1|1] = a_{11}a_{22} \cdots a_{nn},$$

que es lo que queríamos demostrar. $\square$

pro:detA=det(AT) 5.4.12. Proposición. *Si $A \in \mathbb{K}^{n\times n}$ entonces* $\det A = \det A^T$.

Demostración. Si $\sigma \in \mathbb{S}_n$ y $\sigma(i) = j$ entonces $\sigma^{-1}(j) = i$ y además $a_{\sigma(i)i} = a_{j\sigma^{-1}(j)}$. Entonces

$$\begin{aligned} \det A &= \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n} \\ &= \sum_{\sigma^{-1} \in \mathbb{S}_n} \operatorname{sgn}(\sigma^{-1}) a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n} \\ &= \det A^T, \end{aligned}$$

pues $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ y si σ recorre todo $\mathbb{S}_n$ entonces σ^{-1} también recorre todo $\mathbb{S}_n$. □

5.4.13. Ejercicio. Demuestre que si $A = (a_{ij}) \in \mathbb{K}^{n\times n}$ es triangular inferior entonces $\det A = a_{11} \cdots a_{nn}$.

5.4.14. En vista de la forma en que construimos la función determinante y de la proposición 5.4.12 se tiene la siguiente propiedad: si B se obtiene de A al sumar en una fila un múltiplo de otra fila (o al sumar a una columna un múltiplo de otra columna) entonces $\det B = \det A$.

xca:bloques_2x2 5.4.15. Ejercicio. Sean $A \in \mathbb{K}^{r\times r}$, $B \in \mathbb{K}^{r\times s}$, $C \in \mathbb{K}^{s\times s}$. Pruebe que

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = (\det A)(\det C).$$

5.4.16. Corolario (Desarrollo por filas). *Sea $A \in \mathbb{K}^{n\times n}$. Entonces, para cada $i \in \{1, \dots, n\}$,*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}(A),$$

donde $M_{ij}(A) = \det A[i|j]$.

Demostración. Es consecuencia directa de la proposición 5.4.12, donde vimos que $\det A = \det A^T$, y el teorema 5.4.2. □

5.4.17. Ejemplo. Como ejemplo, vamos a demostrar que

$$\det \begin{pmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{pmatrix} = (y-x)(z-x)(z-y).$$

Dado que el determinante en matrices de 3×3 es una función 3-lineal, efectuar las operaciones elementales de filas $F_2 - F_1 \to F_2$ y $F_3 - F_1 \to F_3$ no cambia el valor del determinante. Esto implica que

$$\det \begin{pmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{pmatrix} = \det \begin{pmatrix} 1 & x & x^2 \\ 0 & y-x & y^2-x^2 \\ 0 & z-x & z^2-x^2 \end{pmatrix} = (y-x)(z-x)(z-y),$$

después de desarrollar este último determinante por la primera columna.

5.4.18. Dada una matriz $A \in \mathbb{K}^{n\times n}$ se define la matriz de los **cofactores** de A como

$$C_{ij} = (-1)^{i+j} \det A[i|j]$$

para todo i, j. Se define además la **adjunta** $\operatorname{adj} A$ de A como la traspuesta de la matriz de cofactores de A, es decir:

$$(\operatorname{adj} A)_{ij} = (-1)^{i+j} \det A[j|i] = C_{ji}$$

para todo i, j.

5.4.19. Ejemplo. Observemos que

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{pmatrix}, \quad \operatorname{adj} A = \begin{pmatrix} 0 & 0 & 3 \\ 2 & -1 & -2 \\ -1 & 2 & 1 \end{pmatrix}.$$

5.4.20. Ejemplo.

thm:AadjA=(detA)I

5.4.21. Teorema. *Sea* $A \in \mathbb{K}^{n\times n}$. *Entonces*

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

Demostración. FIXME Si $i \neq j$ y B es la matriz A después de haber reemplazado su i-ésima columna por su j-ésima columna, entonces $A[k|i] = B[k|i]$ para todo k. Luego, como B tiene dos columnas iguales, al desarollar el determinante por la columna i-ésima,

$$0 = \det B = \sum_{k=1}^{n} (-1)^{k+i} b_{ki} \det B[k|i] = \sum_{k=1}^{n} (-1)^{k+i} a_{kj} \det A[k|i].$$

Por otro lado, para cada i,j tenemos que

$$((\operatorname{adj} A)A)_{ij} = \sum_{k=1}^{n} (\operatorname{adj} A)_{ik} a_{kj} = \sum_{k=1}^{n} (-1)^{k+i} a_{kj} \det A[k|i].$$

Observemos que si $i \neq j$ entonces $((\operatorname{adj} A)A)_{ij}$ es el determinante de la matriz B, que tiene dos columnas iguales. En cambio, si $i = j$, entonces $(A \operatorname{adj} A) = \det A$ por la fórmula 5.4.3. Luego

$$(A \operatorname{adj} A)_{ij} = \begin{cases} \det A & \text{si } i = j, \\ 0 & \text{si } i \neq j, \end{cases}$$

tal como queríamos demostrar. $\square$

5.4.22. Ejemplo. Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{K}^{2\times 2}$ entonces $\operatorname{adj} A = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Entonces, si $\det A = ad - bc \neq 0$,

$$A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

xca:adjadjA

5.4.23. Ejercicio. Sean $A \in \mathbb{K}^{n\times n}$ matrices inversibles. Demuestre que $\operatorname{adj}(BAB^{-1}) = B \operatorname{adj}(A) B^{-1}$.

xca:adj(BAB^(-1))

5.4.24. Ejercicio. Sean $A, B \in \mathbb{K}^{n\times n}$. Demuestre que si A es inversible entonces $\operatorname{adj}(\operatorname{adj} A) = (\det A)^{n-2} A$.

5.4.25. Ejercicio. Sea $A \in \mathbb{R}^{3\times 3}$ una matriz de rango 1. Calcule el rango de la adjunta de A.

versible<=>detA=0

5.4.26. Corolario. *Sea* $A \in \mathbb{K}^{n\times n}$. *Entonces* A *es inversible si y sólo si* $\det A \neq 0$.

Demostración. Si A es inversible entonces existe $B \in \mathbb{K}^{n\times n}$ tal que $AB = I$ y luego $(\det A)(\det B) = \det(AB) = \det(I) = 1$, que implica que $\det A \neq 0$. Recíprocamente, si $\det A \neq 0$ entonces $A^{-1} = (\det A)^{-1} \operatorname{adj} A$ por el teorema 5.4.21. $\square$

block:Cramer

5.4.27. Ejercicio (Regla de Cramer). Supongamos que $A \in \mathbb{K}^{n\times n}$ es inversible y sea $b \in \mathbb{K}^{n\times 1}$. Utilice el teorema 5.4.21. para demostrar que la solución $x = (x_1, \dots, x_n)^T$ del sistema lineal $Ax = b$ puede calcularse de la siguiente forma:

$$x_i = \frac{\det B_i}{\det A},$$

donde B_i es la matriz que se obtiene de A después de reemplazar su i-ésima columna por el vector columna b.

5.4.28. Vamos a dar una demostración elemental de la regla de Cramer vista en 5.4.27. Sea $\{e_1,\dots,e_n\}$ la base canónica de $\mathbb{K}^{n\times 1}$ y para cada $i\in\{1,\dots,n\}$ sea X_i la que se obtiene después de reemplazar la i-ésima columna de la matriz identidad por el vector columna $x=(x_1\dots x_n)^T$. Por ejemplo:

$$X_1=\begin{pmatrix} x_1 & 0 & 0 & \cdots & 0\\ x_2 & 1 & 0 & \cdots & 0\\ x_3 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ x_n & 0 & 0 & \cdots & 1\end{pmatrix}$$

Entonces

$$AX_1=(Ax,Ae_2,\dots,Ae_n)=(b,A_2,\dots,A_n).$$

donde A_i es la i-ésima columna de A. Luego

$$\begin{aligned}(\det A)x_1&=(\det A)(\det X_1)=\det(AX_1)\\&=\det(b,A_2,\dots,A_n).\end{aligned}$$

De forma similar puede calcularse el valor de cada x_i.

a:rango_submatriz 5.4.29. Ejercicio. Sea $A\in\mathbb{K}^{n\times n}$. Pruebe que $\operatorname{rg}A\geq r$ si y sólo si existe una submatriz de A inversible y de tamaño $r\times r$. Concluya que $\operatorname{rg}A=r$ si y sólo si existe una submatriz de A inversible y de tamaño $r\times r$ y toda submatriz de A de tamaño $(r+1)\times(r+1)$ es no inversible.

5.5. Algunos determinantes especiales

5.5.1 (Matriz de Vandermonde). Sean $x_1,\dots,x_n\in\mathbb{K}$. Vamos a demostrar que si

$$V(x_1,\dots,x_n)=\det\begin{pmatrix} 1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_n\\ \vdots & \vdots & \ddots & \vdots\\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1}\end{pmatrix}$$

entonces

$$V(x_1,\dots,x_n)=\prod_{1\leq i<j\leq n}(x_j-x_i).$$

Sin pérdida de generalidad podemos suponer que todos los x_j son distintos ya que si $x_i=x_j$ para $i\neq j$ entonces el determinate es cero y la fórmula es válida. Podemos suponer entonces que todos los x_j son distintos. Procederemos por inducción en n. El caso $n=2$ es sencillo y se deja como ejercicio. Supongamos que el resultado es válido para algún $n\geq 2$. Sea

$$f=V(x_1,\dots,x_n,X)=\det\begin{pmatrix} 1 & 1 & \cdots & 1 & 1\\ x_1 & x_2 & \cdots & x_n & X\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} & X^{n-1}\\ x_1^n & x_2^n & \cdots & x_n^n & X^n\end{pmatrix}.$$

Por hipótesis inductiva, $V(x_1,\dots,x_n)\neq 0$. Al desarrollar entonces el determinante por la última columna, se ve claramente que el polinomio $V(x_1,\dots,x_n,X)\in\mathbb{K}[X]$ tiene grado n y que su coeficiente principal es $V(x_1,\dots,x_n)$. Como $f(x_j)=0$ para todo $j\in\{1,\dots,n\}$, entonces

$$f=V(x_1,\dots,x_n)\prod_{i=1}^{n}(X-x_j).$$

Luego

$$
\begin{aligned}
f(x_{n+1}) &= V(x_1,\dots,x_{n+1})\\
&= V(x_1,\dots,x_n)\prod_{i=1}^{n}(x_{n+1}-x_j)\\
&= \prod_{1\le i<j\le n+1}(x_j-x_i).
\end{aligned}
$$

tal como queríamos demostrar.

5.5.2. Ejercicio. Sea V un espacio vectorial de dimensión finita n y sea $\{v_1,\dots,v_n\}$ una base de V. Sean $x_1,\dots,x_n\in\mathbb{K}$ con $x_i\neq x_j$ si $i\neq j$. Demuestre que si

$$
\begin{aligned}
w_1 &= v_1+v_2+\cdots+v_n,\\
w_2 &= x_1v_1+x_2v_2+\cdots+x_nv_n,\\
&\ \vdots\\
w_n &= x_1^{n-1}v_1+x_2^{n-1}v_2+\cdots+x_n^{n-1}v_n,
\end{aligned}
$$

entonces $\{w_1,\dots,w_n\}$ es una base de V.

xca:lagrange 5.5.3. Ejercicio (Polinomio interpolador). Sean $x_1,\dots,x_{n+1}\in\mathbb{K}$ tales que $x_i\neq x_j$ si $i\neq j$ y sean $y_1,\dots,y_{n+1}\in\mathbb{K}$. Demuestre que el polinomio

$$f=\sum_{i=1}^{n+1}y_i\prod_{\substack{j=1\\ j\neq i}}^{n+1}\frac{X-x_j}{x_i-x_j}$$

es el único polinomio $f\in\mathbb{K}[X]$ de grado $\le n$ tal que $f(x_i)=y_i$ para todo $i\in\{1,\dots,n+1\}$.

5.5.4 (Matriz compañera). $\boxed{\text{mover!}}$ Vamos a demostrar que si

$$f=X^n+a_{n-1}X^{n-1}+\cdots+a_2X^2+a_1X+a_0\in\mathbb{K}[X]$$

entonces

eq:C

$$
f=\det\begin{pmatrix}
X & 0 & 0 & 0 & \cdots & a_0\\
-1 & X & 0 & 0 & \cdots & a_1\\
0 & -1 & X & 0 & \cdots & a_2\\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots\\
0 & 0 & 0 & -1 & X & a_{n-2}\\
0 & 0 & 0 & 0 & -1 & X+a_{n-1}
\end{pmatrix}. \tag{5.5.5}
$$

La **matriz compañera** del polinomio f es la matriz $C=(c_{ij})$ que está dada por

$$
c_{ij}=\begin{cases}
1 & \text{si } i=j+1 \text{ y } j<n,\\
-a_{i-1} & \text{si } j=n,\\
0 & \text{en otro caso.}
\end{cases}
$$

La igualdad (5.5.5) puede escribirse entonces como $f=\det(XI-C)$, donde C es la matriz compañera de f. Para demostrar esta igualdad procederemos por inducción en n. El caso $n=2$ es fácil y se deja como ejercicio. Supongamos que la afirmación es válida para $n \geq 2$. Si desarrollamos el determinante de la matriz

$$d(a_0, \dots, a_n) = \begin{pmatrix} X & 0 & 0 & 0 & \cdots & a_0 \\ -1 & X & 0 & 0 & \cdots & a_1 \\ 0 & -1 & X & 0 & \cdots & a_2 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & X & a_{n-1} \\ 0 & 0 & 0 & 0 & -1 & X + a_n \end{pmatrix}$$

por la primera fila, se tiene que

$$\begin{aligned} d(a_0, \dots, a_n) &= X d(a_1, \dots, a_{n-1}) + a_0(-1)^{n+1}(-1)^{n-1} \\ &= X(X^{n-1} + a_{n-1}X^{n-2} + \cdots + a_1) + a_0 \\ &= X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0. \end{aligned}$$

5.6. Anillos conmutativos con unidad

dar definición y un par de ejemplos básicos

matrices sobre anillos conmutativos

determinantes de matrices cuadradas (como ejercicio)

5.7. Ejercicios

determinante:A1000

5.7.1. Ejercicio. Sea $A \in \mathbb{K}^{4\times 4}$ una matriz tal que $\det A = -1$. Demuestre que el conjunto $\{A^{1000}, A^{1001}\}$ es linealmente independiente.

Capítulo 6

Diagonalización

6.1. Autovalores y autovectores

6.1.1. Sean V un espacio vectorial y $f \in \hom(V, V)$. Un escalar $\lambda \in \mathbb{K}$ es un **autovalor** de f si existe $v \in V \setminus \{0\}$ tal que $f(v) = \lambda v$. El vector v se denomina **autovector** de autovalor λ.

6.1.2. Ejemplos.

1) Si $\ker f \neq \{0\}$ entonces $\lambda = 0$ es un autovalor pues para cada $v \in \ker f \setminus \{0\}$ se tiene $f(v) = 0v$.
2) Si f es un proyector no nulo entonces $\lambda = 1$ es autovalor pues para cada $v \operatorname{im} f \setminus \{0\}$ se tiene que $f(v) = 1v$.
3) Si $V = C^\infty(\mathbb{R})$ y $\partial \colon V \to V$ es la aplicación $f \mapsto f'$ entonces todo $\lambda \in \mathbb{R}$ es autovalor pues $e^{\lambda x} \neq 0$ para todo $x \in \mathbb{R}$ y $\partial(e^{\lambda x}) = \lambda e^{\lambda x}$.

pro:autovalores

6.1.3. Proposición. *Sean V un espacio vectorial de dimensión finita sobre $\mathbb{K}$, $f \in \hom(V, V)$ y $\lambda \in \mathbb{K}$. Las siguientes afirmaciones son equivalentes:*

1) *λ es autovalor de f.*
2) *$\lambda \operatorname{id}_V - f$ no es un isomorfismo.*
3) *$\det(\lambda \operatorname{id}_V - f) = 0$.*

Demostración. Para demostrar que (1) implica (2) basta observar que, por definición, como λ es autovalor, existe $v \in V \setminus \{0\}$ tal que $f(v) = \lambda v$. Luego $v \in \ker(\lambda \operatorname{id}_V - f) \setminus \{0\}$ y entonces $\lambda \operatorname{id}_V - f$ no es un isomorfismo.

Demostremos que (2) implica (3). Sean $\mathscr{B}$ una base de V y $A = [f]_{\mathscr{B},\mathscr{B}}$. Si $\lambda \operatorname{id}_V - f$ no es un isomorfismo entonces la matriz

$$\lambda I - A = [\lambda \operatorname{id}_V - f]_{\mathscr{B},\mathscr{B}}$$

es no inversible por el corolario 3.4.8. Luego $\det(\lambda I - A) = 0$ por el corolario 5.4.26.

Demostremos que (3) implica (1). Supongamos que $\mathscr{B} = \{v_1, \dots, v_n\}$. Como $\lambda I - A$ no es inversible, existen $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ no todos cero tales que

$$(\lambda I - A) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Luego $v = \sum_{i=1}^n \alpha_i v_i$ es no nulo y $f(v) = \lambda v$. $\square$

6.1.4. Sea V un espacio vectorial y sea $f \in \hom(V, V)$. El **espectro** $\operatorname{spec} f$ de f es el conjunto formado por los autovalores de f.

spec(fg)=spec(gf)

6.1.5. Ejercicio. Sea V un espacio vectorial de dimensión finita. Demuestre que si $f, g \in \hom(V, V)$ entonces $\operatorname{spec}(fg) = \operatorname{spec}(gf)$.

6.1.6. Ejemplo. Retomemos el ejemplo 5.4.10 y calculemos los autovalores de la matriz $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. El polinomio característico es

$$\det(XI - A) = \det \begin{pmatrix} X & -1 \\ -1 & X - 1 \end{pmatrix} = X^2 - X - 1.$$

Las raíces del polinomio $X^2 - X - 1$, son

$$\lambda_1 = (1+\sqrt{5})/2, \quad \lambda_2 = (1-\sqrt{5})/2.$$

Tenemos entonces que $\operatorname{spec} A = \{\lambda_1, \lambda_2\}$. Calculemos ahora el autoespacio del autovalor λ_i. Si a la matriz $\lambda_i I - A$ le aplicamos la operación de filas $F_1 + \lambda_i F_2 \to F_2$ vemos que el conjunto

$$\left\{ \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix}, \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} \right\}$$

es una base de $\mathbb{K}^{2\times 1}$ formada por autovectores de A. Entonces $A = CDC^{-1}$, donde

$$C = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}, \qquad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \qquad C^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} -\lambda_2 & 1 \\ \lambda_1 & -1 \end{pmatrix}.$$

En particular, como $A^n = (CDC^{-1})^n = CD^nC^{-1}$, se obtiene una fórmula cerrada para la sucesión de Fibonacci:

$$F_n = \frac{1}{\sqrt{5}} (\lambda_1^n - \lambda_2^n) \quad \text{para todo } n \geq 1.$$

6.2. El polinomio característico

6.2.1. Se define el **polinomio característico** de una matriz $A \in \mathbb{K}^{n\times n}$ como

$$\chi_A = \det(XI - A) \in \mathbb{K}[X].$$

De la definición es evidente que χ_f es un polinomio mónico de grado n y que $\chi_A(0) = (-1)^n \det A$.

lem:chiA=chiB

6.2.2. LEMA. *Sean $A, B \in \mathbb{K}^{n\times n}$. Si A y B son semejantes entonces*

$$\chi_A = \chi_B.$$

DEMOSTRACIÓN. Si existe $C \in \mathbb{K}^{n\times n}$ inversible tal que $B = CAC^{-1}$ entonces $\chi_A = \det(XI - A) = \det(C(XI - A)C^{-1}) = \det(XI - B) = \chi_B$. □

6.2.3. Sean V un espacio vectorial de dimensión finita y $f \in \hom(V, V)$. El lema 6.2.2 afirma que matrices semejantes tienen el mismo polinomio característico. Tiene sentido entonces definir el **polinomio característico** de f como el polinomio característico de la matriz $[f]_{\mathscr{B},\mathscr{B}}$, donde $\mathscr{B}$ es alguna base ordenada de V.

6.2.4. OBSERVACIÓN. En la proposición 6.1.3 se demostró entonces lo siguiente: si V es un espacio vectorial de dimensión finita y $f \in \hom(V, V)$ entonces $\lambda \in \mathbb{K}$ es autovalor de f si y sólo si λ es raíz del polinomio característico de f, es decir: $\chi_f(\lambda) = 0$.

6.2.5. EJEMPLO. Sea $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ dada por $f(x, y) = (-y, x)$. Entonces el polinomio característico $\chi_f = 1 + X^2$ no tiene raíces reales y $\operatorname{spec} f = \emptyset$.

6.2.6. EJEMPLO. Sea V un espacio vectorial complejo de dimensión finita n y sea $f \in \hom(V, V)$. Entonces, por el teorema fundamental del álgebra, f tiene exactamente n autovalores (contados con multiplicidad).

6.2.7. EJEMPLO. Sea V un espacio vectorial real de dimensión finita $n = 2k+1$ y sea $f \in \hom(V, V)$. Entonces f tiene al menos un autovalor.

6.2.8. LEMA. *Sean V un espacio vectorial finita. Sea $f \in \hom(V, V)$ y sean $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ autovalores de f tales que $\lambda_i \neq \lambda_j$ si $i \neq j$. Sean $v_1, \dots, v_r$ tales que $f(v_i) = \lambda_i v_i$ para todo $i \in \{1, \dots, r\}$. Entonces $\{v_1, \dots, v_r\}$ es linealmente independiente.*

Demostración. Procederemos por inducción en r. Si $r = 1$ el resultado es evidente pues $v_1 \neq 0$. Si suponemos que el resultado es válido para un cierto $r \geq 1$ demostremos que es válido para $r+1$. Sean $\lambda_1, \dots, \lambda_{r+1} \in \mathbb{K}$ autovalores distintos entre sí y sean $v_1, \dots, v_{r+1} \in V$ tales que $f(v_i) = \lambda_i v_i$ para todo i. Supongamos que $\alpha_1 v_1 + \cdots + \alpha_{r+1} v_{r+1} = 0$. Si $\alpha_{r+1} = 0$ entonces $\alpha_1 v_1 + \cdots + \alpha_r v_r = 0$ y luego $\alpha_i = 0$ para todo i por hipótesis inductiva. En cambio, si $\alpha_{r+1} \neq 0$, entonces $v_{r+1} \in \langle v_1, \dots, v_r \rangle$ pues

$$v_{r+1} = \beta_1 v_1 + \cdots + \beta_r v_r = \left(-\frac{\alpha_1}{\alpha_{r+1}}\right) v_1 + \cdots + \left(-\frac{\alpha_r}{\alpha_{r+1}}\right) v_r.$$

si $\beta_i = -\alpha_i/\alpha_{r+1}$. Como $f(v_{r+1}) = \lambda_{r+1} v_{r+1}$, se tiene que

$$\beta_1(\alpha_1 - \alpha_{r+1})v_1 + \cdots + \beta_r(\alpha_r - \alpha_{r+1})v_r = 0,$$

y luego $\beta_i = 0$ para todo i. Luego $\alpha_i = 0$ para todo $i \in \{1, \dots, r\}$ y por lo tanto $\alpha_{r+1} v_{r+1} = 0$, una contradicción. $\square$

6.2.9. Si V es de dimensión finita, $f \in \hom(V, V)$ y $\lambda \in \mathbb{K}$ es autovalor de f entonces se define el **autoespacio** de f asociado al autovalor λ como

$$S(\lambda) = \{v \in V : f(v) = \lambda v\}.$$

Queda como ejercicio demostrar que $S(\lambda)$ es un subespacio de V. Se define la **multiplicidad algebraica** de λ como el mayor entero positivo k tal que $(X - \lambda)^k$ divide a χ_f. La **multiplicidad geométrica** es el número $\dim S(\lambda)$.

6.2.10. Ejemplo. Sea $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ definida por $f(x, y, z) = (x, x + y, 2z)$. El polinomio característico de f es $\chi_f = (X - 1)^2(X - 2)$. Queda como ejercicio demostrar que $S(1) = \langle (0, 1, 0) \rangle$ y $S(2) = \langle (0, 0, 1) \rangle$. La multiplicidad algebraica del autovalor 1 es 2 y la multiplicidad geométrica es $\dim S(1) = 1$.

6.2.11. Lema. *Sean V un espacio vectorial finita. Sea $f \in \hom(V, V)$ y sean $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ autovalores de f tales que $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces*

$$S(\lambda_1) + \cdots + S(\lambda_r) = S(\lambda_1) \oplus \cdots \oplus S(\lambda_r).$$

6.2.12. Lema. *Sean V un espacio vectorial de dimensión finita, $f \in \hom(V, V)$ y λ un autovalor de f. Sea m la multiplicidad algebraica de λ. Entonces*

$$\dim S(\lambda) \leq m.$$

Demostración. Supongamos que $x_f = (X - \lambda)^m q$, donde $q \in \mathbb{K}[X]$ y $q(\lambda) \neq 0$. Si $\dim S(\lambda) = r \geq m + 1$, sea $\{v_1, \dots, v_r\}$ una base de $S(\lambda)$. La extendemos a una base $\mathscr{B} = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ de V y consideramos la matriz de f con respecto a $\mathscr{B}$:

$$[f]_{\mathscr{B},\mathscr{B}} = \left(\begin{array}{c|c} \lambda I & \star \\ \hline 0 & B \end{array}\right),$$

donde λI es una matriz de $r \times r$ y B es una matriz de $(n - r) \times (n - r)$. Entonces

$$\chi_f = \det(X - \lambda)^r \det(XI - B) = (X - \lambda)^m (X - \lambda)^{r-m} \det(XI - B),$$

donde $r - m \geq 1$. Luego $(X - \lambda)^{r-m} \det(XI - B)$ es un polionimo que se anula en λ, una contradicción. $\square$

6.2.13. Ejemplo. Sea $f\colon \mathbb{R}^{3\times 1} \to \mathbb{R}^{3\times 1}$ dada por

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ x + y \\ 2z \end{pmatrix}.$$

Entonces $X_f = (X - 1)^2(X - 2)$ y además

$$S(1) = \{(x, y, z)^T \in \mathbb{R}^{3\times 1} : x = z = 0\} = \langle (0, 1, 0)^T \rangle,$$
$$S(2) = \{(x, y, z)^T \in \mathbb{R}^{3\times 1} : x = y = 0\} = \langle (0, 0, 1)^T \rangle.$$

Luego $\dim S(1) = \dim S(2) = 1$. VER

:f_diagonalizable

6.2.14. Ejemplo. Sea $f\colon \mathbb{R}^{3\times 1} \to \mathbb{R}^{3\times 1}$ dada por

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

entonces $x_f = X(X-1)(X+1)$ y

$$S(0) = \langle (0,1,-1)^T \rangle, \qquad S(1) = \langle (0,1,0)^T \rangle, \qquad S(-1) = \langle (1,0,0)^T \rangle.$$

VER

6.2.15. Sean V un espacio vectorial de dimensión finita y $f \in \hom(V,V)$. Se dice que f es **diagonalizable** si existe una base $\mathscr{B}$ de V tal que $[f]_{\mathscr{B},\mathscr{B}}$ es una matriz diagonal.

6.2.16. Ejemplo. La transformación lineal del ejemplo 6.2.14 es diagonalizable ya que si $\mathscr{B} = \{(0,1,-1)^T, (0,1,0)^T, (1,0,0)^T\}$ entonces

$$[f]_{\mathscr{B}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

6.2.17. Teorema. *Sea V un espacio vectorial de dimensión finita sobre $\mathbb{K}$, sea $f \in \hom(V,V)$ y sean $\lambda_1, \dots, \lambda_k$ autovalores distintos de f. Entonces f es diagonalizable si y sólo si $\dim S(\lambda_i) = m_i$ para todo $i \in \{1, \dots, k\}$ y*

$$\chi_f = (X - \lambda_1)^{m_1} \cdots (X - \lambda_k)^{m_k}.$$

Demostración. Supongamos primero que $\dim S(\lambda_i) = m_i$ para todo $i \in \{1, \dots, k\}$ y que

$$\chi_f = (X - \lambda_1)^{m_1} \cdots (X - \lambda_k)^{m_k}.$$

Sean $\mathscr{B}_1 = \{u_i, \dots, u_{m_1}\}$ una base de $S(\lambda_1)$, $\mathscr{B}_2 = \{v_1, \dots, v_{m_2}\}$ una base de $S(\lambda_2)$... y $\mathscr{B}_k = \{w_1, \dots, w_{m_k}\}$ una base de $S(\lambda_k)$. Vamos a demostrar que $\mathscr{B}_1 \cup \dots \cup \mathscr{B}_k$ es base de V, y para esto, como $m_1 + \dots + m_k = n$, basta ver que $\mathscr{B}_1 \cup \dots \cup \mathscr{B}_k$ es un conjunto linealmente independiente. Supongamos que

$$\alpha_1 u_1 + \dots + \alpha_{m_1} u_{m_1} + \beta_1 v_1 + \dots + \beta_{m_2} v_{m_2} + \dots + \gamma_1 w_1 + \dots + \gamma_{m_k} w_{m_k} = 0.$$

Observemos que

$$\begin{aligned} &\alpha_1 u_1 + \dots + \alpha_{m_1} u_{m_1} \in S(\lambda_1), \\ &\beta_1 v_1 + \dots + \beta_{m_2} v_{m_2} \in S(\lambda_2), \\ &\quad \vdots \\ &\gamma_1 w_1 + \dots + \gamma_{m_k} w_{m_k} \in S(\lambda_k). \end{aligned}$$

Luego, como los $S(\lambda_i)$ están en suma directa, $\alpha_i = 0$ para todo i, $\beta_j = 0$ para todo $j\dots$, y $\gamma_k = 0$ para todo k.

Recíprocamente, si f es diagonalizable, existe una base $\{v_1, \dots, v_n\}$ de V tal que la matriz de f en esa base es diagonal. Al agrupar los autovalores iguales se obtiene $X_f = \prod_{i=1}^{r} (X - \lambda_i)^{m_i}$. Es claro que $\dim S(\lambda_i) \geq m_i$. Por otro lado, siempre vale que $\dim S(\lambda_i) \leq m_i$. Luego $\dim S(\lambda_i) = m_i$. □

6.2.18. Una matriz $A \in \mathbb{K}^{n\times n}$ es **diagonalizable** si la transformación lineal $f\colon \mathbb{K}^{n\times 1} \to \mathbb{K}^{n\times 1}$ dada por $x \mapsto Ax$ es diagonalizable.

riz_diagonalizable

6.2.19. Ejercicio. Sea $A \in \mathbb{K}^{n\times n}$. Demuestre que las siguientes afirmaciones son equivalentes:

1) A es diagonalizable.

2) $\chi_A = \prod_{i=1}^{r}(X-\lambda_i)^{m_i}$ y $m_i = n - \operatorname{rg}(\lambda_i I - A)$.
3) A es semejante a una matriz diagonal.

6.3. Polinomios e ideales

6.3.1. TEOREMA (Algoritmo de división). *Sean $f,d \in \mathbb{K}[X]$ con $d \neq 0$. Entonces existen únicos $q,r \in \mathbb{K}[X]$ tales que $f = dq + r$ y además $r = 0$ o bien $\deg r < \deg d$.*

6.3.2. Sean $f,g \in \mathbb{K}[X]$. Recordemos que g **divide** a f (o que f es **divisible** por g, o que g es un **múltiplo** de f) si existe $h \in \mathbb{K}[X]$ tal que $f = gh$. Un escalar $\lambda \in \mathbb{K}$ es **raíz** (o **cero**) de f si $f(\lambda) = 0$.

6.3.3. COROLARIO. *Sea $f \in \mathbb{K}[X]$ y sea $\lambda \in \mathbb{K}$. Entonces f es divisible por $X - \lambda$ si y sólo si $f(\lambda) = 0$.*

DEMOSTRACIÓN. Si f es divisible por $X-\lambda$ entonces $f = (X-\lambda)q$ para algún $q \in \mathbb{K}[X]$. Al evaluar en λ se obtiene entonces $f(\lambda) = 0$. Recíprocamente, si utilizamos el algoritmo de división, existen únicos $q,r \in \mathbb{K}[X]$ tales que $f = (X-\lambda)q + r$. Como $f(\lambda) = 0$, entonces $r(\lambda) = 0$ y luego $r = 0$. □

6.3.4. COROLARIO. *Sea $f \in \mathbb{K}[X]$ de grado n. Entonces f tiene a lo sumo n raíces en $\mathbb{K}$.*

DEMOSTRACIÓN. Procedemos por inducción en n. Si $n = 1$ entonces el resultado es trivialmente válido. Suponemos entonces que el resultado vale para algún $n \geq 1$ y sea f es un polinomio de grado $n+1$. Si f no tiene raíces en $\mathbb{K}$ entonces no hay nada que demostrar. Si $x_0 \in \mathbb{K}$ es raiz de f, existe $g \in \mathbb{K}[X]$ con $\deg g = n$ tal que $f = (X - x_0)g$. Por hipótesis inductiva g tiene a lo sumo n raíces y entonces f tiene a lo sumo $n+1$ raíces en $\mathbb{K}$. □

6.3.5. Un **ideal** de $\mathbb{K}[X]$ es un subespacio $I \subseteq \mathbb{K}[X]$ tal que $fg \in I$ si $f \in I$ y $g \in \mathbb{K}[X]$.

6.3.6. EJEMPLOS. El subespacio $\{0\}$ es un ideal de $\mathbb{K}[X]$. Si $g \in \mathbb{K}[X]$ entonces el conjunto

$$\{fg : f \in \mathbb{K}[X]\}$$

es un ideal de $\mathbb{K}[X]$ que se denota por (g) y se denomina el ideal generado por g. Más generalmente, si $g_1,\dots,g_n \in \mathbb{K}[X]$, el conjunto

$$(g_1,\dots,g_n) = \left\{ \sum_{i=1}^{n} f_i g_i : f_1,\dots,f_n \in \mathbb{K}[X] \right\}$$

es un ideal de $\mathbb{K}[X]$ que se denomina **ideal generado** por $g_1,\dots,g_n$.

6.3.7. EJEMPLO. Demostremos que el ideal

$$I = (X-8, X^2+8X+5) \subseteq \mathbb{R}[X]$$

es igual $\mathbb{R}[X]$. Como I es un ideal, $X^2 - 8X = X(X-8) \in I$. De la misma forma, como los polinomios X^2-8X y X^2+8X+5 son elementos de I, se tiene que $5 = X^2 - 8X - (X^2+8X+5) \in I$. Luego, como $1 = (1/5)5 \in I$, se tiene que $I = \mathbb{R}[X]$ ya que para todo $f \in \mathbb{R}[X]$ tenemos $f = f1 \in I$.

6.3.8. TEOREMA. *Sea $I \subseteq \mathbb{K}[X]$ un ideal no nulo. Entonces existe un único polinomio mónico $g \in \mathbb{K}[X]$ tal que $I = (g)$.*

DEMOSTRACIÓN. Demostremos la existencia. Como $I \neq \{0\}$, existe un polinomio no nulo que pertenece a I. Sea entonces g el polinomio mónico de grado mínimo tal que $g \in I$. Para demostrar que $I = (g)$ sea $f \in I$. Como g tiene grado mínimo entre los polinomios de I, $\deg f \geq \deg g$ y entonces, por el algoritmo de división, existen $h,r \in \mathbb{K}[X]$ tal que

$f = gh + r$ donde $r = 0$ o $\deg r < \deg g$. Como I es un ideal, $r = f - gh \in I$. La minimalidad del grado de g implica que $r = 0$ y entonces $f = gh \in (g)$.

Demostremos la unicidad. Si $I = (g_1) = (g_2)$, existen $h_1, h_2 \in \mathbb{K}[X]$ tales que $g_1 = h_1 g_2$ y $g_2 = h_2 g_1$. Entonces $g_1 = (h_1 h_2) g_2$. Como g_1 y g_2 tienen el mismo grado, $0 = \deg(h_1 h_2) = \deg h_1 + \deg h_2$ y entonces $\deg h_1 = \deg h_2 = 0$. Como g_1 y g_2 son mónicos, $h_1 = h_2 = 1$. $\square$

6.3.9. Recordemos que si $f, g \in \mathbb{K}[X]$ entonces se define el **máximo común divisor** entre f y g como el único polinomio mónico $h \in \mathbb{K}[X]$ tal que

1) h divide a f, h divide a g,
2) h es múltiplo de cada polinomio que divida a f y a g.

El máximo común divisor entre f y g se denota por $(f : g)$.

Similarmente se define el **mínimo común múltiplo** de f y g como el único polinomio mónico $h \in \mathbb{K}[X]$ tal que:

1) f divide a h, g divide a h,
2) h divide a cada polinomio que es múltiplo de f y de g.

El mínimo común múltiplo entre f y g se denota por $[f : g]$.

Observemos que vale

$$fg = (f : g)[f : g].$$

6.3.10. Ejercicio. Sean $f_1, \dots, f_n \in \mathbb{K}[X]$. Demuestre que el generador g del ideal $(f_1, \dots, f_n)$ es el **máximo común divisor** de los polinomios $f_1, \dots, f_n$, es decir: $g \in \mathbb{K}[X]$ es el único polinomio mónico tal que

1) $g \in (f_1, \dots, f_n)$,
2) g divide a cada f_i,
3) g es divisible por todo polinomio que divida a cada uno de los f_i.

6.4. El polinomio minimal

6.4.1. Dados un polinomio $p \in \mathbb{K}[X]$,

$$p = a_0 + a_1 X + \cdots + a_d X^d = \sum_{i=0}^{d} a_i X^i,$$

y una matriz $A \in \mathbb{K}^{n \times n}$ se define

$$p(A) = a_0 I + a_1 A + \cdots + a_n A^n \in \mathbb{K}^{n \times n}.$$

Es evidente que si $p, q \in \mathbb{K}[X]$ y $A \in \mathbb{K}^{n \times n}$ entonces

$$(p + q)(A) = p(A) + q(A),$$
$$(pq)(A) = p(A)q(A) = q(A)p(A),$$

para todo $p, q \in \mathbb{K}[X]$ y $A \in \mathbb{K}^{n \times n}$. Más generalmente, puede demostrarse que el conjunto

$$\{B \in \mathbb{K}^{n \times n} : B = p(A) \text{ para algún } p \in \mathbb{K}[X]\}$$

es un anillo conmutativo con unidad.

6.4.2. Ejemplos. Sea $A \in \mathbb{K}^{n \times n}$. Si $p = X^2 + 1 \in \mathbb{K}[X]$ entonces

$$p(A) = A^2 + I.$$

Si $q = X^3 + 2X^2 - 3X + 4$ entonces

$$q(A) = A^3 + 2A^2 - 3A + 4I.$$

6.4.3. Lema. *Sea $A \in \mathbb{K}^{n \times n}$. Entonces existe $p \in \mathbb{K}[X]$ no nulo tal que $p(A) = 0$.*

DEMOSTRACIÓN. El conjunto $\{I, A, A^2, \dots, A^{n^2}\} \subseteq \mathbb{K}^{n\times n}$ tiene n^2+1 elementos. Como $\dim \mathbb{K}^{n\times n} = n^2$, el conjunto S es linealmente dependiente. Luego existen escalares $a_0, \dots, a_{n^2} \in \mathbb{K}$, no todos cero, tales que $a_0 I + a_1 A + \cdots + a_{n^2} A^{n^2} = 0$. En conclusión, si $p = \sum_{i=0}^{n^2} a_i X^i \in \mathbb{K}[X]$, entonces $p \neq 0$ y $p(A) = 0$. $\square$

6.4.4. Sea $A \in \mathbb{K}^{n\times n}$. El conjunto

$$\{p : p \in \mathbb{K}[X] \text{ tal que } p(A) = 0\}$$

es un ideal de $\mathbb{K}[X]$ y entonces existe un único polinomio mónico que lo genera, es decir: existe un único polinomio mónico m_A que lo genera. Este polinomio se denomina **polinomio minimal** de A.

xca:minimal

6.4.5. EJERCICIO. Sea $A \in \mathbb{K}^{n\times n}$ y sea $m \in \{1, \dots, n\}$ el único entero que satisface que $\{I, A, \dots, A^{m-1}\}$ es linealmente independiente y que $A^m \in \langle I, A, \dots, A^{m-1}\rangle$. Sean $a_0, a_1, \dots, a_{m-1} \in \mathbb{K}$ los únicos escalares tales que $A^m = a_0 I + a_1 A + \cdots + a_{m-1} A_{m-1}$. Demuestre que el polinomio

$$X^m - \sum_{i=0}^{m-1} a_i X^i$$

es el polinomio minimal de A.

6.4.6. EJEMPLOS. Es fácil verificar que $m_0 = X$ y que $m_I = X - 1$. Si $A = \lambda I$ entonces $m_A = X - \lambda$. Si $A \notin \{0, I\}$ satisface que $A^2 = A$ entonces $m_A = X^2 - X$.

6.4.7. EJEMPLO. Calculemos el polinomio minimal de

$$A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

Primero observemos que el conjunto $\{I, A\}$ es linealmente independiente. Buscamos entonces $p = X^2 + bX + c \in \mathbb{R}[X]$ que anule a la matriz A. Un cálculo sencillo muestra que $A^2 + bA + cI = 0$ si y sólo si $b = 2$ y $c = 1$. Luego $m_A = X^2 + 2X + 1$.

rem:m_A|p

6.4.8. OBSERVACIÓN. De la definición de polinomio minimal se obtiene inmediatamente que si $A \in \mathbb{K}^{n\times n}$ y $p \in \mathbb{K}[X]$ entonces $p(A) = 0$ si y sólo si m_A divide a p.

_B=>p(A)_sem_p(B)

6.4.9. LEMA. *Sean $A, B \in \mathbb{K}^{n\times n}$ y $p \in \mathbb{K}[X]$. Si A y B son semejantes entonces $p(A)$ y $p(B)$ son semejantes. En particular, si A y B son semejantes entonces $p(A) = 0$ si y sólo si $p(B) = 0$.*

DEMOSTRACIÓN. Supongamos que $p = \sum_{i=0}^{d} a_i X^i$ y que $A = CBC^{-1}$ para alguna matriz inversible $C \in \mathbb{K}^{n\times n}$. Como $(CBC^{-1})^i = CB^iC^{-1}$ para todo $i \geq 0$,

$$\begin{aligned} p(A) = p(CBC^{-1}) &= \sum_{i=0}^{d} a_i (CBC^{-1})^i \\ &= \sum_{i=0}^{d} a_i CB^iC^{-1} = C\left(\sum_{i=0}^{d} a_i B^i\right) C^{-1} = Cp(B)C^{-1}. \end{aligned}$$

Luego $p(A)$ y $p(B)$ son semejantes. Si A y B son semejantes y entonces $p(A)$ y $p(B)$ son semejantes y luego, en particular, $p(A) = 0$ si y sólo si $p(B) = 0$. $\square$

6.4.10. PROPOSICIÓN. *Sean $A, B \in \mathbb{K}^{n\times n}$. Si A y B son semejantes entonces $m_A = m_B$.*

DEMOSTRACIÓN. Por el lema 6.4.9, las matrices $m_A(A)$ y $m_B(A)$ son semejantes y entonces, como $0 = m_A(A)$, se obtiene que $m_B(A) = 0$. La observación 6.4.8 implica que m_A divide a m_B. Similarmente se demuestra que m_B divide a m_A y luego, como m_A y m_B son mónicos, $m_A = m_B$. $\square$

6.4.11. La proposición anterior nos permite definir el **polinomio minimal** de una transformación lineal $f\colon V \to V$, donde V es un espacio vectorial de dimensión finita. En efecto, basta tomar m_f como $m_{[f]_{\mathscr{B}}}$ donde $\mathscr{B}$ es alguna base de V.

6.4.12. Proposición. *Sea $A \in \mathbb{K}^{n\times n}$. Entonces λ es autovalor de A si y sólo si λ es raíz de m_A.*

Demostración. Supongamos primero que λ es autovalor de A. Entonces existe $v \neq 0$ tal que $Av = \lambda v$. Por el algoritmo de división, existen $q, r \in \mathbb{K}[X]$ tales que $m_A = (X-\lambda)q + r$, donde $r = 0$ o bien $\deg(r) = 0$. Al evaluar en la matrix A obtenemos $0 = m_A(A) = (A-\lambda I)q(A) + rI$ y en particular $0 = (A-\lambda I)q(A)v + rv$. Como $A - \lambda I$ y $q(A)$ conmutan, $rv = 0$. Como $v \neq 0$ entonces $r = 0$.

Recíprocamente, si $m_A(\lambda) = 0$ entonces $m_A = (X-\lambda)q$ para algún $q \in \mathbb{K}[X]$. Como $\deg(q) < \deg(m_A)$, $q(A) \neq 0$. Existe entonces $w \in \mathbb{K}^n$ tal que $q(A)w \neq 0$. Sea $v = q(A)w$. Entonces $(A-\lambda I)v = 0$. □

6.5. El polinomio minimal de un vector

6.5.1. Sean $A \in \mathbb{K}^{n\times n}$ y $v \in \mathbb{K}^{n\times 1}$. Para cada polinomio $p \in \mathbb{K}[X]$ se define $p(v) = p(A)v$. En particular, diremos que el polinomio p se anula (con respecto a la matriz A) en v si $p(v) = 0$. El conjunto

$$\{p \in \mathbb{K}[X] : p(v) = 0\}$$

es un ideal de $\mathbb{K}[X]$ y entonces está generado por un único polinomio mónico m_v. Este polinomio se denomina el **polinomio minimal** del vector v con respecto a la matriz A. Observar que m_v es el polinomio mónico de menor grado que especializado en A se anula en v.

6.5.2. Ejemplo. Sean $A \in \mathbb{K}^{n\times n}$ y $v \in \mathbb{K}^{n\times 1}$. Entonces $m_v = X - \lambda$ si y sólo si v es autovector de A de autovalor λ.

6.5.3. Ejemplo. Sean $v = (1,0)^T$ y

$$A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

Calculemos el polinomio minimal de v con respecto a la matriz A. Primero buscamos $p = X + c \in \mathbb{R}[X]$ tal que $p(v) = 0$. Esto equivale a resolver la ecuación $Av^T = -cv^T$, que no tiene solución. Buscamos entonces un polinomio $X^2 + bX + c \in \mathbb{R}[X]$ que anula al vector v. Un cálculo sencillo muestra que $p(v) = 0$ si y sólo si $b = 2$ y $c = 1$. Luego $m_v = X^2 + 2X + 1$.

rem:m_v|m_A

6.5.4. Observación. De la definición se obtiene inmediatamente que si $A \in \mathbb{K}^{n\times n}$, $v \in \mathbb{K}^{n\times 1}$ y $p \in \mathbb{K}[X]$ entonces $p(v) = 0$ si y sólo si m_v divide a p. En particular, m_v divide a m_A.

pro:mA=mcm

6.5.5. Proposición. *Sean $A \in \mathbb{K}^{n\times n}$ y $\{v_1, \dots, v_n\}$ una base de $\mathbb{K}^{n\times 1}$. Entonces $m_A = \operatorname{mcm}(m_{v_1}, \dots, m_{v_n})$.*

Demostración. Sea $p = \operatorname{mcm}(m_{v_1}, \dots, m_{v_n})$. La observación 6.5.4 implica que m_{v_i} divide a m_A para todo $i \in \{1, \dots, n\}$ y entonces p divide a m_A. Por otro lado, cada m_{v_i} divide a p y entonces $p(A)v_i = p(v_i) = 0$ para todo $i \in \{1, \dots, n\}$. Como $\{v_1, \dots, v_n\}$ es una base de $\mathbb{K}^{n\times 1}$, entonces $p(A) = 0$ y luego m_A divide a p. Como p y m_A son polinomios mónicos, $p = m_A$. □

6.5.6. Ejercicio. Sea $\{e_1, e_2, e_3\}$ la base canónica de $\mathbb{R}^{3\times 1}$ y sea

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Demuestre que $m_{e_1} = X^2 - 2X + 1$, $m_{e_2} = X - 1$, $m_3 = X - 2$ y concluya que $m_A = (X-1)^2(X-2)$. ¿Es A diagonalizable?

al_diagonalizable

6.5.7. Proposición. *Sea $A \in \mathbb{K}^{n\times n}$. Entonces A es diagonalizable si y sólo si*

$$m_A = (X-\lambda_1)\cdots(X-\lambda_r)$$

donde $\lambda_i \neq \lambda_j$ si $i \neq j$.

Demostración. Supongamos que A es diagonalizable. Entonces existe una base $\{v_1,\dots,v_n\}$ de $\mathbb{K}^{n\times 1}$ y existen $\lambda_1,\dots,\lambda_n \in \mathbb{K}$ tales que $Av_i = \lambda_i v_i$ para todo i. Luego $m_{v_i} = X - \lambda_i$ para todo i. La proposición 6.5.5 implica que m_A tiene la forma deseada.

Supongamos ahora que $m_A = (X-\lambda_1)\cdots(X-\lambda_r)$ donde los λ_i son todos distintos entre sí. Entonces $\lambda_1,\dots,\lambda_r$ son raíces distintas de χ_f y por lo tanto son autovalores de f. Luego

$$S(\lambda_1)+\cdots+S(\lambda_r) = S(\lambda_1)\oplus\cdots\oplus S(\lambda_r).$$

Vamos a demostrar que $V = \oplus_{i=1}^r S(\lambda_i)$. Para cada $j \in \{1,\dots,r\}$ definimos

$$p_j = \frac{1}{X-\lambda_j}\prod_{i=1}^r (X-\lambda_i).$$

Como $(p_1 : \cdots : p_r) = 1$, existen $h_1,\dots,h_r \in \mathbb{K}[X]$ tales que $1 = \sum_{i=1}^r h_i p_i$. Si evaluamos en la matriz A obtenemos

$$I = \sum_{i=1}^r h_i(A)p_i(A).$$

Al multiplicar a izquierda por v obtenemos $v = \sum_{i=1}^r h_i(A)p_i(A)v$. Veamos que $h_j(A)p_j(A)v \in \ker(A-\lambda_j I)$. Como $(X-\lambda_j)p_j = m_A$, entonces $(X-\lambda_j)h_j p_j = h_j m_A$. Al evaluar en A y utilizar que $m_A(A) = 0$,

$$(A-\lambda_j I)h_j(A)p_j(A) = h_j(A)m_A(A) = 0.$$

Luego $(A-\lambda_j I)h_j(A)p_j(A)v = 0$ y entonces $h_j(A)p_j(A)v \in S(\lambda_j)$. □

6.5.8. Ejemplo. Sea $\{e_1,e_2,e_3\}$ la base canónica de $\mathbb{R}^{3\times 1}$ y sea

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

Calculemos m_{e_1}, m_{e_2} y m_{e_3}. Como $Ae_1 = e_1$ entonces $m_{e_1} = X-1$. Además $Ae_2 = e_1$ y $A^2e_2 = Ae_1 = e_1 = Ae_2$ y entonces $m_{e_2} = X^2 - X$. Como $Ae_3 = e_1 - e_2$ entonces $A^2e_3 = 0$ y luego $m_{e_3} = X^2$. Observemos que $m_A = [m_{e_1} : m_{e_2} : m_{e_3}] = X^2(X-1)$ y entonces A no es diagonalizable.

exa:A^k=I

6.5.9. Ejemplo. Sea $A \in \mathbb{C}^{n\times n}$ y supongamos que existe $k \in \mathbb{N}$ tal que $A^k = I$. Entonces A es diagonalizable. En efecto, como el polinomio $X^k - 1$ anula a la matriz A, sabemos que m_A divide a $X^k - 1$. Como todas las raíces de $X^k - 1$ son simples, todas las raíces de m_A también son simples. Luego A es diagonalizable por la proposición 6.5.7.

6.5.10. Ejemplo. El resultado del ejemplo 6.5.9 no vale para matrices reales. Sea

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Entonces $A^3 = I$ pero A no es diagonalizable ya que

$$m_A = (X-1)(X^2+X+1) \in \mathbb{R}[X].$$

6.5.11. Ejercicio. Sea V un $\mathbb{C}$-espacio vectorial de dimensión finita y sea $f \in \hom(V,V)$. Demuestre que f es diagonalizable si y sólo si m_f y m_f' son coprimos.

6.6. El teorema de Cayley–Hamilton

como ejercicio y para endomorfismos

m:Hamilton_Cayley

6.6.1. TEOREMA (Cayley–Hamilton). *Si* $A \in \mathbb{K}^{n\times n}$ *entonces* $\chi_A(A) = 0$*, es decir:* m_A *divide a* χ_A*.*

DEMOSTRACIÓN. Sea $T\colon \mathbb{K}^{n\times 1} \to \mathbb{K}^{n\times 1}$ la transformación lineal definida por $x \mapsto Ax$ y sea $v \in \mathbb{K}^{n\times 1}$. Tomemos $k \in \mathbb{N}$ tal que $\{v, Av, \dots, A^k v\}$ sea un conjunto linealmente independiente y $A^{k+1}v$ sea combinación lineal de $v, Av, \dots, A^k v$, digamos

$$A^{k+1}v = -a_0 v - a_1 Av - \cdots - a_k A^k v,$$

donde $a_0, \dots, a_k \in \mathbb{K}$. Extendemos el conjunto $\{v, Av, \dots, A^k v\}$ a una base

$$\mathscr{B} = \{v, Av, \dots, A^k v, w_1, \dots, w_m\}$$

de V, donde $k + 1 + m = n$. La matriz de T con respecto a la base $\mathscr{B}$ es una matriz por bloques

$$[T]_{\mathscr{B},\mathscr{B}} = \left(\begin{array}{c|c} C & \star \\ \hline 0 & M \end{array}\right), \qquad C = \begin{pmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & \vdots \\ & & \ddots & 0 & -a_{k-1} \\ & & & 1 & -a_k \end{pmatrix}.$$

El polinomio característico de A es

$$\chi_A = \chi_{[T]_{\mathscr{B},\mathscr{B}}} = \chi_C \chi_M.$$

Pero sabemos que $\chi_C = X^{k+1} + a_k X^k + \cdots + a_1 X + a_0$. Además, por definición, tenemos que $\chi_C = m_v$. Luego $\chi_A = m_v \chi_M$ y entonces m_v divide a χ_A para cualquier $v \in \mathbb{K}^n$.

Si $\{e_1, \dots, e_n\}$ es la base canónica de $\mathbb{K}^{n\times 1}$, en particular, tenemos que m_{e_i} divide a χ_A para cada $i \in \{1, \dots, n\}$. En consecuencia,

$$m_A = \operatorname{mcm}(m_{e_1}, \dots, m_{e_n})$$

divide a χ_A, que es lo que queríamos probar. □

6.6.2. Sea $A \in \mathbb{K}^{n\times n}$. Como corolario del teorema de Hamilton–Cayley se obtienen fácilmente las siguientes afirmaciones:

1) $\deg m_A \le n$.
2) Si $\deg m_A = n$ entonces $m_A = \chi_A$.
3) Si existe $v \in \mathbb{K}^{n\times 1}$ tal que $\deg m_v = n$ entonces $m_v = m_A = \chi_A$.

6.6.3. EJEMPLO. Vamos a utilizar el teorema de Hamilton–Cayley para calcular las potencias de una matriz. Sea

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

Entonces $\chi_A = (X-1)^2$. Por el algoritmo de división sabemos que existen $q \in \mathbb{R}[X]$ y $a_n, b_n \in \mathbb{R}$ tales que

$$X^n = (X-1)^2 q + (a_n X + b_n)$$

para todo $n \in \mathbb{N}$. Al evaluar en $X = 1$ se obtiene $a_n + b_n = 1$. Al derivar

$$nX^{n-1} = 2(X-1)q + (X-1)^2 q' + a_n$$

y al evaluar esto en $X = 1$, $a_n = n$ para todo $n \in \mathbb{N}$. Luego

$$X^n = (X-1)^2 q + nX + (1-n)$$

para todo $n \in \mathbb{N}$. Si evaluamos este polinomio en A y utilizamos el teorema de Hamilton–Cayley, que implica que $(A - I)^2 = 0$, obtenemos

$$A^n = nA + (1-n)I = \begin{pmatrix} 1-n & -n \\ n & n+1 \end{pmatrix}.$$

6.6.4. Corolario. *Sea $A \in \mathbb{K}^{n\times n}$ una matriz inversible. Entonces A^{-1} es combinación lineal de las matrices $I, A, A^2, \dots, A^{n-1}$.*

Demostración. Supongamos que

$$\chi_A = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0.$$

Sabemos que

$$a_0 = \chi_A(0) = \det(0I - A) = \det(-A) = (-1)^n \det A$$

es distinto de cero pues A es inversible. Sea

$$B = \frac{-1}{a_0}\left(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2 A + a_1\right).$$

Como $\chi_A(A) = 0$, se tiene entonces que $AB = BA = I$. Luego $A^{-1} = B$. □

6.6.5. Ejemplo. Sea C la matriz compañera del polinomio

$$p = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \in \mathbb{K}[X].$$

Demostremos que $m_C = p$. Sea $\{e_1, \dots, e_n\}$ la base canónica de $\mathbb{K}^{n\times 1}$. Por inducción se demuestra fácilmente que

$$e_k = Ce_{k-1} = C^{k-1}e_1, \quad k \in \{2, \dots, n\}.$$

Tenemos entonces que $\{e_1, ce_1, \dots, C^{n-1}e_1\}$ es un conjunto linealmente independiente y luego $\{I, A, \dots, A^{n-1}\}$ es también linealmente independiente. Esto nos dice que $\deg m_A \geq n$. Como χ_A es un polinomio de grado n, el minimal m_A divide al característico χ_A por el teorema de Cayley–Hamilton, y m_A y χ_A son polinomios mónicos, $m_A = \chi_A = p$.

Capítulo 7

Forma de Jordan

7.1. Subespacios invariantes

7.1.1. Sea V un espacio vectorial y sea $f \in \hom(V,V)$. Un subespacio $S \subseteq V$ es **f-invariante** si $f(S) \subseteq S$.

7.1.2. Ejemplos. Sea $f \in \hom(V,V)$. Entonces:

1) $\{0\}$ y V entonces son f-invariantes.
2) $\ker f$ e $\operatorname{im} f$ son subespacios f-invariantes.
3) $S \subseteq V$ es un subespacio f-invariante de dimensión 1 si y sólo si $S = \langle v \rangle$ para algún autovector v de f.

7.1.3. Ejemplo. Sea $f\colon \mathbb{R}^4 \to \mathbb{R}^4$ dada por

$$f(x_1,x_2,x_3,x_4) = (x_1, x_1+x_2, 2x_3, x_4),$$

y sea $\{e_1,e_2,e_3,e_4\}$ la base canónica de $\mathbb{R}^4$. Algunos subespacios f-invariantes son $\langle e_1,e_2\rangle$, $\langle e_3\rangle$, $\langle e_4\rangle$ y $\langle e_1,e_2,e_4\rangle$.

7.1.4. Ejercicio. Sea $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ dada por $f(x,y) = (x, 2x+y)$. Demuestre que $S = \langle (0,1)\rangle$ es el único subespacio de dimensión 1 que es f-invariante.

7.1.5. Ejercicio. Sea V un espacio vectorial y sea $f \in \hom(V,V)$. Demuestre que si $S,T \subseteq V$ son subespacios f-invariantes entonces $S\cap T$ y $S+T$ son f-invariantes.

7.1.6. Si V es un espacio vectorial de dimensión finita, $f \in \hom(V,V)$ y $S \subseteq V$ es un subespacio f-invariante entonces la restricción de f a S,

$$f|_S\colon S \to S, \quad s \mapsto f(s)$$

es una transformación lineal, es decir: $f|_S \in \hom(S,S)$.

7.1.7. Proposición (diagonalización simultánea). *Sea V de dimensión finita y sean $f,g \in \hom(V,V)$ tales que f y g son diagonalizables y $fg = gf$. Entonces existe una base de V donde las matrices de f y g son simultáneamente diagonalizables.*

Demostración. Como f es diagonalizable, $V = S(\lambda_1)\oplus\cdots\oplus S(\lambda_k)$, donde $S(\lambda_i) = \{v \in V : f(v) = \lambda_i v\}$. Cada $S(\lambda_i)$ es g-invariante pues

$$f(g(v)) = (fg)(v) = (gf)(v) = g(f(v)) = g(\lambda v) = \lambda g(v).$$

Luego $g|_{S(\lambda_i)} \in \hom(S(\lambda_i), S(\lambda_i))$ y además $g|_{S(\lambda_i)}$ conmuta con f. Como g es diagonalizable, para cada $i \in \{1,\dots,k\}$ existe una base de $S(\lambda_i)$ formada por autovectores de $g|_{S(\lambda_i)}$. Como estos autovectores son también autovectores de f, entonces $f|_{S(\lambda_i)}$ y $g|_{S(\lambda_i}$ son simultáneamente diagonalizables. □

7.1.8. Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f \in \hom(V,V)$. Si $S \subseteq V$ es f-invariante entonces*

1) *$m_{f|_S}$ divide a m_f.*
2) *$\chi_{f|_S}$ divide a χ_f.*

Demostración. Supongamos que $\dim V = n$ y sea $\mathscr{B}_S = \{v_1, \dots, v_s\}$ una base de S. Si extendemos esta base de S a una base

$$\mathscr{B} = \{v_1, \dots, v_s, v_{s+1}, \dots, v_n\}$$

de V entonces $m_{f|_S} = \operatorname{mcm}(m_{v_1}, \dots, m_{v_s})$ y $m_f = \operatorname{mcm}(m_{v_1}, \dots, m_{v_n})$. Como entonces m_{v_i} divide a m_f para cada $i \in \{1, \dots, s\}$, se tiene entonces que $m_{f|_S}$ divide a m_f.

Para demostrar la segunda afirmación, si escribamos a f en la base $\mathscr{B}$ obtenemos la siguiente matriz por bloques:

$$[f]_{\mathscr{B},\mathscr{B}} = \begin{pmatrix} A & \star \\ 0 & \star \end{pmatrix},$$

donde $A = [f|_S]_{\mathscr{B}_S,\mathscr{B}_S}$. Existe entonces un polinomio $q \in \mathbb{K}[X]$ tal que $\chi_f = \chi_{f|_S} q$, tal como queríamos demostrar. □

7.1.9. Sea V un espacio vectorial y sea $f \in \hom(V, V)$. Si $S \subseteq V$ es f-invariante, un subespacio $T \subseteq V$ es un **complemento invariante** para S si T es f-invariante y $S \oplus T = V$.

7.1.10. Ejemplo. No siempre existen complementos invariantes: si

$$f\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad (x, y) \mapsto (0, x)$$

entonces $S = \langle (0, 1) \rangle$ es f-invariante pero no admite complemento invariante pues todo autovector de f pertenece a S.

block:XY 7.1.11. Sean V un espacio vectorial de dimensión finita, $f \in \hom(V, V)$ y $S, T \subseteq V$ subespacios f-invariantes tales que $S \oplus T = V$. Si $\mathscr{B}_S = \{v_1, \dots, v_s\}$ es una base de S y $\mathscr{B}_T = \{w_1, \dots, w_t\}$ es una base de T, entonces

$$\mathscr{B} = \{v_1, \dots, v_s, w_1, \dots, w_t\}$$

es una base de V. La matriz de f en la base $\mathscr{B}$ es la siguiente matriz por bloques:

$$[f]_{\mathscr{B},\mathscr{B}} = \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix},$$

donde $X = [f|_S]_{\mathscr{B}_S,\mathscr{B}_S}$ y $Y = [f|_T]_{\mathscr{B}_T,\mathscr{B}_T}$.

xca:f_invariante 7.1.12. Ejercicio. Sea V un espacio vectorial de dimensión finita, sea $f \in \hom(V, V)$ y sean S y T subespacios f-invariantes tales que $S \oplus T = V$. Demuestre que entonces:

1) $\chi_f = \chi_{f|_S} \chi_{f|_T}$.
2) $m_f = \operatorname{mcm}(m_{f|_S}, m_{f|_T})$.

7.1.13. Sean V un espacio vectorial, $v \in V$ y $f \in \hom(V, V)$. El subespacio $C(v) = \langle v, f(v), f^2(v), \dots \rangle$ se denomina el **subespacio cíclico** con respecto a f generado por v. Queda como ejercicio demostrar que $C(v)$ es un subespacio f-invariante.

em:dimC(v)=degm_v 7.1.14. Lema. *Sea V de dimensión finita y sea $f \in \hom(V, V)$. Si $v \in V \setminus \{0\}$ entonces $\deg m_v = \dim C(v)$. En particular, si $\deg m_v = d$ entonces el conjunto $\{v, f(v), \dots, f^{d-1}(v)\}$ es base de $C(v)$.*

Demostración. Supongamos que $\deg m_v = d$. Todo elemento de $C(v)$ es de la forma $p(f)(v)$ para algún $p \in \mathbb{K}[X]$. Si $p \in \mathbb{K}[X]$, el algoritmo de división implica que existen $r, q \in \mathbb{K}[X]$ tales que $p = q m_v + r$, donde $r = 0$ o $\deg r < \deg m_v$. Al especializar esta igualdad en f y evaluar en v,

$$p(f)(v) = q(f) m_v(f)(v) + r(f)(v) = r(f)(v),$$

y entonces $r(f)(v) \in C(v)$. Si escribimos $r = \sum_{i=0}^{d-1} \gamma_i X^i$ entonces hemos demostrado que todo elemento de $C(v)$ puede escribirse como combinación lineal de $\{v, \dots, f^{d-1}(v)\}$ Para ver que $\{v, f(v), \dots, f^{d-1}(v)\}$ es linealmente independiente basta observar que en caso contrario existiría un polinomio no nulo de grado $< d$ que anula a v con respecto a f, algo que contradice la minimalidad de m_v. □

xca:auxiliar

7.1.15. Ejercicio. Sea V de dimensión finita, sea $f \in \hom(V,V)$ y sea $p \in \mathbb{K}[X]$. Demuestre las siguientes afirmaciones:

1) $p(C(v)) = C(p(v))$.
2) Si $V = V_1 \oplus \cdots \oplus V_k$ donde cada V_i es p-invariante, entonces $p(V) = p(V_1) \oplus \cdots \oplus p(V_k)$.
3) Si $m_v = m_w$ entonces $m_{p(v)} = m_{p(w)}$ y
$$\dim C(p(v)) = \dim C(p(w)).$$

7.1.16. Lema. *Sea V de dimensión finita y sea $f \in \hom(V,V)$. Entonces existe $k \leq \dim V$ tal que $\{v, f(v), f^2(v), \dots, f^{k-1}(v)\}$ es base de $C(v)$. Más aún, la matriz de f con respecto a esa base es:*
$$\begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{k-1} \end{pmatrix},$$
donde $f^k(v) = \sum_{i=0}^{k-1} a_i f^i(v)$.

Demostración. Sea k el menor entero positivo tal que el conjunto $\{v, f(v), \dots, f^{k-1}(v)\}$ es linealmente independiente y
$$f^k(v) \in \langle v, f(v), \dots, f^{k-1}(v)\rangle.$$
Vamos a demostrar que $f^m(v) \in \langle v, f(v), \dots, f^{k-1}(v)\rangle$ para todo $m \geq k$. Procederemos por inducción en m. Como el caso $m = k$ es trivial, suponemos que el resultado es válido para algún $m \geq k$. Si
$$f^m(v) \in \langle v, f(v), \dots, f^{k-1}(v)\rangle$$
entonces existen $\alpha_0, \dots, \alpha_{k-1} \in \mathbb{K}$ tales que
$$f^m(v) = \sum_{i=0}^{k-1} \alpha_i f^i(v).$$
Al aplicar f se obtiene entonces que
$$\begin{aligned} f^{m+1}(v) &= f\left(\sum_{i=0}^{k-1} \alpha_i f^i(v)\right) \\ &= \sum_{i=0}^{k-2} \alpha_i f^{i+1}(v) + \alpha_{k-1} f^k(v) \in \langle v, f(v), \dots, f^{k-1}(v)\rangle. \end{aligned}$$
Luego $k \leq \dim V$. Un cálculo directo muestra que la matriz de f en la base $\{v, f(v), \dots, f^{k-1}(v)\}$ tiene la forma deseada. □

espacios_ciclicos

7.1.17. Teorema (forma racional de Frobenius). *Sea V de dimensión finita y sea $f \in \hom(V,V)$. Entonces existen vectores no nulos $v_1, \dots, v_k \in V$ y polinomios mónicos $p_1, \dots, p_k \in \mathbb{K}[X]$ tales que*
$$V = C(v_1) \oplus \cdots \oplus C(v_k),$$
$\deg p_i = \dim C(v_i)$, $p_i(f)(v_i) = 0$ *para todo $i \in \{1, \dots, k\}$ y además p_i divide a p_{i-1} para todo $i \in \{2, \dots, k\}$. Más aún, los polinomios $p_1, \dots, p_k$ están unívocamente determinados.*

Demostración. Procederemos por inducción en $\dim V$. Como el caso $\dim V = 1$ es trivial, vamos a suponer que el resultado es válido para todo endomorfismo de un espacio vectorial de dimensión $< \dim V$. Sea m la mayor dimensión que puede tener un subespacio cíclico asociado a f. Entonces existe $v_1 \in V$ tal que $\dim C(v_1) = m$ y $\dim C(v) \leq m$ para todo $v \in V$, es decir: $\{v_1, f(v_1), \dots, f^{m-1}(v_1)\}$ es linealmente independiente y $f^m(v) \in \langle v, f(v), \dots, f^{m-1}(v)\rangle$ para todo $v \in V$. Si $m = \dim V$, no hay nada que demostrar. Supongamos entonces que $m < \dim V$. Vamos a construir un complemento de $C(v_1)$ que sea f-invariante. Sea T la transformación lineal definida por

$$T\colon V \to \mathbb{K}^{m\times 1}, \quad v \mapsto \begin{pmatrix} \varphi(v) \\ \varphi(f(v)) \\ \vdots \\ \varphi(f^{m-1}(v)) \end{pmatrix},$$

donde $\varphi\colon V \to \mathbb{K}$ es una funcional lineal que satisface

$$\varphi(f^i(v_1)) = \begin{cases} 1 & \text{si } i = m-1, \\ 0 & \text{si } i \in \{0, \dots, m-2\}. \end{cases}$$

(La existencia de una φ que cumple lo pedido queda garantizada por el siguiente argumento: si extendemos $\{v_1, f(v_1), \dots, f^{m-1}(v_1)\}$ a una base de V podemos tomar como φ el vector de la base dual que vale uno en $f^{m-1}(v_1)$ y cero en el resto de los vectores.)

La restricción $T|_{C(v_1)}\colon C(v_1) \to \mathbb{K}^{m\times 1}$ es un isomorfismo. En efecto, si $\{e_1, \dots, e_m\}$ es la base canónica de $\mathbb{K}^{m\times 1}$, un cálculo sencillo muestra que

$$T|_{C(v_1)}(f^i(v_1)) = e_{m-i} + \sum_{j=0}^{i-1} \gamma_{m-j} e_{m-j}.$$

Luego $T|_{C(v_1)}$ es epimorfismo y como $\dim C(v_1) = m$ entonces $T|_{C(v_1)}$ es un isomorfismo.

El subespacio $\ker T$ es f-invariante. Para demostrar esta afirmación hay que ver que $f(\ker T) \subseteq \ker T$. Si $v \in \ker T$ entonces $T(v) = 0$ y luego $T(f^i(v)) = 0$ para todo $i \in \{0, \dots, m-1\}$. Por otro lado,

$$T(f(v)) = \begin{pmatrix} \varphi(f(v)) \\ \varphi(f^2(v)) \\ \vdots \\ \varphi(f^{m-1}(v)) \\ \varphi(f^m(v)) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \varphi(f^m(v)) \end{pmatrix}$$

y $\varphi(f^m(v)) = 0$ pues $f^m(v) \in \langle v, f(v), \dots, f^{m-1}(v)\rangle$.

Afirmamos que $V = C(v_1) \oplus \ker T$. Veamos que $C(v_1) \cap \ker T = \{0\}$. Si $v \in C(v_1) \cap \ker T$ entonces $v \in \ker T|_{C(v_1)} = \{0\}$ pues la restricción $T|_{C(v_1)}$ es un isomorfismo. Para demostrar que $V = C(v_1) + \ker T$ basta observar que

$$\begin{aligned} \dim V &= \dim \ker T + \dim \operatorname{im} T = \dim \ker T + m \\ &= \dim \ker T + \dim C(v_1) = \dim(\ker T + \dim C(v_1)). \end{aligned}$$

Afirmamos que $p_1(f)(v) = 0$ para todo $v \in V$. Como $V = C(v_1) \oplus \ker T$ y $p_1(f)(f^i(v)) = 0$ para todo $i \geq 0$ pues $p(f) \circ f = f \circ p(f)$, es necesario demostrar que $p_1(f)(w) = 0$ para todo $w \in \ker T$. Si $w \in \ker T$ entonces existen $\gamma_0, \dots, \gamma_{m-1} \in \mathbb{K}$ tales que

$$f^m(v_1 + w) = \sum_{i=0}^{m-1} \gamma_i f^i(v_1 + w) = \sum_{i=0}^{m-1} \gamma_i f^i(v_1) + \sum_{i=0}^{m-1} \gamma_i f^i(w).$$

Luego, como $\ker T$ es f-invariante,

$$f^m(v_1) - \sum_{i=0}^{m-1} \gamma_i f^i(v_1)) = \sum_{i=0}^{m-1} \gamma_i f^i(w) - f^m(w) \in C(v_1) \cap W = \{0\}.$$

Como $\sum_{i=0}^{m-1} \gamma_i f^i(v_1) = f^m(v_1) = \sum_{i=0}^{m-1} \alpha_i f^i(v_1)$, entonces $\alpha_i = \gamma_i$ para todo $i \in \{0, \dots, m-1\}$. Tenemos entonces

$$0 = p_1(f)(v_1 + w) = p_1(f)(v_1) + p_1(f)(w) = p_1(f)(w)$$

para todo $w \in \ker T$.

Hemos demostrado que $p_1 = m_{v_1}$ y que $V = C(v_1) \oplus \ker T$, donde $C(v_1)$ y $\ker T$ son subespacios f-invariantes. Esto nos permite descomponer a f en una matriz por bloques

$$\begin{pmatrix} C_{p_1} & \\ & [f|_{\ker T}] \end{pmatrix}.$$

Además, como $\dim \ker T < \dim V$, la hipótesis inductiva implica que $\ker T = C(v_2) \oplus \cdots \oplus C(v_k)$ y que la matriz $[f|_{\ker T}]$ puede descomponerse como suma directa de matrices compañeras de ciertos polinomios mónicos $p_2, \dots, p_k \in \mathbb{K}[X]$ tales que:

1) $p_i(f|_{kerT})(v_i) = 0$ para cada $i \in \{2, \dots, k\}$,
2) $\deg p_i = \dim C(v_i)$ para cada $i \in \{2, \dots, k\}$,
3) p_j divide a p_{j-1} para cada $j \in \{3, \dots, k\}$.

Demostremos que $p_i(f)(v_i) = 0$ para cada $i \in \{1, \dots, k\}$. Ya vimos que $p_1(f)(v_1) = 0$. Sea entonces $i \in \{2, \dots, k\}$. Como $v_i \in \ker T$,

$$p_i(f)(v_i) = p_i\left(f|_{\ker T}\right)(v_i) = 0$$

por hipótesis inductiva.

Demostraremos que p_2 divide a p_1. Sea $d = \deg p_2 \leq \deg p_1$. Entonces, como $d = \dim C(v_2)$, el conjunto $\{v_2, f(v_2), \dots, f^{d-1}(v_2)\}$ es linealmente independiente. Por el algoritmo de división, existen $q, r \in \mathbb{K}[X]$ tales que $p_1 = qp_2 + r$, donde $\deg r < \deg p_2 = d$ si $r \neq 0$. Al especializar en f y evaluar en v_2 se obtiene:

$$0 = p_1(f)(v_2) = q(f)p_2(f)(v_2) + r(f)(v_2) = r(f)(v_2).$$

En particular, existen entonces $\beta_0, \dots, \beta_{d-1} \in \mathbb{K}$ tales que

$$0 = r(f)(v_2) = \sum_{i=0}^{d-1} \beta_i f^i(v_2).$$

Como el conjunto $\{v_2, f(v_2), \dots, f^{d-1}(v_2)\}$ es linealmente independiente, entonces $\beta_j = 0$ para todo $j \in \{0, \dots, d-1\}$ y luego $r = 0$.

Veamos la unicidad de los polinomios $p_1, \dots, p_k$. Supongamos que se tiene otra descomposición en subespacios cíclicos

$$V = C(w_1) \oplus \cdots \oplus C(w_l)$$

cuyos polinomios $q_1, \dots, q_l \in \mathbb{K}[X]$ satisfacen:

1) $q_i(f)(w_i) = 0$ para cada $i \in \{1, \dots, l\}$,
2) $\deg q_i = \dim C(w_i)$ para cada $i \in \{1, \dots, l\}$,
3) q_j divide a q_{j-1} para cada $j \in \{2, \dots, l\}$.

Vamos a demostrar que $k = l$ y que $p_i = q_i$ para todo $i \in \{1, \dots, k\}$. Como q_j divide a q_{j-1} para todo $j \in \{2, \dots, l\}$ entonces $\dim C(w_1) \geq \dim C(w_j)$ para todo $j \in \{2, \dots, l\}$. Primero observamos que $p_1 = q_1 = m_f$. En efecto, como $p_1(f)(v) = 0$ para todo $v \in V$ entonces $p_1(f) = 0$ y luego $\deg m_f \leq p_1$. Por otro lado, como $\{v_1, f(v_1), \dots, f^{m-1}(v_1)\}$ es linealmente independiente, el conjunto $\{\mathrm{id}_V, f, \dots, f^{m-1}\}$ es también linealmente independiente, y luego $\deg m_f \geq m = \deg p_1$. Como m_f y p_1 son polinomios mónicos, concluimos que $m_f = p_1$. Como q_j divide a q_1 para todo $j \in \{1, \dots, l\}$ entonces $q_1(f)(w_j) = 0$ para todo $j \in \{1, \dots, l\}$. Además, como $q_1(f) \circ f = f \circ q_1(f)$, entonces $q_1(f)(v) = 0$ para todo $v \in V$. Tal como se demostró que $p_1 = m_f$, se demuestra que $q_1 = m_f = p_1$.

Supongamos que $k \geq 2$. Como $p_1 = q_1$,

$$\dim C(w_1) = \dim C(v_1) < \dim V$$

y luego $l \geq 2$. El ejercicio 7.1.15 implica que de la dos descomposiciones que se tienen para V se obtiene que

$$p_2(V) = \bigoplus_{i=1}^{k} p_2(C(v_i)) = \bigoplus_{i=1}^{l} p_2(C(w_i)).$$

Como $\dim C(p_2(v_1)) = \dim C(p_2(w_1))$ y como $p_2(f)(v_i) = 0$ para todo $i \in \{2, \dots, k\}$, entonces $\dim C(p_2(w_j)) = 0$ para todo $j \in \{2, \dots, l\}$. Como en particular $p_2(w_2) = 0$, se concluye que q_2 divide a p_2. Similarmente se demuestra que p_2 divide a q_2. Como los polinomios p_2 y q_2 son mónicos, entonces $p_2 = q_2$. $\square$

7.1.18. La descomposición de V del teorema 7.1.17 implica la existencia de una base de V en la que la matriz de f tiene la forma

$$\begin{pmatrix} C_{p_1} & & & \\ & C_{p_2} & & \\ & & \ddots & \\ & & & C_{p_k} \end{pmatrix},$$

donde las matrices $C_{p_1}, \dots, C_{p_k}$ son las matrices compañeras de los polinomios $p_1, \dots, p_k$ respectivamente. Recordemos que si $p = X^m - \sum_{i=0}^{m-1} a_i X^i$ entonces

$$C_p = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_m \end{pmatrix}.$$

Es evidente además que $\chi_f = p_1 \cdots p_k$. Los polinomios $p_1, \dots, p_k$ del teorema 7.1.17 son los **factores invariantes** de f.

7.1.19. Ejemplo. Sea

$$f\colon \mathbb{R}^3 \to \mathbb{R}^3, \quad f(x, y, z) = (5x - 6y - 6z, -x + 4y + 2z, 3x - 6y - 4z).$$

Un cálculo directo muestra que

$$\chi_f = (X-1)(X-2)^2, \quad m_f = (X-1)(X-2).$$

La forma racional aplicada a f implica que $\mathbb{R}^3 = C(v_1) \oplus C(v_2)$, donde $\dim C(v_1) = 2$ y $\dim C(V_2) = 1$. Los factores invariantes son entonces $p_1 = m_f$ y $p_2 = X - 2$. Luego, con respecto a la base $\{v_1, f(v_1), v_2\}$, la transformación lineal f tendrá la forma

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

Es fácil encontrar el vector v_1. En efecto, basta tomar cualquier v_1 que no sea autovector. Por ejemplo, si $v_1 = (1, 1, 1)$ entonces $f(v_1) = (-7, 5, -7)$ y $\{v_1, f(v_1)\}$ es linealmente independiente. Para encontrar v_2 necesitamos un autovector, por ejemplo $f(v_2) = (2, 1, 0)$. En la base

$$\{(1, 1, 1), (-7, 5, -7), (2, 1, 0)\}$$

la matriz de f tiene la forma deseada.

hm:CayleyHamilton 7.1.20. Corolario (Cayley–Hamilton). *Sea V de dimensión finita y sea $f \in \hom(V, V)$. Entonces $\chi_f(f) = 0$, es decir: m_f divide a χ_f.*

Demostración. Si utilizamos la forma racional sobre f, tenemos que existen $p_1, \dots, p_k \in \mathbb{K}[X]$ tales que $\chi_f = p_1 \dots p_k$ y además $p_1 = m_f$. Luego m_f divide a χ_f. $\square$

7.1.21. Ejemplo. Las matrices $E_{21} \in \mathbb{C}^{4\times 4}$ y $E_{21} + E_{43} \in \mathbb{C}^{4\times 4}$ tienen a X^4 como polinomio característico y a X^2 como polinomio minimal. Sin embargo, tienen disinta forma racional.

7.2. Endomorfismos nilpotentes

7.2.1. Sea V un espacio vectorial y sea $f \in \hom(V, V)$. Entonces f es **nilpotente** si existe $m \in \mathbb{N}$ tal que $f^m = 0$.

Si $f \in \hom(V, V)$ es nilpotente, el número

$$r = \min\{m \in \mathbb{N} : f^m = 0\}$$

se denomina **índice de nilpotencia** de f.

7.2.2. Ejemplo. Sea

$$f\colon \mathbb{R}^3 \to \mathbb{R}^3, \quad (x, y, z) \mapsto (-x + 2y + z, 0, -x + 2y + z).$$

Entonces f es nilpotente con índice de nilpotencia igual a dos.

7.2.3. Ejemplo. Sea V un espacio vectorial de dimensión finita n y sea $\{v_1, \dots, v_n\}$ una base de V. La aplicación $f\colon V \to V$ definida por

$$f(v_i) = \begin{cases} v_{i+1} & \text{si } i < n, \\ 0 & \text{si } i = n, \end{cases}$$

es nilpotente de índice n pues $f^n = 0$ y $f^{n-1} \neq 0$.

r_diagonalizable=0

7.2.4. Ejercicio. Sea V un espacio vectorial de dimensión finita y sea $f \in \hom(V, V)$ nilpotente y diagonalizable. Demuestre que $f = 0$.

tente:autovalores

7.2.5. Ejercicio. Sea V un espacio vectorial y sea $f \in \hom(V, V)$ nilpotente. Demuestre que $\operatorname{spec} f = \{0\}$.

7.2.6. Proposición. *Sea V de dimensión finita y $f \in \hom(V, V)$. Entonces f es nilpotente de índice k si y sólo si $m_f = X^k$.*

Demostración. Si $f^k = 0$ y $f^{k-1} \neq 0$ entonces X^k anula a f. Luego m_f divide a X^k y entonces $m_f = X^j$ para algún $j \in \{1, \dots, k\}$. Como $f^{k-1} \neq 0$ entonces $m_f = X^k$. Recíprocamente, si $m_f = X^k$ entonces, trivialmente, $f^k = 0$ y $f^{k-1} \neq 0$. ☐

7.2.7. Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f \in \hom(V, V)$ nilpotente de índice r. Sea $v \in V$ tal que $f^{r-1}(v) \neq 0$. Entonces $\{v, f(v), \dots, f^{r-1}(v)\}$ es linealmente independiente. En particular, $r \leq \dim V$.*

Demostración. Si $r = 1$ entonces el resultado es válido pues $v \neq 0$. Si $r \geq 2$ entonces sean $a_0, \dots, a_{r-1} \in \mathbb{K}$ tales que $\sum_{i=0}^{r-1} a_i f^i(v) = 0$. Al aplicar f^{r-1} obtenemos

$$0 = f^{r-1}\left(\sum_{i=0}^{r-1} a_i f^i(v)\right) = \sum_{i=0}^{r-1} a_i f^{r-1+i}(v) = a_0 f^{r-1}(v).$$

pues $f^{r-1+i}(v) = 0$ si $i > 0$. Como $f^{r-1}(v) \neq 0$ entonces $a_0 = 0$. Queda entonces $a_1 f(v) + \cdots + a_{r-1} f^{r-1}(v) = 0$. Al aplicar f^{r-2} en esta igualdad se obtiene, tal como se hizo antes, que $a_1 = 0$. Este proceso nos da entonces que $a_i = 0$ para todo i. ☐

7.2.8. Ejemplo. Si V tiene dimensión n y $f \in \hom(V, V)$ es nilpotente de índice n entonces $\{v, f(v), \dots, f^{n-1}(v)\}$ es una base de V. Luego, en esa base, la matriz f es de la forma

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

y se denomina **bloque de Jordan nilpotente** de f.

Jordan:nilpotente

7.2.9. Teorema. *Sea V de dimensión finita y sea $f \in \hom(V,V)$ nilpotente de índice r. Entonces existe una base de V tal que f en esa base es de la forma*

$$\begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{pmatrix}$$

donde cada J_i es un bloque de Jordan nilpotente de tamaño $n_i \times n_i$ y además $r = n_1 \geq n_2 \geq \cdots \geq n_k \geq 1$.

Demostración. La forma racional de Frobenius, teorema 7.1.17, implica que existen $v_1, \dots, v_k \in V$ y polinomios mónicos $p_1, \dots, p_k \in \mathbb{K}[X]$ tales que $V = C(v_1) \oplus \cdots \oplus C(v_k)$, $\deg p_i = \dim C(v_i)$ y $p_i(f)(v_i) = 0$ para todo $i \in \{1, \dots, k\}$ y p_i divide a p_{i-1} para todo $i \in \{2, \dots, k\}$. Como f es nilpotente de índice r, entonces $p_1 = m_f = X^r$. Luego todo p_i es de la forma X^{n_i}, donde $n_1 = r$, $n_1 \geq n_2 \geq \cdots \geq n_k \geq 1$. La matriz de f asociada a esta descomposición en subespacios f-invariantes es entonces de la forma deseada. □

block:rg(J^k)

7.2.10. Observemos que si J es una matriz de Jordan nilpotente de $n \times n$ entonces $\operatorname{rg}(J^i) = n - i$ para todo $i \in \{0, \dots, n\}$. En efecto, sea $\{v_1, \dots, v_n\}$ una base de $\mathbb{K}^{n \times 1}$. Entonces la aplicación lineal $J \colon \mathbb{K}^{n\times 1} \to \mathbb{K}^{n \times 1}$ dada por $x \mapsto Jx$ satisface

$$J(v_i) = \begin{cases} v_{i+1} & \text{si } i < n, \\ 0 & \text{si } i = n. \end{cases}$$

Entonces J tiene rango uno. Por inducción es fácil ver que

$$J^k(v_i) = \begin{cases} v_{i+k} & \text{si } i + k - 1 < n, \\ 0 & \text{en otro caso.} \end{cases}$$

y luego $\dim \operatorname{im} J^k = n - k$ y $\operatorname{rg}(JP^k) = n - k$.

Jordan:nilpotente

7.2.11. Corolario. *Sea V de dimensión finita y $f \in \hom(V,V)$ nilpotente de índice r. Valen las siguientes afirmaciones:*

1) *La cantidad de bloques de Jordan de f es $\dim \ker f = n - \dim \operatorname{im} f$.*
2) *El bloque de Jordan de f de tamaño máximo es de r.*
3) *Para cada $i \in \{1, \dots, r-1\}$ la cantidad de bloques de Jordan de f de tamaño $\geq i$ es $\dim \operatorname{im} f^i - \dim \operatorname{im} f^{i+1}$.*
4) *Para cada $i \in \{1, \dots, r-1\}$, la cantidad de bloques de Jordan de tamaño i es $\dim \operatorname{im} f^{i+1} - 2 \dim \operatorname{im} f^i + \dim f^{i-1}$.*

Demostración. Para demostrar la primera afirmación observemos que si hay k bloques de Jordan entonces $n = \dim \operatorname{im} f + k$ pues cada bloque de Jordan de tamaño $d \times d$ tiene rango igual a $d - 1$.

Para demostrar la segunda afirmación obvervamos que $m_f = X^r$ y que el primer bloque de Jordan tiene tamaño igual a $\deg m_f$ pues el primer factor invariante de f es m_f.

Para la tercera afirmación, para cada $i \in \{1, \dots, r-1\}$. Sea n_i la cantidad de bloques de Jordan de tamaño i. Entonces, por lo visto en 7.2.10, si agrupamos todos los bloques de tamaño j para cada $j \in \{i, \dots, r-1\}$ obtenemos:

$$\dim \operatorname{im} f^i - \dim \operatorname{im} f^{i+1} = \sum_{j=i+1}^{k} n_j(j-i) - \sum_{j=i+2}^{k} n_j(j-i+1) = \sum_{j=i+1}^{k} n_j,$$

que es la cantidad de bloques de tamaño $\geq i$.

Por ultimo, por la fórmula vista en el ítem anteior, la cantidad de bloques de Jordan de f de tamaño i es

$$\begin{aligned} n_i &= \sum_{j=i}^{k} n_j - \sum_{j=i+1}^{k} n_j \\ &= \dim\operatorname{im} f^{i-1} - \dim\operatorname{im} f^{i} - (\dim\operatorname{im} f^{i} - \dim\operatorname{im} f^{i+1}), \end{aligned}$$

que es lo que queríamos demostrar. $\square$

7.2.12. Ejemplo. Demostremos que no existe una matriz $A \in \mathbb{R}^{15\times 15}$ tal que $\operatorname{rg}(A) = 10$, $\operatorname{rg}(A^4) = 3$ y $\operatorname{rg}(A^5) = 0$. Por el corolario anterior, como A es nilpotente, A tiene una forma de Jordan que contiene $15 - \operatorname{rg}(A) = 5$ bloques, donde el bloque más grande es de 5×5. Además A tiene

$$\operatorname{rg}(A^6) - 2\operatorname{rg}(A^5) + \operatorname{rg}(A^4) = 3$$

bloques de tamaño 5, una contradicción.

7.2.13. Ejemplo. Sea

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 \end{pmatrix} \in \mathbb{R}^{6\times 6}.$$

Entonces $\chi_A = X^6$, $m_A = X^3$ y A es nilpotente de índice tres. Como $\operatorname{rg}(A) = 3$ entonces $\dim\ker A = 3$ y luego, por el corolario 7.2.11, la forma de Jordan de A tendrá tres bloques de Jordan. Como A es nilpotente de índice tres, el mayor de estos bloques de Jordan será de 3×3. La forma de Jordan de A será entonces

$$J = \begin{pmatrix} J_1 & & \\ & J_2 & \\ & & J_3 \end{pmatrix}, \qquad J_1 = \begin{pmatrix} 0 & & \\ 1 & 0 & \\ 0 & 1 & 0 \end{pmatrix}, \qquad J_2 = \begin{pmatrix} 0 & \\ 1 & 0 \end{pmatrix}, \qquad J_3 = (0)\,.$$

Sea $\{e_1, \dots, e_6\}$ la base canónica de $\mathbb{R}^{6\times 1}$. Un cálculo directo muestra que $\dim C(e_1) = 3$, $\dim C(e_2) = 2$, $\dim C(e_3) = 1$ y que además $\{e_1, Ae_1, A^2e_1, e_2, Ae_2, a_3\}$ es una base de $\mathbb{R}^{6\times 1}$. En esa base, la matriz de la transformación lineal $\mathbb{R}^{6\times 1} \to \mathbb{R}^{6\times 1}$ dada por $x \mapsto Ax$, es la matriz J.

7.3. Descomposición primaria

7.3.1. Lema. *Sea V de dimensión finita, $f \in \hom(V, V)$ y $p \in \mathbb{K}[X]$. Entonces $\ker p(f)$ es un subespacio f-invariante.*

Demostración. Tenemos que demostrar que $f(\ker p(f)) \subseteq \ker p(f)$. Sea $v \in \ker p(f)$. Entonces

$$p(f)\,(f(v)) = (p(f) \circ f)(v) = (f \circ p(f))(v) = f\,(p(f)(v)) = f(0) = 0,$$

tal como se quería demostrar. $\square$

lem:V=kerp+kerq 7.3.2. Lema. *Sea V de dimensión finita y sea $f \in \hom(V, V)$. Supongamos que $m_f = pq$ donde $\gcd(p, q) = 1$. Entonces*

1) *$V = \ker p(f) \oplus \ker q(f)$.*
2) *Si $f_p = f|_{\ker p(f)}$ y $f_q = f|_{\ker q(f)}$ entonces $m_{f_p} = p$ y $m_{f_q} = q$.*

Demostración. Demostremos la primera afirmación. Sean $r, s \in \mathbb{K}[X]$ tales que $rp + sq = 1$. Al especializar en f obtenemos

eq:rp+sq=1 (7.3.3) $$r(f)p(f) + s(f)q(f) = \mathrm{id}_V\,.$$

Veamos que $V = \ker p(f) + \ker q(f)$. Sea $v \in V$. La fórmula (7.3.3) nos permite escribir $v = v_1 + v_2$, donde $v_1 = r(f)p(f)(v)$ y $v_2 = s(f)q(f)(v)$. Es fácil comprobar que $v_1 \in \ker q(f)$ y que $v_2 \in \ker p(f)$:

$$\begin{aligned} q(f)(v_1) &= q(f)r(f)p(f)(v) = r(f)p(f)q(f)(v) = r(f)m_f(f)(v) = 0, \\ p(f)(v_2) &= p(f)s(f)q(f)(v) = s(f)p(f)q(f)(v) = s(f)m_f(f)(v) = 0. \end{aligned}$$

Veamos ahora que $\ker p(f) \cap \ker q(f) = \{0\}$: si $p(f)(v) = q(f)(v) = 0$ entonces (7.3.3) implica que $v = 0$.

Para demostrar la segunda afirmación primero observemos que, como $p(f_p) = p(f|_{\ker p(f)}) = 0$, entonces el minimal de f_p divide a p. Por otro lado, como

$$V = \ker p(f) \oplus \ker q(f),$$

entonces $m_{f_p}(f)q(f) = 0$ pues $q(f)(v) = 0$ para todo $v \in \ker q(f)$ y

$$m_{f_p}(f)(v) = m_{f_p}(f|_{\ker p(f)})(v) = 0$$

para todo $v \in \ker p(f)$. Esto implica que $\deg m_{f_p} \geq \deg p$ y luego, como p y m_{f_p} son mónicos, $p = m_{f_p}$. De la misma forma se demuestra que $q = m_{f_q}$. □

posicion_primaria

7.3.4. Teorema (descomposición primaria). *Sea V de dimensión finita y sea $f \in \hom(V,V)$. Supongamos que $m = p_1^{r_1} \cdots p_k^{r_k}$ donde los p_i son polinomios mónicos e irreducibles. Entonces*

$$V = \ker p_1(f)^{r_1} \oplus \cdots \oplus \ker p_k(f)^{r_k}.$$

Demostración. Procederemos por inducción en k. Como el caso $k = 2$ es el lema 7.3.2, vamos a suponer que el resultado es válido para $k - 1 \geq 2$. Sean $p = p_1^{r_1} \cdots p_{k-1}^{r_{k-1}}$ y $q = p_k^{r_k}$. Entonces $m_f = pq$ y $\gcd(p, q) = 1$. Además

$$\ker q(f) = \ker p_k(f)^{r_k}.$$

Por el lema 7.3.2, $V = \ker p(f) \oplus \ker q(f)$ y el minimal de $f|_{\ker p(f)}$ es p. Por hipótesis inductiva,

$$\ker p(f) = \ker p_1(f)^{r_1} \oplus \cdots \oplus \ker p_{k-1}(f)^{r_{k-1}},$$

y esto completa la demostración. □

scJordanChevalley

7.3.5. Corolario (descomposición de Jordan–Chevalley). *Sea V de dimensión finita y sea $f \in \hom(V,V)$. Supongamos que $m_f = \prod_{i=1}^{k}(X - \lambda_i)^{m_i}$, donde $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces $f = g + h$, donde g es diagonalizable, h es nilpotente y $gh = hg$.*

Demostración. La descomposición primaria, teorema 7.3.4, implica que $V = V_1 \oplus V_k$, donde $V_i = \ker(f - \lambda_i \operatorname{id}_V)^{m_i}$. Para definir g y h basta con definirlas en cada V_i. Para $i \in \{1, \dots, k\}$ definimos entonces

$$\begin{aligned} g|_{V_i} &= \lambda \operatorname{id}_{V_i}, \\ h|_{V_i} &= f|_{V_i} - \lambda_i \operatorname{id}_{V_i}. \end{aligned}$$

Como para cada $i \in \{1, \dots, k\}$ se tiene que $(g + h)(v_i) = f(v_i)$ y cada v se escribe únívocamente como $v = v_1 + \cdots + v_k$ con $v_i \in V_i$, entonces $f = g + h$. Es evidente que g es diagonalizable y que $gh = hg$. Para ver que h es nilpotente, vamos a demostrar que si $n = \dim V$ entonces $h^n = 0$. En efecto, si $v = v_1 + \cdots + v_k$ con $v_i \in V_i$ para todo $i \in \{1, \dots, k\}$, entonces

$$h^n(v) = \sum_{i=1}^{k} h^n(v_i) = \sum_{i=1}^{k} (h|_{V_i})^n(v_i) = \sum_{i=1}^{k} (f|_{V_i} - \lambda_i \operatorname{id}_{V_i})^n(v_i) = 0$$

pues $m_i \leq n$ para todo $i \in \{1, \dots, k\}$. □

7.4. Forma de Jordan

7.4.1. TEOREMA (forma de Jordan). *Sea V un espacio vectorial de dimensión finita y sea $f \in \hom(V,V)$. Supongamos que $m_f = \prod_{i=1}^{k}(X-\lambda_i)^{r_i}$, donde $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces existe una base de V tal que en esa base f es diagonal por bloques donde cada bloque es de la forma*

$$\begin{pmatrix} \lambda_i & & & & \\ 1 & \lambda_i & & & \\ 0 & 1 & \lambda_i & & \\ \vdots & \vdots & \ddots & \ddots & \\ 0 & 0 & \cdots & 1 & \lambda_i \end{pmatrix}.$$

*Estos bloques se denominan **bloques de Jordan** de f.*

DEMOSTRACIÓN. Por la descomposición de Jordan–Chevalley, corolario 7.3.5, sabemos que existe g diagonalizable y h nilpotente tales que $f = g+h$ y $gh = hg$. Como g es diagonalizable, $V = \oplus_{i=1}^{k} S(\lambda_i)$, donde $S(\lambda) = \{v \in V : g(v) = \lambda v\}$.

Para cada $i \in \{1,\dots,k\}$ consideremos una base de $S(\lambda_i)$ que corresponda a la descomposición cíclica de $h|_{S(\lambda_i)}$. Como $h|_{S(\lambda_i)}$ es nilpotente, en esa base puede escribirse como

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

Entonces, en esa misma base, $f|_{S(\lambda_i)} = (g+h)|_{S(\lambda_i)}$ se escribe como

$$\begin{pmatrix} \lambda_i & & & & \\ 1 & \lambda_i & & & \\ 0 & 1 & \lambda_i & & \\ \vdots & \vdots & \ddots & \ddots & \\ 0 & 0 & \cdots & 1 & \lambda_i \end{pmatrix}.$$

Como $V = \oplus_{i=1}^{k} S(\lambda_i)$ y los $S(\lambda_i)$ son h-invariantes pues

$$g(h(v)) = (gh)(v) = (hg)(v) = h(g(v)) = h(\lambda_i v) = \lambda_i h(v),$$

el teorema queda demostrado. □

7.4.2. Sea V de dimensión finita y sea $f \in \hom(V,V)$. Vamos a demostrar que la forma de Jordan de f queda unívocamente determinada.

Si V tiene una base tal que f en esa base puede escribirse en forma de Jordan, digamos con bloques de Jordan $J_1,\dots,J_k$ tales que J_i es de tamaño $n_i \times n_i$ y tiene al escalar λ_i en su diagonal. Supongamos además que $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces, como $\chi_f = \prod_{i=1}^{k}(X-\lambda_i)^{n_i}$, tenemos que n_i es la multiplicidad de λ_i como raíz de χ_f. Luego los λ_i y los n_i quedan unívocamente determinados salvo por el orden con el que aparecen.

La forma de Jordan de f nos dice que $V = V_1 \oplus \cdots \oplus V_k$ donde los V_i son f-invariantes. Veamos que $V_i = \ker(f - \lambda_i \operatorname{id}_V)^n$, donde $n = \dim V$. Sea $i \in \{1,\dots,k\}$. La matriz $J_i - \lambda_i I$ es nilpotente y entonces $(J_i - \lambda_i I)^n = 0$. Por otro lado, si $j \neq i$, entonces $J_j - \lambda_i I$ es inversible pues es triangular y tiene elementos no nulos en la diagonal. Luego $V_i = \ker(f - \lambda_i \operatorname{id}_V)^n$, donde $n = \dim V$, y entonces los V_i están unívocamente determinados.

Para cada $i \in \{1,\dots,k\}$ sea $f_i = f|_{V_i}$. Entonces el bloque de Jordan J_i queda unívocamente determinado pues es la forma racional del endomorfismo $f_i - \lambda_i \operatorname{id}_{V_i}$.

7.4.3. COROLARIO. *Sea $A \in \mathbb{C}^{n\times n}$. Entonces existen matrices simétricas $B, C \in \mathbb{C}^{n\times n}$ tales que $A = BC$.*

DEMOSTRACIÓN. Existe $P \in \mathbb{C}^{n\times n}$ inversible tal que $A = PJP^{-1}$, donde J es la forma de Jordan de A. Si $J = KL$ con K y L son matrices simétricas, entonces, si $Q = P^T$, tenemos que A es producto de matrices simétricas pues

$$A = PJP^{-1} = PKLP^{-1} = (PKQ)(Q^{-1}LP^{-1}),$$

y las matrices PKQ y $Q^{-1}LP^{-1}$ son simétricas. Veamos entonces que toda matriz de Jordan puede escribirse como producto de dos matrices simétricas. Supongamos que J tiene bloques de Jordan $J_1, \dots, J_k$, donde J_p es de tamaño $n_p \times n_p$. Para cada $p \in \{1, \dots, k\}$ definimos $D_p \in \mathbb{C}^{n\times n}$ por

$$(D_p)_{ij} = \begin{cases} 1 & \text{si } i + j = n_p + 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces D_p es simétrica e inversible con $D_p^{-1} = D_p$. Sean

$$C = \begin{pmatrix} J_1D_1 & & & \\ & J_2D_2 & & \\ & & \ddots & \\ & & & J_kD_k \end{pmatrix}, \quad D = \begin{pmatrix} D_1 & & & \\ & D_2 & & \\ & & \ddots & \\ & & & D_k \end{pmatrix}.$$

Entonces C y D son simétricas y $J = CD$. $\square$

Capítulo 8

Espacios con producto interno

8.1. Definiciones básicas y ejemplos

8.1.1. Sea $\mathbb{K}$ el cuerpo $\mathbb{R}$ de los números reales o el cuerpo $\mathbb{C}$ de los números complejos. Sea V un espacio vectorial sobre $\mathbb{K}$. Una función $\langle\cdot,\cdot\rangle\colon V\times V\to\mathbb{K}$ es un **producto interno** si:

1) $\langle v_1+v_2,w\rangle=\langle v_1,w\rangle+\langle v_2,w\rangle$ para todo $v_1,v_2,v\in V$.
2) $\langle\lambda v,w\rangle=\lambda\langle v,w\rangle$ para todo $v\in V$ y $\lambda\in\mathbb{K}$.
3) $\langle v,w\rangle=\overline{\langle w,v\rangle}$ para todo $v,w\in V$.
4) $\langle v,v\rangle\in\mathbb{R}_{\geq 0}$ para todo $v\in V$ y $\langle v,v\rangle=0$ si y sólo si $v=0$.

Un **espacio vectorial con producto interno** es un par $(V,\langle\cdot,\cdot\rangle)$, donde V es un espacio vectorial (real o complejo) y $\langle\cdot,\cdot\rangle$ es un producto interno en V.

8.1.2. Si V es un espacio vectorial con producto interno entonces

$$\langle v,w_1+w_2\rangle=\langle v,w_1\rangle+\langle v,w_2\rangle,$$
$$\langle v,\lambda w\rangle=\overline{\lambda}\langle v,w\rangle,$$

para todo $v,w,w_1,w_2\in V$ y $\lambda\in\mathbb{K}$.

8.1.3. Ejercicio. Sea V un espacio vectorial real o complejo y sea $p\in V$. Si $\langle\cdot,\cdot\rangle$ es un producto interno en V entonces $\langle v,w\rangle_p=\langle v-p,w-p\rangle$ es un producto interno en V_p.

8.1.4. Ejercicio. FIXME Demuestre que la función

$$\langle\cdot,\cdot\rangle\colon\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R},\quad\langle(x,y),(x',y')\rangle=x(y-y')-x'(y-\lambda y')$$

es un producto interno en $\mathbb{R}^2$ si y sólo si $\lambda>1$.

xca:<v-w,x>=0

8.1.5. Ejercicio. Sea V un espacio vectorial con producto interno y sean $v,w\in V$. Demuestre que si $\langle v,x\rangle=\langle w,x\rangle$ para todo $x\in V$ entonces $v=w$.

8.1.6. Ejemplos.

1) La función $\langle\cdot,\cdot\rangle\colon\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ definida por
$$\langle(x_1,\dots,x_n),(y_1,\dots,y_n)\rangle=\sum_{i=1}^n x_iy_i$$
es un producto interno en $\mathbb{R}^n$.
2) La función $\langle\cdot,\cdot\rangle\colon\mathbb{C}^n\times\mathbb{C}^n\to\mathbb{C}$ definida por
$$\langle(x_1,\dots,x_n),(y_1,\dots,y_n)\rangle=\sum_{i=1}^n x_i\overline{y_i}$$
es un producto interno en $\mathbb{C}^n$.
3) Si V y W son espacios vectoriales con producto interno entonces
$$\langle(v_1,w_1),(v_2,w_2)\rangle=\langle v_1,v_2\rangle_V+\langle w_1,w_2\rangle_W$$
es un producto interno en $V\times W$.
4) Si $A,B\in\mathbb{C}^{m\times n}$ entonces $\langle A,B\rangle=\operatorname{tr}(B^*A)$, donde $(B^*)_{ij}=\overline{B_{ji}}$, es un producto interno en $\mathbb{C}^{m\times n}$.

5) La función

$$\langle f,g\rangle = \int_a^b f(x)g(x)dx$$

es un producto interno en el espacio vectorial (real) de funciones continuas $[a,b]\to\mathbb{R}$.

6) Sea $\ell^2(\mathbb{C})$ el espacio vectorial de sucesiones $(a_n)_{n\in\mathbb{N}}$ tales que $\sum_{n\geq 0}|a_n|^2<\infty$. Entonces

$$\langle (a_1,a_2,\dots),(b_1,b_2,\dots)\rangle = \sum_{n\geq 0} a_n\overline{b_n}$$

es un producto interno de $\ell^2(\mathbb{C})$. En efecto, como

$$0\leq (|a_n|-|b_n|)^2 = |a_n|^2-2|a_n||b_n|+|b_n|^2,$$

entonces $|a_n||b_n|\leq \frac{1}{2}(|a_n|^2+|b_n|^2)$.

pro:CauchySchwarz

8.1.7. Proposición (desigualdad de Cauchy–Schwarz). *Sea V un espacio vectorial con producto interno. Entonces*

$$\langle v,w\rangle^2 \leq \langle v,v\rangle\langle w,w\rangle$$

para todo $v,w\in V$.

Demostración. Si $\lambda=\frac{\langle v,w\rangle}{\langle w,w\rangle}$ entonces $\langle w,v\rangle-\overline{\lambda}\langle w,w\rangle=0$. Luego

$$\begin{aligned}0\leq \langle v-\lambda w, v-\lambda w\rangle &= \langle v,v\rangle - \overline{\lambda}\langle v,w\rangle - \lambda(\langle w,v\rangle-\overline{\lambda}\langle w,w\rangle)\\ &= \langle v,v\rangle-\overline{\lambda}\langle v,w\rangle,\end{aligned}$$

que implica que $\overline{\lambda}\langle v,w\rangle\leq\langle v,v\rangle$. Esto demuestra la proposición. □

8.1.8. Ejemplos. La desigualdad de Cauchy–Schwarz es una rica fuente de desigualdades. Por ejemplo, si $\lambda_1,\dots,\lambda_n\in\mathbb{R}$, la formula

$$\langle (a_1,\dots,a_n),(b_1,\dots,b_n)\rangle = \sum_{i=1}^n \lambda_i a_i b_i$$

define un producto interno en $\mathbb{R}^n$ y con la desigualdad de Cauchy–Schwarz se obtiene

$$\left|\sum_{i=1}^n \lambda_i a_i b_i\right| \leq \sqrt{\sum_{i=1}^n \lambda_i a_i^2}\sqrt{\sum_{i=1}^n \lambda_i b_i^2}.$$

Similarmente, si se aplica al producto interno en $C[0,1]$ dado por

$$\langle f,g\rangle = \int_0^1 f(x)g(x)dx$$

se obtiene

$$\left|\int_0^1 f(x)g(x)dx\right| \leq \sqrt{\int_0^1 f(x)^2dx}\sqrt{\int_0^1 g(x)^2dx}.$$

8.1.9. Si V es un espacio vectorial con producto interno entonces la función $\|\cdot\|\colon V\to\mathbb{R}_{\geq 0}$, dada por $v\mapsto\sqrt{\langle v,v\rangle}$, satisface las siguientes propiedades:

1) $\|v\|=0$ si y sólo si $v=0$.
2) $\|\lambda v\|=|\lambda|\|v\|$ para todo $v\in V$.
3) (desigualdad triangular) $\|v+w\|\leq\|v\|+\|w\|$ para todo $v,w\in V$.

Las primeras dos afirmaciones quedan como ejercicio. Para demostrar la desigualdad triangular utilizamos la proposición 8.1.7:

$$\begin{aligned} 0 \leq \|v+w\|^2 = \langle v+w, v+w\rangle &= \|v\|^2 + \langle v,w\rangle + \langle w,v\rangle + \|w\|^2 \\ &= \|v\|^2 + \langle v,w\rangle + \overline{\langle v,w\rangle} + \|w\|^2 \\ &= \|v\|^2 + 2\,\mathrm{re}\langle v,w\rangle + \|w\|^2 \\ &\leq \|v\|^2 + 2|\langle v,w\rangle| + \|w\|^2 \\ &\leq (\|v\| + \|w\|)^2. \end{aligned}$$

Luego $\|v+w\| \leq \|v\| + \|w\|$, tal como queríamos demostrar.

8.1.10. Ejercicio (identidades de polarización).

1) En un espacio vectorial real con producto interno vale
$$\langle v,w\rangle = \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2\right)$$
para todo $v, w \in V$.
2) En un espacio vectorial complejo con producto interno vale
$$\langle v,w\rangle = \frac{1}{4}\sum_{k=1} i^k \|v - i^k w\|^2$$
para todo $v, w \in V$.

8.1.11. Sea V un espacio vectorial con producto interno y sean $v, w \in V$. Se define la **distancia** entre v y w como $\mathrm{dist}(v,w) = \|v-w\|$. La función $\mathrm{dist}\colon V \times V \to \mathbb{R}$ satisface las siguientes propiedades:

1) $\mathrm{dist}(v,w) \geq 0$ para todo $v, w \in V$.
2) $\mathrm{dist}(v,w) = 0$ si y sólo si $v = w$.
3) $\mathrm{dist}(v,w) = \mathrm{dist}(w,v)$ para todo $v, w \in V$.
4) $\mathrm{dist}(v,w) \leq \mathrm{dist}(v,u) + \mathrm{dist}(u,w)$ para todo $u, v, w \in V$.

8.2. Ortogonalidad

8.2.1. Sea V un espacio vectorial con producto interno y sean $v, w \in V$. Entonces v y w son **ortogonales**, $v \perp w$, si $\langle v,w\rangle = 0$. Observemos que $v \perp w$ si y sólo si $w \perp v$. Si S y T son subespacios de V entonces S y T son ortogonales, $S \perp T$, si $\langle s,t\rangle = 0$ para todo $s \in S$ y $t \in T$. Si $S \perp T$ entonces $T \perp S$. Si $X \subseteq V$ es un subconjunto, se define

$$X^\perp = \{v \in V : \langle v,x\rangle = 0 \text{ para todo } x \in X\}.$$

Queda como ejercicio demostrar que $X^\perp$ es un subespacio de V.

8.2.2. Ejemplo. Todo vector es ortogonal a $0 \in V$. Recíprocamente, si $v \in V$ satisface que $v \perp w$ para todo w entonces $v = 0$.

8.2.3. Ejemplo. Si consideramos a $\mathbb{R}^2$ con el producto interno usual y $S = \langle (1,1)\rangle$ entonces $S^\perp = \langle (1,-1)\rangle$.

8.2.4. Sea V un espacio vectorial con producto interno. Los vectores $v_1, \dots, v_n \in V$ son ortogonales si $\langle v_i, v_j\rangle = 0$ si $i \neq j$. Un vector $v \in V$ es **unitario** si $\|v\| = 1$. Los vectores no nulos $v_1, \dots, v_n \in V$ son **ortonormales** si $\langle v_i, v_j\rangle = \delta_{ij}$

8.2.5. Ejemplos. Si consideramos a $\mathbb{R}^2$ con el producto interno usual, la base canónica es un conjunto ortonormal. El conjunto $\{(1,1), (1,-1)\} \subseteq \mathbb{R}^2$ es ortogonal y no ortonormal.

8.2.6. EJEMPLO. En el espacio vectorial real de funciones continuas $f\colon (-\pi, \pi) \to \mathbb{R}$ con el producto interno

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)dx,$$

el conjunto

$$\left\{\frac{1}{\sqrt{2}}\right\} \cup \{\sin(nx) : n \in \mathbb{N}\} \cup \{\cos(nx) : n \in \mathbb{N}\}$$

es ortonormal.

xca:Pitagoras

8.2.7. EJERCICIO (teorema de Pitágoras). Sea V un espacio vectorial con producto interno y sean $v, w \in V$. Demuestre que si $v \perp w$ entonces

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

xca:Sperp

8.2.8. EJERCICIO. Sea V un espacio vectorial con producto interno y sean $S, T \subseteq V$. Demuestre las siguientes afirmaciones:

1) $\{0\}^\perp = V$ y $V^\perp = \{0\}$.
2) Si $S \subseteq T$ entonces $T^\perp \subseteq S^\perp$.
3) $(S + T)^\perp = S^\perp \cap T^\perp$.

pro:autoadjunto

8.2.9. PROPOSICIÓN. *Sea V un espacio vectorial con producto interno y sean $v_1, \dots, v_n \in V$ vectores no nulos y ortogonales. Valen entonces las siguientes afirmaciones:*

1) *Para cada $v \in \langle v_1, \dots, v_n\rangle$ se tiene*

$$v = \frac{\langle v, v_1\rangle}{\|v_1\|^2} v_1 + \cdots + \frac{\langle v, v_n\rangle}{\|v_n\|^2} v_n.$$

2) *El conjunto $\{v_1, \dots, v_n\}$ es linealmente independiente.*

DEMOSTRACIÓN. Si $v = \sum_{i=1}^n \alpha_i v_i$ y $j \in \{1, \dots, n\}$ entonces

$$\langle v, v_j\rangle = \left\langle \sum_{i=1}^n \alpha_i v_i, v_j \right\rangle = \sum_{i=1}^n \alpha_i \langle v_i, v_j\rangle = \alpha_j \|v_j\|^2,$$

lo que demuestra la primera afirmación. Para demostrar la segunda afirmación basta tomar $v = 0$ en el ítem anterior. $\square$

8.2.10. COROLARIO. *Sea V un espacio vectorial con producto interno y sean $v_1, \dots, v_n \in V$ vectores no nulos y ortonormales. Si $v = \sum_{i=1}^n \alpha_i v_i$ entonces*

$$v = \langle v, v_1\rangle v_1 + \cdots + \langle v, v_n\rangle v_n, \qquad \|v\|^2 = |\langle v, v_1\rangle|^2 + \cdots + |\langle v, v_n\rangle|^2.$$

DEMOSTRACIÓN. La primera fórmula se deduce de la proposición anterior pues $\|v_i\| = 1$ para todo $i \in \{1, \dots, n\}$. Para la segunda afirmación calculamos

$$\|v\|^2 = \langle v, v\rangle = \left\langle \sum_{i=1}^n \alpha_i v_i, \sum_{j=1}^n \alpha_j v_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \overline{\alpha_j} \langle v_i, v_j\rangle, = \sum_{i=1}^n |\alpha_i|^2$$

y utlizamos que $\alpha_i = \langle v, v_i\rangle$ para todo $i \in \{1, \dots, n\}$. $\square$

8.2.11. TEOREMA (proceso de ortonormalización de Gram-Schmidt). *Sea V un espacio vectorial con producto interno y sea $\{v_1, \dots, v_n\} \subseteq V$ un conjunto linealmente independiente. Entonces existe $\{e_1, \dots, e_n\} \subseteq V$ ortonormal tal que para cada $k \in \{1, \dots, n\}$ se tiene que $\langle v_1, \dots, v_k\rangle = \langle e_1, \dots, e_k\rangle$.*

DEMOSTRACIÓN. Procederemos por inducción en n.

El caso $n = 1$ es trivial pues, como $v_1 \neq 0$, basta tomar $e_1 = \frac{1}{\|v_1\|} v_1$. Supongamos que el resultado es válido para $n - 1$. Como $\{v_1, \dots, v_{n-1}\}$ es linealmente independiente, existen $e_1,\dots,e_{n-1}\in V$ ortonormales y tales que $\langle e_1,\dots,e_k\rangle=\langle v_1,\dots,v_k\rangle$ para todo $k\in\{1,\dots,n-1\}$. Sea

$$e_n'=v_n-\sum_{i=1}^{n-1}\langle v_n,e_i\rangle e_i.$$

Como $\{v_1,\dots,v_n\}$ es linealmente independiente, entonces $e_n'\neq 0$. En efecto, si fuera $e_n'=0$ tendríamos que $v_n\in\langle e_1,\dots,e_{n-1}\rangle=\langle v_1,\dots,v_{n-1}\rangle$. Si $k\in\{1,\dots,n-1\}$ entonces

$$\langle e_n',e_k\rangle=\left\langle v_n-\sum_{i=1}^{n-1}\langle v_n,e_i\rangle e_i,e_k\right\rangle=\langle v_n,e_k\rangle-\sum_{i=1}^{n}\langle v_n,e_i\rangle\langle e_i,e_k\rangle=0.$$

Luego, si $e_n=\frac{1}{\|e_n'\|}e_n'$ entonces $\|e_n\|=1$ y el conjunto $\{e_1,\dots,e_n\}$ es ortonormal. Además $\langle e_1,\dots,e_k\rangle=\langle v_1,\dots,v_k\rangle$ para todo $k\in\{1,\dots,n\}$. $\square$

8.2.12. Corolario. *Sea V un espacio vectorial con producto interno y de dimensión finita. Entonces V tiene una base ortonormal.*

plemento_ortogonal

8.2.13. Corolario. *Sea V un espacio vectorial con producto interno y de dimensión finita. Si $S\subset V$ es un subespacio entonces $V=S\oplus S^\perp$. En particular,* $\dim V=\dim S+\dim S^\perp$,

Demostración. Sea $\{v_1,\dots,v_m\}$ una base de S. Extendemos este conjunto linealmente independiente a una base $\{v_1,\dots,v_m,v_{m+1},\dots,v_n\}$ de V. Si utilizamos el proceso de Gram–Schmidt tenemos una base ortornormal $\{e_1,\dots,e_m,e_{m+1},\dots,e_n\}$ de V. Por construcción, $\{e_1,\dots,e_m\}$ es base de S. Veamos que $\{e_{m+1},\dots,e_n\}$ es base de $S^\perp$. Si $v\in S^\perp$, escribimos $v=\sum_{i=1}^n\alpha_ie_i$. Como $v\in S^\perp$, se tiene que $0=\langle v,v_j\rangle$ para todo $j\in\{1,\dots,m\}$ y luego $S^\perp\subseteq\langle e_{m+1},\dots,e_n\rangle$. $\square$

8.2.14. Ejemplo. El corolario 8.2.13 no vale en dimensión infinita. Si V es el espacio vectorial (real) de funciones continuas $[0,1]\to\mathbb{R}$ y

$$S=\{f\in V: f\text{ es derivable en }(0,1)\}$$

entonces $S^\perp=\{0\}$. En efecto, si $h\in S^\perp$ entonces $\int_0^1 f(x)h(x)dx=0$ para todo $f\in S$. En particular, si tomamos la función constantemente igual a uno, se tiene que $\int_0^1 h(x)dx=0$. Sea $g(s)=\int_0^s h(x)dx$, $s\in[0,1]$. Entonces $g\in S$, $g(0)=g(1)=0$ y $g'(x)=h(x)$ para todo $x\in(0,1)$. Entonces, para toda $f\in S$, se tiene

$$\begin{aligned}0=(fg)(1)-(fg)(0)=f(1)g(1)-f(0)g(0)&=\int_0^1(fg)'(x)dx\\&=\int_0^1 f'(x)g(x)dx+\int_0^1 f(x)g'(x)dx\\&=\int_0^1 f'(x)g(x)dx+\int_0^1 f(x)h(x)dx\\&=\int_0^1 f'(x)g(x)dx,\end{aligned}$$

pues $f\in S$ y $h\in S^\perp$. En particular, si $f(s)=\int_0^s g(x)dx$, $s\in[0,1]$, entonces $f'(x)=g(x)$ para todo $x\in[0,1]$,

$$0=\int_0^1 f'(x)g(x)dx=\int_0^1 g^2(x)dx,$$

y luego $g(x)=0$. Por lo tanto $h(x)=g'(x)=0$ para todo $x\in[0,1]$.

xca:perpperp

8.2.15. Ejercicio. Sea V un espacio vectorial con producto interno y de dimensión finita. Si $S\subseteq V$ es un subespacio entonces $(S^\perp)^\perp=S$.

8.3. Proyección ortogonal

8.3.1. Sea V un espacio vectorial con producto interno. Sea $v \in V$ y sea $S \subseteq V$ un subespacio. El vector $v_S \in S$ es una **mejor aproximación** a v por vectores de S si $\operatorname{dist}(v, v_S) \leq \operatorname{dist}(v, w)$ para todo $w \in S$.

ejor_aproximacion

8.3.2. Teorema. *Sean V un espacio vectorial con producto interno y S un subespacio de V. Sean $v \in V$ y $v_s \in S$. Si $v - v_S \perp S$ entonces v_S es una mejor aproximación para v por vectores de S.*

Demostración. Supongamos que $v - v_S \perp S$ y sea $w \in S$. Entonces, como $v_S - w \in S$, se tiene que

$$\begin{aligned} \|v - w\|^2 &= \|v - v_S\|^2 + 2\operatorname{re}\langle v - v_S, v_S - w\rangle + \|v_S - w\|^2 \\ &= \|v - v_S\|^2 + \|v_S - w\|^2 \\ &\geq \|v - v_S\|^2, \end{aligned}$$

que implica la desigualdad que se quería demostrar. □

8.3.3. Corolario. *Sean V un espacio vectorial con producto interno y S un subespacio de V de dimensión finita. Sean $v \in V$ y $\{v_1, \dots, v_m\}$ una base ortonormal de S. Entonces*

$$\sum_{i=1}^{m} \langle v, v_i\rangle v_i$$

es la mejor aproximación a v por vectores de S.

Demostración. Demostremos que $v - \sum_{i=1}^{m} \langle v, v_i\rangle v_i$ es ortogonal a todo elemento de S. Si $v_s = \sum_{i=1}^{m} \langle v, v_i\rangle v_i$ y $w = \sum_{j=1}^{m} \langle w, v_j\rangle v_j \in S$, un cálculo directo muestra que

$$\langle v_S, w\rangle = \sum_{i=1}^{m}\sum_{j=1}^{m} \overline{\langle w, v_j\rangle}\langle v, v_i\rangle\langle v_i, v_j\rangle = \sum_{i=1}^{m} \overline{\langle w, v_i\rangle}\langle v, v_i\rangle = \langle v, w\rangle.$$

Luego, como $\langle v - v_S, w\rangle = 0$, concluimos que v_S es la mejor aproximación a v por vectores de S. □

8.3.4. Un proyector $p\colon V \to V$ es un **proyector ortogonal** si $\operatorname{im} p$ y $\ker p$ son ortogonales.

royector_ortogonal

8.3.5. Ejercicio. Sea V un espacio vectorial con producto interno y sea $p\colon V \to V$ un proyector ortogonal. Demuestre que valen las siguientes afirmaciones:

1) $(\operatorname{im} p)^{\perp} = \ker p$.
2) $(\ker p)^{\perp} = \operatorname{im} p$.
3) $(\operatorname{im} p)^{\perp\perp} = \operatorname{im} p$.

8.3.6. Proposición. *Sea V un espacio vectorial con producto interno y de dimensión finita y sea $S \subseteq V$ un subespacio. Entonces existe un único proyector ortogonal $p\colon V \to V$ tal que $\operatorname{im} p = S$ y $\ker p = S^{\perp}$. Más aún, si $\{v_1, \dots, v_m\}$ es una base ortonormal de S entonces*

$$p(v) = \sum_{i=1}^{n} \langle v, v_i\rangle v_i.$$

Demostración. Es claro que p es una transformación lineal. Además $p^2 = p$ pues, como $p(v_i) = v_i$ para todo $i \in \{1, \dots, n\}$,

$$p^2(v) = \sum_{i=1}^{n} \langle v, v_i\rangle p(v_i) = \sum_{i=1}^{n} \langle v, v_i\rangle v_i = p(v).$$

Demostremos que $\operatorname{im} p = S$. Es evidente que $\operatorname{im} p \subseteq S$. Por otro lado, si $v \in S$ entonces $v = \sum_{i=1}^{n} \langle v, v_i\rangle v_i = p(v)$.

Demostremos ahora que $\ker p = S^\perp$. Para la inclusión $\ker p \subseteq S^\perp$ observamos que si $v \in \ker p$ entonces, como los v_i son base de S, se tiene que $\langle v, v_i\rangle = 0$ para todo i. Por otro lado, si $v \in S^\perp$ entonces $\langle v, v_i\rangle = 0$ para todo $i \in \{1, \dots, n\}$ y luego $p(v) = 0$.

Por último demostremos la unicidad. Sean dos proyectores ortogonales p y q tales que $S = \ker p = \ker q$ y $S^\perp = \operatorname{im} p = \operatorname{im} q$. Como $V = S \oplus S^\perp$, todo $v \in V$ se escribe unívocamente como $v = v_1 + v_2$, donde $v_1 \in S$ y $v_2 \in S^\perp$. Entonces, como $p(w) = q(w) = w$ para todo $w \in S^\perp$,

$$\begin{aligned} p(v) &= p(v_1+v_2) = p(v_1) + p(v_2) \\ &= p(v_2) = v_2 = q(v_2) = q(v_1) + q(v_2) = q(v_1+v_2) = q(v), \end{aligned}$$

tal como se quería demostrar. □

xca:pSpT 8.3.7. Ejercicio. Sea V un espacio vectorial con producto interno y sean p_S p_T proyectores ortogonales tales que $\operatorname{im} p_S \perp \operatorname{im} p_T$. Demuestre que entonces $p_S p_T = 0$.

8.4. Transformaciones lineales adjuntas

thm:interno:Riesz 8.4.1. Teorema (teorema de representación de Riesz). *Sea V un espacio vectorial con producto interno y de dimensión finita. Si $f \in V^*$ entonces existe un único vector $v_f \in V$ tal que $f(v) = \langle v, v_f\rangle$ para todo $v \in V$.*

Demostración. Sea $\{v_1, \dots, v_n\}$ una base ortonormal de V. Definimos $v_f = \sum_{i=1}^n \overline{f(v_i)} v_i$. y sea $v \in V$. Vamos a demostrar que $f(v) = \langle v, v_f\rangle$. En efecto, si escribimos $v = \sum_{i=1}^n \langle v, v_i\rangle v_i$, entonces

$$f(v) = \sum_{i=1}^n \langle v, v_i\rangle f(v_i) = \left\langle v, \sum_{i=1}^n \overline{f(v_i)} v_i \right\rangle = \langle v, v_f\rangle.$$

Veamos la unicidad: sean $v_f, v_f' \in V$ tales que $\langle v, v_f\rangle = f(v) = \langle v, v_f'\rangle$ para todo $v \in V$. Entonces, como $\langle v, v_f - v_f'\rangle = 0$ para todo $v \in v$, se concluye que $v_f - v_f' = 0$. □

8.4.2. Sean V y W espacios vectoriales con producto interno. Una transformación lineal $g \in \hom(W, V)$ es una **adjunta** para $f \in \hom(V, W)$ si $\langle f(v), w\rangle = \langle v, g(w)\rangle$ para todo $v \in V$ y $w \in W$. Observemos que, de existir, la adjunta es única. Si $g_1, g_2 \hom(W, V)$ son adjuntas para f entonces, para cada $v \in V$ se tiene que

$$\langle v, g_1(w) - g_2(w)\rangle = \langle v, g_1(w)\rangle - \langle v, g_2(w)\rangle = \langle f(v), w\rangle - \langle f(v), w\rangle = 0,$$

y se concluye que $g_1(w) = g_2(w)$ para todo $w \in W$.

8.4.3. Ejemplo. Si V es un espacio vectorial con producto interno entonces $\mathrm{id}_V^* = \mathrm{id}_V : V \to V$ pues entonces $\langle v, \mathrm{id}_V(w)\rangle = \langle v, w\rangle = \langle \mathrm{id}_V(v), w\rangle$ para todo $v, w \in V$.

djunta:existencia 8.4.4. Teorema. *Sean V y W espacios vectoriales con producto interno y de dimensión finita y sea $f \in \hom(V, W)$. Entonces existe una única adjunta de f, que será denotada por f^*.*

Demostración. Para cada $w \in W$ definimos

$$\phi_w : V \to \mathbb{K}, \quad v \mapsto \langle f(v), w\rangle.$$

Como $\phi_w \in V^*$, el teorema de representación de Riesz, teorema 8.4.1, implica que existe un único vector $f^*(w) \in V$ tal que $\langle f(v), w\rangle = \langle v, f^*(w)\rangle$ para todo $v \in V$. Para completar la demostración del teorema necesitamos ver que $f^* \in \hom(W, V)$. Si $v \in V$ y $w_1, w_2 \in W$ entonces

$$\begin{aligned} &\langle v, f^*(w_1+w_2) - f^*(w_1) - f^*(w_2)\rangle \\ &\qquad = \langle v, f^*(w_1+w_2)\rangle - \langle v, f^*(w)\rangle - \langle v, f^*(w_2)\rangle \\ &\qquad = \langle f(v), w_1+w_2\rangle - \langle f(v), w_1\rangle - \langle f(v), w_2\rangle = 0. \end{aligned}$$

Luego $f^*(w_1+w_2)-f^*(w_1)-f^*(w_2)=0$ para todo $w_1,w_2\in W$. Similarmente, si $v\in V$, $w\in W$ y $\lambda\in\mathbb{K}$, entonces

$$\langle v, f^*(\lambda w)-\lambda f^*(w)\rangle = 0$$

y luego $f^*(\lambda w)=\lambda f(w)$ para todo $w\in W$ y $\lambda\in\mathbb{K}$. □

8.4.5. Ejemplo. Mostraremos que en espacios de dimensión finita no siempre existe la adjunta de una transformación lineal. Sea $V=\mathbb{R}[X]$ con el producto interno

$$\left\langle \sum_{i=0}^{n} a_i X^i, \sum_{j=0}^{m} b_j X^j \right\rangle = \sum_{i=0}^{r} a_i b_i, \quad r=\text{mín}\{n,m\}.$$

En particular, si $p=\sum_{i=0}^{n} a_i X^i$, entonces

$$\langle p, X^m\rangle = \langle X^m, p\rangle = \begin{cases} 0 & \text{si } n<m,\\ a_m & \text{si } n\geq m.\end{cases}$$

Vamos a demostrar que la transformación lineal $f\colon V\to V$ definida en la base canónica de $\mathbb{R}[X]$ por

$$f(X^k)=1+X+\cdots+X^k, \quad k\geq 0,$$

no tiene adjunta. En efecto, si existiera $f^*\colon V\to V$, entonces, para todo $k,l\geq 0$, tendríamos

$$\langle X^k, f^*(X^l)\rangle = \langle f(X^k), X^l\rangle = \langle 1+X+\cdots+X^k, X^l\rangle = \begin{cases} 1 & \text{si } k\geq l,\\ 0 & \text{si } k<l.\end{cases}$$

Luego $f^*(X^l)=X^l+X^{l+1}+\cdots\notin\mathbb{R}[X]$, una contradicción.

8.4.6. Proposición. *Sean V y W espacios vectoriales con producto interno y de dimensión finita y sea $f\in\hom(V,W)$. Si $\mathscr{B}_V$ es una base ortonormal de V y $\mathscr{B}_W$ es una base ortonormal de W*

$$[f^*]_{\mathscr{B}_W,\mathscr{B}_V} = [f]^*_{\mathscr{B}_V,\mathscr{B}_W}.$$

Demostración. Supongamos que

$$\mathscr{B}_V=\{v_1,\dots,v_n\}, \quad \mathscr{B}_W=\{w_1,\dots,w_m\}.$$

Supongamos además que $[f]_{\mathscr{B}_V,\mathscr{B}_W}=(a_{ij})$ y que $[f^*]_{\mathscr{B},\mathscr{B}}=(b_{ij})$. Entonces

$$\langle f(v_i), w_j\rangle = \left\langle \sum_{k=1}^{m} a_{ki}w_k, w_j\right\rangle = \sum_{k=1}^{m} a_{ki}\delta_{kj} = a_{ji}.$$

Por otro lado,

$$\langle v_i, f^*(w_j)\rangle = \left\langle v_i, \sum_{k=1}^{n} b_{kj}v_k\right\rangle = \sum_{k=1}^{n} \overline{b_{kj}}\delta_{ik} = \overline{b_{ij}}.$$

Luego $a_{ij}=\overline{b_{ji}}$ pues $\langle f(v_i),w_j\rangle = \langle v_i, f^*(w_j)\rangle$ para todo i,j. □

8.4.7. Ejercicio. Sea V un espacio vectorial con producto interno. Sean $f,g\in\hom(V,V)$ y supongamos que existen las transformaciones adjuntas de f y de g. Demuestre las siguientes afirmaciones:

1) $(f+g)^*=f^*+g^*$.
2) $(fg)^*=g^*f^*$.
3) $(\lambda f)^*=\overline{\lambda}f^*$ para todo $\lambda\in\mathbb{K}$.
4) $f^{**}=f$.

8.4.8. Ejercicio. Sean V y W espacios vectoriales con producto interno y sea $f\in\hom(V,W)$ tal que existe f^*. Demuestre que valen las siguientes afirmaciones:

1) $\ker f^*=(\operatorname{im} f)^\perp$.

2) $\operatorname{im} f^* \subseteq (\ker f)^\perp$ y vale la igualdad si V y W son de dimensión finita.

a:autoadjunta:iso

8.4.9. Ejercicio. Sean V y W dos espacios vectoriales con producto interno y de dimensión finita. Si $f \in \hom(V, W)$ es un isomorfismo entonces f^* es un isomorfismo y vale que $(f^*)^{-1} = (f^{-1})^*$.

8.5. El teorema de Schur

8.5.1. Teorema (Schur). *Sea V un espacio vectorial con producto interno y de dimensión finita y sea $f \in \hom(V, V)$. Si χ_f tiene todas sus raíces en $\mathbb{K}$, entonces existe una base ortonormal de V tal que la matriz de f en esa base es triangular superior.*

Demostración. Procederemos por inducción en $n = \dim V$.

Como el caso $n = 1$ es trivial, suponemos que el resultado es válido para todo endomorfismo de un espacio vectorial de dimensión $n-1$. Sean $\lambda \in \mathbb{K}$ y $v \in V$ tal que $\|v\| = 1$ y $f(v) = \lambda v$. Entonces $f^*(v) = \overline{\lambda} v$. Si $S = \langle v \rangle$ entonces $V = S \oplus S^\perp$. Además $S^\perp$ es f-invariante pues si $w \in S^\perp$ entonces

$$\langle f(w), \mu v \rangle = \langle w, f^*(\mu v) \rangle = \langle w, \mu f^*(v) \rangle = \overline{\mu} \langle w, \lambda v \rangle = \overline{\mu \lambda} \langle w, v \rangle = 0$$

para todo $\mu \in \mathbb{K}$. Sea $g = f|_{S^\perp}$. Como χ_g divide a χ_f entonces χ_g también tiene a todas sus raíces en $\mathbb{K}$. Entonces, como $\dim S^\perp = n-1$, por hipótesis inductiva, existe una base ortonormal de $S^\perp$ tal que la matriz de g en esa base es triangular superior. $\square$

8.5.2. Corolario. $\boxed{P}$ *Toda matriz cuadrada compleja es semejante a una matriz triangular superior.*

Demostración. Es consecuencia inmediata del teorema de Schur. $\square$

8.5.3. Recordemos que una matriz $A \in \mathbb{C}^{n \times n}$ es **hermitiana** si $A^* = A$, donde $(A^*)_{ij} = \overline{A_{ji}}$ para todo i, j. Una matriz $A \in \mathbb{C}^{n \times n}$ es **unitaria** si $AA^* = A^*A = I$.

8.5.4. Ejemplo. Las matrices $\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$ y $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ son hermitianas. La matriz $\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$ es unitaria.

cor:Schur

8.5.5. Corolario. *Sea $A \in \mathbb{C}^{n \times n}$ una matriz hermitiana. Entonces existe una matriz unitaria P tal que PAP^{-1} es diagonal.*

Demostración. Por el teorema de Schur, existe una matriz unitaria P tal que PAP^* es triangular superior. Como $PAP^* = (PAP^*)^*$ es también triangular inferior, concluimos que PAP^* es diagonal. $\square$

8.6. Transformaciones lineales autoadjuntas

ineal!autoadjunta

8.6.1. Sea V un espacio vectorial con producto interno. Una transformación lineal $f \in \hom(V, V)$ es **autoadjunta** si existe la adjunta f^* de f y vale $f^* = f$.

xca:autoadjunta

8.6.2. Ejercicio. Sea V un espacio vectorial con producto interno y de dimensión finita. Demuestre que si $f \in \hom(V, V)$ y $\mathscr{B}$ es una base ortonormal de V entonces f es autoadjunta si y sólo si $[f]_{\mathscr{B}} = [f]_{\mathscr{B}}^*$. En particular, si $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$) entonces f es autoadjunta si y sólo si $[f]_{\mathscr{B}}$ es simétrica (resp. hermitiana).

8.6.3. Proposición. *Sea V un espacio vectorial con producto interno y de dimensión finita. Sea $p\colon V \to V$ un proyector. Entonces p es un proyector ortogonal si y sólo si p es autoadjunto.*

Demostración. Supongamos que p es un proyector ortogonal. Como $V = \ker p \oplus \operatorname{im} p$, si $v, w \in V$ entonces $v = v_1 + v_2$ y $w = w_1 + w_2$ con $v_1, w_1 \in \ker p$ y $v_2, w_2 \in \operatorname{im} p$. En particular, $p(v) = v_2$ y $p(w) = w_2$. Como $\ker p \perp \operatorname{im} p$,

$$\langle p(v), w \rangle = \langle v_2, w_1 + w_2 \rangle = \langle v_2, w_2 \rangle = \langle v_1 + v_2, w_2 \rangle = \langle v, p(w) \rangle.$$

Recíprocamente, supongamos que p es un proyector autoadjunto. Si $v \in \ker p$ y $w \in \operatorname{im} p$, entonces $p(v) = 0$ y $p(w) = w$. Luego

$$0 = \langle 0, w\rangle = \langle p(v), w\rangle = \langle v, p(w)\rangle = \langle v, w\rangle,$$

tal como se quería demostrar. □

pro:autoadjunta

8.6.4. Proposición. *Sea V un espacio vectorial con producto interno y sea $f \in \hom(V, V)$. Supongamos que f es autoadjunta. Entonces:*

1) *Todo autovalor de f es real.*
2) *Si $v, w \in V$ son autovectores de v de autovalores distintos entonces $\langle v, w\rangle = 0$.*

Demostración. Demostremos la primera afirmación. Sean $\lambda \in \mathbb{K}$ y $v \in V \setminus \{0\}$ tales que $f(v) = \lambda v$. Como

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle f(v), v\rangle = \langle v, f^*(v)\rangle = \langle v, f(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle,$$

y $v \neq 0$, entonces $\lambda \in \mathbb{R}$ pues $\lambda = \overline{\lambda}$.

Demostremos ahora la segunda afirmación. Sean $\lambda, \mu \in \mathbb{K}$ autovalores distintos con autovectores $v, w \in V$. Entonces, como λ y μ son números reales,

$$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle f(v), w\rangle = \langle v, f^*(w)\rangle = \langle v, f(w)\rangle = \langle v, \mu w\rangle = \mu\langle v, w\rangle.$$

Como $\lambda \neq \mu$, entonces $\langle v, w\rangle = 0$. □

8.6.5. Proposición. *Sea V un espacio vectorial con producto interno y de dimensión finita. Toda $f \in \hom(V, V)$ autoadjunta tiene al menos un autovalor.*

Demostración. Supongamos que V es un espacio vectorial real, ya que si V fuera un espacio vectorial complejo el resultado es válido gracias al teorema fundamental del álgebra. Sea $\mathscr{B} = \{v_1, \dots, v_n\}$ una base ortonormal de V, sea $A = [f]_{\mathscr{B},\mathscr{B}} \in \mathbb{R}^{n\times n}$ y sea $g\colon \mathbb{C}^n \to \mathbb{C}^n$ definida por $g(x) = Ax$. Como f es autoadjunta, A es simétrica y g es autoadjunta. El polinomio característico χ_g de g tiene grado n, se factoriza linealmente en $\mathbb{C}$ y luego existe un autovalor λ de g. Como g es autoadjunta, $\operatorname{spec} f = \operatorname{spec} g \subseteq \mathbb{R}$. □

8.6.6. Teorema. *Sea V un espacio vectorial con producto interno y de dimensión finita. Sea $f \in \hom(V, V)$ autoadjunta. Entonces existe una base ortonormal $\mathscr{B}$ tal que $[f]_{\mathscr{B}}$ es diagonal.*

Demostración. Es consecuencia inmediata del corolario 8.5.5 y del ejercicio 8.6.2. □

8.6.7. Ejercicio. Sea V un espacio vectorial con producto interno y de dimensión finita. Sea $f \in \hom(V, V)$. Demuestre que f es autoadjunta si y sólo si existe una base ortonormal de autovectores con autovalores reales.

8.7. Descomposición en valores singulares

valores_singulares

8.7.1. Teorema. *Sean V y W espacios vectorial con producto interno y de dimensión finita. Sea $f \in \hom(V, W)$ y sean $n = \dim V$ y $m = \dim W$. Entonces existen $k \leq \min\{n, m\}$, escalares $\sigma_1, \dots, \sigma_k \in \mathbb{R}_{>0}$ y bases ortonormales $\{v_1, \dots, v_n\}$ de V y $\{w_1, \dots, w_m\}$ de W tales que*

$$f(v_i) = \begin{cases} \sigma_i w_i & \text{si } i \in \{1, \dots, k\}, \\ 0 & \text{en otro caso,} \end{cases} \qquad f^*(w_i) = \begin{cases} \sigma_i v_i & \text{si } i \in \{1, \dots, k\}, \\ 0 & \text{en otro caso.} \end{cases}$$

En particular,

1) $\ker f = \langle v_{k+1}, \dots, v_n\rangle$.
2) $\operatorname{im} f = \langle w_1, \dots, w_k\rangle$.
3) $\ker f^* = \langle w_{k+1}, \dots, w_m\rangle$.
4) $\operatorname{im} f^* = \langle v_1, \dots, v_k\rangle$.

5) *Para cada* $v \in V$ *se tiene que* $f(v) = \sum_{i=1}^{k} \sigma_i \langle v, v_i \rangle w_i$.

Demostración. Si $f = 0$ no hay nada para demostrar. Supongamos entonces que $f \neq 0$. Sea $g = f^* f$. Como $g \in \hom(V, V)$ es autoadjunto, existe una base ortonormal $\{v_1, \dots, v_n\}$ formada por autovectores v_i de g de autovalor λ_i. Como g es autoadjunta, los λ_i son reales y no negativos pues

$$\lambda_i = \lambda_i \langle v_i, v_i \rangle = \langle \lambda_i v_i, v_i \rangle = \langle g(v_i), v_i \rangle = \langle f(v_i), f(v_i) \rangle \geq 0$$

para todo $i \in \{1, \dots, n\}$. Sin pérdida de generalidad podemos suponer que

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0, \quad \lambda_{k+1} = \cdots = \lambda_n = 0.$$

Para cada $i \in \{1, \dots, k\}$ sean $\sigma_i = \sqrt{\lambda_i}$ y $w_i = \sigma_i^{-1} f(v_i) \in W$. Si $i \neq j$ entonces

$$\begin{aligned} \langle w_i, w_j \rangle &= (\sigma_i \sigma_j)^{-1} \langle f(v_i), f(v_j) \rangle \\ &= (\sigma_i \sigma_j)^{-1} \langle g(v_i), v_j \rangle = (\sigma_i \sigma_j)^{-1} \lambda_i \langle v_i, v_j \rangle = 0. \end{aligned}$$

Además

$$\langle w_i, w_i \rangle = \sigma_i^{-2} \langle f(v_i), f(v_i) \rangle = \sigma_i^{-2} \langle g(v_i), v_i \rangle = \sigma_i^{-2} \lambda_i \langle v_i, v_i \rangle = 1.$$

Luego $\{w_1, \dots, w_k\}$ es un conjunto ortonormal.

Por construcción $f(v_i) = \sigma_i w_i$ para todo $i \in \{1, \dots, k\}$. Además

$$f^*(w_i) = f^*(\sigma_i^{-1} f(v_i)) = \sigma_i^{-1} g(v_i) = \sigma_i^{-1} \lambda_i v_i = \sigma_i v_i$$

para todo $i \in \{1, \dots, k\}$.

Si $i \in \{k+1, \dots, n\}$ entonces $g(v_i) = 0$ y, como

$$0 = \langle g(v_i), v_i \rangle = \langle f(v_i), f(v_i) \rangle,$$

se concluye que $v_i \in \ker f$. Esto demuestra la fórmula que queríamos para los $f(v_i)$. Si $k < m$, sea $\{w_{k+1}, \dots, w_m\}$ una base ortonormal de $\ker f^*$. Entonces, como $\ker f^* = (\operatorname{im} f)^\perp$, el conjunto $\{w_1, \dots, w_k, w_{k+1}, \dots, w_m\}$ es una base ortonormal de W. □

8.7.2. Los escalares $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k$ del teorema 8.7.1 se denominan **valores singulares** de f.

8.7.3. Corolario. *Sea* $A \in \mathbb{C}^{m \times n}$. *Entonces existe* $k \in \min\{n, m\}$ *y existen matrices unitarias* $P \in \mathbb{C}^{m \times m}$, $Q \in \mathbb{C}^{n \times n}$ *y una matriz diagonal real* $D = \operatorname{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0) \in \mathbb{R}^{m \times n}$, *donde* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k$ *son los valores singulares de* A, *tales que*

$$A = PDQ^*.$$

Más aún, valen las siguientes afirmaciones:

1) *Las primeras* k *columnas de* P *dan una base ortonormal de* $\operatorname{rg} A$.
2) *Las últimas* $m - k$ *columnas de* P *dan una base ortonormal de* $\ker A^*$.
3) *Las primeras* k *columnas de* Q *dan una base ortonormal de* $\operatorname{rg} A^*$.
4) *Las últimas* $n - k$ *columnas de* Q *dan una base ortonormal de* $\ker A$.

Demostración. Sea $V = \mathbb{K}^{n \times 1}$, $W = \mathbb{K}^{m \times 1}$ y $f \colon V \to W$ dado por $x \mapsto Ax$. La descomposición en valores singulares aplicada a f nos dice que existe una base ortonormal $\mathscr{B}_V = \{v_1, \dots, v_n\}$ de V y una base ortonormal $\mathscr{B}_W = \{w_1, \dots, w_m\}$ de W tal que

$$[f]_{\mathscr{B}_V, \mathscr{B}_W} = \operatorname{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0) = D,$$

donde $\sigma_1, \dots, \sigma_k$ son los valores singulares de f. Si $\mathscr{E}_V$ y $\mathscr{E}_W$ son las bases canónicas de V y W, respectivamente, entonces la matriz f con respecto a estas bases es A. Las matrices

$$P = C(\mathscr{B}_W, \mathscr{E}_W) = (w_1 | \cdots | w_m), \quad Q = C(\mathscr{B}_V, \mathscr{E}_V) = (v_1 | \cdots | v_n),$$

son unitarias y vale que

$$A = PDQ^*,$$

tal como se quería demostrar. □

valores_singulares

8.7.4. Ejemplo. Sea

$$A = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 3}.$$

Vamos a encontrar la descomposición en valores singulares de A. Queremos entonces $P \in \mathbb{R}^{2\times 2}$ ortogonal , $Q \in \mathbb{R}^{3\times 3}$ ortogonal y $D \in \mathbb{R}^{2\times 3}$ tales que $A = PDQ^T$, donde

$$D = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{pmatrix},$$

con $\sigma_1 \geq \sigma_2 \geq 0$, es la matriz de valores singulares de A. Como A tiene rango dos, tendremos dos valores singulares no nulos, es decir $\sigma_1 \geq \sigma_2 > 0$. Empecemos por calcular la matriz Q. Calculemos entonces los autovalores de la matriz

$$A^TA = \begin{pmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{pmatrix}.$$

Los autovalores de A^TA son las soluciones de la ecuación

$$0 = \det(A^TA - \lambda I) = -\lambda^3 + 22\lambda^2 - 120\lambda = -(\lambda - 12)(\lambda - 10)\lambda.$$

Luego λ es autovalor de A^TA si y sólo si $\lambda \in \{12, 10, 0\}$. Los valores singulares no nulos de A serán entonces $\sigma_1 = \sqrt{12}$ y $\sigma_2 = \sqrt{10}$. Calculemos los autovectores asociados a estos autovalores. Por ejemplo, para $\lambda = 12$ tenemos que resolver el sistema

$$(A^TA - 4I)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 2 \\ 0 & -2 & 4 \\ 2 & 4 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

La solución de este sistema es el espacio vectorial generado por el vector $(1, 2, 1)$. Análogamente, el autoespacio asociado al autovalor $\lambda = 10$ es el espacio generado por $(2, -1, 0)$ y el autoespacio asociado al autovalor $\lambda = 0$ está generado por el vector $(1, 2, -5)$. Ahora usamos el proceso de ortonormalización de Gram-Schmidt en cada autoespacio de A^TA. Dejamos como ejercicio chequear que de esta forma obtenemos la base ortonormal

$$\left\{ w_1 = \frac{1}{\sqrt{6}}(1, 2, 1), w_2 = \frac{1}{\sqrt{5}}(2, -1, 0), w_3 = \frac{1}{\sqrt{30}}(1, 2, -5) \right\}$$

y la diagonalización ortogonal de la matriz A^TA:

$$A^TA = Q\begin{pmatrix} 12 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 0 \end{pmatrix}Q^T, \quad Q = (w_1|w_2|w_3) = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{-5}{\sqrt{30}} \end{pmatrix}.$$

Para obtener entonces la descomposición en valores singulares necesitamos hacer algo similar para la matriz AA^T. Los autovalores de la matriz

$$AA^T = \begin{pmatrix} 11 & 1 \\ 1 & 11 \end{pmatrix}$$

son $\lambda = 12$ y $\lambda = 10$. Tal como hicimos antes, calculamos una base de autovectores para AA^T y usamos Gram-Schmidt en cada autoespacio para obtener una matriz ortogonal P que diagonalice ortogonalmente a la matriz AA^T:

$$AA^T = P\begin{pmatrix} 12 & 0 \\ 0 & 10 \end{pmatrix}P^T, \quad P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

La descomposición en valores singulares de la matriz A es

$$A = P\begin{pmatrix} \sqrt{12} & 0 & 0 \\ 0 & \sqrt{10} & 0 \end{pmatrix}Q^T.$$

8.7.5. Ejemplo. La descomposición en valores singulares de una matriz A nos da bases ortonormales para los subespacios: $\ker(A)$, $\operatorname{im}(A^T)$, $\operatorname{im}(A)$ y $\ker(A^T)$. Por lo visto en el ejemplo 8.7.4, si

$$A = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 3}$$

entonces

$$\begin{aligned}
&\left\{ \frac{1}{\sqrt{2}}(1,1), \frac{1}{\sqrt{2}}(1,-1) \right\} \text{ es una base ortonormal de } \operatorname{im}(A),\\
&\left\{ \frac{1}{\sqrt{6}}(1,2,1), \frac{1}{\sqrt{5}}(2,-1,0) \right\} \text{ es una base ortonormal de } \operatorname{im}(A^T),\\
&\left\{ \frac{1}{\sqrt{30}}(1,2,-5) \right\} \text{ es una base ortonormal de } \ker(A),
\end{aligned}$$

y $\ker(A^T) = \{0\}$.

8.7.6. Corolario (descomposición polar). *Si $A \in \mathbb{K}^{n\times n}$ entonces A se escribe como $A = BC$, donde $B \in \mathbb{K}^{n\times n}$ es unitaria y $C \in \mathbb{K}^{n\times n}$ es autoadjunta con autovalores no negativos.*

Demostración. Por la descomposición en valores singulares, teorema 8.7.1, $A = PDQ^*$, donde P y Q son matrices unitarias. Si $B = PQ^*$ y $C = QDQ^*$ entonces B es unitaria, C es autoadjunta con autovalores no negativos y

$$A = (PQ^*)(QDQ^*),$$

tal como se quería demostrar. □

8.8. Transformaciones lineales normales

8.8.1. Sean V un espacio vectorial con producto interno y $f \in \hom(V,V)$ que una transformación lineal que admite adjunta. Diremos que f es **normal** si $f^*f = ff^*$.

8.8.2. Ejemplo. Sea $\theta \in \mathbb{R}$ tal que θ no es un múltiplo entero de π. Si $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ está dada por

$$f(x,y) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

entonces f es normal pues $f^*f = f^*f = \mathrm{id}$ y no es autoadjunta.

lem:normales

8.8.3. Lema. *Sea V un espacio vectorial con producto interno y de dimensión finita y sea $f \in \hom(V,V)$ una transformación lineal normal. Entonces v es autovector de f de autovalor λ si y sólo si v es autovector de f^* de autovalor $\overline{\lambda}$.*

Demostración. Sea $\lambda \in \mathbb{K}$ y sea $g = f - \lambda\,\mathrm{id}_V$. Como f es normal, g es normal. Además $g^* = f^* - \overline{\lambda}\,\mathrm{id}_V$ y si $v \in V$ entonces

$$\begin{aligned}
\|g(v)\|^2 &= \langle g(v), g(v)\rangle = \langle v, g^*g(v)\rangle \\
&= \langle v, g^*g(v)\rangle = \langle g^*(v), g^*(v)\rangle = \|g^*(v)\|^2.
\end{aligned}$$

Luego $g(v) = 0$ si y sólo si $g^*(v) = 0$, es decir v es autovector de f de autovalor λ sólo si v es autovector de f^* de autovalor $\overline{\lambda}$. □

normales:diagonal

8.8.4. Teorema. *Sea V un espacio vectorial complejo con producto interno y de dimensión finita y sea $f \in \hom(V,V)$ una transformación lineal normal. Entonces existe una base ortonormal de V tal que $[f]_{\mathscr{B},\mathscr{B}}$ es diagonal.*

Demostración. Procederemos por inducción en $n = \dim V$.

El caso $n = 1$ es trivial, así que supongamos que el resultado es válido para todo endomorfismo normal de un espacio vectorial de dimensión $n-1$. Como estamos sobre los números complejos, existe un autovalor $\lambda \in \mathbb{C}$. Sea $v \in V$ un autovector de f de autovalor λ. Como $v \neq 0$, sin pérdida de generalidad podemos suponer que $\|v\| = 1$. Por el lema 8.8.3 sabemos que $f^*(v) = \overline{\lambda}v$. Sea $S = \langle v\rangle^{\perp}$. Entonces S es f-invariante pues si $w \in S$ entonces

$$\langle f(w), v\rangle = \langle w, f^*(v)\rangle = \langle w, \overline{\lambda}v\rangle = \lambda\langle w, v\rangle = 0.$$

Similarmente demostramos que f^* invariante pues si $w \in S$ entonces

$$\langle f^*(w), v\rangle = \langle w, f(v)\rangle = \langle w, \lambda v\rangle = \overline{\lambda}\langle w, v\rangle = 0.$$

Las restricciones $f|_S$ y $f^*|_S$ son entonces endomorfismos de S. Además $f|_S$ es normal pues $f^*|_S = (f|_S)^*$. Como $\dim S = \dim V - 1$, la hipótesis inductiva en $f|_S$ implica que existe una base ortonormal de S tal que la matriz de $f|_S$ en esa base es diagonal. De aquí se deduce el teorema. □

8.8.5. Corolario. *Sea V un espacio vectorial complejo con producto interno y de dimensión finita y sea $f \in \hom(V,V)$. Entonces f es normal si y sólo si existe una base ortonormal formada por autovectores de f.*

Demostración. Una de las implicaciones es el teorema anterior. Recíprocamente, si existe una base ortonormal $\mathscr{B}$ tal que $[f]_{\mathscr{B},\mathscr{B}}$ es diagonal, la matriz $[f^*]_{\mathscr{B},\mathscr{B}} = [f]^*_{\mathscr{B},\mathscr{B}}$ es también diagonal. Luego $[f]_{\mathscr{B},\mathscr{B}}$ y $[f^*]_{\mathscr{B},\mathscr{B}}$ conmutan, lo que implica que f y f^* conmutan. □

thm:espectral

8.8.6. Teorema (teorema espectral). *Sea V un espacio vectorial complejo con producto interno y de dimensión finita y sea $f \in \hom(V,V)$ una transformación lineal normal. Supongamos que $\chi_f = \prod_{i=1}^{k}(X-\lambda_i)^{m_i}$ con $\lambda_i \neq \lambda_j$ si $i \neq j$. Para cada $i \in \{1,\dots,k\}$ sean $V_i = \ker(f - \lambda_i \operatorname{id}_V)$ y $p_i \colon V \to V$ el proyector ortogonal con $\operatorname{im} p_i = V_i$. Entonces:*

1) *$V = V_1 \oplus \dots \oplus V_k$ con $V_i \perp V_j$ si $i \neq j$.*
2) *$p_1 + \dots + p_k = \operatorname{id}_V$ con $p_ip_j = 0$ si $i \neq j$.*
3) *$f = \lambda_1 p_1 + \dots + \lambda_k p_k$.*

Demostración. Veamos la primera afirmación. Sean $i, j \in \{1,\dots,n\}$ con $i \neq j$ y sean $v_i \in V_i \setminus \{0\}$ y $v_j \in V_j \setminus \{0\}$. Entonces, por el lema 8.8.3,

$$\lambda_i\langle v_i, v_j\rangle = \langle \lambda_i v_i, v_j\rangle = \langle f(v_i), v_j\rangle = \langle v_i, f^*(v_j)\rangle = \langle v_i, \overline{\lambda_j}v_j\rangle = \lambda_j\langle v_i, v_j\rangle$$

y luego $\langle v_i, v_j\rangle = 0$ pues $\lambda_i \neq \lambda_j$. Por el teorema 8.8.4, existe una base de V formada por autovalores, luego $V = V_1 \oplus \dots \oplus V_k$.

Para demostrar la segunda afirmación, sea $v \in V$. Como, por el ítem anterior, $v = v_1 + \dots + v_k$ con $v_j \in V_j$ para todo $j \in \{1,\dots,k\}$, entonces $p_i(v) = p_i(v_i) = v_i$ para todo $i \in \{1,\dots,k\}$. Luego

$$(p_1 + \dots + p_k)(v) = \sum_{i=1}^{k}\sum_{j=1}^{k} p_i(v_j) = v_1 + \dots + v_k = v.$$

Para ver que $p_ip_j = 0$ si $i \neq j$ se usa el ejercicio 8.3.7.

Como $p_i|_{V_j} = 0$ si $i \neq j$ y $p_i|_{V_i} = \operatorname{id}_{V_i}$, se tiene que

$$(\lambda_1 p_1 + \dots + \lambda_k p_k)|_{V_j} = \lambda_k \operatorname{id}_{V_j} = f|_{V_j}.$$

Como $V = V_1 \oplus \dots \oplus V_k$, esto implica que $f = \lambda_1 p_1 + \dots + \lambda_k p_k$. □

:normal:polinomio

8.8.7. Lema. *Sea V un espacio vectorial y sean $p_1,\dots,p_n \colon V \to V$ proyectores tales que $p_1 + \dots + p_n = \operatorname{id}_V$ y $p_ip_j = 0$ si $i \neq j$. Si*

$$f = \lambda_1 p_1 + \dots + \lambda_n p_n,$$

donde $\lambda_1, \dots, \lambda_n \in \mathbb{K}$*, entonces, para cada* $p \in \mathbb{K}[X]$*, se tiene que*

$$p(f) = p(\lambda_1)p_1 + \cdots + p(\lambda_n)p_n.$$

Demostración. Basta demostrar el lema para $p \in \{X^m : m \geq 0\}$. Para esto, procederemos por inducción en m.

El caso $m = 1$ es trivial pues $f = \lambda_1 p_1 + \cdots + \lambda_n p_n$. Supongamos entonces que el resultado es válido para $p = X^{m-1}$ y veamos que vale para $p = X^m$. La hipótesis inductiva, $p_i^2 = p_i$ para todo $i \in \{1, \dots, n\}$ y $p_i p_j = 0$ si $i \neq j$ implican que

$$\begin{aligned} f^{m+1} = f^m f &= \left(\sum_{i=1}^n \lambda_i^m p_i\right)\left(\sum_{j=1}^n \lambda_j p_j\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i^m \lambda_j p_i p_j = \sum_{i=1}^m \lambda_i^{m+1} p_i, \end{aligned}$$

que demuestra el lema. □

8.8.8. Teorema. *Sea* V *un espacio vectorial complejo con producto interno y de dimensión finita. Sea* $f \in \hom(V, V)$*. Entonces* f *es normal si y sólo si existe* $p \in \mathbb{C}[X]$ *tal que* $f^* = p(f)$*.*

Demostración. $\boxed{\text{FIXME}}$ Si $f^* = p(f)$ para algún $p \in \mathbb{C}[X]$ entonces f es normal pues $p(f)$ conmuta con f. Supongamos entonces que f es normal. Por el teorema espectral, teorema 8.8.6, existen $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ con $\lambda_i \neq \lambda_j$ si $i \neq j$ y proyectores $p_1, \dots, p_k \colon V \to V$ tales que $f = \lambda_1 p_1 + \cdots + \lambda_k p_k$ y $p_i p_j = 0$ si $i \neq j$. Sea $p \in \mathbb{C}[X]$ un polinomio de grado k tal que $p(\lambda_i) = \overline{\lambda_i}$ para todo $i \in \{1, \dots, k\}$. Por el lema 8.8.7,

$$p(f) = \sum_{i=1}^k p(\lambda_i)p_i = \sum_{i=1}^k \overline{\lambda_i} p_i.$$

Por otro lado, como

$$f^* = (\lambda_1 p_1 + \cdots + \lambda_k p_k)^* = \overline{\lambda_1} p_1 + \cdots + \overline{\lambda_k} p_k,$$

se tiene $p(f) = f^*$, tal como quería demostrar. □

8.9. Clasificación de transformaciones ortogonales

8.9.1. Sea V un espacio vectorial real. Diremos que $f \in \hom(V, V)$ es **ortogonal** si exite la adjunta de f y además $f^* f = f^* f = \mathrm{id}_V$.

`xca:ortogonal` 8.9.2. Ejercicio. Sea V un espacio vectorial real con producto interno y de dimensión finita. Demuestre que son equivalentes:

1) $\|f(v)\| = \|v\|$ para todo $v \in V$.
2) $\langle f(v), f(w)\rangle = \langle v, w\rangle$ para todo $v, w \in V$.
3) Si $\{v_1, \dots, v_n\}$ es base ortonormal de V entonces $\{f(v_1), \dots, f(v_n)\}$ es base ortonormal de V.
4) $f^* f = \mathrm{id}_V$.

`lem:ortogonal` 8.9.3. Lema. *Sea* V *un espacio vectorial real con producto interno y de dimensión finita. Sea* $f \in \hom(V, V)$ *una transformación ortogonal. Valen las siguientes afirmaciones:*

1) *Si* λ *es autovalor de* f *entonces* $\lambda \in \{-1, 1\}$*.*
2) *Si* $S \subseteq V$ *es un subespacio* f*-invariante entonces* $S^\perp$ *es* f*-invariante.*

Demostración. Para demostrar la primera afirmación observamos que si v es autovector de autovalor λ entonces,

$$|\lambda|\langle v,v\rangle = \lambda\overline{\lambda}\langle v,v\rangle = \langle f(v), f(v)\rangle = \langle v, f^*f(v)\rangle = \langle v,v\rangle.$$

Como $v\neq 0$, se obtiene que $|\lambda|=1$. Luego, como $\lambda\in\mathbb{R}$, se concluye que $\lambda\in\{-1,1\}$.

Demostremos ahora la segunda afirmación: si $v\in S^\perp$ y $w\in S$ entonces $\langle v,w\rangle=0$. Como $f^*f=ff^*=\mathrm{id}_V$, f es sobreyectiva. En particular, la restricción $f|_S\colon S\to S$ es también sobreyectiva. Luego w puede escribirse como $w=f(s)$ para algún $s\in S$. Tenemos entonces que $f(v)\in S^\perp$ pues

$$\langle f(v), w\rangle = \langle f(v), f(s)\rangle = \langle v, f^*f(s)\rangle = \langle v,s\rangle = 0,$$

que es lo que se quería demostrar. □

8.9.4. Sea V un espacio vectorial real con producto interno y de dimensión finita. Una **rotación** de V es una transformación lineal $V\to V$ tal que $\det f=1$. Si $S\subseteq V$ es un **hiperplano** (es decir, un subespacio de dimensión $\dim V-1$) entonces $f\colon V\to V$ es una **simetría** respecto de S si $f|_S=\mathrm{id}_S$ y $f|_{S^\perp}=-\mathrm{id}_{S^\perp}$.

ortogonales:dim=2

8.9.5. Proposición. *Sea V un espacio vectorial real con producto interno y tal que $\dim V=2$. Sea $f\colon V\to V$ una transformación ortogonal. Entonces f es una simetría o una rotación.*

Demostración. Como f es ortogonal, entonces $f^*f=ff^*=\mathrm{id}_V$. Si $\mathscr{B}=\{v_1,v_2\}$ es una base ortonormal de V entonces $\{f(v_1),f(v_2)\}$ es una base ortonormal de V pues

$$\langle f(v_i), f(v_j)\rangle = \langle v_i, f^*f(v_j)\rangle = \langle v_i,v_j\rangle = \delta_{ij}$$

para todo i,j. Sean $a,b,c,d\in\mathbb{R}$ tales que

$$f(v_1)=av_1+bv_2,\quad f(v_2)=cv_1+dv_2.$$

Entonces $\{(a,b),(c,d)\}$ es una base ortonormal de $\mathbb{R}^2$ pues

$$a^2+b^2=c^2+d^2=1,\quad ac+bd=0.$$

Esto implica que $a^2+b^2=1$ y $(c,d)\in\{(-b,a),(b,-a)\}$. En efecto, como $a^2+b^2=1$, entonces, al multiplicar por c^2, se tiene que $(ac)^2+(bc)^2=c^2$. Luego, como $ac=-bd$,

$$b^2=b^2(d^2+c^2)=(bd)^2+(bc)^2=c^2,$$

o bien $c\in\{-b,b\}$. Luego, si $c=\pm b$, entonces $(c,d)\in\{(-b,a),(b,-a)\}$ tal como habíamos afirmado.

Se tienen entonces dos posibilidades:

Si $(c,d)=(-b,a)$ entonces

$$[f]_{\mathscr{B},\mathscr{B}}=\begin{pmatrix} a & -b\\ b & a\end{pmatrix},\quad \chi_f=X^2-2aX+1.$$

Si χ_f tiene raíces reales, entonces las raíces están en $\{-1,1\}$. Esto implica que $(a,b)\in\{(1,0),(-1,0)\}$ y luego $f=\pm\mathrm{id}_V$. Supongamos entonces que χ_f no tiene raíces reales. Esto implica que $(2a)^2-4a<0$ y entonces, como $a\in[-1,1]$, se tiene que $a\in(0,1)$. Existe entonces $\theta\in[0,2\pi)$ tal que $a=\cos\theta$ y $b=\sin\theta$. Luego

$$[f]_{\mathscr{B},\mathscr{B}}=\begin{pmatrix} \cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

y f es una rotación de ángulo θ.

Si $(c,d)=(b,-a)$ entonces, como

$$[f]_{\mathscr{B},\mathscr{B}}=\begin{pmatrix} a & b\\ b & -a\end{pmatrix}$$

es una matriz simétrica y $\chi_f=X^2-1$, existe una base ortonormal en cuya base la matriz de f es $\operatorname{diag}(1,-1)$, es decir, f es una simetría. □

ortogonales:dim=3

8.9.6. PROPOSICIÓN. *Sea V un espacio vectorial real con producto interno y tal que $\dim V = 3$. Sea $f\colon V \to V$ una transformación ortogonal. Entonces f es una simetría, una rotación, o una composición de una simetría y una rotación.*

DEMOSTRACIÓN. Como χ_f es de grado tres, entonces tiene una raíz real, es decir f tiene un autovalor real. Por el lema 8.9.3 sabemos que este autovalor está en $\{-1, 1\}$.

Si $\lambda = 1$ es autovalor de f, sea $v_1 \in V$ autovector de autovalor $\lambda = 1$ tal que $\|v_1\| = 1$. Sea $S = \langle v_1 \rangle$. Como v_1 es autovector, S es f-invariante. Además $S^\perp$ es f-invariante por el lema 8.9.3. La restricción $f|_{S^\perp}\colon S^\perp \to S^\perp$ es una transformación ortogonal si consideramos a $S^\perp$ con el producto interno inducido por el producto interno de V. Como $\dim S^\perp = 2$, la proposición 8.9.5 implica que existe una base ortonormal $\{v_2, v_3\}$ de $S^\perp$ tal que, en esa base, la matriz de $f|_{S^\perp}$ es

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ o bien } \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ para algún } \theta \in [0, 2\pi).$$

Luego $\mathscr{B} = \{v_1, v_2, v_3\}$ es una base ortonormal de V y vale que

$$[f]_{\mathscr{B},\mathscr{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

lo que significa que f es una simetría con respecto al subespacio $\langle v_1, v_2 \rangle$, o bien que

$$[f]_{\mathscr{B},\mathscr{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

para algún $\theta \in [0, 2\pi)$.

Si 1 no es autovalor de f entonces, por lo visto anteriormente, $\lambda = -1$ es autovalor de f. Sea $v_1 \in V$ un autovector de autovalor $\lambda = -1$ tal que $\|v_1\| = 1$. Tal como se hizo en el caso anterior, se demuestra que existe una base ortonormal $\mathscr{B} = \{v_1, v_2, v_3\}$ de V tal que

$$[f]_{\mathscr{B},\mathscr{B}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

para algún $\theta \in [0, 2\pi)$. El teorema queda demostrado al observar que

$$[f]_{\mathscr{B},\mathscr{B}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},$$

es decir que $[f]_{\mathscr{B},\mathscr{B}}$ es la composición de una simetría y una rotación. □

8.9.7. TEOREMA. *Sea V un espacio vectorial real y de dimensión finita. Sea $f\colon V \to V$ una transformación ortogonal. Entonces existe una base ortonormal $\mathscr{B}$ de V y existen $\theta_1, \dots, \theta_k \in [0, \pi)$ tales que*

$$[f]_{\mathscr{B},\mathscr{B}} = \begin{pmatrix} I_r & & & & \\ & -I_s & & & \\ & & R_{\theta_1} & & \\ & & & \ddots & \\ & & & & R_{\theta_k} \end{pmatrix},$$

donde $r, s \in \mathbb{N}_0$, I_r es la matriz identidad de $r \times r$, I_s es la matriz identidad de $s \times s$, $\theta_1, \dots, \theta_k \in [0, 2\pi)$ y para cada $j \in \{1, \dots, k\}$

$$R_{\theta_j} = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix}.$$

DEMOSTRACIÓN. Procederemos por inducción en $n = \dim V$.

El caso $n = 2$ fue demostrado en la proposición 8.9.5. Supongamos entonces que $n > 2$ y que el resultado es válido para transformaciones ortogonales definidas en espacios de dimensión $< n$. Sea V un espacio de dimensión n y sea $f: V \to V$ una transformación ortogonal. Por el lema 8.9.3 hay tres casos a considerar:

Supongamos que $\lambda = 1$ es autovalor de f. Sea $v_1 \in V$ autovector de autovalor $\lambda = 1$ tal que $\|v_1\| = 1$. Entonces $S = \langle v_1 \rangle$ es un subespacio f-invariante y $S^{\perp}$ es también f-invariante por. Como $\dim S^{\perp} = n - 1$, la hipótesis inductiva aplicada a la restricción $f|_{S^{\perp}}: S^{\perp} \to S^{\perp}$ implica que existe una base ortonormal $\{v_2, \dots, v_n\}$ de $S^{\perp}$ tal que, en esa base, la matriz de $f|_{S^{\perp}}$ es de la forma deseada. Observemos que $\{v_1, v_2, \dots, v_n\}$ es una base ortonormal de V y que, en esa base, la matriz de f es de la forma buscada.

Si 1 no es autovalor de f y -1 es autovalor de f, se tiene un autovector v_1 de autovalor -1 y tal que $\|v_1\| = 1$. En este caso se procede como se hizo en el caso anterior.

Si f no tiene autovalores reales entonces $m_f = p_1 \cdots p_k$, donde para cada $i \in \{1, \dots, k\}$ se tiene que $p_i \in \mathbb{R}[X]$ es irreducible de grado dos. Sea $q = p_2 \cdots p_k$. Entonces q divide a m_f y además, como $q \neq m_f$, existe $w \in V$ tal que $q(f)(w) \neq 0$. Si $v = q(f)(w)$ entonces

$$0 = m_f(f)(w) = ((p_1 q)(f))(w) = p_1(f)(q(f)(w)) = p_1(f)(v).$$

Si $S = \langle v, f(v) \rangle$ entonces S es f-invariante pues $p_1(f)(v) = 0$ y $\deg p_1 = 2$; además, como v no es autovector de f, se tiene que $\dim S = 2$. Luego, como la restricción $f|_S: S \to S$ es ortogonal, existe una base ortonormal de S tal que, en esa base, la matriz de $f|_S$ es de la forma R_{θ_1} para algún $\theta_1 \in [0, 2\pi)$. Como $S^{\perp}$ es f-invariante y $\dim S^{\perp} = n - 2$, por hipótesis inductiva, existe una base ortonormal $\{v_3, \dots, v_n\}$ de $S^{\perp}$ tal que, en esa base, la matriz de $f|_{S^{\perp}}$ es

$$\begin{pmatrix} R_{\theta_2} & & \\ & \ddots & \\ & & R_{\theta_k} \end{pmatrix},$$

donde $\theta_2, \dots, \theta_k \in [0, 2\pi)$. (Sabemos que las restricción $f|_{S^{\perp}}$ no tiene autovalores reales pues el polinomio característico de $f|_{S^{\perp}}$ divide χ_f.) El conjunto $\mathscr{B} = \{v, f(v), v_3, \dots, v_n\}$ es ortonormal y la matriz de f en la base $\mathscr{B}$ es de la forma deseada. $\square$

Capítulo 9

Cocientes

9.1. Espacio cociente

9.1.1. Sea V un espacio vectorial sobre $\mathbb{K}$ y sea $S \subseteq V$ un subespacio. Se define la relación de equivalencia en V dada por

$$u \equiv v \text{ mód } S \Leftrightarrow u - v \in S.$$

Si $u \equiv v$ mód S diremos que u y v son **equivalentes módulo** S. Cada clase de equivalencia módulo S de V es de la forma $v + S$ con $v \in V$ pues

$$\begin{aligned} &\{u \in V : u \equiv v \text{ mód } S\} \\ &\quad = \{u \in V : u - v \in S\} \\ &\quad = \{u \in V : u = v + s \text{ para algún } s \in S\} \\ &\quad = \{v + s : s \in S\}. \end{aligned}$$

Las clases de equivalencia de V módulo S se denominan **coclases** de S en V. El conjunto de coclases de S en V se denota por $V/S = \{v + S : v \in V\}$.

block:V/S

9.1.2. Si $S \subseteq V$ es un subespacio entonces en V/S pueden definirse las siguientes operaciones:

$$\begin{aligned} &(u + S) + (v + S) = (u + v) + S, && u, v \in V, \\ &\lambda(v + S) = \lambda v + S, && v \in V,\ \lambda \in \mathbb{K}. \end{aligned}$$

Veamos que las operaciones están bien definidas.

Supongamos que $u_1 + S = u_2 + S$ y sea $\lambda \in \mathbb{K}$. Entonces $u_1 - u_2 \in S$ y, como S es un subespacio, $\lambda u_1 - \lambda u_2 - \lambda(u_1 - u_2) \in S$. Luego $\lambda u_1 + S = \lambda u_2 + S$ y entonces la multiplicación por escalares en V/S está bien definida.

Supongamos ahora que $u_1 + S = u_2 + S$ y que $v_1 + S = v_2 + S$. Entonces $u_1 - u_2 \in S$ y $v_1 - v_2 \in S$ y, como S es un subespacio,

$$(u_1 - u_2) + (v_1 - v_2) = (u_1 + v_1) - (u_2 + v_2) \in S.$$

Luego $u_1 + v_1 + S = u_2 + v_2 + S$ y entonces la suma en V/S está bien definida.

9.1.3. Ejercicio. Demostrar que con la suma y el producto por escalares definidos en 9.1.2, el cociente V/S es un espacio vectorial.

9.1.4. Ejemplos. Es evidente que $V/\{0\} \simeq V$ y que $V/V \simeq \{0\}$.

1) $V/0 \simeq V$.
2) $V/V \simeq 0$.
3) $\mathbb{R}^2/\langle(1,1)\rangle$.

9.2. Teoremas de isomorfismo

9.2.1. Proposición. *Sea $S \subseteq V$ un subespacio. La función*

$$p_S \colon V \to V/S, \quad v \mapsto v + S$$

es un epimorfismo y $\ker p_S = S$. La transformación lineal p_S se denomina el ***epimorfismo canónico*** *de V en V/S.*

DEMOSTRACIÓN. La función p_S es una transformación lineal pues

$$\begin{aligned} p_S(u+v) &= (u+v)+S = (u+S)+(v+S) = p_S(u)+p_S(v), \\ p_S(\lambda v) &= (\lambda v)+S = \lambda(v+S) = \lambda p_S(v), \end{aligned}$$

para todo $u, v \in V$ y $\lambda \in \mathbb{K}$. Es evidente que p_S es un epimorfismo. Calculemos $\ker(p_S)$. Es claro que $S \subseteq \ker(p_S)$, y si $v \in \ker(p_S)$ entonces $p_S(v) = v+S = S$ y luego $v \in S$. $\square$

9.2.2. COROLARIO. *Sea V un espacio vectorial de dimensión finita. Entonces V/S es también de dimensión finita y*

$$\dim(V/S) = \dim(V) - \dim(S).$$

La dimensión de V/S se denomina ***codimensión*** *de S en V y se denota por* $\operatorname{codim}(S)$.

DEMOSTRACIÓN. Si V es de dimensión finita entonces, por el teorema de la dimensión aplicado al epimorfismo canónico p_S,

$$\dim(V) = \dim\ker(p_S) + \dim\operatorname{im}(p_S) = \dim(S) + \dim(V/S),$$

que es equivalente a lo que se quería demostrar. $\square$

9.2.3. COROLARIO. *Sea V un espacio vectorial y sean $S, T \subseteq V$ subespacios tales que $V = S \oplus T$. Entonces $T \simeq V/S$, es decir: todo complemento de S en V es isomorfo a V/S.*

DEMOSTRACIÓN. Sea $f\colon T \to V/S$ la transformación lineal dada por $t \mapsto t+S$. Veamos que f es monomorfismo:

$$\ker f = \{t \in T : f(t) = S\} = \{t \in T : t \in S\} = S \cap T = \{0\}$$

pues $V = S \oplus T$. Veamos que f es epimorfismo:

$$\operatorname{im} f = \{f(t) : t \in T\} = \{t+S : t \in T\} = \{v+S : v \in V\}$$

pues todo $v \in V$ se escribe unívocamente como $v = s+t$ con $s \in S$ y $t \in T$. Luego $T \simeq V/S$. $\square$

9.2.4. TEOREMA (teorema de la correspondencia de subespacios). *Sea $S \subseteq V$ un subespacio. La función*

$$\{\textit{subespacios de } V \textit{ que contienen a } S\} \to \{\textit{subespacios de } V/S\},$$

dada por $T \mapsto p(T)$, donde $p\colon V \to V/S$ es el epimorfismo canónico, es biyectiva.

DEMOSTRACIÓN. Observemos que si T es un subespacio de V y $S \subseteq T$ entonces $p(T)$ es un subespacio de V/S y así la función $T \mapsto p(T)$ del enunciado está bien definida. Veamos que la función

$$\begin{gathered} \{\text{subespacios de } V \text{ que contienen a } S\} \to \{\text{subespacios de } V/S\} \\ p^{-1}(L) \hookleftarrow L \end{gathered}$$

está bien definida: si $L \subseteq V/S$ es un subespacio, $p^{-1}(L)$ es un subespacio de V que contiene a S pues si $s \in S$ entonces $p(s) = 0 \in L$.

Veamos que $p^{-1}(L) \hookleftarrow L$ es la inversa de $T \mapsto p(T)$. Primero observemos que $p(p^{-1}(L)) = L$ pues p es sobreyectiva. Por otro lado, debemos demostrar que $p^{-1}(p(T)) = T$. Como es evidente que $p^{-1}(p(T)) \supseteq T$, basta demostrar que $p^{-1}(p(T)) \subseteq T$. Si $v \in p^{-1}(p(T))$, como entonces $p(v) \in p(T)$, se tiene que $p(v) = p(t)$ para algún $t \in T$. Luego $v - t \in \ker p = S \subseteq T$ y por lo tanto $v \in T$. $\square$

opiedad_universal

9.2.5. Teorema (propiedad universal del cociente). *Sean V y W espacios vectoriales, $f \in \hom(V,W)$ y $S \subseteq \ker f$ un subespacio. Entonces existe una única $g \in \hom(V/S, W)$ tal que $gp_S = f$, es decir: existe una única $g \in \hom(V/S,W)$ que hace que el siguiente diagrama conmute:*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "p_S"'] & W \\ V/S \arrow[ur, dashed, "g"'] & \end{tikzcd}$$

Más aún, $\ker g = \ker f/S$ *y* $\operatorname{im} g = \operatorname{im} f$.

Demostración. Definimos a g como $g(v+S) = f(v)$. Veamos que está bien definida: si $v_1 + S = v_2 + S$ entonces $v_1 - v_2 \in S$. Como $S \subseteq \ker f$, $f(v_1 - v_2) = 0$ y luego $g(v_1 + S) = f(v_1) = f(v_2) = g(v_2 + S)$

Como f es una transformación lineal, es fácil ver que g es también una transformación lineal. Calculemos la imagen de g:

$$\operatorname{im} g = \{g(v+S) : v \in V\} = \{f(v) : v \in V\} = \operatorname{im} f.$$

Calculemos ahora el núcleo de g:

$$\begin{aligned} \ker g &= \{v + S : g(v+S) = 0\} \\ &= \{v + S : f(v) = 0\} = \{v + S : v \in \ker f\} = \ker f + S. \end{aligned}$$

Queda demostrar la unicidad: si $h \in \hom(V/S, W)$ cumple que $hp_S = f$ entonces $h(v+S) = hp_S(v) = f(v) = gp_S(v) = g(v+S)$ y luego $h = g$. □

9.2.6. Corolario (primer teorema de isomorfismo). *Sea $f \in \hom(V,W)$. Entonces la transformación lineal $V/\ker f \to W$, $v + \ker f \mapsto f(v)$, es inyectiva. En particular $V/\ker f \simeq \operatorname{im} f$.*

Demostración. Es consecuencia de la propiedad universal del cociente, teorema 9.2.5, aplicada al subespacio $S = \ker f$. □

9.2.7. Corolario (segundo teorema de isomorfismo). *Sean $S, T \subseteq V$ subespacios. Entonces*

$$\frac{S+T}{T} \simeq \frac{S}{S \cap T}.$$

Demostración. Sea $f \colon S + T \to S/S \cap T$ dada por $f(s+t) = s + (S \cap T)$. Primero debemos demostrar que f está bien definida: si $s + t = s' + t'$ con $s, s' \in S$ y $t, t' \in T$ entonces $s' - s \in S \cap T$ y $s + (s' - s) = s'$. Luego $f(s+t) = f(s' + t')$. Como

$$\ker f = \{s + t : s + S \cap T = S \cap T\} = \{s + t : s \in S \cap T\} = \{t + S \cap T\} = T,$$

y f es un epimorfismo, el primer teorema de isomorfismo demuestra el corolario. □

9.2.8. Corolario (tercer teorema de isomorfismo). *Sean $S \subseteq T \subseteq V$ subespacios. Entonces*

$$\frac{V/S}{T/S} \simeq \frac{V}{T}.$$

Demostración. Sea $f \colon V/S \to T/S$ dada por $v + S \mapsto v + T$. Veamos la buena definición: si $v_1 + S = v_2 + S$ entonces $v_1 - v_2 \in S$ y luego, como $S \subseteq T$, $v_1 + T = v_2 + T$. Como

$$\ker f = \{v + S : v + T = T\} = \{v + S : v \in T\} = T,$$

y f es epimorfismo, el primer teorema de isomorfismo nos da el isomorfismo que queríamos demostrar. □

Capítulo 10

Soluciones a los ejercicios

10.1. Capítulo 1

Solución (ejercicio 1.2.26). Sea y tal que $Ay = b$. Entonces $y - p$ es solución de $Ax = 0$. Luego $y = (y - p) + p \in S + p$. Recíprocamente, si y satisface $Ay = 0$ entonces $A(y + p) = Ay + Ap = 0 + b = b$.

Solución (ejercicio 1.2.28).

10.2. Capítulo 2

Solución (ejercicio 2.2.2). Si S es un subespacio entonces es evidente que $v+\lambda w \in S$ para todo $v, w \in S$ y $\lambda \in \mathbb{K}$. Recíprocamente, como S es no vacío, existe un elemento $v_0 \in S$ y entonces $0 = v_0 + (-1)v_0 \in S$. Si $v, w \in S$ entonces $v + w = v + 1w \in S$. Por último, si $v \in S$ y $\lambda \in \mathbb{K}$ entonces $\lambda v = v + (\lambda - 1)v \in S$.

Solución (ejercicio 2.3.8). Si $f(1) = 0$ entonces $X-1$ divide a f y luego $f = (X-1)g$ para algún $g \in \mathbb{R}[X]$.

Solución (ejercicio 2.3.9). Como $f(1) = f(2) = 0$, el polinomio

$$X^2 - 3X + 2 = (X - 1)(X - 2)$$

divide a f. Luego $f = (X^2 - 3x + 2)g$ para algún $g \in \mathbb{R}[X]$.

Solución (ejercicio 2.3.15).

10.3. Capítulo 3

Solución (ejercicio 3.2.19). Para demostrar la primera afirmación, sea $g_1 = g|_{\operatorname{im} f}$. Es evidente que $\ker g_1 \subseteq \ker g$. Entonces, por el teorema de la dimensión,

$$\begin{aligned}
\dim\ker(gf) &= \dim U - \dim\operatorname{im}(gf) \\
&= \dim U - \dim\operatorname{im} f + \dim\operatorname{im} f - \dim\operatorname{im}(gf) \\
&= \dim\ker f + \dim\operatorname{im} f - \dim\operatorname{im}(gf) \\
&= \dim\ker f + \dim\ker g_1 \\
&\leq \dim\ker f + \dim\ker g.
\end{aligned}$$

Para demostrar la segunda desigualdad, primero observamos que

$$\dim\operatorname{im}(gf) \leq \dim\operatorname{im} g$$

pues $\operatorname{im}(gf) \subseteq \operatorname{im} g$. Por otro lado, como $\operatorname{im}(gf) \subseteq \operatorname{im} g_1$, tenemos

$$\dim\operatorname{im}(gf) \leq \dim\operatorname{im}(g_1) = \dim\operatorname{im} f - \dim\ker g_1 \leq \dim\operatorname{im} f.$$

Luego,

$$\dim\operatorname{im}(gf) \leq \operatorname{mín}\{\dim\operatorname{im} g, \dim\operatorname{im} f\}.$$

Para demostrar la tercera igualdad, usamos el teorema de la dimensión en gf y el primer ítem y obtenemos

$$\begin{aligned}\dim \operatorname{im}(gf) &= \dim U - \dim \ker(gf)\\ &\geq \dim U - \dim \ker f - \dim \ker g\\ &= \dim U - \dim U + \dim \operatorname{im} f - \dim V + \dim \operatorname{im} g\\ &= \dim \operatorname{im} f + \dim \operatorname{im} g - \dim V.\end{aligned}$$

Solución (ejercicio 3.3.4). Observemos que $(2f - \mathrm{id}_V)^2 = \mathrm{id}_V$ si y sólo si $(2f - \mathrm{id}_V)(2f - \mathrm{id}_V)(v) = v$ para todo $v \in V$. Esto es equivalente a

$$4f^2(v) - 4f(v) - v = v$$

para todo $v \in V$, que a su vez equivale a $4f^2(v) = 4f(v)$ para todo $v \in V$. Como el cuerpo de base es $\mathbb{R}$, esta última igualdad equivale a $f^2(v) = v$ para todo $v \in V$.

Solución (ejercicio 3.3.3). Si f es un proyector y $w \in \operatorname{im} f$ entonces $w = f(v)$ para algún $v \in V$ y luego $f(w) = f(f(v)) = f(v) = w$. Recíprocamente, si $f(w) = w$ para todo $w \in \operatorname{im} f$ y $v \in V$ entonces $f(f(v)) = f(v)$.

Solución (ejercicio 3.3.6). Supongamos que $f(S) \subseteq S$ y sea $p\colon V \to V$ con $\operatorname{im} p = S$. Sea $v \in V$. Entonces $(fp)(v) \in S = \operatorname{im} p$ y luego existe $w \in V$ tal que $(fp)(v) = p(w)$. Al aplicar p se obtiene

$$(pfp)(v) = p^2(w) = p(w) = (fp)(v),$$

que es lo que se quería demostrar.

Supongamos ahora que $pfp = fp$ para todo proyector $p\colon V \to V$ con $\operatorname{im} p = S$. Si $s \in S$ entonces, como $S = \operatorname{im} p$ y $p(s) = s$, $f(s) = fp(s) = pfp(s) \in S$.

Solución (ejercicio 3.3.8). Para demostrar que (1) implica (2) basta tomar las transformaciones lineales $p_i\colon V \to V$ dado por $v = s_1 + \cdots + s_n \mapsto s_i$. Recíprocamente, si $s_1 + \cdots + s_n = 0$ entonces, al aplicar p_i y utilizar que los p_i son proyectores que satisfacen $p_i p_j = 0$ si $i \neq j$, se obtiene que $s_i = 0$ para todo i.

Solución (ejercicio 3.3.9). Si f es un proyector entonces $V = \ker f \oplus \operatorname{im} f$ y además $f(w) = w$ para todo $w \in \operatorname{im} f$. Sea $\{v_1, \dots, v_r\}$ una base de $\operatorname{im} f$ y $\{v_{r+1}, \dots, v_n\}$ una base de $\ker f$. Luego $\mathscr{B} = \{v_1, \dots, v_n\}$ es base de V y la matriz de f en la base $\mathscr{B}$ tiene la forma deseada.

Recíprocamente, si $\{v_1, \dots, v_r\}$ es base de $\operatorname{im} f$ y $\{v_{r+1}, \dots, v_n\}$ es base de $\ker f$ entonces $f(w) = w$ para todo $w \in \operatorname{im} f$. Luego f es un proyector por el ejercicio 3.3.3.

10.4. Capítulo 4

Solución (ejercicio 4.1.6). Si $i \neq j$ entonces

$$\delta(E^{ij}) = \delta(E^{ii}E^{ij}) = \delta(E^{ij}E^{ii}) = \delta(0) = 0.$$

Por otro lado, para cada $i, k \in \{1, \dots, n\}$ tenemos

$$\delta(E^{ii}) = \delta(E^{ik}E^{ki}) = \delta(E^{ki}E^{ik}) = \delta(E^{kk}).$$

Sea entonces $\lambda = \delta(E^{11})$. Si escribimos a A en la base de los E^{ij} tenemos

$$\delta(A) = \sum_{1 \leq i,j \leq n} a_{ij}\delta(E^{ij}) = \sum_{i=1}^{n} a_{ii}\delta(E^{ii}) = \lambda \sum_{i=1}^{n} a_{ii} = \lambda \operatorname{tr}(A),$$

tal como queríamos demostrar.

Solución (ejercicio 4.1.12). Si $v = 0$ entonces $f(v) = 0$ para todo $f \in V^*$. Recíprocamente, si $v \neq 0$, extendemos el conjunto linealmente independiente $\{v\}$ a una base $\mathscr{B}$ de V. La existencia de la base dual a la base $\mathscr{B}$ garantiza la existencia de un $f \in V^*$ tal que $f(v) = 1$.

Solución (ejercicio 4.1.13). Como $f(v) \neq 0$, todo $w \in V$ puede escribirse como

$$w = w - \frac{f(w)}{f(v)}v + \frac{f(w)}{f(v)}v,$$

donde $f\left(w - \frac{f(w)}{f(v)}v\right) = 0$, y entonces $V = \ker f + \langle v \rangle$. Si $w \in \ker f \cap \langle v \rangle$ entonces $f(w) = 0$ y $w = \lambda v$ para algún $\lambda \in \mathbb{K}$. Luego $\lambda f(v) = 0$ y entonces, como $f(v) \neq 0$, $\lambda = 0$ y $w = 0$.

Solución (ejercicio 4.1.14). Si existe $\lambda \in \mathbb{K}$ tal que $f = \lambda g$ entonces es evidente que $\ker f = \ker g$. Recíprocamente, supongamos que $f \neq 0$ (de lo contrario, el resultado es trivial) y sea $\{v_1, \dots, v_{n-1}\}$ una base de $\ker f$. La extendemos a una base $\{v_1, \dots, v_n\}$ de V y entonces, por construcción, $f(v_n) \neq 0$ y $g(v_n) \neq 0$. Si definimos $\lambda = f(v_n)/g(v_n)$ entonces $\lambda \neq 0$ y luego $f = \lambda g$ pues coinciden en la base $\{v_1, \dots, v_n\}$.

Solución (ejercicio 4.3.2). Si $v \in \langle X \rangle$ entonces existen $x_1, \dots, x_n \in X$ y $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tales que $v = \sum_{i=1}^{n} \alpha_i x_i$. Luego $f(v) = 0$ para toda $f \in \operatorname{Ann} X$ y entonces $\operatorname{Ann} X \subseteq \operatorname{Ann}\langle X \rangle$. La recíproca es cierta pues si $f \in \operatorname{Ann}\langle X \rangle$ entonces $X \subseteq \langle X \rangle \subseteq \ker f$.

Solución (ejercicio 4.4.6). Supongamos que $\{f_1, \dots, f_n\}$ son linealmente independientes y sea $\{v_1, \dots, v_n\}$ una base de V que tiene a la base de las f_i como su base dual. Si escribimos $v = \sum_{i=1}^{n} \alpha_i v_i$ entonces para cada j se tiene

$$0 = f_j(v) = f_j\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i f_j(v_i) = \sum_{i=1}^{n} \alpha_i \delta_{ji} = \alpha_j.$$

Luego $v = 0$, una contradicción.

10.5. Capítulo 5

Solución (ejercicio 5.1.6). Sea $j \in \{1, \dots, n\}$. Si $\alpha(j) \neq j$ entonces $\beta(j) = j$ y además $\beta(\alpha(j)) = \alpha(j)$ (pues de lo contrario tendríamos $\alpha(\alpha(j)) = \alpha(j)$ que implica $\alpha(j) = j$). Luego

$$(\alpha\beta)(j) = \alpha(\beta(j)) = \alpha(j) = \beta(\alpha(j)).$$

Similarmente se hace el caso en que $\beta(j) \neq j$. Por último, si $\alpha(j) = \beta(j) = j$ entonces trivialmente se obtiene $(\alpha\beta)(j) = (\beta\alpha)(j)$.

Solución (ejercicio 5.1.8). Si $k \geq 0$ entonces

$$\sigma^k(i) = (\alpha\beta)^k(i) = \alpha^k\left(\beta^k(i)\right) = \alpha^k(i).$$

Solución (ejercicio 5.4.4). Supongamos que d es una función n-lineal y alternada sea $A \in \mathbb{K}^{n\times n}$. Escribamos a cada fila de A como $A_i = \sum_{j=1}^{n} a_{ij} e_j$ donde $\{e_1, \dots, e_n\}$ es la base canónica de $\mathbb{K}^{1\times n}$. Entonces

$$d(A) = d(A_1, \dots, A_n) = \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1j_1} \cdots a_{nj_n} d(e_{j_1}, \dots, e_{j_n}).$$

Como d es alternada, la suma es no nula únicamente cuando todos los j_k son distintos, es decir cuando $|\{j_1, \dots, j_n\}| = n$. Entonces, la suma anterior se hace sobre todas las n-uplas $(j_1, \dots, j_n)$ de elementos distintos del conjunto $\{1, \dots, n\}$. Luego, por el lema 5.4.1,

$$\begin{aligned} d(A) &= \sum_{\sigma \in \mathbb{S}_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} d(e_{\sigma(i)}, \dots, e_{\sigma(n)}) \\ &= \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} d(e_1, \dots, e_n) \\ &= (\det A) d(I). \end{aligned}$$

Solución (ejercicio 5.4.15). Escribimos

$$\det\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & C \end{pmatrix}\begin{pmatrix} A & B \\ 0 & I \end{pmatrix}$$

y observamos que

$$\det\begin{pmatrix} I & 0 \\ 0 & C \end{pmatrix} = \det C, \qquad \det\begin{pmatrix} A & B \\ 0 & I \end{pmatrix} = \det A.$$

Solución (ejercicio 5.5.3). Si $f = a_0 + a_1X + \cdots + a_nX^n$ y $f(x_i) = y_i$ para todo $i \in \{1, \dots, n+1\}$ entonces

$$\begin{cases} a_0 + a_1x_1 + \cdots + a_nx_1^n = y_1, \\ a_0 + a_1x_2 + \cdots + a_nx_2^n = y_2, \\ \quad\vdots \\ a_0 + a_1x_{n+1} + \cdots + a_nx_{n+1}^n = y_{n+1}. \end{cases}$$

Este sistema tiene solución única pues el determinante de la matriz asociada es el determinante de Vandermonde

$$V(x_1, \dots, x_{n+1}) = \prod_{i<j}(x_i - x_j) \neq 0.$$

Solución (ejercicio 5.4.23). Como $A^{-1} = (\det A)^{-1}(\operatorname{adj} A)$ entonces $(\det A)A = (\operatorname{adj} A)^{-1}$. Por otro lado, al aplicar determinante en la igualdad $(\operatorname{adj} A)A = (\det A)I$, se obtiene $\det \operatorname{adj} A = (\det A)^{n-1}$. Luego

$$\operatorname{adj}(\operatorname{adj} A) = (\det A)^{n-1}(\operatorname{adj} A)^{-1} = (\det A)^{n-2}A.$$

Solución (ejercicio 5.4.24). Sabemos que

$$\operatorname{adj}(BAB^{-1})(BAB^{-1}) = \det(BAB^{-1})I = (\det A)I.$$

Entonces

$$\operatorname{adj}(BAB^{-1}) = (\det A)BA^{-1}B = B(\operatorname{adj} A)B^{-1}$$

pues $A^{-1} = (\det A)^{-1} \operatorname{adj} A$.

Solución (ejercicio 5.7.1). Sean $\alpha, \beta \in \mathbb{R}$ tales que $\alpha A^{1000} + \beta A^{1001} = 0$. Si aplicamos la función determinante a la igualdad $\alpha A^{1000} = -\beta A^{1001}$, obtenemos $\alpha^4(\det A)^{1000} = (-\beta)^4(\det A)^{1001}$. Luego $\alpha^4 + \beta^4 = 0$ y entonces $\alpha = \beta = 0$.

10.6. Capítulo 6

Solución (ejercicio 6.1.5). Sea $\lambda \in \operatorname{spec}(fg)$. Si $\lambda = 0$ entonces $fg \notin \operatorname{Aut} V$. Luego, como $f \notin \operatorname{Aut} V$ o $g \notin AutV$, se tiene que $gf \notin \operatorname{Aut} V$. Si $\lambda \neq 0$ sea $v \in V \setminus \{0\}$ tal que $(fg)(v) = \lambda v \neq 0$. Entonces, como $w = g(v) \neq 0$,

$$(gf)(w) = (fgf)(v) = g(\lambda v) = \lambda g(v) = \lambda w.$$

Luego $\lambda \in \operatorname{spec}(gf)$ y entonces $\operatorname{spec}(fg) \subseteq \operatorname{spec}(gf)$.

10.7. Capítulo 7

Solución (ejercicio 7.1.12). La primera afirmación es consecuencia de lo visto en 7.1.11. Demostremos la segunda afirmación. Sea

$$p = \operatorname{mcm}(m_{f|_S}, m_{f|_T}).$$

Como $m_{f|_S}$ y $m_{f|_T}$ dividen a f, entonces p divide a m_f. Por otro lado, como $V = S \oplus T$, para cada $v \in V$ existe únicos $s \in T$ y $T \in T$ tales que $v = s + t$. Luego

$$p(f)(v) = p(f)(s+t) = p(f)(s) + p(f)(t) = p(f|_S)(s) + p(f|_T)(t).$$

Como $m_{f|_S}$ y $m_{f|_T}$ ambos dividen a p, entonces $p(f|_S)(s) = p(f|_T)(t) = 0$. Luego m_f divide a p y entonces, como m_f y p son mónicos, $m_f = p$.

Solución (ejercicio 7.1.15).

1) Primero observemos que, como $p(f) \circ f = f \circ p(f)$, entonces
$$p(f)(f^i(v)) = f^i(p(f)(v)) \in C(p(f)(v))$$
para todo $i \geq 0$ y luego $p(C(v)) \subseteq C(p(v))$. Recíprocamente,
$$f^i(p(f)(v)) = p(f)(f^i(v)) \in p(f)(C(v))$$
para todo $i \geq 0$ y luego $p(C(v)) \supseteq C(p(v))$.
2) Como $p(f)$ es un endomorfismo de V, entonces
$$p(V) = p(V_1) + \cdots + p(V_k).$$
Para $i \in \{1, \dots, k\}$ sean $v_i \in V_i$ tales que $v_1 + \cdots + v_k = 0$. Entonces $p(f)(v_1) + \cdots + p(f)(v_k) = p(f)(v_1 + \cdots + v_k) = 0$. Como los V_i son p-invariantes, $p(f)(v_i) \in V_i$ y luego, como los V_i están en suma directa, $p(f)(v_i) = 0$ para cada $i \in \{1, \dots, k\}$.
3) Por definición,
$$0 = m_{p(v)}(f)(p(f)(v)) = (m_{p(v)}p)(f)(v)$$
y entonces $m_v = m_w$ divide al polinomio $m_{p(v)}p$. En particular, como $m_{p(v)}p$ da cero al ser evaluado en w, obtememos
$$0 = (m_{p(v)}p)(f)(w) = m_{p(v)}(f)(p(f)(w)).$$
Luego $m_{p(w)}$ divide a $m_{p(v)}$. Análogamente $p_{p(v)}$ divide a $m_{p(w)}$ y entonces, como $m_{p(v)}$ y $m_{p(w)}$ son mónicos, $m_{p(v)} = m_{p(w)}$.
Como $m_{p(v)} = m_{p(w)}$ y que $\deg m_{p(v)} = \dim C(p(v))$ por el lema 7.1.14, entonces $\dim C(p(v)) = \dim C(p(w))$.

Solución (ejercicio 7.2.4). Sean $\lambda_1, \dots, \lambda_n$ los autovalores de f. Sea $\mathscr{B}$ una base de f tal que $[f]_{\mathscr{B},\mathscr{B}}$ es una matriz diagonal cuya diagonal principal es $\lambda_1, \dots, \lambda_n$. Como f es nilpotente, existe $m \in \mathbb{N}$ tal que $f^m = 0$. Entonces
$$0 = [f^m]_{\mathscr{B},\mathscr{B}} = [f]^m_{\mathscr{B},\mathscr{B}}.$$
Esto implica que $\lambda_i^m = 0$ para todo i y por lo tanto $\lambda_i = 0$ para todo i.

Solución (ejercicio 7.2.5).

1) Sea r el índice de nilpotencia de f. Si $r = 1$ entonces $f = 0$ y el ejercicio queda resuelto. Si $r > 1$ entonces sea $v \in V$ tal que $f^r(v) = 0$ y $f^{r-1}(v) \neq 0$. Si $w = f^{r-1}(v)$ entonces $f(w) = 0w$.
2) Sea r el índice de nilpotencia de f y sea $v \neq 0$ tal que $f(v) = \lambda v$. Entonces $0 = f^r(v) = \lambda^r v$. Como $v \neq 0$ entonces $\lambda^r = 0$ y luego $\lambda = 0$.

10.8. Capítulo 8

Solución (ejercicio 8.1.5). Si $\langle v, x\rangle = \langle w, x\rangle$ entonces $\langle v - w, x\rangle = 0$ y luego, al tomar $x = v - w$, se tiene $0 = \|v - w\|^2 = \langle v - w, v - w\rangle$, que implica que $v = w$.

Solución (ejercicio 8.2.7). Si $v \perp w$ entonces $\langle v, w\rangle = \langle w, v\rangle = 0$. Luego
$$\begin{aligned}\|v + w\|^2 &= \langle v + w, v + w\rangle \\ &= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle = \|v\|^2 + \|w\|^2.\end{aligned}$$

Solución (ejercicio 8.2.8).

1) Se deduce trivialmente de $\langle 0, v\rangle = \langle v, 0\rangle = 0$.

2) Si $v \in T^{\perp}$ entonces $\langle v, w\rangle = 0$ para todo $w \in T$. En particular, $v \in S^{\perp}$ pues, como $S \subseteq T$, $\langle v, w\rangle = 0$ para todo $w \in S$.
3) Como $S \subseteq S+T$ y $T \subseteq S+T$, entonces, por lo visto en ítem anterior, $(S+T)^{\perp} \subseteq S^{\perp} \cap T^{\perp}$. Recíprocamente, si $v \in S^{\perp} \cap T^{\perp}$ entonces, dados $x \in S$, $y \in T$, se tiene
$$\langle v, x+y\rangle = \langle v, x\rangle + \langle v, y\rangle = 0 + 0 = 0.$$

SOLUCIÓN (ejercicio 8.2.15). Por definición se tiene que $S \subseteq S^{\perp\perp}$. Por otro lado, como V es de dimensión finita y S y $S^{\perp}$ son subespacios de V, $\dim V = \dim S + \dim S^{\perp} = \dim S^{\perp} + \dim S^{\perp\perp}$. Como $\dim S = \dim S^{\perp\perp}$, entonces $S = S^{\perp\perp}$.

SOLUCIÓN (ejercicio 8.3.5).

1) Sea $v \in (imp)^{\perp}$. Como $v - p(v) \in \ker p$ y $\ker p \perp \operatorname{im} p$, entonces
$$\langle v, p(v)\rangle - \langle p(v), p(v)\rangle = \langle v - p(v), p(v)\rangle = 0.$$
Como $v \in (\operatorname{im} p)^{\perp}$ entonces $\langle v, p(v)\rangle = 0$ y luego $\|p(v)\|^2$. Esto implica que $p(v) = 0$ y que $v \in \ker p$. Por otro lado, como p es un proyector ortogonal, $\ker p \perp \operatorname{im} p$, y entonces es evidente que $\ker p \subseteq (\operatorname{im} p)^{\perp}$.
2) Como $\ker p \perp \operatorname{im} p$ entonces $\operatorname{im} p \subseteq (\ker p)^{\perp}$. Por otro lado, si $v \in (\ker p)^{\perp}$ entonces, como $v - p(v) \in \ker p$ y $\ker p \perp \operatorname{im} p$,
$$\begin{aligned}\|v - p(v)\|^2 &= \langle v - p(v), v - p(v)\rangle \\ &= \langle v, v - p(v)\rangle - \langle p(v), v - p(v)\rangle = 0.\end{aligned}$$
Luego $v = p(v) \in \operatorname{im} p$.
3) Se deduce trivialmente de los ítems anteriores.

SOLUCIÓN (ejercicio 8.3.7). Como $\operatorname{im} p_S \perp \operatorname{im} p_T$, entonces por el ejercicio 8.3.5, $\operatorname{im} p_T \subseteq (\operatorname{im} p_S)^{\perp} = \ker p_S$. Luego $p_S p_T = 0$.

Índice alfabético