Algebra Lineal: Aplicaciones a la Física

Resumen del curso 2012 para Lic. en Física (2º año), Depto. de Física, UNLP. Prof.: R. Rossignoli

0. Repaso de estructuras algebraicas básicas

Un sistema algebraico es un conjunto no vacío de elementos, A, junto con un conjunto de operaciones **cerradas** sobre A que satisfacen ciertas propiedades.

Un **monoide** es un sistema algebraico $\{A, *\}$ formado por un conjunto A y una operación binaria cerrada $*: A \times A \to A$ (también llamada ley de composición interna).
Ejemplos bien conocidos son $\{\mathbb{R}, +\}$ (el conj. de num. reales con la suma), $\{\mathbb{R}, \cdot\}$ (los reales con el prod.) $\{\mathbb{R}_+, \cdot\}$ (los reales positivos con el producto). En cambio, $\{\mathbb{R}_-, \cdot\}$, donde $\mathbb{R}_-$ denota los reales negativos, no es un monoide pues la operación no es cerrada.

Un **grupo** es un monoide $\{A, *\}$ donde A y la operación $*$ satisfacen
1) Asociatividad: $a * (b * c) = (a * b) * c \;\forall\; a, b, c \in A$
2) Existencia de elemento identidad: $\exists\; I \in A$ t.q. $\forall\; a \in A$, $a * I = I * a = a$
3) Existencia de elemento inverso: $\forall\; a \in A \;\exists\; a^{-1} \in A$ t.q. $a * a^{-1} = a^{-1} * a = I$
Un **semigrupo** es un monoide $\{A, *\}$ que satisface la prop. 1), aunque no necesariamente 2) y 3).

Recordemos la *unicidad* de la identidad y de la inversa:
i) La identidad I, si existe, es única: Supongamos que $\exists\; I$, I' t.q. $I * a = a$, $a * I' = a \;\forall\; a \in A$. Entonces $I * I' = I'$ (caso $a = I'$) y $I * I' = I$ (caso $a = I$) por lo que $I' = I$. Esto muestra también que si existe una identidad a izquierda I y una identidad a derecha I', necesariamente son coincidentes!
ii) La inversa $b = a^{-1}$ de a, si existe, es única: Supongamos que $\exists$ otra inversa b'. Entonces $b * (a * b') = b * I = b$, pero por asociatividad, $b * (a * b') = (b * a) * b' = I * b' = b'$, de donde $b = b'$.

Esto muestra también que si $\exists$ una inversa a izquierda b y una inversa a derecha b' necesariamente son coincidentes. Puede ocurrir, no obstante, que exista inversa a izquierda pero no a derecha, y viceversa, en cuyo caso pueden no ser únicas, como veremos posteriormente.

Si la operación binaria satisface $a * a' = a' * a \;\forall\; a, a' \in A$, se dice que es *conmutativa*. En tal caso, el grupo se denomina conmutativo o *abeliano*. En caso contrario el grupo es no abeliano o no conmutativo.

Ejemplos de grupos abelianos son:
1) $\{\mathbb{R}, +\}$, donde la identidad es el 0 y el inverso de $a \in \mathbb{R}$ el opuesto $-a$
2) $\{\mathbb{R}_0, \cdot\}$, donde $\mathbb{R}_0$ denota los reales sin el 0, la identidad es el 1 y el elemento inverso $a^{-1} = 1/a$
3)$\{\mathbb{R}^2, +\}$, donde la identidad es el vector nulo $(0, 0)$ y el inverso de $a = (x, y)$ el vector opuesto $(-x, -y)$
4) $\{\mathbb{R}^{2\times 2}, +\}$, donde la identidad es la matriz nula $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ y el inverso de $a = \left(\begin{smallmatrix} x & y \\ z & t \end{smallmatrix}\right)$ la matriz opuesta $\left(\begin{smallmatrix} -x & -y \\ -z & -t \end{smallmatrix}\right)$.

Nótese en cambio que $\{\mathbb{R}^3, \times\}$, donde $\times$ denota el producto vectorial, es un monoide pero no es grupo ni semigrupo pues el producto vectorial no es asociatvo (por ej., $\vec{i} \times (\vec{i} \times \vec{j}) = \vec{i} \times \vec{k} = -\vec{j} \neq (\vec{i} \times \vec{i}) \times \vec{j} = \vec{0}$).

Ejemplos de grupos no abelianos son:
1) $GL(n)$: $\{M, \cdot\}$, donde $M = \{A \in \mathbb{R}^{n\times n}, \mathrm{Det}[A] \neq 0\}$ es el conjunto de matrices reales de $n \times n$ de determinante no nulo y $\cdot$ la operación de multiplicación matricial usual (Grupo gral. lineal). En efecto,
i) la operación $\cdot$ es cerrada en el conjunto pues si $A, B \in M$, $A.B \in \mathbb{R}^{n\times n}$ y $\mathrm{Det}[A \cdot B] = \mathrm{Det}[A]\mathrm{Det}[B] \neq 0$
ii) el producto matricial es asociativo
iii) La identidad $I \in M$ ($\mathrm{Det}[I] = 1 \neq 0$)
iv) Si $A \in M \Rightarrow \mathrm{Det}\, A \neq 0$ y por lo tanto $\exists$ la matriz inversa A^{-1} t.q. $A \cdot A^{-1} = A^{-1} \cdot A = I$. Por su puesto, $A^{-1} \in M$ pues $\mathrm{Det}[A^{-1}] = 1/\mathrm{Det}[A] \neq 0$.
Ejercicio: Explicar, considerando la multiplicación matricial usual, por qué
1a) el conjunto de matrices reales de 2×2 de determinante 2 no es un grupo.
1b) el conjunto de todas las matrices reales de 2×2 sin la matriz nula $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, tampoco es grupo.
1c) el conjunto de matrices reales de 2×2 de determinante 1 si es un grupo.

2) $O(n)$: Grupo de matrices reales ortogonales de $n \times n$ (con la multiplicación matricial usual), donde ortogonal significa que $A^{-1} = A^t$ (matriz traspuesta), es decir, $AA^t = A^tA = I$. En efecto,
i) Si $A, B \in O(n)$, $(AB)^{-1} = B^{-1}A^{-1} = B^tA^t = (AB)^t$, por lo que $AB \in O(n)$
ii) El producto matricial es asociativo
iii) $I \in A$, ya que $I^{-1} = I = I^t$. Aquí I denota la matriz identidad, de elementos $I_{ij} = \delta_{ij} = \{^{1,\ i=j}_{0,\ i\neq j}$.
iv) si $A \in O(n) \Rightarrow A^{-1} = A^t \in O(n)$ pues $(A^{-1})^{-1} = (A^t)^{-1} = (A^{-1})^t$.

Notemos que si $AA^t = I \Rightarrow \mathrm{Det}[AA^t] = \mathrm{Det}[A]^2 = \mathrm{Det}[I] = 1$, por lo que $\mathrm{Det}[A] = \pm 1$.

3) $SO(n)$: grupo de matrices reales ortogonales de $n \times n$ de determinante 1.
Este subconjunto de $O(n)$ es también un grupo, pues
i) Si $A, B \in SO(n)$, $\mathrm{Det}[AB] = \mathrm{Det}[A]\mathrm{Det}[B] = 1$
iii) $I \in SO(n)$ pues $\mathrm{Det}[I] = 1$
iv) $A^{-1} \in SO(n)$ pues $\mathrm{Det}[A^{-1}] = 1/\mathrm{Det}[A] = 1$.
Veremos luego que $SO(n)$ es el grupo de rotaciones en $\mathbb{R}^n$ (rotaciones alrededor de ejes que pasen por el origen, es decir, que dejan el origen fijo).

4) $U(n)$: Grupo de matrices complejas unitarias de $n \times n$, con la operación de multiplicación matricial usual, donde unitario significa que $A^{-1} = A^\dagger$, siendo $A^\dagger \equiv (A^t)^*$ (matriz traspuesta conjugada).
Se deja como ejercicio mostrar que es grupo.
5) $SU(n)$: Grupo de matrices complejas unitarias de $n \times n$ de determinante 1.
Se deja también como ejercicio mostrar que es grupo.

Los grupos anteriores constan de un número infinito de elementos. Un grupo puede también constar de un número finito de elementos (grupo finito). Por ejemplo, $\{\{1, -1\}, .\}$ es un grupo con el producto usual, y también lo es $\{\{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right),\ \left(\begin{smallmatrix}-1&0\\0&-1\end{smallmatrix}\right)\}, \cdot\}$ con el producto matricial usual.

Dado que el conjunto de operaciones geométricas (rotaciones, reflexiones, etc.) que dejan invariante un cierto sistema físico forma un grupo con respecto a la operación de composición, los grupos juegan un rol fundamental en Física, especialmente en Mecánica Cuántica, caracterizando las simetrías y determinando sus consecuencias.

Un **anillo** $\{A, +, *\}$ es un conjunto A munido de dos operaciones binarias que satisface :
1) $\{A, +\}$ es un grupo abeliano
2) $a * (b * c) = (a * b) * c\ \forall\ a, b, c \in A$ (Asociatividad de $*$)
3) $a * (b + c) = a * b + a * c$, $\ (b + c) * a = b * a + c * a$ (Distributividad)
Si además $a * b = b * a\ \forall\ a, b \in A$ el anillo es conmutativo.
Ej.: $\{\mathbb{Z}, +, \cdot\}$ es anillo conmutativo.
El inverso respecto de $+$ se denomina opuesto, y la identidad respecto de $+$ elemento neutro o 0.

Un **cuerpo** o **campo** $\{F, +, *\}$ es un conjunto F munido de dos operaciones binarias $+, *$ que satisface:
1) $\{F, +\}$ es grupo abeliano
2) $\{F_0, *\}$ es grupo abeliano, donde F_0 es el conjunto de elementos de F distintos de 0 (elem. neutro)
3) $*$ es distributiva con respecto a $+$.
La identidad respecto de $*$ se denomina 1 (unidad). Un cuerpo es pues un anillo conmutativo con unidad donde $\forall\ a \in F_0\ \exists\ a^{-1} \in F_0$ tal que $a^{-1} * a = a * a^{-1} = 1$.

Ejemplos: $\{\mathbb{R}, +, \cdot\}$, $\{\mathbb{C}, +, \cdot\}$ son cuerpos. En cambio, $\{\mathbb{Z}, +, \cdot\}$ no es cuerpo pues la inversa de un entero no es necesariamente entero.
Los cuerpos pueden constar también de un número *finito* de elementos. El menor es $Z_2 = \{A = \{0, 1\}, +, \cdot\}$, donde $+$ y $\cdot$ denotan la suma y producto módulo 2 (el resto de dividir la suma y mult. ordinarias por 2): $0 + 0 = 0$, $0 + 1 = 1$, $1 + 1 = 0$, $0 \cdot 0 = 0 \cdot 1 = 0$, $1 \cdot 1 = 1$.
En general, $Z_p = \{(0, 1, \ldots, p - 1), +, \cdot\}$, con $+$ y $\cdot$ la suma y producto módulo p, es cuerpo para p primo. Esto puede demostrarse a partir del "pequeño teorema de Fermat": Si p es primo $\Rightarrow a^p = a \pmod p\ \forall\ a$ entero.

1. Espacios vectoriales

Partiendo del concepto intuitivo de vectores de $\mathbb{R}^2$ o $\mathbb{R}^3$, extenderemos el concepto de vector a elementos de un sistema algebraico abstracto, llamado *espacio vectorial* (o *lineal*), en el que se cumplen propiedades análogas a las de $\mathbb{R}^2$ o $\mathbb{R}^3$ con respecto a la suma de vectores y a la multiplicación de un vector por un número real. Remarquemos que estas dos operaciones son **cerradas** en $\mathbb{R}^2$ y en $\mathbb{R}^3$.

Definición: Sea $\{K, +, \cdot\}$ un cuerpo, y sea $\{V, \oplus\}$ un grupo abeliano. Un espacio vectorial V sobre el cuerpo K, denotado por $V(K)$, es una estructura algebraica $\{K, +, \cdot, V, \oplus, *\}$ donde $* : V \times K \to V$ denota una multiplicación de elementos de K por elementos de V que da como resultado un elemento de V y que satistface: $\forall\ \alpha, \beta \in K$ y $\forall\ v, w \in V$:
1) $\alpha * (v \oplus w) = (\alpha * v) \oplus (\alpha * w)$
2) $(\alpha + \beta) * v = (\alpha * v) \oplus (\beta * v)$
3) $(\alpha \cdot \beta) * v = \alpha * (\beta * v)$
4) $1 * v = v$
donde 1 denota la identidad del cuerpo K respecto del producto.
Los elementos de V se denominan **vectores** y los de K **escalares**.

En el caso de $\mathbb{R}^2$, $\{V, \oplus\}$ es el grupo $\{\mathbb{R}^2, +\}$, con $\oplus = +$ la suma usual de vectores, y $\{K, +, \cdot\}$ el cuerpo de los reales $\{\mathbb{R}, +, \cdot\}$ con la suma y producto usual. La operación $*$ es el producto de un vector por un número real.
La definición general extiende pues $\{\mathbb{R}^2, +\}$ a un grupo abeliano arbitrario $\{V, \oplus\}$ y $\{\mathbb{R}, +, \cdot\}$ a un cuerpo arbitrario $\{K, +, \cdot\}$. Si este es el cuerpo de los reales $\{\mathbb{R}, +, \cdot\}$, el espacio vectorial se dice *real*, y si es el cuerpo de los complejos $\{\mathbb{C}, +, \cdot\}$, el espacio vectorial se dice *complejo*.

En los sucesivo, para aligerar la notación seguiremos la costumbre universal de denotar la operación $\oplus$ (suma de vectores) también con $+$ y de omitir los símbolos $\cdot$ y $*$, quedando la multiplicación de escalares y de escalares por vectores automáticamente asumida. Las 4 condiciones anteriores se reescriben como:
1) $\alpha(v + w) = \alpha v + \alpha w$
2) $(\alpha + \beta)v = \alpha v + \beta v$
3) $(\alpha\beta)v = \alpha(\beta v)$
4) $1v = v$

De la definición de espacio vectorial se desprende que si $v, w \in V$ y $\alpha, \beta \in K$, la *combinación lineal*

$$u = \alpha v + \beta w$$

queda automáticamente definida y *pertenece también* a V, para todo par de elementos α y β del cuerpo y v, w de V. Esta, podemos afirmar, es la característica principal de un espacio vectorial. Es decir, es posbile multiplicar un vector por un escalar, lo cual es siempre otro vector de V, y también es posible sumar dos vectores cualesquiera, siendo la suma también un vector de V.
En general, si $v_i \in V$, $\alpha_i \in K$, $i = 1, \ldots, n$,

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

se denomina *combinación lineal* de los vectores $v_1, \ldots, v_n$, y es un vector $\in V$. Los escalares $\alpha_1, \ldots, \alpha_n$ se denominan los coeficientes de la combinación lineal.

Por ejemplo, el conjunto de todos los vectores del plano forma un espacio vectorial sobre $\mathbb{R}$ con la suma usual, pero el conjunto de los vectores del plano superior $\{(x, y) \in \mathbb{R}^2,\ y \geq 0\}$ NO es un espacio vectorial pues $-1 * (x, y) = (-x, -y)$ pertenece al plano inferior.

Debe destacarse además que el producto de vectores (escalar, vectorial u otro) no juega absolutamente ningún rol en la definición de espacio vectorial, y puede no estar definido en el mismo.

Demostremos ahora cuatro propiedades básicas válidas en cualquier espacio vectorial:
a) $0v = 0 \ \forall \ v \in V$
donde el primer 0 denota el 0 del cuerpo K (el elemento neutro respecto de la operación + para escalares) y el segundo el cero de V (la identidad respecto de la operación + para vectores).
En efecto, $0v = (0+0)v = 0v+0v$ por (2). Sumando el inverso $-(0v)$ ($-(0v)+(0v) = 0$) en ambos miembros obtenemos $0 = (0v) + 0 = 0v$, por lo que $0v = 0$.
b) $\alpha 0 = 0 \ \forall \ \alpha \in K$
donde 0 denota el cero de V. Tenemos $\alpha 0 = \alpha(0 + 0) = \alpha 0 + \alpha 0$, por 1). Sumando el inverso $-(\alpha 0)$ en ambos miembros se obtiene $0 = (\alpha 0) + 0 = \alpha 0$, por lo que $\alpha 0 = 0$.
c) $(-\alpha)v = -1(\alpha v) = -(\alpha v) = \alpha(-v)$
Tenemos, por 3), $(-\alpha)v = (-1\alpha)v = -1(\alpha v)$.
Además, $(-\alpha)v + \alpha v = (-\alpha + \alpha)v = 0v = 0$ por 2) y a), por lo que $(-\alpha)v = -(\alpha v)$ (opuesto de αv).
Finalmente, de b) y 1), $0 = \alpha 0 = \alpha(v + (-v)) = \alpha v + (\alpha(-v))$, por lo que $\alpha(-v)$ es también el opuesto de αv y por lo tanto coincide con $(-\alpha)v$ (unicidad del opuesto!).
d) Si $\alpha v = 0 \Rightarrow \alpha = 0$ o $v = 0$.
En efecto, por b), 3) y 4), si $\alpha \neq 0 \Rightarrow 0 = \alpha^{-1}0 = \alpha^{-1}(\alpha v) = (\alpha^{-1}\alpha)v = 1v = v$ por lo que $v = 0$. Por a), también se cumple si $\alpha = 0$.

Ejemplos de espacios vectoriales:
1) $K^n = \{(x_1, \ldots, x_n), x_i \in K, i = 1, \ldots, n\}$
Es el conjunto de n-uplas de elementos del cuerpo K. La suma de vectores se define como

$$(x_1, \ldots, x_n) + (x'_1, \ldots, x'_n) = (x_1 + x'_1, \ldots, x_n + x'_n)$$

y la multiplicación por un escalar $\alpha \in K$ como

$$\alpha(x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n)$$

El elemento neutro es $(0, \ldots, 0)$ y el opuesto de $(x_1, \ldots, x_n)$ es $-(x_1, \ldots, x_n) = (-x_1, \ldots, -x_n)$.
El caso $n = 1$ corresponde a tomar los elementos del cuerpo como vectores y escalares al mismo tiempo (es decir, como grupo abeliano $\{K, +\}$ y cuerpo $\{K, +, \cdot\}$).
El caso $K = \mathbb{R}$ con la suma y producto usual de números reales, se denomina espacio cartesiano. $\mathbb{R}^1$ consiste en tomar los reales como vectores y escalares al mismo tiempo, y corresponde geométricamente a una recta. $\mathbb{R}^2$ corresponde al plano y $\mathbb{R}^3$ al espacio cartesiano tridimensional.
Si $K = \mathbb{C}$, con + y $\cdot$ la suma y producto usual de números complejos, se obtiene el espacio $\mathbb{C}^n$, es decir, el conjunto de n-uplas complejas $\{(z_1, \ldots, z_n),\ z_i \in \mathbb{C}, i = 1, \ldots, n\}$, con escalares también complejos.
Se lo denota usualmente por $\mathbb{C}^n(\mathbb{C})$, para distinguirlo del espacio $\mathbb{C}^n(\mathbb{R})$, que es similar al anterior pero con los escalares restringidos a números reales, es decir, $\alpha_i \in \mathbb{R}$. Puede verse fácilmente que $\mathbb{C}^n(\mathbb{R})$ es también un espacio vectorial.
2) $\mathbb{K}^{m\times n}$. Es el conjunto de matrices A de $m \times n$ con elementos $A_{ij} = x_{ij}$ pertenecientes al cuerpo K, $i = 1, \ldots, m$, $j = 1, \ldots, n$, consideradas con la suma usual y el producto usual por un escalar. La suma se define por

$$\begin{pmatrix} x_{11} & \ldots & x_{1n} \\ & \ldots & \\ x_{m1} & \ldots & x_{mn} \end{pmatrix} + \begin{pmatrix} y_{11} & \ldots & y_{1n} \\ & \ldots & \\ y_{m1} & \ldots & y_{mn} \end{pmatrix} = \begin{pmatrix} x_{11}+y_{11} & \ldots & x_{1n}+y_{1n} \\ & \ldots & \\ x_{m1}+y_{m1} & \ldots & x_{mn}+y_{mn} \end{pmatrix}$$

y la multiplicación por un escalar como

$$\alpha \begin{pmatrix} x_{11} & \ldots & x_{1n} \\ & \ldots & \\ x_{m1} & \ldots & x_{mn} \end{pmatrix} = \begin{pmatrix} \alpha x_{11} & \ldots & \alpha x_{1n} \\ & \ldots & \\ \alpha x_{m1} & \ldots & \alpha_{mn} \end{pmatrix}$$

El elemento neutro es la matriz nula $0 = \begin{pmatrix} 0 & \ldots & 0 \\ & \ldots & \\ 0 & \ldots & 0 \end{pmatrix}$, y el opuesto de A es $-A = \begin{pmatrix} -x_{11} & \ldots & -x_{1n} \\ & \ldots & \\ -x_{m1} & \ldots & -x_{mn} \end{pmatrix}$.
Si $K = \mathbb{R}$ se obtiene el espacio de matrices reales de $m \times n$, y si $K = \mathbb{C}$, el de matrices complejas de $m \times n$. Este último puede considerarse con escalares complejos ($C^{m\times n}(\mathbb{C})$) o reales ($C^{m\times n}(\mathbb{R})$).

3) En general, si D es un conjunto no vacío, puede definirse el espacio vectorial

$$K^D = \{f \mid f \text{ es funcion de } D \text{ en } K\}$$

es decir, el conjunto de las funciones $f : D \to K$. La suma y producto de funciones se define como

$$(f+g)(x) = f(x) + g(x), \quad (\alpha f)(x) = \alpha f(x)$$

$\forall\, x \in D$, siendo el cero la función nula $0(x) = 0\ \forall\, x \in D$. Se verifican fácilmente que se satisfacen todas las condiciones de espacio vectorial. Por ejemplo,
$[\alpha(f+g)](x) = \alpha[(f+g)(x)] = \alpha(f(x)+g(x)) = \alpha f(x) + \alpha g(x) = (\alpha f + \alpha g)(x)$.
$[(\alpha+\beta)f](x) = (\alpha+\beta)f(x) = \alpha f(x) + \beta f(x) = (\alpha f + \beta f)(x)$
(verificación de las restantes a cargo del lector si lo considera necesario).

Así, si D es el conjunto $\{1, \ldots, n\}$, K^n es equivalente al espacio de n-uplas K^n anterior.

Si $K = D = \mathbb{R}$, se obtiene el espacio vectorial de funciones reales $\mathbb{R}^{\mathbb{R}} = \{f | f : \mathbb{R} \to \mathbb{R}\}$, y si $K = D = \mathbb{C}$, el espacio de funciones complejas $\mathbb{C}^{\mathbb{C}} = \{f | f : \mathbb{C} \to \mathbb{C}\}$.
Si $D = \mathbb{N}$, se obtiene el espacio vectorial de sucesiones de elementos de K, $\{x_n\} \equiv (x_1, x_2, \ldots)$.

2. Subespacios

Un subconjunto de vectores $S \subset V$ es un subespacio de V si es también un espacio vectorial.
Como consecuencia, un subconjunto de vectores $S \subset V$ *no vacío* es un subespacio si y sólo si S es *cerrado* bajo las operaciones de suma de vectores y multiplicación por escalar. Debe cumplirse entonces
0) $0 \in S$ (asegura que no sea vacío)
1) Si $v, w \in S \Rightarrow v + w \in S$
2) Si $v \in S$ y $\alpha \in K \Rightarrow \alpha v \in S$
Si se sabe que es no vacío bastan 1) y 2), pues si $\exists v \in S$, $0 = 0v \in S$ por 2).

Dem.: Es evidente que estas condiciones son necesarias. Para probar la suficiencia, podemos ver que por 1), la operación de suma es cerrada y asociativa en S, que por 0) o 2) $0 \in S$ y que $\forall\, v \in S\ \exists$ el elemento opuesto $-v = -1v \in S$ por 2), de modo que $\{S, +\}$ es grupo abeliano. Además, el producto de un vector de S por un escalar es siempre otro vector de S, por 2), por lo que la combinación lineal $\alpha v + \beta w$ pertence siempre a S. Las demás condiciones 1-4 se heredan de V, pues las operaciones son las mismas.
Cualquier espacio vectorial contiene siempre dos subespacios triviales: $S = V$ y $S = \{0\}$ (el vector nulo).

Ejemplos:
1) Si $V = \mathbb{R}^2$, $S = \{(x,y) \mid ax + by = 0,\ a, b \in \mathbb{R},\ a \neq 0 \text{ o } b \neq 0\}$ es siempre un subespacio de $\mathbb{R}^2$, que representa geométricamente una recta que pasa el origen. En efecto, si $(x,y), (x',y') \in S$, $(x,y) + (x',y') = (x+x', y+y') \in S$ pues $a(x+x') + b(y+y') = (ax+by) + (ax'+by') = 0 + 0 = 0$, y $\alpha(x,y) = (\alpha x, \alpha y) \in S$ pues $a\alpha x + b\alpha y = \alpha(ax+by) = 0$.
Geométricamente, los subespacios no triviales de $\mathbb{R}^2$ son pues rectas que pasan por el origen.
2) Si $V = \mathbb{R}^3$, se prueba en forma análoga que los subespacios no triviales son planos o rectas que pasan por el origen, es decir, $S = \{(x,y,z) \mid ax + by + cz = 0, \text{ con } (a,b,c) \text{ vector no nulo}\}$ (plano $\perp$ a (a,b,c)) o
$S = \{(x,y,z) \mid ax + by + cz = 0,\ dx + ey + fz = 0, \text{ con } (a,b,c),\ (d,e,f) \text{ vectores no nulos y no paralelos}\}$ (rectas).
Por ejemplo, $S = \{(x,y,z) \,|\, x+y+z = 0\}$ es un subespacio de $\mathbb{R}^3$ que representa geométricamente un plano ($\perp$ a $(1,1,1)$) que contiene al origen, $S = \{(x,y,z) \,|\, x+y+z = 0,\ x-y = 0\}$ es un subespacio que representa a una recta que pasa por el origen ($\|$ a $(1,1,-2)$) pero $C = \{(x,y,z) \,|\, x+y+z = 1\}$, $D = \{(x,y,z) \,|\, x \geq 0\}$, y $E = \{(x,y,z) \,|\, x^2 + y = 0\}$ NO son subespacios (probar!).
En general, si $V = \mathbb{K}^n$, $S = \{(x_1, \ldots, x_n) \,|\, a_{i1}x_1 + \ldots + a_{in}x_n = 0, i = 1, \ldots, m,\ a_{in} \in K\}$ es siempre un subespacio de V (que puede ser $\{0\}$ si la única solución al sistema es $x_i = 0$ para $i = 1, \ldots, n$, o V si todos los coeficientes a_{ij} son nulos). Corresponde en general a un hiperplano que pasa por el origen. Se prueba de la misma manera anterior (hecho en clase y se deja como ejercicio).

3) Si $V=\mathbb{R}^{n\times n}$, son subespacios:
El conjunto de matrices diagonales ($A_{ij}=0$ si $i\neq j$)
El conjunto de matrices simétricas ($A_{ij}=A_{ji}\ \forall\ i,j$)
El conjunto de matrices antisimétricas ($A_{ij}=-A_{ji}\ \forall\ i,j$)
El conjunto de matrices donde los coeficientes satisfacen un conjunto de ecuaciones lineales homogéneas $\sum_{i,j} a_{kij}A_{ij}=0$, $k=1,\ldots,p$, que incluye como casos particulares todos los anteriores.

4) Si $V=\mathbb{R}^{\mathbb{R}}$ (funciones reales de $\mathbb{R}$ en $\mathbb{R}$), el conjunto de los polinomios, es decir, de las funciones

$$f(x)=a_0+a_1x+\ldots+a_nx^n$$

con $a_i\in\mathbb{R}$, es un subespacio de V. Es claro que la suma es también un polinomio, que la función nula 0 es un polinomio (de grado 0) y que el producto de un polinomio por un escalar es un polinomio.
En cambio, el conjunto de los polinomios de grado fijo $n>0$ NO es un subespacio, ya que en particular 0 no pertenece al mismo (y la suma no es cerrada).
El conjunto de polinomios de grado $\leq n$ si es en cambio un subespacio.
También lo son, por ejemplo (probarlo como ejercicio):
i) el conjunto de funciones reales continuas
ii) el i) el conjunto de funciones reales derivables
iii) el de funciones que satisfacen $f(a)=0$ para un cierto $a\in\mathbb{R}$ (o en general, $\sum_{i=1}^m \alpha_i f(a_i)=0$)
iv) el conjunto de funciones de período L ($f(x+L)=f(x)\ \forall\ x\in\mathbb{R}$).
v) el conjunto de funciones pares ($f(x)=f(-x)$) y el de funciones impares ($f(x)=-f(-x)$).

Subespacio generado por un conjunto de vectores

Dado un conjunto $M=\{v_1,\ldots,v_m\}$ de vectores $\subset V$, el conjunto de combinaciones lineales

$$\overline{M}=\{\alpha_1v_1+\ldots+\alpha_mv_m,\ \ \alpha_i\in K,\ i=1,\ldots,m\}$$

es un subespacio de V denominado subespacio generado por M. Es fácil ver que es un subespacio, ya que
0) $0=0v_1+\ldots 0v_m\in\overline{M}$
1) $(\alpha_1v_1+\ldots+\alpha_mv_m)+(\alpha'_1v_1+\ldots+\alpha'_mv_m)=(\alpha_1+\alpha'_1)v_1+\ldots+(\alpha_m+\alpha'_m)v_m\in\overline{M}$
2) $\beta(\alpha_1v_1+\ldots+\alpha_mv_m)=(\beta\alpha_1)v_1+\ldots+(\beta\alpha_m)v_m\in\overline{M}$
donde, para $i=1,\ldots,m$, $\alpha_i,\alpha'_i,\beta\in K$.
Los vectores de M se denominan *generadores* de $\overline{M}$.
En general, para un conjunto aribtrario $M\subset V$, podemos definir $\overline{M}$ como el conjunto de todas las combinaciones lineales finitas de vectores de M. En particular, si S es un subespacio $\Rightarrow\overline{S}=S$, pues un subespacio debe contener todas las combinaciones lineales de sus vectores.
Como consecuencia, $\overline{M}$ es el *menor* subespacio que contiene a M: Si S es un subespacio y $M\subset S\Rightarrow\overline{M}\subset S$, ya que S debe contener a toda combinación lineal de sus elementos.
Ejemplo: En $V=\mathbb{R}^3$, si $M=\{(1,0,0),(1,1,0\}\}\Rightarrow\overline{M}=\{(x+y,y,0)|\,x,y\in\mathbb{R}\}$ es el plano determinado por los vectores de M.

Un espacio vectorial V se llama **finitamente generado** si existe un conjunto finito de vectores M tal que $\overline{M}=V$.
Por ejemplo, $\mathbb{R}^2$ puede ser generado por los vectores $(1,0)$ y $(0,1)$, ya que $(x,y)=x(1,0)+y(0,1)$, y también por los vectores $(1,1)$ y $(0,1)$, ya que $(x,y)=x(1,1)+(y-x)(0,1)$. También puede ser generado por los vectores $(1,0)$, $(0,1)$ y $(1,1)$, ya que $(x,y)=(x-z)(1,0)+(y-z)(0,1)+z(1,1)$, con z arbitrario.
El espacio $\mathbb{R}^{\mathbb{R}}$ de funciones reales $f:\mathbb{R}\to\mathbb{R}$ no puede ser en cambio generado por conjunto finito de vectores.

Intersección, Unión y Suma de subespacios

1) La **Intersección** $S=S_1\cap S_2$ de dos subespacios S_1, S_2, es un subespacio.
En efecto, $0\in S_1$ y $0\in S_2$, por lo que $0\in S_1\cap S_2$. Además, si $u,v,\in S$, $u,v\in S_1$ y $u,v,\in S_2$, por lo que $u+v\in S_1$ y $u+v\in S_2$, y por lo tanto $u+v\in S$. Análogamente, $\alpha u\in S_1$ y $\alpha u\in S_2$, de modo que $\alpha u\in S$.
Por ejemplo, en $V=\mathbb{R}^2$, si $S_1=\{(x,y),x+y=0\}$ y $S_2=\{(x,y),x-y=0\}\Rightarrow S_1\cap S_2=\{(0,0)\}$

(Geométricamente la intersección de dos rectas distintas que pasan por el origen es $(0,0)$).
Y en $V=\mathbb{R}^3$, si $S_1=\{(x,y,z)|x+y+z=0\}$ y $S_2=\{(x,y,z)|x-y-z=0\}\Rightarrow S_1\cap S_2=\{(0,y,z)|y+z=0\}$ (Geométricamente la intersección de dos planos distintos que pasan por el origen es una recta).

2) La **Unión** de dos subespacios No es en general un subespacio. Por ejemplo, en el caso pre-anterior, $S_1\cup S_2=\{(x,y)\,|\,y=\pm|x|\}$ no es un subespacio, ya que no es cerrado por la operación de suma de vectores (aunque sí lo es por el producto por escalares!).

3) La **Suma** de subespacios $S=S_1+S_2$, definida por

$$S_1+S_2=\{v=v_1+v_2,\ v_1\in S_1,\ v_2\in S_2\}$$

es un subespacio. En efecto, 1) $0=0+0\in S$, 2) si $v=v_1+v_2$ y $u=u_1+u_2$, con $u_i,v_i\in S_i\Rightarrow v+u=(v_1+v_2)+(u_1+u_2)=(v_1+u_1)+(v_2+u_2)\in S$, y 3) si $v=v_1+v_2$, $\alpha v=\alpha v_1+\alpha v_2\in S$.

Obviamente $S_1\cup S_2\subset S_1+S_2$ pues $v_1=v_1+0$, $v_2=0+v_2$.

En realidad, es fácil demostrar que $S_1+S_2=\overline{S_1\cup S_2}$ (probarlo!).
En el ejemplo anterior, $S_1+S_2=\{(x+x',y+y')\,|\,x+y=0,x'-y'=0\}=\{(x+x',x'-x),x,x'\in\mathbb{R}\}=\mathbb{R}^2$.

Suma directa: Si $S_1\cap S_2=\{0\}$, se dice que la suma S_1+S_2 es *directa* y se la escribe como $S_1\oplus S_2$.
Si $S_1\cap S_2=\{0\}$, todo vector $v\in S_1+S_2$ puede escribirse de manera **única** como suma de un vector de S_1 y un vector de S_2, y viceversa. En efecto, si

$$\begin{aligned} v &= v_1+v_2\\ &= v_1'+v_2' \end{aligned}$$

con $v_i,v_i'\in S_i\Rightarrow 0=(v_1-v_1')+(v_2-v_2')$, por lo que $(v_1-v_1')=-(v_2-v_2')$, lo que implica, como $v_2-v_2'\in S_2$, que $v_1-v_1'\in$ también a S_2 y por lo tanto a $S_1\cap S_2$. Si $S_1\cap S_2=\{0\}\Rightarrow v_1-v_1'=0=v_2-v_2'$, por lo que $v_1=v_1'$, $v_2=v_2'$.
Análogamente, si todo vector $v\in S_1+S_2$ puede escribirse de manera única como v_1+v_2 y $v\in S_1\cap S_2\Rightarrow v=v+0=0+v$, por lo que la única posibilidad es $v=0$.
Demotraremos luego que dado un subespacio $S_1\subset V$, siempre existe un subespacio $S_2\subset V$ tal que $V=S_1\oplus S_2$ (se demostrará luego de introducir bases).
Ejemplos:
1) $\mathbb{R}^2=S_1\oplus S_2$, donde $S_1=\{(x,0)|x\in\mathbb{R}\}$ y $S_2=\{(0,y),|y\in\mathbb{R}\}$. En efecto, $S_1\cap S_2=\{0\}$ y $\forall\ v\in\mathbb{R}^2$ se cumple $v=(x,y)=v_1+v_2$, donde $v_1=(x,0)\in S_1$, $v_2=(0,y)\in S_2$.
Notemos, sin embargo, que también $\mathbb{R}^2=S_1\oplus S_2'$, donde nuevamente $S_1=\{(x,0)|x\in\mathbb{R}\}$ pero $S_2'=\{(x,x)|x\in\mathbb{R}\}$. En efecto, $S_1\cap S_2'=\{0\}$ y $\forall\ v\in V$ se cumple $v=(x,y)=v_1+v_2'$, donde $v_1=(x-y,0)\in S_1$ y $v_2=(y,y)\in S_2'$.

2) $\mathbb{R}^{n\times n}=\mathbb{R}^{n\times n}_s\oplus\mathbb{R}^{n\times n}_a$, donde $\mathbb{R}^{n\times n}_s$, $\mathbb{R}^{n\times n}_a$ denotan los subespacios de matrices simétricas y antisimétricas respectivamente. En efecto, $\mathbb{R}^{n\times n}_s\cap\mathbb{R}^{n\times n}_a=\{0\}$ (pues si $A\in\mathbb{R}^{n\times n}_s$ y $A\in\mathbb{R}^{n\times n}_a\Rightarrow A_{ij}=A_{ji}=-A_{ji}\ \forall i,j$, por lo que $A_{ij}=0\ \forall\ i,j$). Además, toda matriz A puede escribirse como

$$A=A_s+A_a,\ \ A_s=\frac{1}{2}(A+A^t)\in\mathbb{R}^{n\times n}_s,\ \ \ A_a=\frac{1}{2}(A-A^t)\in\mathbb{R}^{n\times n}_a$$

donde A^t es la matriz traspuesta, de modo que $\mathbb{R}^{n\times n}_s\oplus\mathbb{R}^{n\times n}_a=\mathbb{R}^{n\times n}$.

3) $\mathbb{R}^{\mathbb{R}}=\mathbb{R}^{\mathbb{R}}_p\oplus\mathbb{R}^{\mathbb{R}}_i$, donde $\mathbb{R}^{\mathbb{R}}_p$, $\mathbb{R}^{\mathbb{R}}_i$ denotan los subespacios de funciones pares e impares. En efecto, si $f(x)=f(-x)=-f(-x)\ \forall\ x\Rightarrow f(x)=0\ \forall\ x$. Además, toda función puede escribirse como

$$f(x)=f_p(x)+f_i(x),\ \ f_p(x)=\frac{1}{2}(f(x)+f(-x))\in\mathbb{R}^{\mathbb{R}}_p,\ \ f_i(x)=\frac{1}{2}(f(x)-f(-x))\in\mathbb{R}^{\mathbb{R}}_i$$

Los desarrollos de Taylor alrededor del origen de f_p y f_i, si existen, contienen sólo potencias pares o impares respect. Por ejemplo, si $f(x)=e^x$, $f_p(x)=\cosh(x)$, $f_i(x)=\sinh(x)$.

3. Independencia lineal, bases y dimensión

Los vectores $v_1,\ldots,v_n\in V$ son **linealmente independientes** (LI) si y sólo si (sii) la ecuación

$$\alpha_1 v_1+\ldots+\alpha_n v_n=0$$

implica

$$\alpha_1=\alpha_2=\ldots=\alpha_n=0$$

De lo contrario, los vectores son **linealmente dependientes** (LD).
Para $n=1$, esta definición implica que v_1 es LI sii es un vector no nulo (Prop. básica d).
Si $n>1$, los vectores son LD sii al menos uno de ellos puede escribirse como combinación lineal de los restantes, es decir, si pertence al espacio generado por los restantes. En efecto, si son LD existe al menos un α_i, por ej., α_1, que es no nulo ($\alpha_1\neq 0$). En tal caso,

$$v_1=-(\alpha_2 v_2+\ldots+\alpha_n v_n)/\alpha_1$$

Análogamente, si $v_1=\alpha_2 v_2+\ldots+\alpha_n v_n\Rightarrow v_1-\alpha_2 v_2-\ldots-\alpha_n v_n=0$, siendo $\alpha_1=1\neq 0$, por lo que son LD.

Para $n=2$, esto implica que dos vectores no nulos son LI sii no son proporcionales (es decir sii $\nexists\ \alpha\in K$ t.q. $v_2=\alpha v_1$). En $V=\mathbb{R}^3$, tres vectores no nulos y no paralelos son LI sii ninguno de ellos pertenece al plano generado por los otros dos.
Si uno de los vectores es nulo, los vectores $v_1,\ldots,v_n$ son LD: Por ejemplo, si $v_1=0\Rightarrow\alpha_1 v_1+0v_2+\ldots+0v_n=0$ para $\alpha_1\neq 0$, lo que implica que son LD.
En general, si el conjunto $\{v_1,\ldots,v_n\}$ contiene un subconjunto de vectores LD entonces el conjunto total es LD (Probar).

Teorema. Sean $\{b_1,\ldots,b_n\}$ n vectores LI. Entonces los n vectores

$$v_j=\sum_{i=1}^{n}S_{ij}b_i=S_{1j}b_1+\ldots+S_{nj}b_n,\;\; j=1,\ldots,n,\;\; S_{ij}\in K \qquad (1)$$

son LI si y sólo si la matriz S de coeficientes S_{ij} (de $n\times n$) es *no singular* ($\mathrm{Det}[S]\neq 0$).
Dem.: Si

$$0=\sum_{j=1}^{n}\alpha_j v_j=\sum_{j=1}^{n}\alpha_j(\sum_{i=1}^{n}S_{ij}b_i)=\sum_{i=1}^{n}(\sum_{j=1}^{n}S_{ij}\alpha_j)b_i=\sum_{i=1}^{n}\beta_i b_i,\quad \beta_i=\sum_{j=1}^{n}S_{ij}\alpha_j$$

debe ser $\beta_i=0$ para $i=1,\ldots,n$ por ser los b_i LI. Es decir, $\sum_{j=1}^{n}S_{ij}\alpha_j=0,\ i=1,\ldots,n$, o en forma matricial,

$$\begin{pmatrix}S_{11}&\ldots&S_{1n}\\ &\ldots&\\ S_{n1}&\ldots&S_{nn}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \ldots\\ \alpha_n\end{pmatrix}=\begin{pmatrix}0\\ \ldots\\ 0\end{pmatrix}$$

Esto constituye un sistema de n ecuaciones lineales homogéneas con n incógnitas α_j, $j=1,\ldots,n$. Si los v_j son LI, la única solución de este sistema debe ser $\alpha_j=0\ \forall j$ y por lo tanto la matriz S debe ser *no singular*. Por otro lado, si S es no singular, la única solución del sistema es $\alpha_j=0$ para $j=1,\ldots,n$ y por lo tanto los vectores v_j son LI.

Corolario: Si S es no singular, el subespacio generado por $M=\{b_1,\ldots,b_n\}$ y $M'=\{v_1,\ldots,v_n\}$ **es el mismo** ($\overline{M}=\overline{M'}$).
Dem.: Si S es no singular, *existe la matriz inversa* S^{-1}, t.q. $SS^{-1}=S^{-1}S=I$, es decir, $\sum_{j=1}^{n}(S^{-1})_{kj}S_{ji}=\delta_{ki}$. Esto implica que **existe la transformación inversa de** (1), dada por

$$b_i=\sum_{j=1}^{n}(S^{-1})_{ji}v_j,\;\; i=1,\ldots,n \qquad (2)$$

ya que $\sum_{j=1}^{n}(S^{-1})_{ji}v_j=\sum_{j=1}^{n}(S^{-1})_{ji}(\sum_{k=1}^{n}S_{kj}b_k)=\sum_{k=1}^{n}(\sum_{j=1}^{n}S_{kj}(S^{-1})_{ji})b_k=\sum_{k=1}^{n}\delta_{ki}b_k=b_i$.
Por lo tanto, de (1) es obvio que $\overline{M'}\subset\overline{M}$ (pues $\sum_{j=1}^{n}\alpha_j v_j=\sum_{i=1}^{n}\beta_i b_i$, con $\beta_i=\sum_{j=1}^{n}S_{ij}\alpha_j$), y de (2) es obvio que $\overline{M}\subset\overline{M'}$ (pues $\sum_{i=1}^{n}\beta_i b_i=\sum_{j=1}^{n}\alpha_j v_j$, con $\alpha_j=\sum_{i=1}^{n}(S^{-1})_{ji}\beta_i$), por lo que $\overline{M}=\overline{M'}$.

Ejemplo: Si $\{e_1, e_2, e_3\}$ es un conj. LI en un cierto espacio, los vectores $v_1 = e_1$, $v_2 = e_1 + e_2$ y $v_3 = e_1 + e_2 + e_3$ son LI ya que $\mathrm{Det}[S] = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 1 \neq 0$. Como $S^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$, la transformación inversa está dada por $e_1 = v_1$, $e_2 = v_2 - v_1$, $e_3 = v_3 - v_2$, como es fácil comprobar. Los vectores $\{v_1, v_2, v_3\}$ generan pues el mismo subespacio que $\{e_1, e_2, e_3\}$.

Bases

Sea V un espacio vectorial, que supondremos distinto del subespacio trivial $S = \{0\}$. Un conjunto finito $B = \{b_1, \ldots, b_n\} \subset V$ es una base de V si los vectores de B
1) Son LI
2) Generan V $\quad(\overline{B} = V)$.

Existen muchas bases diferentes de un espacio vectorial.
Ejemplo 1: Si $V = \mathbb{R}^n$, el conjunto $B = \{e_1, e_2, \ldots, e_n\}$, con $e_i = (0, \ldots, 0, 1_{(i)}, 0, \ldots, 0)$, es una base, denominada *base canónica* de $\mathbb{R}^n$.
En efecto, son LI pues si $0 = \alpha_1 e_1 + \ldots + \alpha_n e_n = (\alpha_1, \ldots, \alpha_n) \Rightarrow \alpha_1 = \ldots = \alpha_n = 0$.
Y generan V pues $(x_1, \ldots, x_n) = x_1 e_1 + \ldots x_n e_n$.
Pero también es base el conjunto $\{(1, 0, \ldots, 0), (1, 1, 0, \ldots, 0), \ldots, (1, 1, \ldots, 1)\}$. (Probarlo!).
Ejemplo 2: Escribir la base canónica de $\mathbb{R}^{m \times n}$.
Ejemplo 3: Si V es el subespacio de polinomios de grado ≤ 2, una base es $\{e_1, e_2, e_3\}$, donde $e_1 = 1$, $e_2 = x$, $e_3 = x^2$, denominada también base canónica.
También es base $\{b_1, b_2, b_3\}$, con $b_1 = e_1$, $b_2 = e_1 + e_2$, $b_3 = e_1 + e_2 + e_3$, como el lector podrá fácilmente probar.

Si V es generado por un conjunto finito de vectores $M = \{v_1 \ldots, v_m\}$ y $V \neq \{0\} \Rightarrow$ existe una base $B = \{b_1, \ldots, b_n\}$ de V incluida en M.
Dem.: Sea $B = \{b_1, \ldots, b_n\}$ un subconjunto de M tal que los vectores de B sean LI y el número n de elementos de B sea máximo. Obviamente $n \geq 1$, pues $\overline{M} = V$ y $V \neq \{0\}$, por lo que existe al menos un vector no nulo en M. Si $v \in M \Rightarrow v \in \overline{B}$, pues los vectores $\{v, b_1, \ldots, b_n\}$ son necesariamente LD (pues son $n + 1$) y por lo tanto, existe una combinación

$$0 = \alpha v + \alpha_1 b_1 + \ldots + \alpha_n b_n$$

con coeficientes no todos nulos. Si $\alpha = 0 \Rightarrow 0 = \alpha_1 b_1 + \ldots + \alpha_n b_n$, pero en tal caso $\alpha_i = 0 \; \forall \, i$ por por ser los b_i LI. Por consiguiente, $\alpha \neq 0$ y $v = -(\alpha_1 b_1 + \ldots + \alpha_n b_n)/\alpha \in \overline{B}$. Por lo tanto, $M \subset \overline{B}$ y entonces $V = \overline{M} = \overline{B}$.

Del teorema de la sección anterior se desprenden ahora las sig. propiedades fundamentales.
Si $B = \{b_1, \ldots, b_n\}$ es una base de V, entonces:

1) Cualquier conjunto de n vectores LI $\{v_1, \ldots, v_n\} \subset V$ es también una base de B.
Dem.: Como B es base, los n vectores v_j pueden ser escritos en la forma (1), con S no singular dado que son LI. Pero en tal caso el espacio generado es el mismo, por lo que forman también una base de V.
En particular, los n vectores $v_j = \sum_{i=1}^{n} S_{ij} b_i$, $j = 1, \ldots, n$, forman una base de V sii S es no singular.

2) Todo conjunto de $n + 1$ vectores $\{v_1, \ldots, v_{n+1}\} \subset V$ es LD.
Dem.: Supongamos que son LI. $\Rightarrow$ los primeros n vectores son LI. Pero en tal caso forman una base, por el corolario anterior, por lo que v_{n+1} pertenece al espacio generado por ellos y el conjunto es entonces LD.
Todo conjunto con $m > n$ vectores es por lo tanto también LD. Y un conjunto con $m < n$ vectores no podría ser base, pues en tal caso B no sería base.

Como consecuencia, todas las bases de un espacio V tienen **el mismo número de elementos**, n. A ese número se lo denomina **dimensión** del espacio V: $n = \mathrm{dim} V$. Representa el máximo número de vectores LI. Un espacio en el que $\exists$ un N^o arbitrariam. grande de vectores LI se dice que tiene dimensión infinita.
Ejemplo: La dimensión de $\mathbb{R}^n$ es n, y la de $\mathbb{R}^{m \times n}$, $m \cdot n$. La de $\mathbb{R}^{\mathbb{R}}$ es ∞.
La dimensión de $\mathbb{C}^n(\mathbb{C})$ es también n (una base es $\{e_1, \ldots, e_n\}$, con $e_j = (0, \ldots, 1_{(j)}, \ldots, 0)$, $j = 1, \ldots, n$), mientras que la dimensión de $\mathbb{C}^n(\mathbb{R})$ es $2n$ (una base es $\{e_1, \ldots, e_n, \tilde{e}_1, \ldots \tilde{e}_n\}$, con $\tilde{e}_j = (0, \ldots, i_{(j)}, \ldots, 0)$).

4. Coordenadas de un vector en una base y cambio de base

Si $B = \{b_1, \ldots, b_n\}$ es una base de V, todo vector $v \in V$ puede escribirse en forma **única** como combinación lineal de elementos de B. Dem.: Si $v \in V$ y

$$\begin{aligned} v &= \alpha_1 b_1 + \ldots + \alpha_n b_n \\ &= \alpha'_1 b_1 + \ldots + \alpha'_n b_n \end{aligned}$$

entonces

$$0 = (\alpha_1 - \alpha'_1) b_1 + \ldots + (\alpha_n - \alpha'_n) b_n$$

por lo que $\alpha_i = \alpha'_i$ para $i = 1, \ldots, n$ por ser los vectores LI.
Análogamente, si todo vector de V puede escribirse en forma única como comb. lineal de los b_i, estos son LI pues en particular, la única forma de escribir el vector nulo será $0 = 0b_1 + \ldots 0b_n$.
Los coeficientes $\alpha_1, \ldots, \alpha_n$ que determinan el vector v son pues únicos y reciben el nombre de **coordenadas** del vector v en la base dada.

Cambo de base

Consideremos en lo sucesivo *bases ordenadas* $B = (b_1, \ldots, b_n)$, con el objeto de asignar un orden determinado a las componentes de un vector. Si B es una base de V, todo $v \in B$ puede representarse como

$$v = \sum_{i=1}^{n} \alpha_i b_i, \quad \alpha_i \in K$$

Consideremos ahora otra base $B' = (b'_1, \ldots, b'_n)$ de V. Por ser B base podemos también escribir

$$b'_j = \sum_{i=1}^{n} S_{ij} b_i, \quad j = 1, \ldots, n$$

donde los elementos S_{ij}, $i = 1, \ldots, n$ (columna j de S) son las componentes de b'_j en la base B. La matriz

$$S = \begin{pmatrix} S_{11} & \ldots & S_{1n} \\ \ldots & \ldots & \ldots \\ S_{n1} & \ldots & S_{nn} \end{pmatrix} = \begin{pmatrix} . & . & . \\ [b'_1]_B & \ldots & [b'_n]_B \\ . & . & . \end{pmatrix}$$

se denomina *matriz de cambio de base* y debe ser *no singular* ($\mathrm{Det}[S] \neq 0$), por lo demostrado anteriormente. Podemos ahora escribir v en la base B' como

$$v = \sum_{j=1}^{n} \alpha'_j b'_j$$

donde α'_j son las componentes de v en la base B'. Escribiendo b'_j en términos de los b_i, obtenemos

$$v = \sum_{j=1}^{n} \alpha'_j (\sum_{i=1}^{n} S_{ij} b_i) = \sum_{i=1}^{n} \alpha_i b_i, \quad \alpha_i = \sum_{j=1}^{n} S_{ij} \alpha'_j, \ i = 1, \ldots, n$$

En forma matricial, esto puede escribirse como

$$\begin{pmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} S_{11} & \ldots & S_{1n} \\ \ldots & \ldots & \ldots \\ S_{n1} & \ldots & S_{nn} \end{pmatrix} \begin{pmatrix} \alpha'_1 \\ \ldots \\ \alpha'_n \end{pmatrix}$$

o, en forma concisa,

$$[v]_B = S\,[v]_{B'}$$

donde

$$[v]_B \equiv \begin{pmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{pmatrix}, \ [v]_{B'} = \begin{pmatrix} \alpha'_1 \\ \ldots \\ \alpha'_n \end{pmatrix}$$

denotan las matrices columna de componentes de v en las bases B y B' respectivamente. Podemos entonces determinar $[v]_{B'}$ a partir de $[v]_B$ como

$$[v]_{B'} = S^{-1}[v]_B$$

donde S^{-1} es la matriz inversa de S. Remarquemos que la forma de construir S es notando que su columna i es la matriz columna de componentes de b'_i en la base B, es decir, $[b'_i]_B$. Notemos también que la columna i de S^{-1} es la matriz de componentes de b_i en la base B' ($[b_i]_{B'}$).
Fialmente, notemos que si $v_1, v_2 \in V$ y $\alpha \in K$, se tiene obviamente

$$[v_1 + v_2]_B = [v_1]_B + [v_2]_B, \;\; [\alpha v]_B = \alpha[v]_B$$

Ejemplo 1: Sea $B = (e_1, e_2)$, con $e_1 = (1, 0)$, $e_2 = (0, 1)$ la base canónica en R^2. Consideremos ahora la nueva base $B' = (e'_1, e'_2)$, donde

$$e'_1 = (1, 0), \;\; e'_2 = (1, 1)$$

o sea, $e'_1 = e_1$, $e'_2 = e_1 + e_2$. En este caso,

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

Por lo tanto, si $v = (x, y) = xe_1 + ye_2$, podemos escribir también $v = x'e'_1 + y'e'_2$ con

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

Se verifica que $x'e'_1 + y'e'_2 = (x - y)e_1 + y(e_1 + e_2) = x_1e_1 + ye_2$. Notemos además que las columnas de S^{-1} son las coordenadas de la base canónica en la nueva base: $e_1 = e'_1$, $e_2 = -e'_1 + e'_2$.
Ej. sugerido: Hallar las coordendadas de $v = (x, y)$ en la base formada por $e'_1 = (1, 0)$, $e'_2 = (1, \varepsilon)$, con $\varepsilon \neq 0$, y analizar el límite $\varepsilon \to 0$.
Ejemplo 2: Rotación en el plano. Sean nuevamente $e_1 = (1, 0)$, $e_2 = (0, 1)$ los vectores de la base canónica en R^2 y sean $e'_1 = \cos(\theta)e_1 + \sin(\theta)e_2$, $e'_2 = -\sin(\theta)e_1 + \cos(\theta)e_2$. Estos vectores son los vectores e_1, e_2 rotados un ángulo θ en sentido antihorario respecto del eje x (recordar dibujo hecho en clase). Tenemos

$$S = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \;\; S^{-1} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix},$$

(o sea, $S^{-1}(\theta) = S(-\theta) = S(\theta)^t$). Por lo tanto, las componentes x', y' en la base rotada de un vector $v = (x, y) = xe_1 + ye_2$ son

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos(\theta) + y\sin(\theta) \\ -x\sin(\theta) + y\cos(\theta) \end{pmatrix}$$

de forma que $v = x'e'_1 + y'e'_2$. (Verificar que $x'e'_1 + y'e'_2 = xe_1 + ye_2$!).

Ejemplo 3: Ecuación de una elipse rotada un ángulo θ (antihorario) respecto del eje x. Respecto del sistema rotado tenemos la ecuación

$$\frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 1$$

con a, b los semiejes de la elipse. Reemplazando $x' = x\cos(\theta) + y\sin(\theta)$, $y' = -x\sin(\theta) + y\cos(\theta)$, obtenemos

$$x^2(\frac{\cos^2\theta}{a^2} + \frac{\sin^2\theta}{b^2}) + y^2(\frac{\sin^2\theta}{a^2} + \frac{\cos^2\theta}{b^2}) + xy\sin(2\theta)(\frac{1}{a^2} - \frac{1}{b^2}) = 1$$

Si $a = b$ (circunferencia) la forma de la ecuación permanece invariante.

Ejemplo 4: Producto escalar usual en R^2 expresado en base arbitraria: El producto escalar usual en la base canónica puede expresarse como

$$v_1 \cdot v_2 = x_1x_2 + y_1y_2 = [v_1]_e^t \cdot [v_2]_e$$

donde x_i, y_i, $i = 1, 2$ son las componentes de v_1, v_2 en la base canónica ($v_i = (x_i, y_i)$) y t denota traspuesto. Reemplazando $[v_i]_e = S[v_i]_{e'}$, obtenemos, para una base arbitraria e',

$$v_1 \cdot v_2 = (S[v_1]_{e'})^t(S[v_2]_{e'}) = [v_1]_{e'}^t(S^tS)[v_2]_{e'}$$

El producto escalar queda entonces determinado por la matriz *simétrica* S^tS y tendrá en general términos "cruzados" $\propto x_1y_2$ y x_2y_1 además de "diagonales" proporcionales a x_1x_2 y y_1y_2. En el caso de rotaciones, $S^t = S^{-1}$ y por lo tanto la forma del producto escalar usual permanece *invariante.*

Algebra Lineal: Aplicaciones a la Física, Curso 2012

5. Transformaciones lineales

Una transformación lineal (TL) es una función $F : V \to V'$ entre dos espacios vectoriales V, V' *sobre el mismo cuerpo* K que satisface

$$i)\;\; F(v_1 + v_2) = F(v_1) + F(v_2) \;\; \forall\, v_1, v_2 \in V$$

$$ii)\;\; F(\alpha v) = \alpha F(v) \;\; \forall \alpha \in K,\; v \in V$$

es decir, $F(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 F(v_1) + \alpha_2 F(v_2)\;\forall\; \alpha_1, \alpha_2 \in K$ y $v_1, v_2 \in V$. F es pues un morfismo entre espacios vectoriales. Si $V' = V$, F es un endomorfismo y en tal caso se denomina también *operador lineal.*

Propiedades fundamentales:

I) $F(0) = 0$ (el primer 0 es el vector nulo de V y el segundo el vector nulo de V')
Dem.: $F(0) = F(0 + 0) = F(0) + F(0) \Rightarrow F(0) = 0$
II) $F(-v) = -F(v)$
Dem.: $0 = F(0) = F(v + (-v)) = F(v) + F(-v)$, de donde $F(-v) = F(v)$.
(También pueden demostrarse utilizando ii))

Ejemplos (probar como ej.):
1) $F : V \to V$ definida por $F(v) = \alpha v$, con $\alpha \in K$, es TL $\forall \alpha \in K$ (incluso $\alpha = 0$).
Si $\alpha = 1 \Rightarrow F$ es la *identidad* ($F(v) = v\; \forall v \in V$), denotada por I.
Si $\alpha = 0 \Rightarrow F$ es la TL *nula* ($F(v) = 0\; \forall\; v \in V$).
2) $F : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (x + y, 2y + x)$ es TL.
También lo es $F : \mathbb{R}^2 \to \mathbb{R}^3$ dada por $F(x, y) = (x + y, 2y + x, 3x + y)$.
3) $F : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (x + y^2, 2y + x)$ no es TL. Tampoco lo es $F(x, y) = (1 + x + y, 2y)$.
4) $F : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (ax + by, cx + dy)$ es TL $\forall\; a, b, c, d \in \mathbb{R}$.
5) $F : \mathbb{R}^n \to \mathbb{R}^m$ dada por $F(x_1, \ldots, x_n) = (a_{11}x_1 + \ldots + a_{1n}x_n, \ldots, a_{m1}x_1 + \ldots + a_{mn}x_n)$ es TL $\forall\; a_{ij} \in \mathbb{R}$.
6) $F : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ dada por $F(A) = A^t$ (traspuesta) es TL.
También lo es $F(A) = B \cdot A$, con B matriz real fija de $n \times n$, y $G(A) = B \cdot A + A \cdot C$.
7) $F : C \to C$ ($C(\mathbb{R})$ es el subespacio de funciones reales continuas) dada por $[F(f)](x) = \int_0^x f(t)dt$, es TL.
8) Si $D \subset \mathbb{R}^{\mathbb{R}}$ es el subespacio de funciones reales derivables, $F : D \to \mathbb{R}^{\mathbb{R}}$ dada por $F(f) = f'$, es también TL.

Más propiedades fundamentales:
III) Si S es un subespacio de V, la imagen de S por f, $F(S) = \{F(v)|v \in S\}$, es un subespacio de V'.
En particular, si $S = V$, el subespacio $F(V)$ se denomina *imagen* de F y se denota $I(F)$.
Dem.: Si $v_1, v_2 \in S \Rightarrow F(v_1) + F(v_2) = F(v_1 + v_2) \in F(S)$ pues $v_1 + v_2 \in S$.
Si $\alpha \in K$ y $v \in S \Rightarrow \alpha F(v) = F(\alpha v) \in F(S)$ pues $\alpha v \in S$.
En particular, $0 \in F(S)$ ($F(0) = 0$)

IV) La pre-imagen de un subespacio S' de V', $S = \{v \in V|F(v) \in S'\}$ es un subespacio de V.
Dem.: $0 \in S$ pues $F(0) = 0 \in S'$.
Si $v_1, v_2 \in S \Rightarrow v_1 + v_2 \in S$ pues $F(v_1 + v_2) = F(v_1) + F(v_2) \in S'$
Si $\alpha \in K$ y $v \in S \Rightarrow \alpha v \in S$ pues $F(\alpha v) = \alpha F(v) \in S'$.
En particular, la pre-imagen de $\{0\}$ (subespacio nulo de V') se denomina *núcleo* o *espacio nulo* de F y se denota $N(F)$: $N(F) = \{v \in V|F(v) = 0\}$.

V) Si $v = \sum_{i=1}^m \alpha_i v_i$, con $\alpha_i \in K$ y $v_i \in V \Rightarrow F(v) = F(\sum_{i=1}^m \alpha_i v_i) = \sum_{i=1}^m \alpha_i F(v_i)$
Esto puede demostrarse fácilmente por inducción (para los que no lo ven obvio).
Esta propiedad implica que la imagen del subespacio $\overline{C}$ generado por un subconjunto de vectores $C = \{v_1, \ldots, v_m\} \subset V$ es el subespacio $\overline{F(C)}$ generado por la imagen $F(C) = \{F(v_1), \ldots, F(v_m)\} \subset V'$:

$$F(\overline{C}) = \overline{F(C)}$$

En particular, si V es finitamente generado y $B = \{B_1, \ldots, B_n\}$ es una base de V, $I(F) = F(\overline{B}) = \overline{F(B)}$.
En otras palabras, la imagen $I(F)$ es el subespacio generado por los n vectores $F(b_1), \ldots, F(b_n)$ de V'.
Esto implica que una transformación lineal *queda completamente determinada por los vectores que asigna a los elementos de una base*, es decir, por los n vectores $F(b_1), \ldots, F(b_n)$:

Si $v \in V \Rightarrow v = \sum_{i=1}^n \alpha_i e_i$ y $F(v) = \sum_{i=1}^n \alpha_i F(e_i)$.
Nótese también que si $v_1, \ldots, v_m$ son L.D. (linealmente *dependientes*) $\Rightarrow$ los vectores $F(v_1), \ldots, F(v_m)$ son también L.D.: Si $0 = \sum_{i=1}^m \alpha_i v_i$, con algún $\alpha_i \neq 0 \Rightarrow 0 = F(\sum_{i=1}^m \alpha_i v_i) = \sum_{i=1}^m \alpha_i F(v_i)$.

VI) Si V es finitamente generado entonces

$$\dim N(F) + \dim I(F) = \dim V$$

Dem.: Sea $\{b_1, \ldots, b_m, b_{m+1}, \ldots, b_n\}$ base de V tal que $\{b_1, \ldots, b_m\}$ sea base de $N(F)$ ($F(b_i) = 0$ si $i \leq m$). Si $v = \sum_{i=1}^n \alpha_i b_i \in V \Rightarrow$

$$F(v) = \sum_{i=1}^n \alpha_i F(b_i) = \sum_{i=m+1}^n \alpha_i F(b_i)$$

pertenece al espacio generado por $F(b_{m+1}) \ldots, F(b_n)$. Además, $F(b_{m+1}), \ldots, F(b_n)$ son L.I. pues si

$$0 = \sum_{i=m+1}^n \alpha_i F(b_i) = F(\sum_{i=m+1}^n \alpha_i b_i)$$

el vector $\sum_{i=m+1}^n \alpha_i b_i \in N(F)$ y por tanto, $\sum_{i=m+1}^n \alpha_i b_i = \sum_{i=1}^m \alpha_i b_i$. Pero por independencia lineal de los b_i, debe ser $\alpha_i = 0$ para $i = 1, \ldots, n$, por lo que $F(b_{m+1}), \ldots, F(b_n)$ son L.I.
La dimensión de la imagen es por lo tanto $n-m$, y se cumple entonces $\dim N(F) + \dim I(F) = m + (n-m) = n = \dim V$. La dimensión de la imagen $I(F)$ se denomina *rango* de F y la dimensión del espacio nulo $N(F)$ *nulidad* de F.

Ejemplos simples:
1) $F: V \to V$ dada por $F(v) = \alpha v$
Si $\alpha = 0$, $N(F) = V$, $I(F) = \{0\}$. Si dim $V = n$, dim $N(F)$+dim $I(F) = n + 0 = n$.
Si $\alpha \neq 0$, $N(F) = \{0\}$, $I(F) = V$. Si dim $V = n$, dim $N(F)$+dim $I(F) = 0 + n = n$.
2) $F: \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (x, 0)$. $N(F) = \{(0, y) | y \in \Re\}$, $I(F) = \{(x, 0) | x \in \Re\}$.
dim $N(Ff)$+dim $I(F)$=1+1=2=dim V.
3) $F: \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ dada por $F(A) = A^t$. $N(F) = \{0 \equiv \binom{0\,0}{0\,0}$, $I(F) = \mathbb{R}^2$; dim $N(F)$+dim$I(F)$=0+4=4

5.1 Representación matricial de funciones lineales:

Sea $F: V \to V'$, con V, V' finitamente generados de dimensión n y m respect. Sea $B = (b_1, \ldots, b_n)$ una base ordenada de V y $B' = (b'_1, \ldots, b'_m)$ una base ordenada de V'. Podemos escribir

$$F(b_i) = \sum_{j=1}^m T_{ji} b'_j, \quad i = 1, \ldots, n$$

donde $T_{ji} \in K$ son las coordenadas de $F(b_i)$ en la base B' de V'. Por lo tanto, si $v \in V$, $v = \sum_{i=1}^n \alpha_i b_i$ y

$$F(v) = \sum_{i=1}^n \alpha_i F(b_i) = \sum_{i=1}^n \alpha_i \sum_{j=1}^m T_{ji} b'_j = \sum_{j=1}^m (\sum_{i=1}^n T_{ji} \alpha_i) b'_j = \sum_{j=1}^m \alpha'_j b'_j$$

con $\alpha'_j = \sum_{i=1}^n T_{ji} \alpha_i$, $j = 1, \ldots, m$. Es decir, en forma matricial,

$$\begin{pmatrix} \alpha'_1 \\ \ldots \\ \alpha'_m \end{pmatrix} = \begin{pmatrix} T_{11} & \ldots & T_{1n} \\ \ldots & \ldots & \ldots \\ T_{m1} & \ldots & T_{mn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{pmatrix},$$

que se escribe en forma concisa como

$$[F(v)]_{B'} = [F]_{B'}^B [v]_B$$

donde

$$[F(v)]_{B'} = \begin{pmatrix} \alpha'_1 \\ \ldots \\ \alpha'_m \end{pmatrix}, \quad [v]_B = \begin{pmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{pmatrix},$$

son las coordenadas de $F(v)$ en la base B' y de v en la base B, y

$$[F]_{B'}^{B} = \begin{pmatrix} T_{11} & \dots & T_{1n} \\ \dots & \dots & \dots \\ T_{m1} & \dots & T_{mn} \end{pmatrix} = \begin{pmatrix} . & . & . \\ [F(b_1)]_{B'} & \dots & [F(b_n)]_{B'} \\ . & . & . \end{pmatrix}$$

es la matriz de $m \times n$ que representa la transformación lineal B respecto de las bases B y B' de V y V'. Esta matriz depende de las bases elegidas, pero una vez elegida las bases es claramente *única*, ya que la función lineal queda completamente determinada por los vectores que asigna a los elementos de una base. En particular, si $V = V'$ y $B = B'$,

$$[F(v)]_B = [F]_B^B [v]_B$$

Notemos que la función identidad $I : V \to V$ definida por $I(v) = v$ queda representada por la matriz identidad I_n: $[I]_B^B = I_n$. Por simplicidad, denotaremos a $[F]_B^B$ también como $[F]_B$ cuando quede claro que estamos trabajando con operadores lineales representados en una misma base.

Ejemplo 1: Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (2x + y, 4y + 3x)$. En la base canónica $B = (b_1, b_2)$, $b_1 = (1, 0)$, $b_2 = (0, 1)$, tenemos $F(b_1) = (2, 3) = 2b_1 + 3b_2$, $F(b_2) = (1, 4) = b_1 + 4b_2$, y la matriz que representa a F en esta base es

$$[F]_B^B = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$$

Ejemplo 2: Reflexión respecto del eje x en $\mathbb{R}^2$. Si $F(v)$ es el vector obtenido al reflejar v respecto del eje x, tenemos (recordar dibujo) $F(b_1) = b_1$, $F(b_2) = -b_2$ y por lo tanto

$$[F]_B^B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Ejemplo 3: Reflexión respecto de la recta de ec. $y = x$ en $\mathbb{R}^2$: Si $F(v)$ es el vector obtenido al reflejar v respecto de la recta $y = x$, tenemos $F(b_1) = b_2$, $F(b_2) = b_1$ y por lo tanto

$$[F]_B^B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Ejemplo 4: Rotación de ángulo θ en $\mathbb{R}^2$. Si $F(v)$ es el vector obtenido al rotar v un ángulo θ antihorario, tenemos (recordar dibujo) $F(b_1) = \cos(\theta)b_1 + \sin(\theta)b_2$, $F(b_2) = -\sin(\theta)b_1 + \cos(\theta)b_2$ y por lo tanto

$$[F]_B^B = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

Ejemplo 5: Rotación de ángulo θ en $\mathbb{R}^3$ alrededor del eje z. Si $F(v)$ es el vector obtenido al rotar v un ángulo θ antihorario alrededor del eje z, tenemos, en la base canónica de $\mathbb{R}^3$, $B = ((1, 0, 0), (0, 1, 0), (0, 0, 1))$, $F(b_1) = \cos(\theta)b_1 + \sin(\theta)b_2$, $F(b_2) = -\sin(\theta)b_1 + \cos(\theta)b_2$ y $F(b_3) = b_3$. Por lo tanto

$$[F]_B^B = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Ejemplo 6: Sea P_n el espacio de polinomios de grado $\leq n$, y sea $D : P_2 \to P_1$ el operador derivación restringido a polinomios de grado ≤ 2 con codominio P_1. Sea $(b_1 = 1, b_2 = t, b_3 = t^2)$ la base "canónica" de P_2 y $(b'_1 = 1, b'_2 = t)$ la de P_1. Tenemos $D(b_1) = 0$, $D(b_2) = b'_1$, $D(b_3) = 2b'_2$ y por lo tanto

$$[D]_{B'}^{B} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Notemos que estas representaciones implican $F(x, y) = (x, -y)$ en (2), $F(x, y) = (y, x)$ en (3), $F(x, y) = (x\cos(\theta) + y\sin(\theta), -x\sin(\theta) + y\cos(\theta))$ en (4), $F(x, y, z) = (x\cos(\theta) + y\sin(\theta), -x\sin(\theta) + y\cos(\theta), z)$ en (5) y $D(x.1 + yt + zt^2) = y.1 + 2zt$ en (6).

5.2 Cambio de base

Consideremos primero el caso de endomorfismos $F : V \to V$, y sean $B = (b_1, \ldots, b_n)$, $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_n)$ dos bases ordenadas de V. Tenemos $[v]_B = S[v]_{\tilde{B}}$, $[F(v)]_B = S[F(v)]_{\tilde{B}}$, siendo S la matriz de cambio de base (su columna i es el vector de coordenadas $[\tilde{b}_i]_B$). Por lo tanto, $\forall\ v \in V$,

$$[F(v)]_{\tilde{B}} = S^{-1}[F(v)]_B = S^{-1}([F]_B^B[v]_B) = S^{-1}([F]_B^B S[v]_{\tilde{B}}) = (S^{-1}[F]_B^B S)[v]_{\tilde{B}}$$

es decir, comparando con $[F(v)]_{\tilde{B}} = [F]_{\tilde{B}}^{\tilde{B}}[v]_{\tilde{B}}$,

$$[F(v)]_{\tilde{B}}^{\tilde{B}} = S^{-1}[F]_B^B S$$

que implica también $[F]_B^B = S[f]_{\tilde{B}}^{\tilde{B}}S^{-1}$.
La matrices que representan a un endomorfismo F en diferentes bases son entonces *semejantes* (A de $n \times n$ es semejante a B de $n \times n$ si $A = S^{-1}BS$, con S de $n \times n$ no singular). Las matrices semejantes poseen **el mismo detereminante** y **la misma traza**:

$$\mathrm{Det}[A] = \mathrm{Det}[S^{-1}BS] = \mathrm{Det}[S^{-1}]\mathrm{Det}[B]\mathrm{Det}[S] = \mathrm{Det}[B]$$

$$\mathrm{Tr}[A] = \mathrm{Tr}[S^{-1}BS] = \mathrm{Tr}[BSS^{-1}] = \mathrm{Tr}[B]$$

Recordemos que la traza se define como **la suma de todos los elementos diagonales de una matriz**:

$$\mathrm{Tr}[A] = \sum_{i=1}^{n} A_{ii}$$

y satisface $\mathrm{Tr}[AB] = \mathrm{Tr}[BA]\ \forall\ A, B$ de $n \times n$: $\sum_{i,j} A_{ij}B_{ji} = \sum_{i,j} B_{ij}A_{ji}$.

Las matrices que representan a un mismo operador lineal en distintas bases tienen pues *el mismo determinante y la misma traza*. Estas cantidades **permanecen invariantes** frente a cambios de fase y constituyen pues propiedades del operador, no dependientes de la representación.

La matriz S de cambio de base puede reescribirse en este contexto como la matriz que representa la identidad I ($I : V \to V$ definida por $I(v) = v\ \forall V \in V$) entre dos bases diferentes:

$$[v]_{\tilde{B}} = [I(v)]_{\tilde{B}} = [I]_{\tilde{B}}^B[v]_B$$

Comparando con $[v]_{\tilde{B}} = S^{-1}[v]_B$, tenemos pues

$$S^{-1} = [I]_{\tilde{B}}^B, \ \ S = [I]_B^{\tilde{B}}$$

Por lo tanto, en el caso de endomorfismos podemos escribir

$$[F]_{\tilde{B}}^{\tilde{B}} = [I]_{\tilde{B}}^B[F]_B^B[I]_B^{\tilde{B}}$$

Nótese también que la transformación lineal $G : V \to V$ definida por $G[b_i] = \tilde{b}_i$, $i = 1, \ldots, n$, puede representarse en la base B por la matriz

$$[G]_B^B = [I]_{\tilde{B}}^B = S$$

mientras que $[G]_{\tilde{B}}^B$ es obviamente la matriz identidad I_n.

Ejemplo 1: La matriz que representa a una reflexión F respecto de la recta de ec. $y = x$ en $\mathbb{R}^2$, obtenida anteriormente, se relaciona con aquella que representa a la reflexión respecto del eje x mediante un cambio de base, y son por lo tanto semejantes. Si $B = (b_1, b_2)$ es la base canónica, respecto de la base $\tilde{b}_1 = (b_1 + b_2)/\sqrt{2}$, $\tilde{b}_2 = (-b_1 + b_2)/\sqrt{2}$ (vectores unitarios paralelos a las rectas de ec. $y = x$ y $y = -x$) tenemos $F(\tilde{b}_1) = \tilde{b}_1$, $F(\tilde{b}_2) = -\tilde{b}_2$. Por lo tanto,

$$[F]_{\tilde{B}}^{\tilde{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

La base $\tilde{B}$ se relaciona con la base canónica mediante la matriz

$$S = [I]_B^{\tilde{B}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

con $S^{-1} = S^t = [I]_{\tilde{B}}^{B}$. Por lo tanto,

$$[F]_B^B = S[F]_{\tilde{B}}^{\tilde{B}} S^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

que es el resultado obtenido anteriormente. Nótese que la matriz no es diagonal en la base canónica, pero si lo es en la base $\tilde{B}$.

Ejemplo 2: Construir la matriz que representa a una reflexión F respecto de una recta inclinada un ángulo θ (antihorario) respecto del eje x, en R^2. Respecto de la base formada por $\tilde{b}_1 = \cos(\theta)b_1 + \sin(\theta)b_2$, $\tilde{b}_2 = -\sin(\theta)b_1 + \cos(\theta)b_2$, tenemos nuevamente y por definición de reflexión,

$$[F]_{\tilde{B}}^{\tilde{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

La base $\tilde{B}$ se relaciona con la base canónica B mediante la matriz de rotación

$$S = [I]_B^{\tilde{B}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

con $S^{-1} = S^t = [I]_{\tilde{B}}^{B}$. Por lo tanto,

$$[F]_B^B = S[f]_{\tilde{B}}^{\tilde{B}} S^{-1} = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}$$

Nótese que existe una base ($\tilde{B}$) donde la transformación queda representada por una simple matriz diagonal.

Caso general: Llamemos $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_n)$ una nueva base de V y $\tilde{B}' = (\tilde{b}'_1, \ldots, \tilde{b}'_m)$ una nueva base de V', definidas por matrices de cambio de base S y S' respectivamente ($S = [I]_B^{\tilde{B}}$, $S' = [I]_{B'}^{\tilde{B}'}$). Dado que $[v]_B = S[v]_{\tilde{B}}$ y $[F(v)]_{B'} = S'[F(v)]_{\tilde{B}'}$, tenemos

$$[F(v)]_{\tilde{B}'} = S'^{-1}[F(v)]_{B'} = S'^{-1}[F]_{B'}^{B}(S[v]_{\tilde{B}}) = (S'^{-1}[F]_{B'}^{B}S)[v]_{\tilde{B}}$$

y por lo tanto, $[F(v)]_{\tilde{B}'} = [F]_{\tilde{B}'}^{\tilde{B}}[v]_{\tilde{B}}$, con

$$[F]_{\tilde{B}'}^{\tilde{B}} = S'^{-1}[F]_{B'}^{B}S = [I]_{\tilde{B}'}^{B'}[F]_{B'}^{B}[I]_B^{\tilde{B}}$$

Nótese que $[F]_{\tilde{B}'}^{\tilde{B}}$ y $[F]_{B'}^{B}$ son de $m \times n$, S' es de $m \times m$ y S de $n \times n$.

Ejemplo: Sea $D : P_2 \to P_1$ la función derivación restringida a polinomios de grado ≤ 2 con codominio P_1 y sea $B = (1, t, t^2)$ la base canónica de P_2, $B' = (1, t)$ la base canónica de P_1, $\tilde{B} = (1, 1+t, 1+t+t^2/2)$ y $\tilde{B}' = (1, 1+t)$. Tenemos

$$[D]_{B'}^{B} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, S = [I]_B^{\tilde{B}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1/2 \end{pmatrix}, S' = [I]_{B'}^{\tilde{B}'} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

con $S'^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. Por lo tanto,

$$[D]_{\tilde{B}'}^{\tilde{B}} = S'^{-1}[D]_{B'}^{B}S = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

lo cual es también obvio a partir de la definición de D.

5.3 Composición (Producto) de operadores lineales

Sea $F: V \to V'$ y $G: V' \to V''$ dos transformaciones lineales. La composición o producto $(GF): V \to V''$ se define por

$$(GF)(v) = (G \circ F)(v) = G(F(v))$$

El producto de transformaciones lineales es una transformacíon lineal:

$$(GF)(v_1 + v_2) = G(F(v_1 + v_2)) = G(F(v_1) + F(v_2)) = G(F(v_1)) + G(F(v_2)) = (GF)(v_1) + (GF)(v_2)$$

$$(GF)(\alpha v) = G(F(\alpha v)) = G(\alpha F(v)) = \alpha G(F(v)) = \alpha (GF)(v)$$

Para espacios finitamente generados, la matriz $[GF]^B_{B''}$ que representa a GF en las bases B, B'' de V y V'', es el *producto* de las matrices $[G]^{B'}_{B''}$ y $[F]^B_{B'}$ que representan a F y G en bases B', B'' y B, B', siendo B' una base de V':

$$[GF]^B_{B''} = [G]^{B'}_{B''}[F]^B_{B'}$$

Notemos que si las dimensiones de V, V', V'' son n, m, p respect. $\Rightarrow$ $[GF]^B_{B''}$ es de $p \times n$, $[G]^{B'}_{B''}$ es de $p \times m$ y $[F]^B_{B'}$ es de $m \times n$.

Dem.:

$$[(GF)(v)]_{B''} = [G(F(v))]_{B''} = [G]^{B'}_{B''}[F(v)]_{B'} = [G]^{B'}_{B''}([F]^B_{B'}[v]_B) = ([G]^{B'}_{B''}[F]^B_{B'})[v]_B$$

En particular, si $V = V' = V''$, con $B = B' = B''$,

$$[(GF)]^B_B = [G]^B_B[F]^B_B$$

El producto de funciones así definido es obviamente asociativo ($H(GF) = (HG)G$ para $Gf : V \to V'$, $G: V' \to V''$ y $H: V'' \to V'''$), pero en general, *no conmutativo*, aún cuando $V = V' = V''$.

Ejemplo: Consideremos la composición en $\mathbb{R}^2$ de una rotación F de $\pi/2$ antihoraria seguida de una reflexión G respecto del eje x: Tenemos, en la base canónica $B = ((1,0),(0,1))$:

$$[GF]^B_B = [G]^B_B[F]^B_B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = -[H]^B_B$$

con H la reflexión respecto de la recta de ec. $y = x$ ($-H$ es la reflexión respecto de la recta de ec. $y = -x$) Por otro lado, la composición en sentido inverso, es decir una reflexión respecto del eje x seguida de una rotación de $\pi/2$, da como resultado

$$[FG]^B_B = [F]^B_B[G]^B_B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = [H]^B_B$$

Este sencillo ejemplo muestra que el producto de operadores lineales no es en general conmutativo.

En general, se define el *conmutador* de dos operadores lineales $F: V \to V$, $G: V \to V$ como

$$[F, G] = FG - GF$$

La matriz que representa el conmutador es el conmutador de las matrices que representan F y G:

$$[[F, G]]^B_B = [F]^B_B[F]^B_B - [G]^B_B[F]^B_B$$

En el ejemplo anterior, $[F, G] = 2H$ ya que $[[F, G]]^B_B = 2[H]^B_B$.

5.4 El espacio vectorial de transformaciones lineales

Consideremos dos operadores lineales $F: V \to V'$, $G: V \to V'$. La suma $(F + G): V \to V'$ se define como

$$(F + G)(v) = F(v) + G(v)$$

y es claramente una función lineal:

$$(F+G)(v_1+v_2) = F(v_1+v_2)+G(v_1+v_2) = F(v_1)+F(v_2)+G(v_1)+Gg(v_2) = (F+G)(v_1)+(F+G)(v_2)$$

$$(F+G)(\alpha v) = F(\alpha v) + G(\alpha v) = \alpha F(v) + \alpha G(v) = \alpha(F(v)+G(v)) = \alpha(F+G)(v)$$

Es fácil verificar que la suma es conmutativa ($F{+}G = G{+}F$) y asociativa ($(F{+}G){+}H = F{+}(G{+}H)$). Existe además un elementro neutro 0, que es la función nula definida por $0(v) = 0\ \forall\ v \in V$, con $F{+}0 = 0{+}F = F$. El elemento opuesto de F es entonces $-F$, definido por $-F(v) = -(F(v))$, que es también lineal. El conjunto de las funciones lineales $\{F : V \to V',\ F \text{ lineal}\}$ es pues un grupo abeliano con la operación de suma.

El producto por un escalar $\alpha \in K$, $(\alpha F) : V \to V'$ se define obviamente como

$$(\alpha F)(v) = \alpha F(v)$$

y es también una función lineal:

$$(\alpha F)(v_1+v_2) = \alpha F(v_1+v_2) = \alpha(F(v_1)+F(v_2)) = \alpha F(v_1) + \alpha F(v_2) = (\alpha F)(v_1) + (\alpha F)(v_2)$$

$$(\alpha F)(\beta v) = \alpha F(\beta v) = \alpha(\beta F(v)) = (\alpha\beta)F(v) = (\beta\alpha)F(v) = \beta(\alpha F)(v)$$

Es fácil verificar además que $\alpha(\beta F) = (\alpha\beta)F$, $(\alpha+\beta)F = \alpha F + \beta F$, $\alpha(F+G) = \alpha F + \alpha G$, $1F = F$.

El conjunto de todas las transformaciones lineales $F : V \to V'$ es entonces *un espacio vectorial* sobre K, denominado $Hom(V, V')$ (Homomorfismos de V en V').

Notemos también que con respecto al producto (composición) de funciones, la suma verifica las propiedades distributivas $(G+H)F = GF + HF$ para $F : V \to V'$, y $G, H : V' \to V''$ y $H(F+G) = HF + HG$ para $H : V' \to V''$ y $F, G : V \to V'$. Además, por ser lineales, $\alpha(GF) = (\alpha G)F = G(\alpha F)$ para $\alpha \in K$.

La matriz que representa a $(F+G)$ en las bases B, B' es claramente la suma de matrices,

$$[F+G]^B_{B'} = [F]^B_{B'} + [G]^B_{B'}$$

y la que representa a (αF) es obviamente

$$[\alpha F]^B_{B'} = \alpha[F]^B B'$$

En efecto, $[(F{+}G)(v)]_{B'} = [F(v){+}g(v)]_{B'} = [F(v)]_{B'}{+}[G(v)]_{B'} = [F]^B_{B'}[v]_B{+}[G]^B_{B'}[v]_B = ([F]^B_{B'}{+}[G]^B_{B'})[v]_B$. $[(\alpha F)(v)]_{B'} = [\alpha F(v)]_{B'} = \alpha[F(v)]_{B'} = \alpha([F]^B_{B'}[v]_B) = (\alpha[F]^B_{B'})[v]_B$.

6. Monomorfismos, Epimorfismos e Isomorfismos

I) Un monomorfismo es una TL $F : V \to V'$ inyectiva (o sea, $F(v_1) \neq F(v_2)$ si $v_1 \neq v_2$).
F es un monomorfismo *si y sólo si* $N(F) = \{0\}$.
Dem.: Si F es un monomorfismo y $v \neq 0$, $F(v) \neq F(0) = 0 \Rightarrow N(F) = \{0\}$.
Si $N(F) = \{0\} \Rightarrow F(v_1) - F(v_2) = F(v_1 - v_2) \neq 0$ si $v_1 - v_2 \neq 0$, o sea, $F(v_1) \neq F(v_2)$ si $v_1 \neq v_2$.

Como consecuencia, dim $N(F) = 0$. Por lo tanto, si V es de dimensión finita, dim $I(F) =$ dim V.
Y como $I(F) \subset V'$, F puede ser un monomorfismo sólo si dim $V \leq$ dim V'.
Los monomorfismos **conservan la independencia lineal**: Si $\{v_1, \ldots, v_m\}$ son vectores LI de V y F es un monomorfismo $\Rightarrow$ $\{F(v_1), \ldots, F(v_m)\}$ son vectores LI de V'. Dem.: Si

$$0 = \sum_{i=1}^{m} \alpha_i F(v_i) = F(\sum_{i=1}^{m} \alpha_i v_i)$$

entonces $\sum_{i=1}^{m} \alpha_i v_i \in N(F)$. Como $N(F) = \{0\} \Rightarrow \sum_{i=1}^{m} \alpha_i v_i = 0$, lo que implica $\alpha_i = 0$ para $i = 1, \ldots, m$ por ser los v_i LI. Por lo tanto, $\{F(v_1), \ldots, F(v_m)\}$ son LI
En particular, si $B = (b_1, \ldots, b_n)$ es una base de un espacio V finitamente generado y $F : V \to V'$ es un monomorfismo, $(F(b_1), \ldots, F(b_n))$ es una **base** de $I(F)$.

II) Un epimorfismo es una TL $F : V \Rightarrow V'$ suryectiva ($I(F) = V'$). Si V' es de dimensión finita $\Rightarrow$ F es un epimorfismo si y sólo si dim $I(F) =$ dim V'.
Como dim $I(F) \leq$ dim V, F puede ser un epimorfismo sólo si dim $V' \leq$ dim V.

III) Un isomorfismo es una TL biyectiva, es decir, es a la vez un monomorfismo y un epimorfismo.
En espacios de dimensión finita, un isomorfismo puede entonces existir sólo si dim V=dim V' (pues por ser monomorfismo, dim $V \leq$ dim V' y por ser epimorfismo, dim $V' \leq$ dim V.)
Si $B = (b_1, \ldots, b_n)$ es una base de V y $F : V \to V'$ es un isomorfismo, $\{F(b_1), \ldots, F(b_n)\}$ es una base de V', pues son LI (por ser F monomorfismo) y generan V' (por ser F epimorfismo). Los isomorfismos transforman entonces bases en bases.
Si $V = V'$, un operador que es un isomorfismo se dice *no singular* o automorfismo. En caso contrario se dice *singular*.

Dados dos espacios V, V' sobre el mismo cuerpo K, se dice que V es *isomorfo* a V' si existe un isomorfismo $F : V \to V'$. En tal caso existe pues una correspondencia uno a uno entre los elementos de V y V'.
Dos espacios vectoriales V, V' de dimensión finita sobre el mismo cuerpo K son isomorfos *si y sólo si* dim $V =$ dim V'.
Dem.: Ya hemos demostrado que si $\exists$ un isomorfismo entre V y V' $\Rightarrow$ dim V=dim V'.
Por otro lado, si dim $V =$ dim $V' = n$ y $B = \{b_1, \ldots, b_n\}$, $B' = \{b'_1, \ldots, b'_n\}$ son bases de V y V', la TL definida por

$$F(b_i) = b'_i, \quad i = 1, \ldots, n$$

(o sea $F(B) = B'$) es un isomorfismo. En efecto, es suryectiva, pues $I(F) = F(\overline{B}) = \overline{F(B)} = \overline{B'} = V'$
(o sea, si $v' \in V' \Rightarrow v' = \sum_{i=1}^{n} \alpha_i b'_i = \sum_{i=1}^{n} \alpha_i F(b_i) = F(\sum_{i=1}^{n} \alpha_i b_i) \in I(F)$), por lo que $I(F) = V'$).
Esto implica que F es también un monomorfismo pues dim $N(F) =$ dim V-dim $I(F)$ =dim V-dim $V' = 0$, lo que implica $N(F) = \{0\}$.

Notemos finalmente que si dim $V =$ dim $V' = n$ y $F : V \to V'$ es una TL, son equivalentes:
a) F es un isomorfismo; b) F es monomorfismo; c) F es epimorfismo
como consecuencia de la relación $\dim N(F) + \dim I(F) = \dim V$. En efecto, a) implica b)+c) (y viceversa) por definición, b) implica c) ya que en tal caso $N(F) = \{0\}$ y por lo tanto dim $I(F) = n$, y c) implica b) por la misma propiedad (pues si dim $I(F) = n \Rightarrow$ dim $N(F) = 0$ y por lo tanto $N(F) = \{0\}$).

No obstante, si V es de dimensión infinita, un operador lineal $F : V \to V'$ puede ser monomorfismo sin ser epimorfismo y viceversa. Pro ejemplo, si P es el espacio de polinomios reales, $D : P \to P$ definido por $D(p(x)) = p'(x)$ (derivada) no es monomorfismo ($D(1) = 0$) pero es epimorfismo (probar como ejercicio). Y el operador $S : P \to P$ definido por $S(p(x)) = \int_0^x p(t)dt$ es monomorfismo (pues $N(S) = \{0\}$) pero no es epimorfismo (probar como ej.).

IV) Si $F: V \to V'$ es un isomorfismo $\Rightarrow$ la transformación inversa $F^{-1}: V' \to V$, definida por $F^{-1}(v') = v$, con v el único vector $\in V$ tal que $F(v) = v'$, es lineal y es un isomorfismo.
Dem.: Si F es isomorfismo, la inversa F^{-1} es obviamente una función bien definida.
Si $F(v_1) = v'_1$, $F(v_2) = v'_2 \Rightarrow F(v_1 + v_2) = F(v_1) + F(v_2) = v'_1 + v'_2$, lo que implica $F^{-1}(v'_1 + v'_2) = v_1 + v_2 = F^{-1}(v'_1) + Ff^{-1}(v'_2)$.
Si $F(v) = v'$ y $\alpha \in K \Rightarrow F(\alpha v) = \alpha F(v) = \alpha v'$, lo que implica $F^{-1}(\alpha v') = \alpha v = \alpha F^{-1}(v')$.
F^{-1} es por lo tanto una TL.
Además, F^{-1} es un monomorfismo, pues $N(F^{-1}) = \{0\}$ y es un epimorfismo pues si $v \in V$, $v = F^{-1}(v')$, con $v' = F(v)$, por lo que $I(F^{-1}) = V$.
Como consecuencia de la definición, $F(F^{-1}(v')) = v' \ \forall\ v' \in V'$ y $F^{-1}(F(v)) = v \ \forall\ v \in V$. Por lo tanto,

$$FF^{-1} = I_{V'}, \quad F^{-1}F = I_V$$

donde $I_{V'}: V' \to V'$ denota la identidad en V' y $I_V: V \to V$ aquella en V. Para V y V' de dimensión finita, las matrices correspondientes satisfacen

$$[FF^{-1}]_{B'} = [F]^{B}_{B'}[F^{-1}]^{B'}_{B} = I_n, \qquad [F^{-1}F]_B = [F^{-1}]^{B'}_{B}[F]^{B}_{B'} = I_n$$

donde $I_n = [I_{V'}]_{B'} = [I_V]_B$ es la matriz identidad de $n \times n$. Por lo tanto, $[F^{-1}]^{B'}_{B}$ *es la matriz inversa* de $[F]^{B}_{B'}$:

$$[F^{-1}]^{B'}_{B} = ([F]^{B}_{B'})^{-1}$$

Una TL $F: V \to V'$ entre espacios de dimensión finita es pues un isomorfismo *si y sólo si* está representada por matrices $[F]^{B}_{B'}$ cuadradas *no singulares* ($\mathrm{Det}[F]^{B}_{B'} \neq 0$).
Dem.: Si es isomorfismo, por lo visto anteriormente $[F]^{B}_{B'}$ es cuadrada e invertible y por lo tanto no singular. Y si $[F]^{B}_{B'}$ es cuadrada no singular, la única solución de $[F]^{B}_{B'}[v]_B = 0$ es $[v]_B = 0$, es decir, $v = 0$. Esto implica que $N(F) = \{0\}$ y $\Rightarrow F$ es monomorfismo y por lo tanto isomorfismo.

Si $V = V'$, F es un operador no singular sii $[F]_B$ es una matriz no singular. En tal caso $[F^{-1}]_B = [F]^{-1}_B$.

Recordemos que la inversa de una matriz no singular A ($\mathrm{Det}[A] \neq 0$) puede obtenerse como

$$A^{-1} = C/\mathrm{Det}[A], \text{ con } C_{ij} = (-1)^{i+j}\mathrm{Det}[M_{ji}]$$

donde Det denota el determinante y C la matriz de cofactores traspuesta, siendo M_{ji} la matriz de $(n-1) \times (n-1)$ obtenida al suprimir la fila j y columna i de A.
Por ejemplo, si A es de 2×2,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Rightarrow \quad A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Ejemplo 1): Si V es un espacio vectorial sobre el cuerpo K con $\dim V = n$ y $B = (b_1, \ldots, b_n)$ es una base ordenada de V, la función $R: V \to K^n$ dada por

$$R(v) = [v]_B = \begin{pmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{pmatrix}$$

donde $v = \sum_{i=1}^{n} \alpha_i b_i$, es un *isomorfismo*.
En efecto, hemos demostrado anteriormente que $R(v)$ es lineal ($[\alpha v]_B = \alpha[v]_B$ y $[v_1 + v_2]_B = [v_1]_B + [v_2]_B$). Además $N(R) = \{0\}$ (pues $[v]_B = 0$ (vector columna nulo) si y sólo si $v = 0$). Por lo tanto es un isomorfismo. Esto implica en particular que un conjunto de vectores $v_1, \ldots, v_m \in V$ serán LI si y sólo si los correspondientes vectores columna $[v_1]_B, \ldots, [v_m]_B \in K^n$ son LI.

Ejemplo 2): Si $\dim V = n$, $\dim V' = m$ y B, B' son bases ordenadas de V y V', la función $H: Hom(V, V') \to K^{m \times n}$ dada por $H(F) = [F]^{B}_{B'}$ es un isomorfismo.
En efecto, H es lineal (ya que $[(F+G)]^{B}_{B'} = [F]^{B}_{B'} + [G]^{B}_{B'}$ y $[\alpha F]^{B}_{B'} = \alpha[F]^{B}_{B'}$). Además, H es inyectiva, pues $N(H) = \{0\}$ (0 denota la función nula, representada en cualquier base por la matriz nula de $m \times n$ (sus elementos son todos 0)) y es también suryectiva, pues una matriz arbitraria $T \in K^{m \times n}$ corresponde a la transformación lineal definida por $F(b_i) = \sum_{j=1}^{m} T_{ji} b'_j$, $i = 1, \ldots, n$. Es decir, $I(H) = K^{m \times n}$. Esto implica que $\dim Hom(V, V') = \dim K^{m \times n} = m.n$.

La dimensión de la imagen de F puede calcularse evaluando el **rango** de la matriz $T=[F]^B_{B'}$ de $m\times n$ (es decir, el número de columnas (o filas) LI) en *cualquier* par de bases B, B', ya que esto será equivalente al número de vectores $F(b_i)$ LI. Del mismo modo, la imagen $I(F)$ puede obtenerse a partir del espacio columna de T (es decir, el espacio generado por las columnas de T) y el núcleo $N(F)$ a partir del espacio nulo de T (este último es el conjunto de vectores $[v]_B\in K^{n\times 1}$ que satisfacen $T[v]_B=0$, y que son por tanto ortogonales a todas las filas de T).
Como base de $K^{m\times n}$ pueden elegirse las matrices E^{ij} cuyo único elemento no nulo es el ij, definidas por $E^{ij}_{kl}=1$ si $k=i$ y $l=j$ y $E^{ij}_{kl}=0$ en caso contrario. Como base de $Hom(V,V')$ pueden elegirse las correspondientes transformaciones lineales F^{ij} definidas por $F^{ij}(e_l)=0$ si $l\neq j$ y $F^{ij}(e_l)=e'_i$ si $l=j$, tal que $[F^{ij}]^e_{e'}=E^{ij}$. Aquí e y e' denotan las bases canónicas de K^n y K^m respectivamente.

6.1 Rango, espacio columna y espacio fila de una matriz.
Recordemos que el **espacio columna** (e.c.) de una matriz T de $m\times n$ es el subespacio $S_C\subset K^{m\times 1}$ generado por las n columnas de T, y el **espacio fila** (e.f.) de T el subespacio $S_F\subset K^{1\times n}$ generado por las m filas de T. Estos subespacios son en general diferentes, pero tienen siempre **la misma dimensión** (véase demostración abajo). El **rango** de una matriz T es la dimensión del e.f. o e.c. de la misma.
Recordemos también que el **espacio nulo** de T es el subespacio $S_N\subset K^{n\times 1}$ formado por las soluciones $x\in K^{n\times 1}$ de la ecuación homogénea $Tx=0$. Su dimensión se denomina nulidad de T.

Consideremos ahora las relaciones entre F y la matriz $T=[F]^B_{B'}$ de $m\times n$ que representa a F en bases $B=(b_1,\ldots,b_n)$, $B'=(b'_1,\ldots,b'_m)$ de V, V' La columna i de T es la matriz columna de componentes $[F(b_i)]_{B'}$, con $[F(v)]_{B'}=T[v]_B$.
a) $\dim I(F)=\text{rango}(T)$.
En efecto, $I(F)=\overline{\{F(b_1),\ldots,F(b_n)\}}$, por lo que su dimensión es el número de vectores $F(b_i)$ LI. Pero esto coincide con el número de columnas $[F(b_i)]_{B'}$ LI de T (véase Ej. 1), que es la dimensión del e.c. de T, es decir su rango.
b) Como $\dim I(F)$ es independiente de las bases B y B' $\Rightarrow$ rango(T)=rango(T') si $T'=S'^{-1}TS$, con S' de $m\times m$ y S de $n\times n$ no singulares, ya que T' corresponde a la representación de F en otro par de bases ($T'=[F]^{\tilde{B}}_{\tilde{B}'}$, con $S=[I]^{\tilde{B}}_B$, $S'=[I]^{\tilde{B}'}_{B'}$).
c) $\dim N(F)=$ nulidad (T). En efecto, $N(F)$ es el subespacio formado por los vectores $v\in V$ que satisfacen $F(v)=0$, y por lo tanto, $[F(v)]_{B'}=T[v]_B=0$, de modo que $[v]_B$ pertence al espacio nulo de T. Y si x pertenece al espacio nulo de T ($Tx=0$) entonces $F(v)=0$, con $[v]_B=x$. Ambos subespacios son pues isomorfos y tienen la misma dimensión. Esto también implica que nulidad(T) =nulidad(T') si $T'=S'^{-1}TS$ con S' y S no singulares.
La relación $\dim N(F)+\dim I(F)=\dim V=n$ implica entonces nulidad (T)+rango $(T)=n$.

d) La dimensión del e.f. y del e.c. de una matriz arbitraria T de $m\times n$ coinciden.
Dem.: Podemos considerar a T como la representación de una TL $F:V\rightarrow V'$ en bases B y B' de V y V', con $\dim V=n$, $\dim V'=m$, tal que $T=[F]^B_{B'}$. En nuevas bases definidas por $\tilde{B}=(\tilde{b}_1,\ldots,\tilde{b}_k,\tilde{b}_{k+1},\ldots,\tilde{b}_n)$ tal que $(\tilde{b}_{k+1},\ldots,\tilde{b}_n)$ sea base de $N(F)$, y $\tilde{B}'=(\tilde{b}'_1,\ldots,\tilde{b}'_k,\tilde{b}'_{k+1},\ldots,\tilde{b}'_m)$ tal que $\tilde{b}'_i=F(b_i)$ para $i\leq k$ (recordar que estos son LI !) se tendrá $[F(\tilde{b}_i)]_{\tilde{B}'}=0$ si $i>k$ y $([F(\tilde{b}_i)]_{\tilde{B}'})_j=\delta_{ij}$ si $i\leq k$, por lo que la matriz $T'=[F]^{\tilde{B}}_{\tilde{B}'}=S'^{-1}TS$, con $S=[I]^{\tilde{B}}_B$ y $S'=[I]^{\tilde{B}'}_{B'}$ no singulares, contendrá una submatriz identidad de $k\times k$ en las primeras k filas, siendo las restantes nulas. La dim. del e.c. y del e.f. de esta matriz es por lo tanto k.
Entonces k es la dimensión de la imagen de F, por lo que la dim. del e.c. de $T=S'T'S^{-1}$ será k.
A su vez, el e.f. de T es el e.c. de T^t. Como la dim. del e.c. de T'^t es también k, la dim. del e.c. de $T^t=S^{-1\,t}T'^tS'^t$ es entonces k, por ser S'^t, $S^{-1\,t}$ no singulares. Pero esta es la dim. del e.f. de T.

Una demostración alternativa que permite hallar una base del e.f. y del e.c. de T es la siguiente: Aplicando un número finito de operaciones elementales por fila, se puede llevar T a la forma de escalonada de

Gauss-Jordan,

$$T' = S'^{-1}T = \left(\begin{array}{ccccccccc} 1 & x & \dots & 0 & x & \dots & 0 & x & \dots \\ 0 & 0 & \dots & 1 & x & \dots & 0 & x & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 & x & \dots \\ & & \dots & & & \dots & & & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots \\ & & \dots & & & \dots & & & \dots \end{array}\right)$$

donde x representa elementos no necesariamente nulos, y S'^{-1} una matriz no singular que es el producto de las operaciones elementales. T' posee k filas no nulas que son LI y por lo tanto la dimensión del e.f. de T' (idéntico al espacio fila de T, por ser las filas de T' combinaciones lineales de las de T) es k. Una base del e.f. de T son pues las k filas no nulas de T'.

k es también el número de columnas LI de T', ya que las columnas con pivotes (primer elemento no nulo de c/fila no nula) son LI y generan el e.c. de T'. Por lo tanto, la dimensión del e.c. de T' es también k. Pero esta es entonces la dimensión del e.c. de T por ser S'^{-1} no singular (véase b)).

Considerando a T como la representación de una TL $F : V \to V'$ en bases B, B' de V y V' tal que $T = [F]^B_{B'}$, la matriz $T' = S'^{-1}T = [F]^B_{\tilde{B}'}$ corresponde a un cambio de base en V', con $S' = [I]^{\tilde{B}'}_{B'}$. Las columnas con pivotes de T', $[F(b_{i_p})]_{\tilde{B}'}$, forman una base del e.c. de T', por lo que los correspondientes vectores $F(b_{i_p})$ forman una base de $I(F)$. Las correspondientes columnas de T, $[F(b_{i_p})]_{B'}$, forman entonces una base del e.c. de T. Nótese que en general, e.c. $(T) \neq$ e.c. (T'), aunque las dimensiones sean iguales.

Ejemplo 3) Sea $D : P_2 \to P_2$ el operador derivada restringido al subespacio de polinomios de grado ≤ 2. Es obvio que $N(D) = P_0$ (el subespacio de polinomios de grado 0, es decir, constantes) y que $I(D) = P_1$ (el subespacio de polinomios de grado ≤ 1), con dim $N(D)+$ dim $I(D) = 1 + 2 = 3 =$dim P_2.
En forma matricial, en la base "canónica" $e = (1, t, t^2)$, tenemos

$$[D]_e = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

El rango de esta matriz es 2, ya que posee dos filas (o columnas) LI, que coincide con dim $I(D)$.
Además, el espacio columna de $[D]_e$ es $\overline{\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}\}} = \overline{\{[e_1]_e, 2[e_2]_e\}}$, de modo que $I(D)$ será el subespacio generado por $e_1 = 1$ y $e_2 = t$, es decir, P_1.
El espacio nulo de $[D]_e$ es $\overline{\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\}} = \overline{\{[e_1]_e\}}$, y $N(D)$ es por lo tanto el subespacio generado por e_1, es decir, P_0.

Ejemplo 4) Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (2x + y, 3x - y)$. Mostrar que F es no singular y hallar su inversa.
F es un isomorfismo pues $I(F) = \{x(2, 3) + y(1, -1),\ x, y \in \mathbb{R}\} = \mathbb{R}^2$, ya que $(2, 3)$ y $(1, -1)$ son LI y por lo tanto base de $\mathbb{R}^2$. Puede llegarse al mismo resultando notando que $N(F) = \{(0, 0)\}$. Y también, notando que la matriz que representa a F en la base canónica $e = ((1, 0), (0, 1))$ es

$$[F]_e = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$$

Como $[F]_e$ es no singular (Det$[[F_e] = -5 \neq 0$), $[F]_e$ es invertible y por lo tanto F es un isomorfismo. La matriz que representa a su inversa es

$$[F^{-1}]_e = ([F]_e)^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}$$

por lo que la función inversa está dada por $F^{-1}(x, y) = (x + y, 3x - 2y)/5$.

Ejemplo 5) Sea $F:\mathbb{R}^3\to\mathbb{R}^3$ dada por $F(x,y,z)=(x+y+z,2x+z,3x-y+z)$. Hallar $N(F)$ y $I(F)$. En este caso conviene pasar directamente a la representación matricial. La matriz que representa a F en la base canónica $e=((1,0,0),(0,1,0),(0,0,1))$ es

$$[F]_e=\begin{pmatrix}1&1&1\\2&0&1\\3&-1&1\end{pmatrix}$$

F no es un isomorfismo pues $[F]_e$ es una matriz singular ($\mathrm{Det}[[F]_e]=0$) y por lo tanto no posee inversa. Las columnas b_i de $[F]_{ee}$ están vinculadas por $b_3=(b_2+b_1)/2$ (y las filas a_i de $[F]_e$ por $a_3=2a_2-a_1$), siendo b_2 y b_1 L.I.. El rango de $[F]_e$ es por lo tanto 2. Esto implica que $\dim I(F)=2$ y que $\dim N(F)=3-2=1$. Para hallar $N(F)$ se resuelve el sistema de ecuaciones que resulta de $F(x,y,z)=(0,0,0)$, o sea, $x+y+z=0$, $2x+z=0$, $3x-y+z=0$, que puede reescribirse en forma matricial como

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

es decir, $[F]_e[v]_e=0$ (vector columna nulo). Puede verse fácilmente que $[F]_e$ es equivalente por filas a la matriz $\begin{pmatrix}2&0&1\\0&2&1\\0&0&0\end{pmatrix}$. La solución al sistema homogéneo está entonces dada por $x=-z/2$, $y=-z/2$, con z arbitrario, por lo que el espacio nulo de la matriz es el conjunto $\overline{\{(-1/2,-1/2,1)^t\}}$.
El núcleo de F es pues el subespacio generado por el vector $v_0=(-1/2,-1/2,1)$
Una base del espacio columna de $[F]_e$ es, por ej., el conj. formado por las dos primeras columnas.
Por lo tanto, $I(F)$ es el espacio generado por $v_1=(1,2,3)=f(e_1)$, $v_2=(1,0,-1)=f(e_2)$.
Podemos escribir entonces $V=\overline{\{e_1,e_2\}}\oplus\overline{v_0}$ (la barra sobre vectores indica el espacio generado por dichos vectores), con v_0 base del núcleo y $(F(e_1),F(e_2))$ base de $I(F)$.

Ejemplo 6) Sea $f:\mathbb{R}^2\to\mathbb{R}^2$, definida por $F(1,1)=(2,1)$, $F(1,-1)=(-1,0)$. Hallar $F(x,y)$.
Los datos alcanzan para definir F pues $e'_1=(1,1)$, $e'_2=(1,-1)$ son LI y por lo tanto base de $\mathbb{R}^2$ (y toda transformación lineal queda completamente determinada por los vectores que asigna a los elementos de una base). Tenemos, a partir de los datos,

$$[F]_e^{e'}=\begin{pmatrix}2&-1\\1&0\end{pmatrix}$$

Además,

$$[I]_e^{e'}=S=\begin{pmatrix}1&1\\1&-1\end{pmatrix},\quad [I]_{e'}^{e}=S^{-1}=\frac{1}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}$$

Por lo tanto,

$$[F]_e=[F]_e^{e'}[I]_{e'}^{e}=[F]_e^{e'}S^{-1}=\begin{pmatrix}2&-1\\1&0\end{pmatrix}\begin{pmatrix}1&1\\1&-1\end{pmatrix}/2=\frac{1}{2}\begin{pmatrix}1&3\\1&1\end{pmatrix}$$

que implica

$$F(x,y)=(x+3y,x+y)/2$$

Puede llegarse al mismo resultado a partir de la relación $e_1=(e'_1+e'_2)/2$, $e_2=(e'_1-e'_2)/2$, con $Ff[e_1]=[F(e'_1)+F(e'_2)]/2=(1,1)/2$, $F[e_2]=[F(e'_1)-F(e'_2)]/2=(3,1)/2$ y por lo tanto $F(x,y)=xF(e_1)+yF(e_2)=(x+3y,x+y)/2$. El método matricial es, no obstante, más directo y apto para ser aplicado a sistemas de grandes dimensiones.

7- Inversas a Izquierda y Derecha

Sea $F : V \to V'$ una transformación lineal. $G : V' \to V$ lineal se denomina inversa a izquierda de F si

$$GF = I_V$$

donde $I_V : V \to V$ denota el operador identidad en V. En tal caso F es la inversa a derecha de G.

Teorema: Una transformación lineal posee inversa a izquierda si y sólo si es un monomorfismo, e inversa a derecha si y sólo si es un epimorfismo.
(Recordar esquema gráfico hecho en clase).
Dem.: a) Si F es un monomorfismo $\Rightarrow \forall\, v' \in I(F)\ \exists$ un y sólo un $v \in V$ tal que $F(v) = v'$. Podemos escribir en general $V' = I(F) \oplus Q$, donde Q es un suplemento de $I(F)$. Todo vector $v' \in V'$ puede pues escribirse en forma *única* como $v' = v'_1 + v'_2$, donde $v'_1 = F(v_1) \in I(F)$ y $v'_2 \in Q$. Definimos entonces $G : V' \to V$ como

$$G(v') = v_1$$

De esta forma, $G(v'_1) = v_1$ y $G(v'_2) = 0$ si $v'_1 \in I(F)$ y $v'_2 \in Q$. Es fácil comprobar que G es lineal (pues $G(v' + u') = v_1 + u_1 = G(v') + G(u')$ si $v' = v_1 + v_2$ y $u' = u_1 + u_2$, y $G(\alpha v') = \alpha v_1 = \alpha G(v')$). Es además un epimorfismo y satisface $GF = I_V$.
Notemos que si $I(F) \neq V'$, la inversa a izquierda no es única, pues podemos sumar a G cualquier función lineal $H : V' \to V$ no nula que satisfaga $H(v') = 0$ si $v' \in I(F)$ (o sea, $I(F) \subset N(H)$) tal que $HF = 0$ y por lo tanto $(G + H)F = GF$.
Además, si F no es monomorfismo, $\exists\ v \in V$, con $v \neq 0$, tal que $F(v) = 0$ y por lo tanto, $(GF)(v) = G(F(v)) = G(0) = 0 \neq v$, por lo que F no puede tener inversa a izquierda en tal caso.

b) Si $G : V' \to V$ es un epimorfismo, sea $N(G)$ su espacio nulo y sea Q un suplemento tal que $V' = N(G) \oplus Q$. G restringido a Q ($G : Q \to V$) es un isomorfismo, ya que sigue siendo epimorfismo y además, si $v' \in Q$ y $v' \neq 0$, $G(v') \neq 0$. Definamos $F : V \to V'$ tal que $F(v)$ es el único vector v' de Q que satisface $G(v') = v$ ($I(F) = Q$). Es fácil ver que F es lineal, es monomorfismo y satisface $GF = I_V$.
No obstante, la inversa a derecha no es única si $N(G) \neq \{0\}$, pues podemos sumar a F cualquier función no nula $H : V \to V'$ con $I(H) \subset N(G)$, tal que $GH = 0$.
Además, si G no es epimorfismo, $\exists\, v \in V$ tal que v no pertence a $I(G)$ y por lo tanto $(GF)(v) = G(F(v)) \neq v$, pues $G(F(v)) \in I(G)$. No puede pues existir inversa a derecha en este caso.

Si V es de dimensión n y V' de dimensión m, las matrices que representan a F y G en bases ordenadas B, B' de V y V' satisfacen

$$[G]_B^{B'}[F]_{B'}^{B} = I_n$$

con I_n la identidad de $n \times n$, $[F]_{B'}^{B}$ de $m \times n$ y $[G]_B^{B'}$ de $n \times m$. La matriz $[G]_B^{B'}$ se dice que es inversa a izquierda de la matriz $[F]_{B'}^{B}$, y $[F]_{B'}^{B}$ la matriz inversa a derecha de $[G]_B^{B'}$. Si F posee una inversa a izquierda G y a derecha H debe ser entonces un isomorfismo, con $m = n$ en el caso de dimensión finita. En tal caso $G = H$, ya que $G = GI_{V'} = G(FH) = (GF)H = I_V H = H$.

Este teorema implica que una matriz A de $m \times n$ (m filas, n columnas) **tiene inversa a izquierda** B ($BA = I_n$, con B de $n \times m$) **si y sólo si Rango** $(A) = n$ (y por lo tanto $m \geq n$), en cuyo caso A representa a un **monomorfismo**. Y una matrix B de $m \times n$ tiene **inversa a derecha** A ($BA = I_m$, con A de $n \times m$) **si y sólo si Rango** $(B) = m$ (y por lo tanto $m \leq n$), en cuyo caso representa un epimorfismo.
Si una matriz posee inversa izquierda y a derecha entonces debe ser necesariamente cuadrada y representar un isomorfismo, siendo pues no singular. En tal caso la inversa a izquierda y a derecha coinciden.

Si $A \in \mathbb{C}^{m\times n}$ tiene rango n (lo que implica $m \geq n$) sus columnas son LI. Es fácil mostrar que $A \in \mathbb{C}^{m\times n}$ **tiene rango** n **si y sólo si la matriz** $A^\dagger A \in \mathbb{C}^{n\times n}$ **es no singular (**$\text{Det}(A^\dagger A) \neq 0$**)**. Recordemos que $A^\dagger = (A^t)^*$, es decir, $A^\dagger_{ij} = A^*_{ji}\ \forall\ i, j$.
Dem.: Si las columnas son independientes, la única solución de $AX = 0$ (con $X \in \mathbb{C}^{n\times 1}$) es $X = 0$ (en otras palabras, A representa a un monomorfismo y por lo tanto su espacio nulo es $\{0\}$). Si existe X tal que $A^\dagger AX = 0$ entonces $X^\dagger A^\dagger AX = (AX)^\dagger(AX) = |AX|^2 = 0$ y por lo tanto $AX = 0$. Esto implica entonces $X = 0$, por lo que $A^\dagger A$ es no singular: $\text{Det}(A^\dagger A) \neq 0$.

Análogamente, Si $A^\dagger A$, es no singular, la única solución de $A^\dagger AX = 0$ es $X = 0$, por lo que la única solución de $AX = 0$ es $X = 0$, indicando que las columnas de A son linealmente independientes, es decir, que A tiene rango n.
Esto permite pues construir en forma inmediata una inversa a izquierda B $\forall$ $A \in \mathbb{C}^{m\times n}$ con rango n:

$$B = (A^\dagger A)^{-1}A^\dagger \qquad (1)$$

ya que $BA = I_n$ (notar que $B \in \mathbb{C}^{n\times m}$).

Análogamente, si $B \in \mathbb{C}^{m\times n}$ tiene rango m, B representa a un epimorfismo. En tal caso $B^\dagger \in \mathbb{C}^{n\times m}$ tiene rango m y por lo tanto $BB^\dagger$ es no singular. Una inversa a derecha de B es pues

$$A = B^\dagger(BB^\dagger)^{-1} \qquad (2)$$

ya que $BA = I_m$ (notar que $A \in \mathbb{C}^{n\times m}$). Recordemos, no obstante, que si $m \neq n$, existen otras inversas a izquierda y a derecha respectivamente, aunque las inversas (1) y (2) poseen ciertas propiedades especiales que discutiremos más adelante. Por otro lado, si $m = n \Rightarrow B = A^{-1}$ en (1) y $A = B^{-1}$ en (2), como el lector puede fácilmente comprobar.

Ejemplo 1) Sea $D : P \to P$ la derivación considerada en el espacio vectorial P de todos los polinomios, de dimensión infinita. D no es un monomorfismo, pues $D(1) = 0$ (la derivada de cualquier polinomio de grado 0 es el polinomio nulo), por lo que $N(D) = \overline{\{1\}}$, pero sí es un epimorfismo, pues $\forall q(t) \in P$ $\exists$ $p(t) \in P$ tal que $D(p(t)) = q(t)$ (Por ejemplo, $p(t) = \int_0^t q(t')dt'$, que es un polinomio).
Por lo tanto, D tendrá inversa a derecha. Una inversa es precisamente la integración $S : P \to P$ definida por $S(q(t)) = \int_0^t q(t')dt'$, que es lineal y que satisface

$$DS = I_P$$

(I_P es la identidad en P). En efecto, $(DS)(t^n) = D(S(t^n)) = D(t^{n+1}/(n+1)) = t^n$ $\forall$ $n \geq 0$. No obstante, $SD \neq I_P$ pues $SD(1) = S(0) = 0 \neq 1$, o sea que S es inversa sólo a derecha de D.
S es un monomorfismo ($N(S) = \{0\}$), pero no un epimorfismo, ya que la imagen de S no contiene a los polinomios de grado 0.
Notemos también que S' definido por $S'(q(t)) = \int_a^t q(t')dt'$ es también una inversa a derecha de D para cualquier a real, y también lo es T dada por $T(q(t)) = \int_0^t q(t')dt' + cq(0)$, siendo c cualquier constante $\in K$.

Ejemplo 2) Sea $D : P_2 \to P_1$ la derivación restringida a P_2 (Polinomios de grado ≤ 2) con codominio P_1. Tenemos, en las bases $e = (1, t, t^2)$, $e' = (1, t)$ de P_2 y P_1 respectivamente,

$$[D]^e_{e'} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

ya que $D(e_1) = 0$, $D(e_2) = 1 = e'_1$, $D(e_3) = 2t = 2e'_2$. D así definido es claramente un epimorfismo, pues $I(D) = P_1$. Una inversa a derecha de D es la transformación $S : P_1 \to P_2$ definida como la integral $S(p(t)) = \int_0^t p(t')dt'$, con $S(e'_1) = t = e_2$, $S(e'_2) = t^2/2 = e_3/2$, representada por la matriz

$$[S]^{e'}_e = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1/2 \end{pmatrix}$$

Es fácil ver que S es un monomorfismo y que es una inversa a derecha de D, pues

$$[D]^e_{e'}[S]^{e'}_e = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

es decir,

$$DS = I_{P_1}$$

Precisamente, si $A = [S]^{e'}_e \Rightarrow$ la ec. (1) nos da $B = [D]^e_{e'}$. Y si $B = [D]^e_{e'}$, la ec. (2) nos da $A = [S]^{e'}_e$, como el lector puede fácilmente verificar.

No obstante, notemos que $[S]_e^{e'}[D]_{e'}^e = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \neq I_3$, por lo que $SD \neq I_{P_2}$.

Notemos también que $S' : P_1 \to P_2$ definida por $S'(e_1) = t + a$, $S'(e_2) = t^2/2 + b$, y representada por

$$[S']_e^{e'} = \begin{pmatrix} a & b \\ 1 & 0 \\ 0 & 1/2 \end{pmatrix}$$

es *también* inversa a derecha de $[D]_{e'}^e$ $\forall\ a, b$, como puede verificarse fácilmente, y que

$$[D']_{e'}^e = \begin{pmatrix} c & 1 & 0 \\ d & 0 & 2 \end{pmatrix}$$

es también una inversa a izquierda de $[S]_e^{e'}$ $\forall\ c, d$, de modo que éstas no son únicas.

8- Solución general de una ecuación lineal

Como aplicación fundamental, consideremos, para $F : V \to V'$ lineal, la ecuación general

$$F(v) = v'$$

donde se trata de encontrar el conjunto de vectores $v \in V$ que satisfacen $F(v) = v'$.
Si $v' = 0$, la ecuación se denomina homogénea y el conjunto de soluciones es el espacio nulo $N(F)$.
Si $v' \neq 0$, la ecuación se denomina no homogénea. En tal caso el conjunto de soluciones de la ecuación no homogénea *no es un espacio vectorial* (por ej. 0 no pertenece al conjunto pues $F(0) = 0 \neq v'$). Se cumple en cambio que si $F(v_1) = v'_1$ y $F(v_2) = v'_2 \Rightarrow$

$$F(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 v'_1 + \alpha_2 v'_2$$

o sea, la combinación lineal $\alpha_1 v_1 + \alpha_2 v_2$ es solución de la ecuación $F(v) = \alpha_1 v'_1 + \alpha_2 v'_2$.

Es obvio además que existirá solución si y sólo si $v' \in I(F)$.
Supongamos ahora que existan dos soluciones v_1, v_2, tal que $F(v_1) = F(v_2) = v'$. Entonces

$$0 = F(v_1) - F(v_2) = F(v_1 - v_2)$$

por lo que $v_2 - v_1 \in N(F)$. Por lo tanto, $v_2 = v_1 + v_n$, donde $v_n \in N(F)$. Es decir, si v_p es una solución particular que satisface $F(v_p) = v' \Rightarrow$ cualquier otra solución v es de la forma

$$v = v_p + v_n$$

donde $v_n \in N(F)$ es un vector del espacio nulo de F, es decir, una solución de la ec. *homogénea* ($F(v_n) = 0$). La solución *general* estará dada entonces por la suma de una solución particular v_p de la ecuación no homogénea y de una solución v_n de la ecuación homogénea.

Resulta claro entonces que si $v' \in I(F)$, la solución será única si y sólo si $N(F) = 0$, o sea, si y sólo si F es un monomorfismo. En tal caso, la única solución de $F(v) = v'$ puede encontrarse como

$$v = G(v')$$

donde *G es una inversa a izquierda de F*. En efecto, si $F(v) = v' \Rightarrow G(F(v)) = (GF)(v) = I_V(v) = v = G(v')$. Puede utilizarse cualquier inversa a izquierda ya que difieren entre sí sólo para vectores que no pertenecen a $I(F)$ ($G_2(v') = G_1(v')$ si $v' \in I(F)$).

Por otro lado, si $F : V \to V'$ es un epimorfismo $\Rightarrow$ la ecuación $F(v) = v'$ tendrá siempre solución, pero no será única a no ser que $N(F) = \{0\}$ (en cuyo caso F es isomorfismo). La solución general será

$$v = G(v') + v_n$$

donde G es una inversa a derecha de F y v_n un vector arbitrario de $N(F)$. En efecto, $F(G(v')+v_n) = F(G(v'))+F(v_n) = (FG)(v')+0 = I_{V'}(v') = v'$. Aquí $G(v')$ representa la solución particular v_p, la cual, remarquemos, *es una función lineal* de v'. Puede utilizarse cualquier inversa a derecha pues estas difieren sólo en un vector de v_n ($G_2(v') = G_1(v') + v_n$ con $v_n \in N(F)$).

Finalmente, si F es un isomorfismo, existe una única solución $\forall\ v' \in V'$ dada por

$$v = F^{-1}(v')$$

con F^{-1} la (única) inversa de F, que es a la vez inversa a izquierda y derecha.

Ejemplo 1): Para el caso de sistemas de m ecuaciones lineales con n incógnitas, dados por

$$AX = Y$$

con A de $m \times n$, X de $n \times 1$, Y de $m \times 1$ y A y Y de elementos reales (que corresponde a la función $F: \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ dada por $F(X) = AX$) los resultados anteriores implican que:
1) El sistema posee solución si y sólo si Y pertenece al espacio columna de A.
En tal caso, la solución general será de la forma

$$X = X_p + X_n$$

donde X_p es una solución particular ($AX_p = Y$) y X_n una solución general del sistema homogéneo ($AX_n = 0$, con 0 el vector columna nulo).

2) La solución será única si y sólo si el espacio nulo de A es el vector columna nulo ($X_n = 0$), es decir, si y sólo si Rango $(A) = n$ (y por lo tanto, $m \geq n$). En este caso F es un monomorfismo y la única solución (en el caso que Y pertenezca al espacio columna de A) puede encontrarse como $X = BY$, con B de $n \times m$ una inversa a izquierda de A ($BA = I_n$).

3) Si la dimensión del espacio columna es m (en cuyo caso Rango $(A) = m$ y por lo tanto, $m \leq n$) existirá solución para cualquier Y de $m \times 1$. En este caso F es un epimorfismo y la solución general puede escribirse como $X = CY + X_n$, con C de $n \times m$ una inversa a derecha de A ($AC = I_m$) y X_n solución del sistema homogéneo $AX_n = 0$.

4) Si $m = n$ y Rango $(A) = n \Rightarrow$ existe siempre una única solución dada por $X = A^{-1}Y$, con A^{-1} la inversa de A. En este caso F representa un isomorfismo.

Ejemplo 2): Resolver el sistema $x + y = a$, $4x + 2z = b$, en forma matricial. Corresponde a

$$\begin{pmatrix} 1 & 1 & 0 \\ 4 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

Como el rango de la matriz es 2, existirá solución $\forall\, a, b$. Reduciendo por filas el sistema ampliado y llevándolo a la forma de Gauss-Jordan, se obtiene

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & a \\ 4 & 0 & 2 & b \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{1}{4}b \\ 0 & 1 & -\frac{1}{2} & a - \frac{1}{4}b \end{array}\right)$$

de donde se lee la solución general $x = b/4 - z/2$, $y = a - b/4 + z/2$, con z libre. Podemos escrbir esta solución como

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{4} \\ 1 & -\frac{1}{4} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + t \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

con $t \in \mathbb{R}$ arbitrario (parámetro libre), donde $B = \begin{pmatrix} 0 & \frac{1}{4} \\ 1 & -\frac{1}{4} \\ 0 & 0 \end{pmatrix}$ es precisamente *una inversa a derecha de la matriz de coeficientes* $A = \begin{pmatrix} 1 & 1 & 0 \\ 4 & 0 & 2 \end{pmatrix}$ y $X = t \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$ un elemento arbitrario del espacio nulo de A

$(AX = 0)$. Nótese que

$$\begin{pmatrix} 1 & 1 & 0 \\ 4 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{4} \\ 1 & -\frac{1}{4} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

pero

$$\begin{pmatrix} 0 & \frac{1}{4} \\ 1 & -\frac{1}{4} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 4 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

Anexo: Operadores de Proyección:

Recordemos que si S_1 es un subespacio de V y S_2 un suplemento tal que $V = S_1 \oplus S_2$, todo vector $v \in V$ puede escribirse en forma única como $v = v_1 + v_2$, con $v_1 \in S_1$, $v_2 \in S_2$.
El proyector P_{S_1/S_2} sobre S_1 en la dirección de S_2 (recordar la interpretación geométrica dada en clase) queda entonces definido por

$$P_{S_1/S_2}(v) = v_1$$

y satisface por lo tanto $P^2_{S_1/S_2} = P_{S_1/S_2}$, pues $P^2_{S_1/S_2}(v) = P_{S_1/S_2}(P_{S_1/S_2}(v)) = P_{S_1/S_2}(v_1) = v_1$. Obviamente, su núcleo e imagen son $N(P_{S_1/S_2}) = S_2$, $I(P_{S_1/S_2}) = S_1$. Depende de S_1 y S_2.

Análogamente, si $P^2 = P$, P es un proyector sobre $S_1 = I(P)$ en la dirección de $S_2 = N(P)$, con $S_1 \oplus S_2 = V$, es decir $P = P_{I(P)/N(P)}$.
En efecto, para cualquier $v \in V$, $v = P(v) + v - P(v) = v_1 + v_2$, con $v_1 = P(v) \in I(P)$ y $v_2 = v - P(v) \in N(P)$ pues $P(v - P(v)) = P(v) - P^2(v) = P(v) - P(v) = 0$. Además $I(P) \cap N(P) = \{0\}$ pues si $v = P(v')$ y $P(v) = 0 \Rightarrow 0 = P^2(v') = P(v') = v$. Por lo tanto $V = I(P) \oplus N(P)$.
Finalmente $P(v) = P(v_1 + v_2) = P(v_1) + P(v_2) = P^2(v) + 0 = P(v) = v_1$, por lo que P es proyector sobre $S_1 = I(P)$ en la dirección de $S_2 = N(P)$. Esto incluye los casos triviales $I(P) = V$, en cuyo caso $N(P) = \{0\}$ y por lo tanto $P_{S_1/S_2} = I_V$ (operador identidad), y $I(P) = \{0\}$, en cuyo caso $N(P) = V$ y $P = 0$ (operador nulo). Dado que $v = v_1 + v_2 = P_{S_1/S_2}(v) + P_{S_2/S_1}(v)$, se cumple siempre

$$P_{S_1/S_2} + P_{S_2/S_1} = I_V$$

Si $\dim V = n$, la representación matricial de P_{S_1/S_2} en una base ordenada $B = (b_1, \ldots, b_k, b_{k+1}, \ldots, b_n)$, en la que $(b_1, \ldots, b_k)$ es base de S_1 y $(b_{k+1}, \ldots, b_n)$ es base de S_2, es de la forma

$$[P_{S_1/S_2}]^B_B = \begin{pmatrix} I_{k\times k} & 0_{k\times m} \\ 0_{m\times k} & 0_{m\times m} \end{pmatrix}$$

donde $I_{k\times k}$ denota la matriz identidad de $k \times k$, $m = n - k$ y 0_{pq} la matriz nula de $p \times q$. Es claro que P_{S_1/S_2} es un operador singular ($\mathrm{Det}[P_{S_1/S_2}] = 0$), salvo en el caso $S_1 = V$ ($S_2 = \{0\}$, $P = I_V$).

Los proyectores ortogonales son aquellos en los que S_2 es el subespacio ortogonal a S_1, tema que veremos en detalle más adelante. En general, S_2 puede ser cualquier suplemento de S_1, no necesariamente el ortogonal, por lo que la matriz que representa a P_{S_1/S_2} puede no ser simétrica o hermítica en una base ortonormal.

Ej.: Consideremos, para $V = \mathbb{R}^2$, el proyector sobre el subespacio generado por $b_1 = (1, 0)$ en la dirección del subespacio generado por $b_2 = (1, 1)$. Como $P(b_1) = b_1$, $P(b_2) = 0$, en la base $b = \{b_1, b_2\}$ se obtiene

$$[P_{S_1/S_2}]^b_b = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

En la base canónica $e = (e_1, e_2)$, con $e_1 = (1, 0)$, $e_2 = (0, 1)$, se obtiene entonces

$$[P_{S_1/S_2}]^e_e = S[P_{S_1/S_2}]^b_b S^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

donde $S = [I]^b_e = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ y $S^{-1} = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right) = [I]^e_b$. Las columnas de $[P_{S_1/S_2}]^e_e$ son proporcionales a $[b_1]^e_e = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$.
Para $v = (x, y)$ obtenemos entonces $[P_{S_1/S_2}(v)]_e = [P_{S_1/S_2}]^e_e \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} x-y \\ 0 \end{smallmatrix}\right)$, es decir, $P_{S_1/S_2}(x, y) = (x - y, 0)$ en acuerdo con $(x, y) = (x - y)(1, 0) + y(1, 1)$ (Recordar dibujo).
El proyector ortogonal usual sobre el subespacio S_1 generado por $e_1 = b_1$ es $P_{S_1} \equiv P_{S_1/S_1^\perp}$, donde $S_1^\perp$ es el subespacio ortogonal, generado por ej. por e_2. Obtenemos $[P_{S_1}]^e_e = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ y por lo tanto $P(x, y) = (x, 0)$, de acuerdo a $(x, y) = x(1, 0) + y(0, 1)$.

9. Autovalores y Autovectores

Sea V un espacio vectorial sobre el cuerpo K y sea $F : V \to V$ un operador lineal. Un escalar $\lambda \in K$ es un *autovalor* de F si existe $v \in V$, con $v \neq 0$, tal que

$$F(v) = \lambda v \qquad (v \neq 0)$$

En tal caso v es un *autovector* de F correspondiente al autovalor λ. Sinónimo de autovalor es *valor propio* (en inglés "eigenvalue", donde "eigen" proviene del alemán y significa propio) y sinónimo de autovector es *vector propio* ("eigenvector" en inglés).
La acción de F sobre un autovector es pues la multiplicación por un escalar. Por ejemplo, si $V = \mathbb{R}^n$, esto implica que $v \neq 0$ es autovector de F si y sólo si $w = F(v)$ tiene la misma dirección que v (aunque no necesariamente el mismo sentido) o es nulo. Como veremos, el conocimiento del conjunto de los autovalores y autovectores de un operador permite conocer su estructura en detalle.

El concepto de autovalor y autovector de un operador lineal tiene una importancia fundamental en Física, especialmente en Mecánica Cuántica. Mencionemos que en la misma los observables físicos corresponden a operadores lineales (de características determinadas) en un cierto espacio (el de los vectores que representan los estados del sistema) y las cantidades medibles son precisamente los autovalores de dichos operadores. E inmediatamente después de una medición, el sistema cuántico queda en un estado que es autovector del operador correspondiente al autovalor medido.
El concepto de autovalor y autovector es también fundamental para el tratamiento de sistemas descriptos por ecuaciones diferenciales lineales acopladas, tales como osciladores armónicos acoplados, ya sean clásicos o cuánticos, donde permite entender el concepto de desacoplamiento y modo normal. Otras aplicaciones incluyen, como veremos, la determinación de los ejes principales de inercia en un cuerpo rígido y la resolución de sucesiones definidas mediante relaciones recursivas lineales, tales como la famosa sucesión de Fibonacci.

Veamos algunas consecuencias inmediatas de tal definición:
10.1) El conjunto de autovectores de F correspondientes a un determinado autovalor λ, junto con el vector nulo 0, forma un subespacio de V denominado *espacio propio* o autoespacio ("eigenspace" en inglés) de F correspondiente al autovalor λ, que denotaremos como $V_F(\lambda)$ o simplemente, $V(\lambda)$.
En efecto, si $v \neq 0$ es autovector de F corresp. al autovalor λ, y $\alpha \in K$, $F(\alpha v) = \alpha F(v) = \alpha\lambda v = \lambda(\alpha v)$ $\forall$ α, por lo que αv es también autovector de F corresp. a λ si $\alpha \neq 0$
Y si v_1 y v_2 son autovectores corresp. al mismo autovalor λ, $F(v_1 + v_2) = F(v_1) + F(v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$, por lo que $v_1 + v_2$ es también autovector corresp. al autovalor λ (si $v_1 + v_2 \neq 0$).
La dimensión de $V_F(\lambda)$ es como mínimo uno (ya que al menos existe un autovector no nulo).

9.2) El espacio propio $V_F(\lambda)$, con λ autovalor de F, es el núcleo del operador $F - \lambda I$:

$$V_F(\lambda) = N[F - \lambda I]$$

donde I denota el operador identidad en V ($I(v) = v$ $\forall$ $v \in V$).
En efecto, si $F(v) = \lambda v \Rightarrow 0 = F(v) - \lambda v = (F - \lambda I)(v)$, y si $(F - \lambda I)(v) = 0 \Rightarrow F(v) = \lambda v$.
Por lo tanto $\lambda \in K$ es autovalor de F si y sólo si

$$N[F - \lambda I] \neq \{0\}$$

En particular, $\lambda = 0$ es autovalor de F si y sólo si $N(F) \neq \{0\}$, es decir, si y sólo si F no es monomorfisomo. En tal caso $V_F(0) = N(F)$.

9.3) Si V es de dimensión finita n, $\lambda \in K$ es autovalor si y sólo si

$$\mathrm{Det}[F - \lambda I] = 0$$

En efecto, si λ es autovalor, $N[F - \lambda I] \neq \{0\}$, por lo que $F - \lambda I$ no es monomorfismo. La matriz que representa a $F - \lambda I$ debe ser entonces singular en cualquier base, y por lo tanto, su determinante nulo. Por otro lado, si $\mathrm{Det}[F - \lambda I] = 0$, $F - \lambda I$ no es monomorfismo, y por lo tanto existe $v \neq 0$ tal que $(F - \lambda I)(v) = 0$.
Recordemos que si e es una base de V,

$$\mathrm{Det}[F - \lambda I] = |[F]_e - \lambda I_n|$$

donde $[F]_e \equiv [F]_e^e$ es la matriz que representa a F en dicha base, $I_n = [I]_e$ es la matriz identidad de $n \times n$ y $|A| = \text{Det}A$ denota el determinante de la matriz A. $\text{Det}[F - \lambda I]$ es *independiente* de la base elegida e:

$$|[F]_{e'} - \lambda I_n| = |S^{-1}[F]_e S - \lambda I_n| = |S^{-1}([F]_e - \lambda I_n)S| = |S^{-1}||[F]_e - \lambda I_n||S| = |[F]_e - \lambda I_n|$$

ya que $[I]_{e'} = [I]_e = I_n$ y $|S^{-1}| = 1/|S|$. Los autovalores de F en un espacio de dimensión finita n se obtienen entonces como las raíces pertenecientes al cuerpo K del polinomio

$$P(\lambda) = \text{Det}[F - \lambda I]$$

denominado *polinomio característico*, que es de grado n (pues $[F - \lambda I]_e$ es de $n \times n$). La ecuación $P(\lambda) = 0$ se denomina *ecuación caracerística* y posee, por lo tanto, a lo sumo n *raíces distintas*, que en general pueden ser complejas. Serán autovalores si pertenecen al cuerpo K. Si $K = \mathbb{C} \Rightarrow$ *toda raíz* de $P(\lambda)$ es autovalor.

9.4) *Teorema: Los autovectores de F correspondientes a autovalores distintos son LI*
Demostraremos el teorema por inducción. Para $n = 1$, v_1 autovector es LI pues es no nulo (mejor comprensión se logra comenzando con $n = 2$: Si $v_1 \neq 0$ y $v_2 \neq 0$ son autovectores correspondientes a autovalores *distintos* $\Rightarrow$ son LI pues de lo contrario $v_2 = \alpha v_1$, y correspondería por 9.1 al mismo autovalor que v_1). Supongamos ahora que $v_1, \ldots, v_{k-1}$ son autovectores LI, con autovalores $\lambda_1, \ldots, \lambda_{k-1}$, y que $v_k \neq 0$ es autovector con autovalor λ_k, siendo λ_k distinto a todos los anteriores. Si

$$0 = \alpha_1 v_1 + \ldots + \alpha_{k-1} v_{k-1} + \alpha_k v_k$$

entonces

$$0 = F(0) = \alpha_1 F(v_1) + \ldots + \alpha_{k-1} F(v_{k-1}) + \alpha_k F(v_k) = \alpha_1 \lambda_1 v_1 + \ldots \alpha_{k-1} \lambda_{k-1} v_{k-1} + \alpha_k \lambda_k v_k$$

Restando ahora la primera ecuación mult. por λ_k,

$$0 = (\lambda_1 - \lambda_k)\alpha_1 v_1 + \ldots + (\lambda_{k-1} - \lambda_k)\alpha_{k-1} v_{k-1}$$

que implica $\alpha_1 = \ldots = \alpha_{k-1} = 0$ por ser $v_1, \ldots v_{k-1}$ LI y $\lambda_k \neq \lambda_i$ para $i = 1, \ldots, k-1$. Por lo tanto, también $\alpha_k = 0$ y entonces $v_1, \ldots, v_k$ son L.I.

Nótese que la demostración es igualmente válida si los $\lambda_1, \ldots, \lambda_{k-1}$ no son todos distintos (pero sí distintos a λ_k) siempre que los $v_1, \ldots, v_{k-1}$ sean LI
Por lo tanto, *ningún elemento $v_k \neq 0$ de $V(\lambda_k)$ puede ser generado por autovectores correspondientes a autovalores distintos de λ_k.*

9.5) Si $\lambda_1 \neq \lambda_2 \Rightarrow V_F(\lambda_1) \cap V_F(\lambda_2) = \{0\}$
Es consecuencia inmediata del último párrafo. Si $v \in V_F(\lambda_2)$ y $v \in V_F(\lambda_1)$, entonces $F(v) = \lambda_1 v = \lambda_2 v$, por lo que $(\lambda_1 - \lambda_2)v = 0$ y por lo tanto $v = 0$. Esto implica que $V_F(\lambda_1) + V_F(\lambda_2) = V_F(\lambda_1) \oplus V_F(\lambda_2)$.
Del mismo modo, la intersección de $V_F(\lambda_k)$ con la suma $V_F(\lambda_1) + \ldots + V_F(\lambda_{k-1})$ es $\{0\}$ si λ_k es distinto a todos los autovalores anteriores. Si así no fuese existiría un vector $v \in V_F(\lambda_k)$ con $v \neq 0$ que puede ser escrito como combinación lineal de autovectores correspondientes a autovalores distintos, pero esto es imposible por el teorema 9.4 anterior.
Podemos entonces escribir la suma de espacios propios como suma directa $V_F(\lambda_1) \oplus \ldots \oplus V_F(\lambda_k)$.

9.6) Un operador F en un espacio V de dimensión finita se dice *diagonalizable* si existe una base formada por autovectores de F.
En tal caso, denotando la base como $e' = (e'_1, \ldots, e'_n)$, con $F(e'_i) = \lambda_i e'_i$, $i = 1, \ldots, n$, la matriz que representa a F en dicha base es *diagonal*:

$$[F]_{e'} = \left(\begin{array}{ccc} & \ldots & \\ [F(e'_1)]_{e'} & \ldots & [F(e'_n)]_{e'} \\ & \ldots & \end{array} \right) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

Recíprocamente, si $[F]_{e'}$ es diagonal $\Rightarrow F(e'_i) = \lambda_i e'_i$ y e' es necesariamente una base de autovectores. Si F es diagonalizable y e es una base arbitraria de V, tenemos

$$[F]_{e'} = S^{-1}[F]_e S$$

con $[F]_{e'}$ *diagonal* y $S = [I]_e^{e'}$ la matriz de cambio de base, *por lo que existe una matriz no singular* S *tal que* $S^{-1}[F]_e S$ *es diagonal.* La columna i de S es el vector de componentes $[e'_i]_e$ del autovector e'_i corresp. al autovalor λ_i en la base original e. Esta es la forma de construir la matriz diagonalizante S.

Consecuencia Importante de 9.4: Si $P(\lambda)$ *posee* n *raíces distintas* $\lambda_i \in K$, $\Rightarrow F$ *es diagonalizable.*
En efecto, en tal caso existirán n autovectores LI (uno por cada autovalor) que serán por tanto base de V. La dimensión de cada espacio propio $V_F(\lambda_i)$ es en este caso 1, que es igual a la multiplicidad de cada raíz.

9.7) Teorema: *F es diagonalizable si y sólo si i) todos las raíces de $P(\lambda)$ pertenecen al cuerpo y ii) la dimensión del espacio propio $V_F(\lambda_i)$ correspondiente a la raíz λ_i es igual a la multiplicidad m_i de dicha raíz.*
La dimensión del espacio propio $d_i = \dim V_F(\lambda_i)$ es el máximo número de autovectores LI que pueden obtenerse para un mismo autovalor λ_i, y se denomina también *multiplicidad geométrica* de λ_i.
Demostración: Supongamos F diagonalizable. En una base e' en la que $[F]_{e'}$ es diagonal, tenemos

$$P(\lambda) = (\lambda_1 - \lambda) \ldots (\lambda_n - \lambda)$$

La multiplicidad m_i de una raíz λ_i será pues igual al número de veces que λ_i se repite en la diagonal. Pero este número es igual al número de vectores de la base e' que tienen a λ_i como autovalor, que es precisamente la dimensión del espacio propio. Nótese también que d_i es el número de filas nulas de $[F]_{e'} - \lambda_i I_n$.
Por otro lado, si la dimensión de cada espacio propio es igual a la multiplicidad m_i de la raíz λ_i, la suma directa de todos los espacios propios correspondientes a autovalores *distintos*, $V_F(\lambda_1) \oplus \ldots \oplus V_F(\lambda_k)$ tendrá dimensión $d_1 + \ldots + d_k = m_1 + \ldots + m_k = n$ (ya que la suma de todas las multiplicidades es igual al grado del polinomio), por lo que será igual al espacio V. Existe entonces una base formada por autovectores de F.

9.8) En general, la dimensión del espacio propio $V_F(\lambda_i)$ puede ser igual o menor que la multiplicidad de la raíz λ_i: dim $V_F(\lambda_i) \leq m_i$.
En efecto, eligiendo una base e donde los primeros $d_i = \dim V_F(\lambda_i)$ elementos formen una base de $V_F(\lambda_i)$, las primeras d_i columnas de $[F]_e$ tendrán elementos no nulos sólo en la diagonal y por lo tanto $P(\lambda) = \text{Det}[[F]_e - \lambda I_n] = (\lambda_i - \lambda)^{d_i} Q(\lambda)$, con $Q(\lambda)$ un polinomio de grado $n - d_i$, por lo que m_i será como mínimo d_i ($m_i = d_i$ si $Q(\lambda_i) \neq 0$ y $m_i > d_i$ si $Q(\lambda_i) = 0$).
Si $d_i < m_i$, F *no es diagonalizable* (aún tomando como cuerpo $\mathbb{C}$).

Ejemplo 1 (Casos Triviales): Si $I : V \to V$ es el operador identidad $\Rightarrow I(v) = v \ \forall \ v \in V$ por lo que su único autovalor es 1. El espacio propio correspondiente es V ($V_F(1) = V$).
Si V es de dimensión finita n, puede llegarse a la misma conclusión notando que

$$P(\lambda) = \text{Det}[I - \lambda I] = |I_n - \lambda I_n| = |(1 - \lambda) I_n| = (1 - \lambda)^n$$

$\lambda = 1$ es pués la única raíz de $P(\lambda)$, y posee multiplicidad n, que es igual a la dimensión del espacio propio (el mismo V). I es por lo tanto diagonalizable (caso trivial).

Si $0 : V \to V$ es el operador nulo $\Rightarrow 0(v) = 0 = 0v \ \forall \ v \in V$ por lo que su único autovalor es 0, y el espacio propio correspondiente es V ($V_F(0) = V$).
Si V es de dimensión $n \Rightarrow [0]_e = 0$ (la matriz nula) en cualquier base y por lo tanto $P(\lambda) = \text{Det}[0 - \lambda I_n] = |-\lambda I_n| = (-\lambda)^n$, por lo que 0 es la única raíz, con multiplicidad n. 0 es también trivialm. diagonalizable.

Ejemplo 2: Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ la reflexión respecto del eje x. Si $e = (e_1, e_2)$ es la base canónica, con $e_1 = (1, 0)$, $e_2 = (0, 1)$, sabemos que $F(e_1) = e_1$, $F(e_2) = -e_2$. Por lo tanto e_1 es autovector de F con autovalor 1 y e_2 autovector de F con autovalor -1. No pueden existir otros autovalores pues la dimensión de V es 2. $V_F(1)$ es entonces el espacio generado por e_1, es decir, el conjunto de vectores $(x, 0)$, con $x \in \mathbb{R}$, sobre los que F actúa como identidad, y $V_F(-1)$ el generado por e_2, es decir, el conjunto de vectores $(0, y)$, con $y \in \mathbb{R}$, para los que la acción de F es la inversión de sentido.
Podemos obtener el mismo resultado a partir de la representación matricial

$$[F]_e = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

que ya es diagonal, por lo que los autovalores son 1 y -1: Tenemos $P(\lambda) = |[F]_e - \lambda I_2| = (1 - \lambda)(-1 - \lambda)$, siendo entonces las raíces ± 1.

Ejemplo 3: Sea $F:\mathbb{R}^2\to\mathbb{R}^2$ la reflexión respecto de la recta de ecuación $y=x$, dada por (recordar ejemplo dado) $F(x,y)=(y,x)$. Si $e'=(e'_1,e'_2)$, con $e'_1=(1,1)$, $e'_2=(-1,1)$, tenemos $F(e'_1)=e'_1$, $F(e'_2)=-e'_2$, por lo que los autovalores son nuevamente 1 y -1, con $V_F(1)$ el espacio generado por e'_1 y $V_F(-1)$ el espacio generado por e'_2. Se obtiene entonces

$$[F]_e=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad [F]_{e'}=\begin{pmatrix}1&0\\0&-1\end{pmatrix}=S^{-1}[F]_eS\,,\quad S=\begin{pmatrix}1&-1\\1&1\end{pmatrix}$$

El pol. característico es $P(\lambda)=|[F]_e-\lambda I_2|=\lambda^2-1=(1-\lambda)(-1-\lambda)=|[F]_{e'}-\lambda I_2|$ y sus raíces ±1.

Ejemplo 4: **Operador de Proyección:** Si $P^2=P\Rightarrow$ los únicos autovalores posibles de P son 0 y 1, con $V_P(0)=N(P)$ si $N(P)\neq\{0\}$ y $V_P(1)=I(P)$ si $I(P)\neq\{0\}$.
Dem.: Hemos visto que si $P^2=P\Rightarrow P$ es un proyector sobre $I(P)$ en la dirección de $N(P)$, con $I(P)\oplus N(P)=V$ y $P(v)=v$ si $v\in I(P)$ y $P(v)=0=0v$ si $v\in N(P)$. Por lo tanto, los autovalores son: 1 (si $I(P)\neq\{0\}$, es decir, si $P\neq0$) y 0 (si $N(P)\neq\{0\}$, es decir, si $P\neq I_V$). P es pues dagonalizable, siendo una base de autovectores la formada por la unión de una base de la imagen $I(P)$ y una del núcleo $N(P)$. P es pues siempre diagonalizable.
Estas propiedades pueden también demostrarse directamente: Si $P(v)=\lambda v$, con $v\neq0\Rightarrow P^2(v)=P(P(v))=P(\lambda v)=\lambda P(v)=\lambda^2v$, pero también $P^2(v)=P(v)=\lambda v$, por lo que $\lambda^2=\lambda$, es decir $\lambda(\lambda-1)=0$. Esto implica $\lambda=0$ o $\lambda=1$. Si $\exists\, v\neq0$ tal que $P(v)=0\Rightarrow 0$ es autovalor y $V_P(0)=N(P)$. Si $\exists\, v\neq0$ tal que $P(v)=v\Rightarrow1$ es autovalor y $V_P(1)=I(P)$. En efecto, si $v\neq0$ y $P(v)=v\Rightarrow v\in I(P)$, y si $v\in I(P)$ $\Rightarrow v=P(v')$ y $P(v)=P(P(v'))=P^2(v')=P(v')=v$, por lo que $v\in V_P(1)$.

Ejemplo 6: **Potencias de operadores lineales.** Si v es autovector de F con autovalor $\lambda\Rightarrow v$ es también autovector de F^k con autovalor λ^k para cualquier $k>0$ natural, y también para cualquier $k<0$ entero si F es invertible (o sea, automorfismo), en cuyo caso $\lambda\neq0$ (pues necesariamente $N(F)=\{0\}$; ver 9.4).
Dem.: Si $k=2$, $F^2(v)=F(F(v))=F(\lambda v)=\lambda F(v)=\lambda^2v$. La demostración para $k>2$ es análoga y puede hacerse fácilmente por inducción: $F^k(v)=F(F^{k-1}(v))=F(\lambda^{k-1}v)=\lambda^{k-1}F(v)=\lambda^kv$.
Si $k=-1$, $F^k=F^{-1}$ es la inversa de F y $v=F^{-1}F(v)=F^{-1}(\lambda v)=\lambda F^{-1}(v)$, por lo que $F^{-1}(v)=\lambda^{-1}v$.
El resultado para $F^{-k}\equiv(F^{-1})^k$ y $k>1$ es entonces obvio: $F^k(v)=\lambda^kv\ \forall\ k\in\mathbb{Z}$ (véase también 9.4)

Ejemplo 7: Si λ es un autovalor de $F\Rightarrow\alpha\lambda+c$ es autovalor del operador $\alpha F+cI$, con $\alpha,I\in K$ e I el op. identidad: En efecto, si $F(v)=\lambda v$, $v\neq0\Rightarrow(\alpha F+cI)(v)=\alpha F(v)+cv=(\alpha\lambda+c)v$.

Ejemplo 8: Sea $D^2:V\to V$ el operador derivada segunda en el espacio V de funciones $f:[0,a]\to\mathbb{R}$, $a>0$, que son derivables a cualquier orden y satisfacen $f(0)=f(a)=0$ (V es de dimensión infinita).
La ecuación $D^2(f)=\lambda f$ conduce a la ecuación diferencial $f''=\lambda f$, cuya solución es $f(x)=c_1\cos(sx)+c_2\sin(sx)$ con $\lambda=-s^2$. La condición de contorno $f(0)=f(1)=0$ implica $c_1=0$ y $\sin(sa)=0$, o sea, $s=n\pi/a$, con $n>0$ natural. Por lo tanto, los autovalores son $\lambda_n=-n^2\pi^2/a^2$ y los autovectores (llamados en este caso autofunciones) $f_n(x)=c_n\sin(n\pi x/a)$, con $n=1,2,\ldots$ y $c_n\neq0$ arbitrario.
Ejemplo 9: La **ecuación de Schrödinger** estacionaria de la mecánica cuántica es

$$H|\Psi\rangle=E|\Psi\rangle$$

donde H es un operador lineal, $|\Psi\rangle$ un vector no nulo que representa el estado del sistema y E la energía del estado. Es pues una ecuación de autovalores: La energía E representa un autovalor de H y el vector $|\Psi\rangle$ el autovector correspondiente de H. Para una partícula en una dimensión, H toma la forma

$$H=-\frac{\hbar^2}{2m}D^2+V(x)$$

donde D es el operador derivada anterior, $\hbar=h/(2\pi)$, con h la constante de Planck, m la masa de la partícula y $V(x)$ el potencial, con $|\Psi\rangle\to\psi(x)$, siendo $\psi(x)$ la "función de onda" (en realidad, $\psi(x)=\langle x|\Psi\rangle$, y la forma anterior de H es su representación efectiva en la base de autoestados $|x\rangle$ del operador posición: $\langle x|H|\Psi\rangle=(-\frac{\hbar^2}{2m}D^2+V(x))\psi(x))$. H es el operador energía pues el impulso está representado por el operador $p=-i\hbar D$, por lo que el primer término de H es el operador energía cinética $P^2/(2m)$.
Si $V(x)=0$ para $x\in[0,a]$ y $V(x)=\infty$ para $x>a$ o $x<0$, la Ec. de Schrödinger se reduce al problema del ej. 8: Tenemos $H=-\frac{\hbar^2}{2m}D^2$ para $x\in[0,a]$, con $\psi(x)=0$ para $x\geq a$ o $x\leq0$, debiendo ser continua. Los autovalores son por lo tanto $E_n=\frac{\hbar^2n^2\pi^2}{2ma^2}$ y los autovectores $\psi_n(x)=c_n\sin(n\pi x/a)$.

10. Autovalores y Autovectores de Matrices

Todas las definiciones y propiedades anteriores se aplican igualmente al cálculo de autovalores y autovectores de matrices cuadradas, que pueden ser siempre consideradas como la representación de un cierto operador lineal (en un espacio vectorial de dimensión n) en una cierta base. Consideraremos en lo sucesivo $K = \mathbb{C}$.
Dada una matriz A de $n \times n$, λ es autovalor de A sii

$$|A - \lambda I_n| = 0$$

donde $|\ldots|$ denota el determinante. El polinomio

$$P(\lambda) = |A - \lambda I_n| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ldots & a_{2n} \\ & & \ldots & \\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda \end{vmatrix} \qquad (10.1)$$

se denomina *polinomio característico* de la matriz A y es de grado n en λ, por lo que posee a lo sumo n raíces distintas.
Un vector columna $X = \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix}$ de $n \times 1$ es *autovector* de A correspondiente al autovalor λ si $X \neq 0$ y

$$AX = \lambda X$$

El conjunto de autovectores X corresp. a λ puede obtenerse resolviendo el sistema homogéneo

$$(A - \lambda I_n)X = 0$$

10.1) Las matrices semejantes poseen el mismo polinomio característico y por lo tanto los mismos autovalores. Recordemos que A es semejante o similar a B si $A = S^{-1}BS$, con S de $n \times n$ no singular.
En efecto $P_A(\lambda) = |A - \lambda I_n| = |S^{-1}BS - \lambda I_n| = |S^{-1}(B - \lambda I_n)S| = |B - \lambda I_n| = P_B(\lambda)$

Los autovectores no son en general los mismos: Si $AX = \lambda X$ y $A = S^{-1}BS$, el correspondiente autovector de B es SX, como el lector podrá fácilmente demostrar. Esto puede verse también directamente a partir del cambio de base asociado.

10.2) A y A^t (t denota traspuesta) poseen el mismo polinomio característico y por lo tanto los mismos autovalores (pero no necesariamente los mismos autovectores). Como $|B| = |B^t|$ $\forall$ B de $n \times n$,

$$P_{A^t}(\lambda) = |A^t - \lambda I_n| = |(A - \lambda I_n)^t| = |A - \lambda I_n| = P_A(\lambda)$$

10.3) Si A es real $\Rightarrow$ $P(\lambda)$ es real y por lo tanto sus raíces complejas aparecerán en pares conjugados:
Si λ es una raíz compleja, $0 = [P(\lambda)]^* = P(\lambda^*)$.
Además, si X es autovector de A con autovalor λ y A es real, entonces X^* es autovector de A con autovalor λ^*: Como $AX = \lambda X \Rightarrow (AX)^* = AX^* = \lambda^* X^*$.

10.4) A es una matriz singular si y sólo si A posee al menos un autovalor nulo.
Si A es singular $\Rightarrow$ $|A| = 0$ y por lo tanto, $|A - 0I_n| = 0$, por lo que 0 es autovalor. Análogamente, si $|A - 0I_n| = 0 \Rightarrow |A| = 0$ y por lo tanto, A es singular. Esto puede también deducirse directamente de 10.5

Un operador $F : V \to V$ en un espacio de dimensión finita tendrá pues un autovalor nulo si y sólo si no es un automorfismo. Notemos también que si $\mathrm{Det}[F] = 0$, el núcleo $N(F)$ no es otra cosa que el espacio propio correspondiente el autovalor 0: $N(F) = \{v|F(v) = 0\} = \{v|F(v) = 0.v\} = V_F(0)$

10.5) El determinante de una matriz es igual al producto de *todos* sus autovalores (reales y complejos, y elevados a sus respectivas multiplicidades):

$$|A| = \lambda_1 \lambda_2 \ldots \lambda_n$$

En efecto, si $\lambda_1,\ldots,\lambda_n$ son las raíces de $P(\lambda)$, podemos escribir (utilizando la factorización en términos de raíces y notando que el término de grado n es $(-1)^n\lambda^n$)

$$P(\lambda)=|A-\lambda I_n|=(\lambda_1-\lambda)(\lambda_2-\lambda)\ldots(\lambda_n-\lambda) \qquad (10.2)$$

Por lo tanto, $|A|=P(0)=\lambda_1.\lambda_2\ldots\lambda_n$.
Esto implica que en un espacio de dimensión finita *el determinante de un operador lineal F es el producto de todos sus autovalores.*

10.6) La traza de una matriz es igual a la suma de todos sus autovalores (reales y complejos, y repetidos tantas veces como indica su multiplicidad):

$$\mathrm{Tr}A=\sum_{i=1}^{n}a_{ii}=\sum_{i=1}^{n}\lambda_i$$

A partir de la expresión (9.2) para $P(\lambda)$, vemos que el término de grado $n-1$ en λ es $(-\lambda)^{n-1}(\lambda_1+\ldots+\lambda_n)$, mientras que a partir de (9.1) vemos que el mismo es $(-\lambda)^{n-1}(a_{11}+\ldots+a_{nn})$. Como ambos son idénticos, se obtiene el resultado deseado.
Esto implica que la traza de un operador F es la suma de todos sus autovalores.

10.7 Diagonalización de Matrices

Una matriz A es diagonalizable si existe S no singular tal que

$$S^{-1}AS=A',\ \text{ con }\ A'\text{ diagonal : }\ A'=\begin{pmatrix}\lambda_1 & 0 & \ldots & 0\\ 0 & \lambda_2 & \ldots & 0\\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n\end{pmatrix}$$

es decir, si es semejante a una matriz diagonal.
En tal caso los elementos diagonales son necesariamente los autovalores de A, ya que

$$|A-\lambda I_n|=|A'-\lambda I_n|=(\lambda_1-\lambda)(\lambda_2-\lambda)\ldots(\lambda_n-\lambda)$$

y la columna i de S es autovector correspondiente al autovalor λ_i, pues $AS=SA'$:

$$S=\begin{pmatrix} & \ldots & \\ X_1 & \ldots & X_n\\ & \ldots & \end{pmatrix},\ \text{ con }\ AX_i=\lambda_iX_i,\ i=1,\ldots,n,\ \ X_i=\begin{pmatrix}x_{1i}\\ \ldots\\ x_{ni}\end{pmatrix}$$

Análogamente, si existen n vectores columna X_i LI tales que $AX_i=\lambda_iX_i,\quad i=1,\ldots,n$ entonces A es diagonalizable, con S la matriz de columnas X_i (que será invertible pues los X_i son LI y por lo tanto $|S|\neq 0$)

La dimensión del espacio propio correspondiente al autovalor λ_i es la dimensión del espacio nulo de $|A-\lambda_iI_n|$:

$$d_i=\mathrm{dim}V(\lambda_i)=\dim N[A-\lambda_iI_n]=n-R(A-\lambda_iI_n)$$

donde R denota el rango.
Notemos que si $P(\lambda)$ posee n raíces distintas $\Rightarrow$ A es diagonalizable, pues en tal caso existirán n vectores columna X_i LI tales que $AX_i=\lambda_iX_i$.
Si A es diagonalizable, resulta evidente que $|A|=|A'|=\lambda_1\ldots\lambda_n$ y que $\mathrm{Tr}A=\mathrm{Tr}A'=\lambda_1+\ldots+\lambda_n$, ya que el determinante y la traza de matrices semejantes son idénticas ($|S^{-1}AS|=|A|$, $\mathrm{Tr}S^{-1}AS=\mathrm{Tr}ASS^{-1}=\mathrm{Tr}A$).

Ejemplo 1: Consideremos la matriz

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

que corresponde a la representación en la base canónica de la reflexión respecto de la recta $y=x$. Los autovalores se obtienen de la ecuación

$$|A-\lambda I_2|=\begin{vmatrix}-\lambda & 1\\ 1 & -\lambda\end{vmatrix}=1-\lambda^2=0$$

Por lo tanto, $\lambda=\pm 1$. El autovector X_1 correspondiente a $\lambda_1=1$ se obtiene resolviendo el sistema homogéneo $(A-\lambda_1 I_2)X_1=0$, es decir,

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

que conduce a $x=y$. Los autovectores son entonces de la forma $x\binom{1}{1}$, con $x\neq 0$ y $V(1)$ es el espacio generado por $\binom{1}{1}$. Este corresponde al espacio generado por e'_1 en el ej. 3 anterior ($[e'_1]_e=\binom{1}{1}$).
Notemos que la matriz anterior tiene rango 1, por lo que $\dim V(1)=2-1=1$.
El autovector correspondiente a $\lambda_2=-1$ se obtiene resolviendo el sistema $(A-\lambda_2 I_2)X_2=0$, es decir,

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

que conduce a $x=-y$. Los autovectores son entonces de la forma $x\binom{1}{-1}$, con $x\neq 0$, y $V(-1)$ es el espacio generado por $\binom{1}{-1}$. Este corresponde al espacio generado por e'_2 en el ej. 3 anterior ($[e'_2]_e=\binom{1}{-1}$).
Una matriz de autovectores es entonces

$$S=\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{con} \quad S^{-1}=\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

y se verifica

$$S^{-1}AS=\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}=\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Notemos que se cumple $|A|=-1=\lambda_1\lambda_2$ y $\mathrm{Tr}A=0=\lambda_1+\lambda_2$.

Ejemplo 2: Consideremos

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

La ec. característica es

$$|A-\lambda I_2|=\begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix}=(1-\lambda)^2=0$$

por lo que el único autovalor es $\lambda=1$ con multiplicidad $m=2$. No obstante, la matriz

$$A-1I_2=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

posee rango 1, por lo que $\dim V(1)=2-1=1<2$. Por lo tanto, esta matriz *no es diagonalizable*, ya que no existe una base de autovectores de la misma. La ecuación

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

conduce a $y=0$, por lo que los autovectores son de la forma $x\binom{1}{0}$ y $V(1)$ es el espacio generado por $\binom{1}{0}$. No existe otro autovector LI de $\binom{1}{0}$. De todos modos, se cumple $|A|=1=1.1$ y $\mathrm{Tr}A=2=1+1$. Nótese que A es no singular ($|A|=1\neq 0$). La condición de no diagonalizable nada tiene que ver con la singularidad.

Cabe destacar, no obstante, que la matriz

$$B=\begin{pmatrix} 1 & 1 \\ \varepsilon & 1 \end{pmatrix}$$

es diagonalizable $\forall\, \varepsilon\neq 0$, ya que en tal caso la ecuación

$$|B-\lambda I_2|=(1-\lambda)^2-\varepsilon=0$$

posee siempre 2 raíces distintas: $\lambda=1\pm\sqrt{\varepsilon}$.
Esta conclusión es general: Si A no es diagonalizable podemos siempre encontrar una matriz B arbitrariamente próxima a A (es decir, cuyos elementos difieran de los de A en menos de ε, con $\varepsilon>0$ arbitrario) tal que B es diagonalizable.

Ejemplo 3: Sea $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Tenemos

$$|A - \lambda I_3| = \begin{pmatrix} 1-\lambda & 1 & 1 \\ 0 & 2-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda)(1-\lambda)$$

por lo que las raíces de $P(\lambda)$ son $\lambda_1 = 1$ y $\lambda_2 = 2$, con multiplicidades 2 y 1 respectivamente. Si $\lambda = \lambda_1$,

$$|A - 1I_3| = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

posee rango 1, por lo que $\dim V(1) = \dim N(A - 1I_3) = 3 - 1 = 2$, igual a la multiplicidad de λ_1. La matriz es por lo tanto *diagonalizable* ya que necesariamente $\dim V(2) = 1$. El sistema $(A - 1I_3)X = 0$ conduce a $y + z = 0$, es decir, $y = -z$, con z y x arbitrarios, por lo que los autovectores para $\lambda_1 = 1$ son de la forma $\begin{pmatrix} x \\ -z \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$. Para $\lambda_2 = 2$,

$$|A - 2I_3| = \begin{pmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

posee rango 2. El sistema $(A - 2I_3)X = 0$ conduce a $x = y$, con $z = 0$, por lo que los autovectores son de la forma $x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. Una matriz de autovectores es por lo tanto

$$S = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \text{con} \quad S^{-1} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

Se verifica entonces

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

Nótese que el orden de los autovalores en A' corresponde al orden de los autovectores (columnas) en S.

Ejemplo 4: (Para hacer en la práctica) Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ el operador de rotación de ángulo θ (antihorario), con $[F]_e = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$, siendo e la base canónica. Es obvio que F no puede tener autovalores reales, ya que no existe $v \neq 0$ tal que $F(v) = \lambda v$. No es por lo tanto diagonalizable para $K = \mathbb{R}$.
Sin embargo, $[F]_e$ tiene autovalores complejos $\lambda = e^{\pm i\theta} = \cos\theta \pm i \sin\theta$ y es por lo tanto *diagonalizable* si se lo considera como $F : \mathbb{C}^2 \to \mathbb{C}^2$, con $K = \mathbb{C}$. Determine el lector los autovectores y compruebe que existe S tal que $S^{-1}[F]_e S$ es diagonal!

Ejemplo 5: Consideremos el operador de proyección ortogonal P sobre el plano $x + y = 0$ en $\mathbb{R}^3$ (o sea, proyección sobre este plano en la dirección del eje z). Si $e' = (e'_1, e'_2, e'_3)$, con $e'_1 = (e_1 + e_2)/\sqrt{2}$, $e'_2 = (-e_1 + e_2)/\sqrt{2}$, $e'_3 = e_3$, con $e = (e_1, e_2, e_3)$ la base canónica, entonces $P(e'_1) = 0$, $P(e'_2) = e'_2$, $P(e'_3) = e'_3$ y por lo tanto, $[P]_{e'} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Esta es pues una base de autovectores de P.
En la base canónica, obtenemos en cambio

$$[P]_e = S[P]_{e'} S^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \text{con} \quad S = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} / \sqrt{2}$$

y $S^{-1}=S^t$. En esta base $[P]_e$ no es diagonal, aunque sigue cumpliendo que $[P]_e^2=[P]_e$. Se deja como ejercicio verificar explícitamente que los autovalores de la matriz $[P]_e$ son 0 y 1, y que una base de autovectores es precisamente e' (aunque por su puesto no es la única), de modo que S es una matriz diagonalizante de $[P]_e$, que verifica $S^{-1}[P]_eS=[P]_{e'}$, con $[P]_{e'}$ diagonal.

Ejemplo 6: **Autovalores de una matrix general de** 2×2. Si $A=\begin{pmatrix} a & b\\ c & d\end{pmatrix}=\frac{a+d}{2}I_2+\begin{pmatrix}\frac{a-d}{2} & b\\ c & -\frac{a-d}{2}\end{pmatrix}$, obtenemos fácilmente, a partir de $|A-\lambda I_2|=0$, que los autovalores son

$$\lambda_\pm=\frac{a+d}{2}\pm\sqrt{(\frac{a-d}{2})^2+bc}=\frac{1}{2}\mathrm{Tr}[A]\pm\sqrt{(\frac{\mathrm{Tr}[A]}{2})^2-\mathrm{Det}[A]}$$

donde $\mathrm{Tr}[A]=a+d$ es la traza de A y $\mathrm{Det}[A]=ad-bc$ su determinante. La última expresión puede obtenerse directamente de resolver el sistema $\left\{\begin{array}{rcl}\lambda_++\lambda_- & = & \mathrm{Tr}[A]\\ \lambda_+\lambda_- & = & \mathrm{Det}[A]\end{array}\right.$
Los dos autovalores quedan pues completamente determinados por la traza y el determinante.

Ejemplo 7: **Corrección de primer orden en los autovalores**.
Consideremos una matriz $B=A+\delta A$ de $n\times n$, siendo A diagonalizable y $\delta A=\varepsilon M$ una perturbación (ε es un parámetro suficientemente pequeño y M una matriz de $n\times n$ arbitraria).
Sea $S=(X_1,\ldots,X_n)$ una matriz de autovectores de A, tal que $S^{-1}AS=A'$ con A' diagonal ($A'_{ij}=\lambda_i\delta_{ij}$, con $AX_i=\lambda_iX_i$). Definiendo $\delta A'=S^{-1}(\delta A)S=\varepsilon S^{-1}MS$ (perturbación en la base en que A es diagonal) y notando que $|B-\lambda I|=|S^{-1}BS-\lambda I|$, obtenemos

$$|B-\lambda I|=|D+\delta A'-\lambda I|=\left|\begin{array}{cccc}\lambda_1+\delta A'_{11}-\lambda & \delta A'_{12} & \ldots & \delta A'_{1n}\\ \delta A'_{21} & \lambda_2+\delta A'_{22}-\lambda & \ldots & \delta A'_{2n}\\ & & \ldots & \\ \delta A'_{n1} & \delta A'_{n2} & \ldots & \lambda_n+\delta A'_{nn}-\lambda\end{array}\right|$$

Consideremos primero el caso en que los autovalores λ_i de A son todos distintos. Para $\lambda=\lambda_i+\delta\lambda_i$ con $\delta\lambda_i$ una corrección de orden ε al autovalor λ_i, obtenemos entonces

$$|B-(\lambda_i+\delta\lambda_i)I|=(\delta A'_{ii}-\delta\lambda_i)\prod_{j\neq i}(\lambda_j-\lambda_i)+O(\varepsilon^2)$$

donde el primer término es el de mayor orden ($O(\varepsilon)$), y los restantes de orden $O(\varepsilon^2)$ o mayor. Por lo tanto, la ec. $|B-\lambda I|=0$ conduce a

$$\delta\lambda_i=\delta A'_{ii}+O(\varepsilon^2),\quad \text{con }\ \delta A'_{ii}=(S^{-1}\delta AS)_{ii}=\varepsilon\sum_{j,k}S^{-1}_{ij}M_{jk}S_{ki}$$

es decir, los $\delta\lambda_i$ son los términos diagonales de δA en la base en que A es diagonal. Como $\sum_j S^{-1}_{ij}S_{ji}=1$, si la columna i de S (el autovector X_i) se multiplica por α, la fila de i de S^{-1} se multiplica por $1/\alpha$, para mantener la igualdad anterior. Por lo tanto, la corrección $\delta A'_{ii}$ es, como debe ser, independiente de la base elegida del espacio propio, es decir de la elección del autovector $X_i\neq 0$ en el espacio propio.

En el caso general, si el espacio propio asociado a un autovalor λ_i tiene dimensión d_i (se dice entonces que tiene degeneración d_i), la corrección a λ_i son los *autovalores* de δA en el espacio propio asociado a λ_i, pues $|A-\lambda I|=|(\delta A')_i-\delta\lambda_iI_{d_i}|\prod_{\lambda_j\neq\lambda_i}(\lambda_j-\lambda_i)^{m_j}+O(\varepsilon^{d_i+1})$, con $\delta A'_i$ la matriz $\delta A'$ restringida al espacio propio asociado a λ_i y m_j la multiplicidad (algebraica) del autovalor λ_j. Se deben pues obtener los autovalores de $\delta A'_i$ (matriz de $d_i\times d_i$). El nivel degenerado λ_i se desdobla normalmente en varios niveles. Se dice entonces que se rompe la degeneración.
Es importante que A sea diagonalizable. De lo contrario, el ej. 2 anterior muestra que en el caso no-diagonalizable, la corrección puede ser por ej. de orden $\sqrt{\varepsilon}$.

10.8 Evaluación de Potencias y Series de Matrices

La diagonalización es muy conveniente para evaluar potencias y series de matrices (de $n \times n$) u operadores. En primer lugar, si

$$A = SA'S^{-1} \qquad (10.3)$$

(A semejante a A') se cumple, para k natural,

$$A^k = SA'^k S^{-1}$$

ya que $A^2 = (SA'S^{-1})(SA'S^{-1}) = SA'^2S^{-1}$ y en general (por inducción) $A^k = AA^{k-1} = SA'S^{-1}SA'^{k-1}S^{-1} = SA'^kS^{-1}$. Análogamente, para funciones definidas por series de potencias $f(u) = \sum_{k=0}^{\infty} a_k u^k$ convergentes $\forall\ u \in \mathbb{C}$,

$$f(A) = \sum_{k=0}^{\infty} a_k A^k = \sum_{k=0}^{\infty} a_k SA'^k S^{-1} = S[\sum_{k=0}^{\infty} a_k A'^k]S^{-1} = Sf(A')S^{-1}$$

Notemos que $f(A)$ está bien definido pues $|(A^k)_{ij}| \leq (mn)^k/n$, donde m el mayor elemento de la matriz ($|A_{ij}| \leq m\ \forall\ i,j$) y n la dimensión. Esto implica que la serie matricial converge absolutamente si la serie converge absolutamente $\forall\ u$ ($|(f(A))_{ij}| \leq f(mn)/n$). En particular,

$$\exp[At] = \sum_{k=0}^{\infty} \frac{(At)^k}{k!} = S\exp[A't]S^{-1}$$

Finalmente, si A es invertible (en cuyo caso A' es también invertible, como el lector debe reconocer inmediatamente) se cumple $A^{-1} = SA'^{-1}S^{-1}$ y en general

$$A^{-k} = SA'^{-k}S^{-1}$$

Si A es *diagonalizable*, $S^{-1}AS = A'$ con A' diagonal. Por lo tanto $A = SA'S^{-1}$. Podemos entonces utilizar las expresiones anteriores con A' diagonal y S una matriz de autovectores, en cuyo caso la evaluación resulta inmediata pues

$$A' = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \quad \Rightarrow \quad (A')^k = \begin{pmatrix} \lambda_1^k & 0 & \ldots & 0 \\ 0 & \lambda_2^k & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n^k \end{pmatrix}$$

para k natural. Esto implica

$$f(A') = \sum_{k=0}^{\infty} a_k (A')^k = \begin{pmatrix} f(\lambda_1) & 0 & \ldots & 0 \\ 0 & f(\lambda_2) & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & f(\lambda_n) \end{pmatrix}$$

En particular,

$$\exp[A't] = \begin{pmatrix} e^{\lambda_1 t} & 0 & \ldots & 0 \\ 0 & e^{\lambda_2 t} & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & e^{\lambda_n t} \end{pmatrix}$$

Además, si A es invertible, sus autovalores son todos no nulos y es fácil ver que

$$(A')^{-1} = \begin{pmatrix} \lambda_1^{-1} & 0 & \ldots & 0 \\ 0 & \lambda_2^{-1} & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n^{-1} \end{pmatrix}$$

y por lo tanto

$$(A')^{-k} = \begin{pmatrix} \lambda_1^{-k} & 0 & \ldots & 0 \\ 0 & \lambda_2^{-k} & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n^{-k} \end{pmatrix}$$

Por ejemplo, en el caso del ej. 1 anterior se obtiene

$$\exp[At] = \exp[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} t] = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix}$$

y en el ej. 3 anterior,

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

$$\exp[At] = S\begin{pmatrix} e^t & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^{2t} \end{pmatrix} S^{-1} = \begin{pmatrix} e^t & e^{2t}-e^t & e^{2t}-e^t \\ 0 & e^{2t} & e^{2t}-e^t \\ 0 & 0 & e^t \end{pmatrix}$$

Ejemplo: **Sucesión de Fibonacci**. Está definida por la relación recursiva *lineal*

$$a_{n+1} = a_n + a_{n-1}, \quad n \geq 1$$

con $a_0 = 0$, $a_1 = 1$.
La expresión explícita de a_n puede obtenerse fácilmente planteando el problema en forma matricial. Resolveremos en realidad el problema para valores iniciales generales a_0, a_1. Tenemos, para $n \geq 1$,

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}$$

Por lo tanto, para $n \geq 1$ y definiendo $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$,

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = A^n \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}$$

La evaluación de A^n puede efectuarse mediante su diagonalización. Los autovalores de A son los números aureos

$$\lambda_\pm = \frac{1 \pm \sqrt{5}}{2}$$

que satisfacen $\lambda^2 = \lambda + 1$, con autovectores $v_\pm \propto \binom{\lambda_\pm}{1}$. Podemos entonces escribir $A = SA'S^{-1}$ con $S = \binom{\lambda_+ \, \lambda_-}{1 \;\; 1}$ y $A' = \binom{\lambda_+ \; 0}{0 \; \lambda_-}$. Por lo tanto, $A^n = S(A')^n S^{-1}$ y se obtiene finalmente (se dejan las cuentas para el lector)

$$a_n = [(\lambda_+^n - \lambda_-^n)a_1 - (\lambda_+^n \lambda_- - \lambda_-^n \lambda_+)a_0]/\sqrt{5}$$

En el caso usual de Fibonacci, $a_0 = 0$, $a_1 = 1$ y $a_n = (\lambda_+^n - \lambda_-^n)/\sqrt{5}$. Como $\lambda_+ = 1.618$, $\lambda_- = -0.618$, el término dominante para n grande es el proporcional a λ_+^n.

Un tratamiento equivalente consiste en expresar el vector inicial $\binom{a_1}{a_0}$ como combinación lineal de los autovectores de A: $A^n\binom{a_1}{a_0} = A^n[c_+\binom{\lambda_+}{1} + c_-\binom{\lambda_-}{1}] = c_+\lambda_+^n\binom{\lambda_+}{1} + c_-\lambda_-^n\binom{\lambda_-}{1}$, de donde $a_n = \lambda_+^n c_+ + \lambda_-^n c_-$. Como $\binom{c_+}{c_-} = S^{-1}\binom{a_1}{a_0}$, se obtiene $c_+ = a_1 - \lambda_- a_0$, $c_- = -a_1 + \lambda_+ a_0$, obteniéndose el resultado anterior.

El mismo método se puede aplicar para toda sucesión definida por una relación recursiva fija *lineal*:

$$a_{n+1} = \alpha_0 a_n + \alpha_1 a_{n-1} + \ldots + \alpha_k a_{n-k}$$

para $n \geq k$, con $a_0, \ldots, a_k$ dados, que conduce a

$$\begin{pmatrix} a_{n+1} \\ a_n \\ \ldots \\ a_{n-k+1} \end{pmatrix} = \begin{pmatrix} \alpha_0 & \alpha_1 & \ldots & \alpha_k \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ & & \ldots & \\ 0 & \ldots & 1 & 0 \end{pmatrix}\begin{pmatrix} a_n \\ a_{n-1} \\ \ldots \\ a_{n-k} \end{pmatrix} = \begin{pmatrix} \alpha_0 & \alpha_1 & \ldots & \alpha_k \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ & & \ldots & \\ 0 & \ldots & 1 & 0 \end{pmatrix}^{n-k+1}\begin{pmatrix} a_k \\ a_{k-1} \\ \ldots \\ a_0 \end{pmatrix}$$

La sucesión geométrica elemental $a_n = \alpha_0^n a_0$ corresponde al caso $k = 0$ ($a_{n+1} = \alpha_0 a_n$ si $n \geq 0$).

10.9 Desacoplamiento de Sistemas de Ecuaciones Diferenciales Lineales

Como otra aplicación, consideremos por ejemplo el sistema de ecuaciones diferenciales ordinarias lineales de primer orden

$$\frac{dX}{dt} = AX$$

con X de $n \times 1$ y A de $n \times n$, con elementos *constantes* (o sea, independientes del tiempo). Suponiendo A diagonalizable, tenemos $A = SA'S^{-1}$, con A' diagonal y S la matriz de autovectores. Por lo tanto $dX/dt = SA'S^{-1}X$, lo que implica

$$dX'/dt = A'X', \quad X' = S^{-1}X,$$

Como A' es diagonal, el sistema en las variables X' está *desacoplado*, y es de fácil resolución. Tenemos, para las componentes x'_i de X', las ecuaciones desacopladas

$$dx'_i/dt = \lambda_i x'_i, \quad i = 1, \ldots, n$$

donde λ_i son los autovalores de A, cuya solución es $x'_i = c_i e^{\lambda_i t}$. Finalmente, se obtiene

$$X(t) = SX'(t) = \sum_{i=1}^{n} c_i V_i e^{\lambda_i t},$$

donde V_i denota los autovectores de A (las columnas de S). Esto constituye la solución *general* del sistema de primer orden, conteniendo n constantes arbitrarias c_i que pueden determinarse a partir de las condiciones iniciales $x_i(0)$.

El procedimiento usualmente utilizado en Física e Ingeniería para llegar a esta solución es plantear una solución del tipo $X(t) = Ve^{\lambda t}$ con V constante. La ec. $dX/dt = AX$ implica entonces $\lambda V = AV$, por lo que V debe ser autovector de A con autovalor λ. La solución general se obtiene luego como combinación lineal arbitraria de estas soluciones particulares. Este procedimiento es en realidad correcto para encontrar la solución general sólo en el caso de matrices A diagonalizables.

Nótese también que el mismo método puede utilizarse para resolver sistemas análogos de segundo orden

$$\frac{d^2X}{dt^2} = AX$$

Sólo es necesario reemplazar $c_i e^{\lambda_i t}$ por $c_i^+ e^{\sqrt{\lambda_i} t} + c_i^- e^{-\sqrt{\lambda_i} t}$ en la solución general anterior.

Ejemplo 1: Consideremos el sistema de tres ecuaciones diferenciales acopladas de primer orden,

$$dx/dt = x + y + z, \quad dy/dt = 2y + z, \quad dz/dt = z$$

donde x, y, z son funciones de t, el cual puede escribirse en forma matricial como

$$\frac{d}{dt}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

o sea, $dv/dt = Av$, siendo A la matriz del ej. 3 anterior y $v = (x, y, z)^t$. Por lo tanto, utilizando las matrices S y S^{-1} de dicho ejemplo,

$$\frac{d}{dt}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = S\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}S^{-1}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

y mult. a izq. por S^{-1}, se llega a

$$\frac{d}{dt}\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \quad \text{con} \quad \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = S^{-1}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y - z \\ -z \\ y + z \end{pmatrix}$$

el cual es un sistema de tres ecuaciones dif. lineales *desacopladas*:

$$dx'/dt = x', \quad dy'/dt = y', \quad dz'/dt = 2z'$$

que es equivalente al original. La solución del sistema desacoplado es muy fácil de obtener:

$$x' = c_1 e^t,\ y' = c_2 e^t,\ z' = c_3 e^{2t}$$

Finalmente

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = S \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x'+z' \\ -y'+z' \\ y' \end{pmatrix} = c_1 e^t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 e^t \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + c_3 e^{2t} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 e^t + c_3 e^{2t} \\ -c_2 e^t + c_3 e^{2t} \\ c_2 e^t \end{pmatrix}$$

Ejemplo 2: Sistema de dos resortes acoplados (recordar dibujo). Ecuación de movimiento:

$$\frac{d^2}{dt^2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -\frac{1}{m} \begin{pmatrix} k_1+k_2 & -k_2 \\ -k_2 & k_1+k_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Resuelto en clase. Detalles a cargo del lector. Sólo recordamos que las frecuencias propias $\omega_i = \sqrt{\lambda_i}$ (con λ_i los autovalores de la matriz $\left(\begin{smallmatrix} k_1+k_2 & -k2 \\ -k2 & k_1+k_2 \end{smallmatrix}\right)/m$ son $\omega_1 = \sqrt{(k_1+2k_2)/m}$, $\omega_2 = \sqrt{k_1/m}$, con $V_1 \propto \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$, $V_2 \propto \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$.

11. Matrices hermíticas y reales simétricas

Un caso especial muy importante para la física es aquel de matrices de $n \times n$ *hermíticas* (o hermitianas o autoadjuntas), que son las que satisfacen $A^\dagger = A$, donde $A^\dagger \equiv A^{t*}$ denota la matriz traspuesta y conjugada (matriz adjunta):

$$A = A^\dagger \;\Leftrightarrow\; A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{12}^* & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{1n}^* & a_{2n}^* & \ldots & a_{nn} \end{pmatrix}$$

con a_{ii} real para $i = 1, \ldots, n$. Si todos los elementos de A son *reales* $\Rightarrow$ $A^\dagger = A^t$ y la condición de A hermítica equivale a A *simétrica* ($A^t = A$).

11.1) *Teorema:* Si A de $n \times n$ es una matriz hermítica sus autovalores λ_i son todos reales y los autovectores X_i correspondientes a autovalores distintos son ortogonales respecto del producto escalar usual para vectores complejos: Si $AX_i = \lambda_i X_i$, $AX_j = \lambda_j X_j \Rightarrow X_i^\dagger X_j = 0$ si $\lambda_i \neq \lambda_j$, donde

$$X_i^\dagger X_j = (x_{1i}^* \ldots x_{ni}^*) \begin{pmatrix} x_{1j} \\ \ldots \\ x_{nj} \end{pmatrix} = x_{1i}^* x_{1j} + \ldots x_{ni}^* x_{nj}$$

Demostración: Sea X_i de $n \times 1$ autovector de A con autovalor λ_i (por lo tanto $X_i \neq 0$). Multiplicando la igualdad $AX_i = \lambda_i X_i$ a izquierda por $X_i^\dagger = X_i^{t*}$ se obtiene

$$X_i^\dagger A X_i = \lambda_i X_i^\dagger X_i \qquad (11.1)$$

con

$$X_i^\dagger X_i = (x_{1i}^* \ldots x_{ni}^*) \begin{pmatrix} x_{1i} \\ \ldots \\ x_{ni} \end{pmatrix} = x_{1i}^* x_{1i} + \ldots x_{ni}^* x_{ni} = |x_{1i}|^2 + \ldots + |x_{ni}|^2 > 0$$

Trasponiendo y conjugando la igualdad (11.1) se obtiene, notando que $(AB)^\dagger = B^\dagger A^\dagger$,

$$X_i^\dagger A^\dagger X_i = \lambda_i^* X_i^\dagger X_i \qquad (11.2)$$

Pero como $A^\dagger = A$, esto implica, comparando con (11.1), que $\lambda_i X_i^\dagger X_i = \lambda_i^* X_i^\dagger X_i$, o sea,

$$0 = (\lambda_i - \lambda_i^*) X_i^\dagger X_i$$

Como $X_i^\dagger X_i > 0 \Rightarrow \lambda_i = \lambda_i^*$, es decir, λ_i real.
Del mismo modo, si $AX_i = \lambda_i X_i$, $AX_j = \lambda_j X_j$, multiplicando a izquierda la primer ecuación por $X_j^\dagger$ y la segunda por $X_i^\dagger$ se obtiene

$$X_j^\dagger A X_i = \lambda_i X_j^\dagger X_i, \quad X_i^\dagger A X_j = \lambda_j X_i^\dagger X_j$$

Trasponiendo y conjugando la pimera de estas ecuaciones se obtiene $X_i^\dagger A^\dagger X_j = \lambda_i^* X_i^\dagger X_j$, es decir, $X_i^\dagger A X_j = \lambda_i X_i^\dagger X_j$ pues $A^\dagger = A$ y $\lambda_i = \lambda_i^*$ (ya demostrado). Por lo tanto, $\lambda_i X_i^\dagger X_j = \lambda_j X_i^\dagger X_j$, o sea,

$$0 = (\lambda_i - \lambda_j) X_i^\dagger X_j$$

por lo que $X_i^\dagger X_j = 0$ si $\lambda_i \neq \lambda_j$.

Puede demostrarse (lo haremos más adelante) que *toda* matriz hermítica A es *diagonalizable*, y que siempre existen n autovectores X_i LI y además ortogonales que satisfacen $AX_i = \lambda_i X_i$, con $X_i^\dagger X_j = 0$ si $i \neq j$ (aún cuando $\lambda_i = \lambda_j$).

En tal caso, eligiendo autovectores normalizados tales que $X_i^\dagger X_i = 1$ para $i = 1, \ldots, n$, se obtiene

$$X_i^\dagger X_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Por lo tanto la matriz de autovectores $S = (X_1 \ldots X_n)$ satisface

$$S^\dagger S = I_n$$

pues $(S^\dagger S)_{ij} = X_i^\dagger X_j = \delta_{ij}$. La inversa de S es pues directamente la matriz adjunta: $S^{-1} = S^\dagger$.

En resumen, si $A = A^\dagger$, $\exists$ S tal que $S^{-1} = S^\dagger$ y $S^\dagger A S = A'$, con A' diagonal y real: $A_{ij} = \lambda_i \delta_{ij}$

Matrices antihermíticas: Si $A^\dagger = -A$, se dice que A es antihermítica. En tal caso, $B = -iA$ resulta hermítica (pues $B^\dagger = iA^\dagger = -iA = B$), lo que implica, como $A = iB$, que A será también diagonalizable, con autovectores ortogonales si corresponden a autovalores distintos, pero con autovalores *imaginarios* en lugar de reales: Si $BX_i = \lambda_i X_i \Rightarrow AX_i = (i\lambda_i) X_i$.

Matrices reales simétricas: Para el caso particular de *matrices reales*, los resultados anteriores implican que *los autovalores de matrices reales simétricas ($A^\dagger = A^t = A$) son todos reales.* Los autovectores pueden pués elegirse reales, y por lo tanto, serán ortogonales respecto del producto escalar usual: Si $AX_i = \lambda_i X_i$ y $AX_j = \lambda_j X_j \Rightarrow$

$$X_i^t X_j = x_{1i} x_{1j} + \ldots + x_{ni} x_{nj} = 0 \text{ si } \lambda_i \neq \lambda_j$$

En tal caso, eligiendo autovectores normalizados (tales que $X_i^t X_i = 1$) la inversa de la matriz $S = (X_1, \ldots, X_n)$ será directamente la traspuesta:

$$S^t S = I_n \qquad (A = A^t \text{ real}, \ X_i \text{ real para } i = 1, \ldots, n)$$

En resumen, si $A = A^t$, con A real, $\exists$ S real tal que $S^{-1} = S^t$ y $S^t A S = A'$, con A' diagonal y real: $A'_{ij} = \lambda_i \delta_{ij}$. Más aun, puede elegirse siempre S tal que det$S = +1$ (Si $S^{-1} = S^t \Rightarrow$ det$S = \pm 1$) en cuyo caso S corresponde a una rotación, como veremos más adelante

Matrices reales antisimétricas: Si A es real y $A^t = -A$, nuevamente $B = -iA$ resulta hermítica ($B^\dagger = iA^t = -iA = B$) y por lo tanto, $A = iB$ será diagonalizable en el cuerpo de los complejos, con autovalores *imaginarios* y autovectores ortogonales si corresponden a autovalores distintos.

Ejemplo 1: Sea

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

A es una matriz real simétrica si v es real. Tenemos

$$|A - \lambda I_2| = (1 - \lambda)^2 - v^2$$

por lo que los autovalores son $\lambda = 1 \pm v$, reales. Para $\lambda_1 = 1 + v$, puede verse fácilmente que el autovector es de la forma $X_1 = x_1 \binom{1}{1}$, mientras que para $\lambda_2 = 1 - v$, es de la forma $X_2 = x_2 \binom{-1}{1}$. Por lo tanto, podemos elegir $x_1 = x_2 = 1/\sqrt{2}$, para que $X_1^t X_1 = X_2^t X_2 = 1$. Se verifica además $X_1^t X_2 = \frac{1}{2}(1, 1)\binom{-1}{1} = 0$. Por lo tanto

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad S^{-1} = S^t = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

con
$$S^t A S = \begin{pmatrix} 1+v & 0 \\ 0 & 1-v \end{pmatrix}$$

Ejemplo 2: Sea
$$A = \begin{pmatrix} 1 & iv \\ -iv & 1 \end{pmatrix}$$

con v real. A es una matriz hermítica ($A^\dagger = A$). Tenemos
$$|A - \lambda I_2| = (1-\lambda)^2 - (iv)(-iv) = (1-\lambda^2) - v^2$$

por lo que los autovalores son nuevamente $\lambda = 1 \pm v$, reales. Para $\lambda_1 = 1+v$, puede verse fácilmente que el autovector es de la forma $X_1 = x_1\binom{i}{1}$, mientras que para $\lambda_2 = 1-v$, es de la forma $X_2 = x_2\binom{-i}{1}$. Por lo tanto, podemos elegir $x_1 = x_2 = 1/\sqrt{2}$, para que $X_1^\dagger X_1 = X_2^\dagger X_2 = 1$. Se verifica además $X_1^\dagger X_2 = \frac{1}{2}(-i,1)\binom{-i}{1} = (-1+1)/2 = 0$. Por lo tanto
$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}, \quad S^{-1} = S^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix}$$

con
$$S^\dagger A S = \begin{pmatrix} 1+v & 0 \\ 0 & 1-v \end{pmatrix}$$

Ejemplo 3: Consideremos la ecuación
$$ax^2 + 2bxy + cy^2 = d$$

con coeficientes y variables reales. Podemos escribirla en forma matricial como
$$X^t A X = d, \quad X = \binom{x}{y}, \quad X^t = (x, y), \quad A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

La matriz A es real y simétrica, siendo por lo tanto siempre diagonalizable. Existe entonces una matriz ortogonal de autovectores S ($S^{-1} = S^t$), con $\mathrm{Det}\, S = 1$, tal que $S^{-1} A S = S^t A S = A'$, con A' diagonal: $A' = \left(\begin{smallmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{smallmatrix}\right)$, siendo $\lambda_\pm$ los autovalores de A (las raíces de $(a-\lambda)(b-\lambda) - b^2 = 0$). En tal caso, si $X = SX'$, tenemos
$$X^t A X = X'^t S^t A S X' = X'^t A' X' = \lambda_+ x'^2 + \lambda_- y'^2 = d$$

Si $d > 0$, vemos entonces que la gráfica de la ecuación en las variables x', y' será una elipse si $\lambda_\pm$ son ambos mayores que 0, y una hipérbola si $\lambda_+\lambda_- < 0$, con ejes principales x', y' en ambos casos. Como la transformación corresponde a una rotación (eligiendo el orden de autovectores tal que $\mathrm{Det}S = +1$), la ecuación original corresponderá si $|A| \neq 0$ a una elipse o hipérbola con ejes principales rotados (como consecuencia del término cruzado $2bxy$). El ángulo de inclinación θ entre los ejes x' y x puede obtenerse a partir de la matriz S escribiéndola en la forma $S = [I]_e^{e'} = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$. Para más detalles ver ejemplo resuelto en clase o en práctica.

Ejemplo 4: **Tensor de Inercia**. El tensor de inercia de un cuerpo rígido respecto de un origen O es
$$I_O = \sum_\nu m_\nu \begin{pmatrix} r_\nu^2 - x_\nu^2 & -x_\nu y_\nu & -x_\nu z_\nu \\ -y_\nu x_\nu & r_\nu^2 - y_\nu^2 & -y_\nu z_\nu \\ -z_\nu x_\nu & -z_\nu y_\nu & r_\nu^2 - z_\nu^2 \end{pmatrix} = \sum_\nu m_\nu (X_\nu^t X_\nu I_3 - X_\nu X_\nu^t)$$

donde $X_\nu^t = (x_\nu, y_\nu, z_\nu)$ y $r_\nu^2 = X_\nu^t X_\nu = x_\nu^2 + y_\nu^2 + z_\nu^2$. I_O queda pues representado por una matriz real *simétrica*. Frente a una rotación del sistema de coordenadas, $X_\nu = SX'_\nu$, con $\mathrm{Det}S = 1$, $S^t S = I_3$, se obtiene
$$I_O = \sum_\nu m_\nu (X'^t_\nu S^t S X'_\nu I_3 - SX'_\nu X'^t_\nu S^t) = S[\sum_\nu m_\nu (X'^t_\nu X'_\nu I_3 - X'_\nu X'^t_\nu)]S^t = S I'_O S^t$$

o sea, $I'_O = S^t I_O S$ con I'_O el tensor de inercia en el sistema rotado. Como I_O es real simétrica, existirá una matriz ortogonal de rotación S (matriz de autovectores normalizados y ordenados tal que $S^t S = I_3$ y $|S| = 1$) tal que I'_O sea *diagonal.* Esta matriz determinará los 3 *ejes principales de inercia*, y los autovalores de I_O serán los *momentos principales de inercia.* Si el vector velocidad angular Ω coincide con alguna de estas direcciones, el vector momento angular (dado en general por $L_O = I_O \Omega$) será proporcional a Ω.

Ejemplo 5: **Sistema general de n resortes acoplados**. El movimiento de tal conjunto esta descripto por un sistema de ecuaciones de segundo orden del tipo (recordar discusión de clase)

$$\frac{m_i d^2 x_i}{dt^2} = -\sum_{j=1}^{n} k_{ij} x_j, \quad i = 1, \ldots, n$$

donde x_i es la posición de la partícula i (medida desde la posición de equilibrio), $m_i > 0$ su masa y $k_{ij} = k_{ji}$. Podemos reescribir tal sistema en forma matricial como

$$M \frac{d^2 X}{dt^2} = -KX$$

donde M es una matriz diagonal de elementos m_i ($M_{ij} = m_i \delta_{ij}$), $X = (x_1, \ldots, x_n)^t$ y K la matriz de elementos k_{ij}. Definiendo la matriz diagonal $M^{1/2}$ de elementos $\sqrt{m_i}$ ($(M^{1/2})_{ij} = \sqrt{m_i}\delta_{ij}$) podemos reescribir tal sistema como $M^{1/2} M^{1/2} \frac{d^2X}{dt^2} = -KX$, y por lo tanto, multiplicando a izquierda por $M^{-1/2} = (M^{1/2})^{-1}$ (matriz diagonal de elementos $(M^{-1/2})_{ij} = \frac{1}{\sqrt{m_i}}\delta_{ij}$), como

$$\frac{d^2 Y}{dt^2} = -\tilde{K} Y, \ \text{ donde } \ \tilde{K} = M^{-1/2} K M^{-1/2}, \ \ Y = M^{1/2} X$$

(de forma que $y_i = \sqrt{m_i} x_i$). La ventaja esta forma matricial es que la matriz $\tilde{K}$ es real **simétrica** ($\tilde{K}^t = \tilde{K}$) y por lo tanto siempre **diagonalizable**. Existe entonces una matriz ortogonal S tal que $S^t \tilde{K} S = \tilde{K}'$, con K' **diagonal**, de elementos $\tilde{K}'_{ij} = \lambda_i \delta_{ij}$, siendo λ_i los autovalores de $\tilde{K}$. Por lo tanto, escribiendo $\tilde{K} = S \tilde{K}' S^t$, el sistema original resulta equivalente a

$$\frac{d^2 Y'}{dt^2} = -\tilde{K}' Y', \quad \text{donde } \ Y' = S^t Y$$

Esto representa, dado que $\tilde{K}'$ es diagonal, un sistema de n resortes **desacoplados**:

$$\frac{d^2 y'_i}{dt^2} = -\lambda_i y'_i, \ \ i = 1, \ldots, n$$

La solución general de c/u de estas ecuaciones es, para $\lambda_i \neq 0$, $y'_i(t) = A e^{i\omega_i t} + B e^{i\omega_i t} = C\cos(\omega_i t + \phi)$, donde $\omega_i = \sqrt{\lambda_i}$ **son las frecuencias propias de vibración del sistema**. Las variables $y'_i = \sum_{j=1}^{n} S_{ji}\sqrt{m_j} x_j$ (o sea, $y'_i = (S^t Y)_i$) se denominan **modos normales de vibración**. Notemos que las frecuencias propias son las raíces de los autovalores de la matriz $M^{-1/2} K M^{-1/2}$, los cuales, en virtud de la propiedad 12.1 siguiente, coinciden con los de la matriz $M^{-1}K$. Por lo tanto, la conocida fórmula $\omega = \sqrt{k/m}$ para la frecuencia angular de un oscilador armónico se generaliza a $\omega_i = \sqrt{(M^{-1}K)_i}$, donde $(M^{-1}K)_i$ denota aquí el iésimo autovalor de la matriz $M^{-1}K$. Puede demostrarse que si la matriz K es definida positiva (definición que veremos luego y que corresponde a un sistema estable) entonces todos los autovalores de $\tilde{K}$ son positivos.

Ejemplo 6: **Problema generalizado de autovalores:** La ecuación

$$AX = \lambda BX$$

donde A y B son matrices de $n \times n$, B es no singular y $X \neq 0$ es un vector columna, define un problema generalizado de autovalores. Es obviamente equivalente al problema estándar $B^{-1}AX = \lambda X$, es decir, a la obtención de los autovalores y autovectores de la matriz $B^{-1}A$. Los autovalores pueden pues obtenerse como $|B^{-1}A - \lambda I| = 0$, equivalente a

$$|A - \lambda B| = 0$$

y los espacios propios pueden obtenerse como $\mathrm{Nu}(B^{-1}A - \lambda I)$, equivalente a $\mathrm{Nu}(A - \lambda B)$.

Si $A^\dagger = A$ y $B^\dagger = B$, $B^{-1}A$ no es general hermítica ($(B^{-1}A)^\dagger = AB^{-1}$). Si B es definida positiva (es decir, todos sus autovalores λ_i^B son positivos), es posible reescribir el problema generalizado como un problema de autovalores *hermítico y estándar*, mediante el mismo método utilizado en el ej. 5:
Si $AX = \lambda BX \Rightarrow B^{-1/2} A B^{-1/2} B^{1/2} X = \lambda B^{1/2} X$, por lo que el problema se reduce al problema estándar

$$\tilde{A}\tilde{X} = \lambda \tilde{X}$$

con $\tilde{A} = B^{-1/2} A B^{-1/2}$ una matriz *hermítica* ($\tilde{A}^\dagger = \tilde{A}$) y $\tilde{X} = B^{1/2}\tilde{X}$. Aqui, si $B = S_B B' S_B^\dagger$, con B' diagonal ($B'_{ij} = \lambda_i^B \delta_{ij}$, $\lambda_i^B > 0$), $B^{1/2} = S_B B'^{1/2} S_B^\dagger$, con $(B'^{1/2})_{ij} = \sqrt{\lambda_i^B}\delta_{ij}$, siendo $B^{-1/2}$ su inversa.

Por lo tanto, los autovalores λ serán *reales*, y pueden obtenerse de $|\tilde{A}-\lambda I|=0$, equivalente a $|A-\lambda B|=0$, mientras que los autovectores $\tilde{X}$ correspondientes a autovalores distintos serán ortogonales: $\tilde{X}_i^\dagger \tilde{X}_j=0$ si $\lambda_i \neq j\lambda_j$. Esto implica que los autovectores X del problema original serán ortogonales para *un producto escalar modificado dependiente de B:* $\tilde{X}_i^\dagger \tilde{X}_j = X_i^\dagger B X_j = 0$ si $\lambda_i \neq \lambda_j$ y $\tilde{X}_i = B^{1/2}X_i$. Veremos en las próximas secciones la definición de producto escalar con más detalle. Al ser $\tilde{A}$ diagonalizable y $B^{1/2}$ no singular, tanto el conjunto $\{\tilde{X}_1,\ldots,\tilde{X}_n\}$ como $\{X_1,\ldots,X_n\}$ formarán una base de $\mathbb{C}^n$.

Más aun, dado que $\tilde{A}$ es hermítica, existirá una matriz no singular de autovectores normalizados y ortogonales $\tilde{S}=(\tilde{X}_1,\ldots,\tilde{X}_n)$ tal que $\tilde{S}^\dagger\tilde{S}=I$ y $\tilde{S}^\dagger\tilde{A}S=\tilde{A}'$ con $\tilde{A}'$ diagonal: $\tilde{A}'_{ij}=\lambda_i\delta_{ij}$. Esto implica que $S=B^{-1/2}\tilde{S}=(X_1,\ldots,X_n)$ satisface simultáneamente

$$S^\dagger AS = A', \quad S^\dagger BS = I$$

con $A'=\tilde{A}'$ diagonal: $A'_{ij}=\lambda_i\delta_{ij}$. Los autovalores generalizados λ_i pueden obtenerse directamente como las raíces de $|A-\lambda B|=0$, mientras que los correspondientes autovectores X_i (columnas de S) de la ecuación $(A-\lambda_i B)X_i=0$. La existencia de tal S queda pues garantizada en el caso $A^\dagger=A$ y $B^\dagger=B$, con B definida positiva ($\lambda_i^B>0\ \forall\ i$).

12. Otras propiedades importantes

12.1) Sean $F:V\to V$, $G:V\to V$ dos operadores sobre el mismo espacio V de dimensión finita. Entonces los autovalores de FG *son los mismos que los de* GF, aún cuando $FG\neq GF$

Demostración: En efecto, si $FG(v)=\lambda v$ con $v\neq 0 \Rightarrow GF(G(v))=G(FG(v))=G(\lambda v)=\lambda G(v)$. Si $G(v)\neq 0\Rightarrow\lambda$ es también autovalor de GF con autovector $G(v)$. Si $G(v)=0\Rightarrow$ necesariamente $\lambda=0$ (pues $FG(v)=F(G(v))=F(0)=0$) y por lo tanto $0=\text{Det}[FG]=\text{Det}[GF]$. Esto implica que 0 es también autovalor de GF. Análogamente se muestra que todo autovalor de GF es autovalor de FG.

Esta propiedad es entonces también válida para matrices. Si A,B son dos matrices de $n\times n\Rightarrow$ los autovalores de AB son los mismos que los de BA, aún si $AB\neq BA$.

12.2) Si $[F,G]=FG-GF=0$ y $v\neq 0$ es autovector de F con autovalor $\lambda\Rightarrow G(v)\in V_F(\lambda)$, es decir, $G(v)$ es también autovector de F con el mismo autovalor si $G(v)\neq 0$.

Demostración: $F(G(v))=FG(v)=GF(v)=G(\lambda v)=\lambda G(v)$, por lo que $G(v)\in V_F(\lambda)$.

Además, si V es de dimensión finita n y los autovalores de F *son todos distintos* $\Rightarrow$ todo autovector de F es también autovector de G, pues en tal caso $V_F(\lambda)$ es de dimensión 1 para todo autovalor λ y por lo tanto, necesariamente $G(v)=\alpha v$.

Todos los operadores que conmutan con F quedan en este caso directamente representados por matrices *diagonales* en la base en que $[F]_e$ es diagonal, y son por lo tanto diagonalizables.

En general, si F y G son ambos diagonalizables y $[F,G]=0\Rightarrow$ existe una base común donde F y G son simultáneamente diagonales. Solo hay que diagonalizar F, y luego diagonalizar G dentro de cada espacio propio de F.

13. Subespacios Invariantes

Sea $F: V \to V$ un operador lineal y sea $W \subset V$ un subespacio de V. W es un subespacio invariante bajo la acción de F (se dice también invariante bajo F o por F) si $F(W) \subset W$, es decir, si $\forall\, v \in W$, $F(v) \in W$.

Ejemplos triviales son $W = V$ y $W = \{0\}$, que son subespacios invariantes para todo operador lineal $F: V \to V$ (pues $F(V) \subset V$ y $F(0) = 0$).
También el núcleo $N(F)$ y la imagen $I(F)$ son siempre invariantes por F ($F(N(F)) = \{0\} \subset N(F)$, y si $v \in I(F)$, $F(v) \in I(F)$).
Resulta asimismo trivial reconocer que si $F = \alpha I$, con I el operador identidad $\Rightarrow$ cualquier subespacio $W \subset V$ es invariante por F, ya que si $v \in W$, $F(v) = \alpha v \in W$.
Como otro ejemplo común consideremos el proyector P sobre S_1 en la dirección de S_2, con $S_1 \oplus S_2 = V$.
Es obvio que S_1 es invariante por P pues si $v_1 \in S_1 \Rightarrow P(v_1) = v_1 \in S_1$ ($P(S_1) = S_1 = I(P)$). Más aún, cualquier subespacio de S_1 es también invariante por P.

13.1) Si λ es autovalor de F $\Rightarrow$ el espacio propio $V(\lambda)$ es un subespacio invariante por F: Si $v \in V(\lambda)$ $\Rightarrow F(v) = \lambda v \in V(\lambda)$.
13.2) La suma de espacios propios $V(\lambda_1) \oplus V(\lambda_2)$ de un mismo operador F es también invariante por F.
Si $v \in V(\lambda_1) \oplus V(\lambda_2)$, $v = v_1 + v_2$, con $F(v_i) = \lambda_i v_i$, $i = 1, 2$ y por lo tanto $F(v) = \lambda_1 v_1 + \lambda_2 v_2 \in V(\lambda_1) \oplus V(\lambda_2)$.
Análogamente, la suma de espacios propios $V(\lambda_1) \oplus \ldots \oplus V(\lambda_k)$ es invariante por F.
13.3) Si $[F, G] = 0$ y W es un subespacio invariante por F, $G(W)$ es también invariante por F.
En efecto, si $v \in W$, $w = F(v) \in W$ y por lo tanto, $F(G(v)) = FG(v) = GF(v) = G(F(v)) = G(w) \in G(W)$.
13.4) Si se escribe a V como suma directa de subespacios invariantes,
$V = S_1 \oplus S_2 \oplus \ldots \oplus S_k$, existe una base (aquella formada por la unión de las bases de cada subespacio invariante) en la que la matriz que representa a $[F]_e$ estará bloqueada en la forma

$$[F]_e = \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & A_k \end{pmatrix}$$

donde A_i es una matriz de dimensión $d_i \times d_i$, con $d_i = \dim S_i$, $i = 1, \ldots, k$ y $\sum_{i=1}^{k} d_i = n$.
En efecto, si $(e_{i1}, \ldots, e_{id_i})$ es una base de S_i, $F(e_{ij}) \in S_i$, por lo que $F(e_{ij}) = \sum_{l=1}^{d_i} (A_i)_{lj} e_{il}$ para $j = 1, \ldots, d_i$.
Análogamente, si existe una base en la que $[F]_e$ tiene la forma de bloques anterior $\Rightarrow V = S_1 \oplus S_2 \oplus \ldots \oplus S_k$, con S_i invariante por F para $i = 1, \ldots, k$. Basta con considerar S_i como el subespacio generado por los vectores de la base correspondientes a cada bloque.

Los autovalores de F pueden pues obtenerse directamente diagonalizando cada bloque A_i: Como $\det F = \prod_i \det A_i$ (demostración dada en clase), el polinomio característico resulta $|F - \lambda I| = \prod_{i=1}^{k} |A_i - \lambda I_{d_i}|$, por lo que sus racies serán las raíces de cada término, es decir, los autovalores de cada bloque A_i. Y los autovectores correspondientes pertenecerán al subespacio invariante asociado a A_i (detalles dados en clase).
El conocimiento de subespacios invariantes posibilita pues grandes simplificaciones cuando se tiene que diagonalizar matrices de grandes dimensiones.

14. Forma Canónica de Jordan

Surge ahora la pregunta sobre cuál es la forma más simple en que pueden escribirse los operadores (o matrices) no diagonalizables. El siguiente teorema nos da la respuesta:
Teorema: Sea $F : V \to V$ un operador lineal en un espacio vectorial V de dimensión finita n sobre $\mathbb{C}$. Entonces existe una base $e = (e_{11}, e_{12}, \ldots, e_{1d_1}, \ldots, e_{k1}, e_{k2}, \ldots, e_{kd_k})$, con $\sum_{i=1}^{k} d_i = n$, en la que

$$\begin{array}{rcl} F(e_{i1}) & = & \lambda_i e_{i1} \\ F(e_{ij}) & = & \lambda_i e_{ij} + e_{i,j-1}, \;\; j = 2, \ldots, d_i \end{array} \;,\;\; i = 1, \ldots, k$$

o sea, $F(e_{11}) = \lambda_1 e_{11}$, $F(e_{12}) = \lambda_1 e_{12} + e_{11}$, $\ldots$, $F(e_{1d_1}) = \lambda_1 e_{1d_1} + e_{1,d_1-1}$ y similar para $i \geq 1$. El caso diagonalizable corresponde a $d_i = 1 \;\forall\; k$, en cuyo caso $k = n$. Los parámetros λ_i no son necesariamente distintos y son los autovalores de F, como demostraremos a continuación.
La matriz $[F]_e \equiv [F]_e^e$ en esta base toma entonces la forma de bloques

$$[F]_e = A = \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & A_k \end{pmatrix}$$

donde A_i son matrices de $d_i \times d_i$ de la forma

$$A_i = \begin{pmatrix} \lambda_i & 1 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & \ldots & 0 \\ & & & \ldots & \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda_i \end{pmatrix} = \lambda_i I_{d_i} + J_{d_i}\,, \quad J_{d_i} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ & & & \ldots & \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

con I_{d_i} la matriz identidad de $d_i \times d_i$.
Cada subespacio $S_i = \overline{(e_{i1}, e_{i2}, \ldots, e_{id_i})}$ es claramente invariante por F, ya que $F(e_{ij}) \in S(\lambda_i)$.
Es claro también que los escalares λ_i, $i = 1, \ldots, k$, son los autovalores de F, pues

$$P(\lambda) = \mathrm{Det}[F - \lambda I] = |A_1 - \lambda I_{d_1}| \ldots |A_k - \lambda I_{d_k}| = (\lambda_1 - \lambda)^{d_1} \ldots (\lambda_k - \lambda)^{d_k}$$

posee como únicas raíces a $\lambda_1, \ldots, \lambda_k$.

Cada submatriz A_i posee un único autovalor λ_i de multiplicidad d_i ($|A_i - \lambda I_{d_i}| = (\lambda_i - \lambda)^{d_i}$), pero el espacio propio correspondiente es *de dimensión 1*: $\dim N[A_i - \lambda_i I_{d_i}] = \dim N[J_{d_i}] = 1$ pues $\mathrm{Rango}(J_{d_i}) = d_i - 1$. Por lo tanto, la submatriz A_i no es diagonalizable si $d_i > 1$. Cada subespacio S_i contiene entonces *un único subespacio propio* de dimensión 1, que es el generado por e_{i1}, y el número total de autovectores LI de F es $k \leq n$ (uno por cada S_i).

Notemos que $(F - \lambda_i I)e_{ij} = e_{i,j-1}$ para $j > 1$, con $(F - \lambda_i I)e_{i1} = 0$, por lo que aplicando m veces el op. $(F - \lambda_i I)$ sobre e_{ij} resulta

$$(F - \lambda_i I)^m e_{ij} = \begin{cases} e_{i,j-m} & m < j \\ 0 & m \geq j \end{cases}$$

Los operadores no diagonalizables en espacios finitos se caracterizan pues por la existencia de vectores e_{ij} no nulos tales que $(F - \lambda_i I)^j e_{ij} = 0$ pero $(F - \lambda_i I)e_{ij} \neq 0$ si $j > 1$. Si F es diagonalizable tales vectores no existen. Notemos que conociendo e_{id_i}, los restantes vectores e_{ij} pueden obtenerse como

$$e_{ij} = (F - \lambda_i I)^{d_i - j} e_{id_i} = (F - \lambda_i I)e_{i,j+1}, \;\; j = 1, \ldots, d_i - 1$$

La ecuación previa implica también que $(F - \lambda_i I)^{d_i}(e_{ij}) = 0, \;\; j = 1, \ldots, d_i$. Por lo tanto, la matriz $J_{d_i} = A_i - \lambda_i I_{d_i}$ es *nilpotente*:

$$(J_{d_i})^{d_i} = 0$$

donde 0 denota la mariz nula de $d_i \times d_i$. Por ejemplo,

$$J_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\; J_2^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},\; J_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},\; J_3^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\; J_3^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

La evaluación de potencias del operador puede entonces realizarse sin mayor dificultad, ya que

$$A^m = \begin{pmatrix} A_1^m & 0 & \dots & 0 \\ 0 & A_2^m & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & A_k^m \end{pmatrix}$$

donde, teniendo en cuenta que I_{d_i} conmuta con J_{d_i},

$$A_i^m = (\lambda_i I_{d_i} + J_{d_i})^m = \lambda_i^m I_{d_i} + m\lambda_i^{m-1} J_{d_i} + \frac{m(m-1)}{2!}\lambda_i^{m-2} J_{d_i}^2 + \dots + J_{d_i}^m$$

para $i = 1, \dots, k$. Esta expansión contiene a lo sumo d_i términos ya que $(J_{d_i})^r = 0$ si $r \geq d_i$.
En general, si $p(t)$ es un polinomio de grado m,

$$p(t) = p(\lambda_i)1 + p'(\lambda_i)(t - \lambda_i) + \dots + \frac{p^{(m)}(\lambda_i)}{m!}(t - \lambda_i)^m$$

se obtiene

$$p(A_i) = p(\lambda_i) I_{d_i} + p'(\lambda_i) J_{d_i} + \dots + \frac{p^{(d_i-1)}(\lambda_i)}{(d_i - 1)!} J_{d_i}^{d_i-1}$$

Además, la forma de Jordan es muy conveniente para la evaluación de exponenciales:

$$\exp[A_i t] = \exp[\lambda_i I_{d_i} t + J_{d_i} t] = \exp[\lambda_i I_{d_i} t]\exp[J_{d_i} t] = e^{\lambda_i t}\left(I_{d_i} + J_{d_i} t + \dots + J_{d_i}^{d_i-1}\frac{t^{d_i-1}}{(d_i-1)!}\right)$$

Por lo tanto, $B(t) = \exp[A_i t]$ será una matriz triangular con elementos $B_{kj} = e^{\lambda_i t} t^{j-k}/(j-k)!$ si $k \leq j$ y $B_{kj} = 0$ si $k > j$.

Para obtener la representación de Jordan se puede, una vez obtenidos los k autovalores λ_i y autovectores e_{i1}, $i = 1, \dots, k$, resolver las ecuaciones inhomogéneas $F(e_{ij}) = \lambda_i e_{ij} + e_{i,j-1} \quad j = 2, \dots, d_i$, es decir, trabajando en forma matricial en la base e,

$$AX_{i1} = \lambda_i X_{i1}, \;\; AX_{ij} = \lambda_i X_{ij} + X_{i,j-1}, \quad j = 2, \dots, d_i, \;\; i = 1, \dots, k$$

que no poseen solución única. Otra forma más eficiente es partir de e_{id_i}, es decir, encontrar un vector X_{id_i} que satisfaga

$$(A - \lambda_i I)^{d_i} X_{id_i} = 0, \;\; (A - \lambda_i I)^{d_i-1} X_{id_i} \neq 0$$

Los vectores restantes del bloque pueden obtenerse como

$$X_{ij} = (A - \lambda_i I)^{d_i - j} X_{id_i}, \;\; j = 1, \dots, d_i - 1$$

Ejemplo:

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

Tenemos $|A - \lambda I_3| = (1-\lambda)(2-\lambda)^2$, por lo que las raíces son $\lambda_1 = 1$, con multiplicidad 1, y $\lambda_2 = 2$, con multiplicidad 2. Como

$$A - 2I_3 = \begin{pmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

posee rango 2, $N[A - 2I_3]$ posee dimensión 1, por lo que A no es diagonalizable.
Para $\lambda_2 = 2$ el sistema homogéneo $(A - \lambda I_3)X = 0$ posee la solución general $x = y$, $z = 0$, de modo que el autovector es de la forma $x(1,1,0)^t$. Eligiendo $X_{11} = (1,1,0)^t$, el vector X_{12} puede obtenerse resolviendo

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

que da como resultado $z=1$, $x=y$. Podemos elegir entonces $X_{12}=(0,0,1)^t$. Finalmente, $A-1I_3=\begin{pmatrix}0&1&1\\0&1&1\\0&0&1\end{pmatrix}$, por lo que $(A-I_3)X_{31}=0$ conduce a $X_{31}=x(1,0,0)^t$. Obtenemos entonces

$$S=\begin{pmatrix}1&0&1\\1&0&0\\0&1&0\end{pmatrix},\quad \text{con } S^{-1}=\begin{pmatrix}0&1&0\\0&0&1\\1&-1&0\end{pmatrix}$$

y finalmente la forma de Jordan

$$A'=S^{-1}AS=\begin{pmatrix}2&1&0\\0&2&0\\0&0&1\end{pmatrix}$$

Puede comenzarse también determinando X_{12} a partir de las condiciones $(A-2I_3)X_{12}\neq 0$, $(A-2I_3)^2X_{12}=0$, y obtener luego X_{11} como $(A-2I_3)X_{12}$ (hecho así en clase).

Se obtiene también

$$\exp[A't]=\begin{pmatrix}\exp[t\begin{pmatrix}2&1\\0&2\end{pmatrix}] & 0\\ 0 & \exp(t)\end{pmatrix}=\begin{pmatrix}e^{2t}&te^{2t}&0\\0&e^{2t}&0\\0&0&e^t\end{pmatrix}$$

ya que $\exp[t\begin{pmatrix}2&1\\0&2\end{pmatrix}]=\exp[t(2I_2+J_2)]=\exp[2tI_2]\exp[tJ_2]=e^{2t}(I_2+tJ_2)=e^{2t}\begin{pmatrix}1&t\\0&1\end{pmatrix}$, pues $J_2^2=0$.

Resolución general de sistemas de ecuaciones lineales de primer orden. La solución del sistema

$$\frac{dX}{dt}=AX$$

con A constante pero no diagonalizable, puede obtenerse a partir de la forma canónica de Jordan. Tenemos, para $X(0)=X_0$ y $A=SA'S^{-1}$,

$$X=\exp[At]X_0=S\exp[A't]C=\sum_{i=1}^{k}e^{\lambda_i t}\sum_{m=1}^{d_i}c_{im}\sum_{j=0}^{m-1}V_{i,m-j}\frac{t^j}{j!}$$

donde $V_{i,m-j}$ denota las columnas de S [$S=(V_{11},V_{12},\ldots,V_{1,d_1},\ldots,V_{k,1},\ldots,V_{k,d_k})$, de forma que $S\exp[A't]=(e^{\lambda_1 t}V_{11},e^{\lambda_1 t}(V_{12}+tV_{11}),e^{\lambda_1 t}(V_{13}+tV_{12}+\frac{t^2}{2!}V_{11}),\ldots)$] y $C=S^{-1}X_0=(c_{11},c_{12},\ldots,c_{1d_1},\ldots)^t$ un vector de constantes determinadas por el vector inicial $X_0=X(0)$. Por ejemplo,

$$X=e^{\lambda t}[c_1V_1+c_2(V_2+V_1t)+c_3(V_3+V_2t+V_1t^2/2!)+\ldots]$$

en el caso de un solo bloque, con $V_j=(A-\lambda I)^{d-j}V_d$ y d la dimensión del bloque.

15. Polinomios Anuladores y Teorema de Cayley-Hamilton

Sea $F:V\to V$ un operador lineal en un espacio vectorial V de dimensión finita n.
Si $p(\lambda)=a_0+a_1\lambda+\ldots+a_m\lambda^m$ es un polinomio de grado m, podemos asociar a $p(\lambda)$ el operador lineal

$$p(F)=a_0I+a_1F+\ldots+a_mF^m$$

La matriz que representa al operador $p(F)$ en una base e es

$$\begin{aligned}[p(F)]_e &= a_0[I]_e+a_1[F]_e+\ldots+a_m[F^m]_e=a_0I_n+a_1[F]_e+\ldots+a_m[F]_e^m\\ &= p([F]_e)\end{aligned}$$

donde hemos asociado $F^0=I$ y $[F]_e^0=I_n$. Además, si escribimos $p(\lambda)$ en términos de sus m raíces λ_i

$$p(\lambda)=a_m(\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_m)$$

entonces

$$p(F) = a_m(F-\lambda_1 I)(F-\lambda_2 I)\ldots(F-\lambda_m I)$$

ya que las potencias F^k de F conmutan todas entre si $\forall\ j \geq 0$.

Un polinomio p se dice que es *anulador* de F si $p(F)=0$ (o sea, si $p(F)$ es el operador nulo). Dado que la dimension del espacio de operadores lineales $H = \{F : V \to V,\ F \text{ lineal}\}$ en un espacio vectorial V de dimensión finita n es n^2, es claro que el conjunto $(I = F^0, F, F^2, \ldots, F^{n^2})$ es LD (pues son $n+1$). y que por lo tanto, existen siempre n^2+1 constantes $c_0, \ldots, c_n$ no todas nulas tales que $c_0 I + c_1 F + \ldots c_n F^{n^2} = 0$. Esto muestra en forma básica que siempre existe un polinomio anulador de F.

No obstante, el siguiente teorema (denominado teorema de Cayley-Hamilton) muestra que el mismo *polinomio característico asociado a F, que es de grado n, es siempre un polinomio anulador de F.*

Teorema: Si $p(\lambda) = \mathrm{Det}[F-\lambda I]$ es el polinomio característico de $F \Rightarrow p(F) = 0$.
Como el polinomio característico es de grado n ($p(\lambda) = a_0 + a_1\lambda + \ldots + a_n\lambda^n$ con $a_n = (-1)^n \neq 0$), el teorema implica que es siempre posible expresar F^n en términos de potencias de grado $\leq n-1$: Como $p(F) = 0 \Rightarrow$

$$F^n = -(a_0 + a_1 F + \ldots a_{n-1}F^{n-1})/a_n$$

Por lo tanto, *cualquier* potencia F^k con $k \geq n$ puede expresarse en términos de potencias de grado $\leq n-1$.

Demostraremos el teorema a partir de la forma canónica de Jordan. No obstante, en el caso en que F es diagonalizable, el teorema es obvio, ya que en tal caso existe una base e en la que $[F]_e$ es diagonal,

$$[F]_e = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

donde λ_i, $i = 1, \ldots, n$ son los autovalores de F, y

$$[p(F)]_e = p([F])_e = \begin{pmatrix} p(\lambda_1) & 0 & \ldots & 0 \\ 0 & p(\lambda_2) & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & p(\lambda_n) \end{pmatrix} \qquad (15.1)$$

para cualquier polinomio p. Pero si p es el polinomio característico, $p(\lambda_i) = 0$ para $i = 1, \ldots, n$ y por lo tanto $[p(F)]_e = p([F]_e) = 0$ (matriz nula). Esto implica a su vez $p(F) = 0$ (operador nulo).

En el caso general, utilizando la base en la que $[F]_e$ tiene la forma canónica de Jordan, tenemos, para cualquier polinomio p,

$$[p(F)]_e = p([F]_e) = \begin{pmatrix} p(A_1) & 0 & \ldots & 0 \\ 0 & p(A_2) & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & p(A_k) \end{pmatrix}$$

donde A_i, $i = 1, \ldots, k$ son matrices de $d_i \times d_i$ que satisfacen $(A_i - \lambda_i I_{d_i})^{d_i} = 0$ (matriz nula).
Recordemos ahora que el polinomio característico de F está dado por

$$p(\lambda) = |[F]_e - \lambda I_n| = |A_1 - \lambda I_{d_1}|\ldots|A_k - \lambda I_{d_k}| = (\lambda_1 - \lambda)^{d_1}\ldots(\lambda_k-\lambda)^{d_k}$$

Por lo tanto,

$$p(A_i) = (\lambda_1 I_{d_i} - A_i)^{d_1}\ldots(\lambda_i I_{d_i} - A_i)^{d_i}\ldots(\lambda_k I_{d_i} - A_i)^{d_k} = 0$$

pues $(\lambda_i I_{d_i} - A_i)^{d_i} = 0$. Esto implica $[p(F)]_e = 0$ y entonces $p(F) = 0$. Se cumple pues, para cualquier base e', $p([F]_{e'}) = [p(F)]_{e'} = [0]_{e'} = 0$.

El teorema vale por consiguiente para matrices A generales de $n \times n$. Si $p(\lambda) = |A - \lambda I_n|$ es el polinomio característico asociado a A (de grado n) $\Rightarrow p(A) = 0$ (la matriz nula de $n \times n$).
Para matrices A diagonalizables el resultado es evidente, ya que en tal caso $A = SA'S^{-1}$, con A' diagonal, y por lo tanto $p(A) = p(SA'S^{-1}) = Sp(A')S^{-1}$, pero $p(A')$ tiene la forma 15.1 y es por lo tanto la matriz nula.

Escribiendo
$$p(F) = c_n F^n + c_{n-1}F^{n-1} + \ldots + c_1 F + c_0 I$$
en el caso del polinomio característico tenemos $c_n = (-1)^n \neq 0$ y $c_0 = \mathrm{Det}[F]$. Por lo tanto, como $p(F) = 0$, podemos escribir
$$F^n = -(c_{n-1}F^{n-1} + \ldots c_1 F + c_0 I)/c_n$$
de modo que F^n (y por lo tanto cualquier potencia F^k, con $k \geq n$ natural) puede escribirse en términos de los operadores $F^{n-1}, \ldots, F, I$. Más aún, si F es invertible, $c_0 \neq 0$ y multiplicando la expresión anterior por F^{-1} se obtiene
$$F^{-1} = -(c_n F^{n-1} + c_{n-1}F^{n-2} + \ldots c_1 I)/c_0$$
de modo que también F^{-1} (y por tanto cualquier potencia F^{-k}, $k > 0$ natural) puede escribirse en términos de $F^{n-1}, \ldots, F, I$.

Cabe destacar que el polinomio característico no es necesariamente el polinomio anulador de grado mínimo. Sí lo es en el caso de autovalores todos distintos o, en general, en el caso de bloques de Jordan con autovalores todos distintos.
Si F es diagonalizable, el polinomio anulador de grado mínimo es simplemente $P_m(\lambda) = \prod_i(\lambda - \lambda_i)$, donde la productoria es sobre *autovalores distintos*.
En el caso general, el polinomio anulador de grado mínimo será $P_m(\lambda) = \prod_i(\lambda - \lambda_i)^{d_i}$, donde la productoria es nuevamente sobre autovalores distintos y d_i es la dimensión del mayor bloque de Jordan asociado a λ_i. $P_m(\lambda)$ es pues de grado $\leq n$.

Ejemplo 1: Sea
$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
El polinomio característico es
$$p(\lambda) = |A - \lambda I_2| = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1$$
Se cumple entonces
$$p(A) = A^2 - I_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$
Esto muestra simplemente que $A^2 = I_2$ y que por lo tanto, $A^k = I_2$ si k es par y $A^k = A$ si k impar.

Ejemplo 2:
$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$
El polinomio característico es
$$p(\lambda) = |A - \lambda I_3| = (1-\lambda)(2-\lambda)^2 = -\lambda^3 + 5\lambda^2 - 8\lambda + 4$$
El teorema implica entonces que
$$-A^3 + 5A^2 - 8A + 4I_3 = 0$$
donde $A^2 = A.A$, $A^3 = A.A.A$ (producto matricial), como es fácil verificar. Por lo tanto, $A^3 = 5A^2 - 8A + 4I_3$ y $A^{-1} = (A^2 - 5A + 8I)/4$. Cualquier potencia A^k con k entero puede expresarse en términos de I_3, A y A^2.

Ej. 3: Matriz A de $n \times n$ de rango 1, con $n \geq 2$. Dado que A y $A' = S^{-1}AS$ poseen el mismo rango y traza, tenemos, si A' es la forma canónica de Jordan de A, $r(A') = 1$ y $\mathrm{Tr}\, A' = \mathrm{Tr}\, A$, por lo que A' tiene sólo una fila no nula. Si $\mathrm{Tr}\, A \neq 0$, la única posibilidad es que A sea diagonalizable y tenga un único autovalor no nulo igual a $\mathrm{Tr}\, A$, siendo los restantes nulos, mientras que si $\mathrm{Tr}\, A = 0$, la única posibilidad para A' es un bloque de Jordan de dimensión 2 de la forma $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, siendo todos los autovalores nulos y A no diagonalizable. Toda matriz de rango 1 es necesariamente de la forma $A = bc^t$, con b y c vectores columna no nulos (de $n \times 1$), como el lector puede fácilmente demostrar, con $\mathrm{Tr}\, A = c^t b = \mathbf{c} \cdot \mathbf{b}$. Se deja como problema hallar los autovectores asociados y el polinomio minimal en ambos casos.

16. Demostración

Daremos aquí un resumen de la demostración de la forma canónica de Jordan. En primer lugar, sabemos que todo operador lineal $F : V \to V$, con V de dimensión finita n, posee un polinomio anulador $P(x)$, tal que $P(F) = 0$ (o sea, $P(F)(v) = 0 \ \forall \ v \in V$). Existirá entonces un polinomio anulador de grado mínimo $P_m(x) = a_0x + a_1x + \ldots + a_mx^m$ (polinomio minimal), tal que $P_m(F) = a_0F + a_1F + \ldots + a_mF^m = 0$.
1) λ es raíz de $P_m(F)$ si y sólo si λ es autovalor de F. Esto indica que las raíces del polinomio minimal y el polinomio característico son las mismas. Sólo la multiplicidad puede ser diferente.
Dem.: Si λ es autovalor de $F \Rightarrow \exists \ v \neq 0$ tal que $F(v) = \lambda v$, y en tal caso $P_m(F)(v) = P_m(\lambda)v = 0$, por lo que $P_m(\lambda) = 0$, es decir, λ es raíz de $P_m(x)$.
Si λ es raíz de $P_m(x) \Rightarrow P_m(x) = Q_{m-1}(x)(x - \lambda)$. En tal caso $P_m(F)(v) = Q_{m-1}(F)(F - \lambda I)(v) = 0$ $\forall \ v \in V$, por lo que necesariamente $\exists \ v \neq 0$ tal que $(F - \lambda I)(v) = 0$, es decir, λ es autovalor de F y v autovector asociado (si tal vector no existiese tendríamos $Q_{m-1}(F)(v) = 0 \ \forall \ v \in V$ y el polinomio minimal sería $Q_{m-1}(F)$, de grado $m - 1 < m$, en contradicción con la hipótesis).
2) Si $P_m(x) = Q_1(x)Q_2(x)$, con $Q_1(x)$ y $Q_2(x)$ polinomios sin raíces comunes y $P_m(F) = Q_1(F)Q_2(F) = 0$ $\Rightarrow V = N(Q_1(F)) \oplus N(Q_2(F))$, donde $N(Q_i(F))$ $(i = 1, 2)$ denota el núcleo de $Q_i(F)$. Los subespacios $N(Q_i(F))$ son además invariantes por F.
Dem.: Al no tener raíces comunes, existen polinomios $A_1(x)$, $A_2(x)$ t.q. $1 = A_1(x)Q_1(x) + A_2(x)Q_2(x)$, o sea,

$$I = A_1(F)Q_1(F) + A_2(F)Q_2(F)$$

Por lo tanto, $\forall \ v \in V$, $v = A_1(F)Q_1(F)(v) + A_2(F)Q_2(F)(v) = v_1 + v_2$, con $v_i = A_i(F)Q_i(F)$. Pero $v_1 \in N(Q_2(F))$ pues $Q_2(F)A_1(F)Q_1(F)(v) = A_1(F)Q_1(F)Q_2(F)(v) = 0$, y análogamente, $v_2 \in N(Q_1(F))$. Esto muestra que $V = N(Q_2(F)) + N(Q_1(F))$.
Además, si $v \in N(Q_1(F))$ y $v \in N(Q_2(F)) \Rightarrow v = A_1(F)Q_1(F)(v) + Q_2(F)Q_2(F)(v) = 0$. Esto muestra que $V = N(Q_1(F)) \oplus N(Q_2(F))$.
Finalmente, si $v \in N(Q_1(F)) \Rightarrow v = A_2(F)Q_2(F)(v)$ y en tal caso $Q_1(F)F(v) = A_2(F)Q_1(F)Q_2(F)F(v) = 0$, por lo que $F(v) \in N(Q_1(F))$. Análogamente, si $v \in N(Q_2(F)) \Rightarrow F(v) \in N(Q_2(F))$, por lo que ambos núcleos son invariantes por F.
3) Generalizando, si

$$P_m(x) = (x - \lambda_1)^{d_1} \ldots (x - \lambda_k)^{d_k}$$

con $\lambda_i \neq \lambda_j$ si $i \neq j$ (las raíces distintas de $P_m(x)$) y $P_m(F) = 0 \Rightarrow V = V_1 \oplus \ldots \oplus V_k$, con $V_i = N(F - \lambda_i I)^{d_i}$. El espacio completo V puede pues subdividirse en k subespacios invariantes, núcleos de $\tilde{F}_i^{d_i}$, donde $\tilde{F}_i = (F - \lambda_i I)$. Podemos pues construir una base de V formada por las bases de V_i.
4) Para construir una base de V_i, notemos que debe existir un vector $v \neq 0$ tal que $\tilde{F}_i^{d_i}(v) = 0$ pero $\tilde{F}_i^{d_i-1}(v) \neq 0$ (de lo contrario $P_m(F)$ no sería el polinomio minimal). En tal caso, los d_i vectores no nulos

$$e_{ij} = \tilde{F}_i^{d_i-j}(v), \quad j = 1, \ldots, d_i$$

(o sea $e_{id_i} = v$, $e_{ij} = \tilde{F}_i(e_{i,j+1})$, $j = 1, \ldots, d_i - 1$) son LI. Dem.: Si

$$c_1\tilde{F}_i^{d_i-1}(v) + c_2\tilde{F}_i^{d_i-2}(v) + \ldots + c_{d_i-1}\tilde{F}_i(v) + c_{d_i}v = 0$$

aplicando $\tilde{F}_i^{d_i-1}$ al segundo miembro obtenemos $c_{d_i}\tilde{F}_i^{d_i-1}(v) = 0$, por lo que $c_{d_i} = 0$. Luego, aplicando sucesivamente $\tilde{F}_i^{j}$, con $j = d_i - 1, \ldots, 0$, vemos que $c_j = 0$ para $j = 1, \ldots, d_i$. Notemos además que $\tilde{F}_i(e_{i1}) = \tilde{F}_i^{d_i}(v) = 0$, o sea, $(F - \lambda_i I)e_{i1} = 0$, por lo que e_{i1} es autovector de F con autovalor λ_i. Tenemos pues, en el subespacio S_i generado por los d_i vectores $B_i = (e_{i1}, \ldots, e_{id_i})$,

$$[\tilde{F}_i]_{B_i} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ & & & \ldots & \\ 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & \ldots & 0 & 0 \end{pmatrix} = J_{d_i}, \ \text{ es decir, } \ [F_i]_{B_i} = [\tilde{F}_i]_{B_i} + \lambda_I I_{d_i} = \begin{pmatrix} \lambda_i & 1 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & \ldots & 0 \\ & & & \ldots & \\ 0 & 0 & \ldots & \lambda_i & 1 \\ 0 & 0 & \ldots & 0 & \lambda_i \end{pmatrix}$$

Se obtiene así un bloque de Jordan de dimensión d_i (grado del término correspondiente $(x - \lambda_i)^{d_i}$ del polinomio minimal).

Puede existir otro vector $v \in N(\tilde{F}_i^{d_i})$ tal que $\tilde{F}^{d_i-1}(v) \neq 0$ pero $\tilde{F}_i^{d_i}(v) = 0$ *y que no pertenezca al espacio generado por los vectores de* B_i. Este vector generaría otro bloque de Jordan de la misma dimensión con el mismo autovalor λ_i. En general, pueden surgir vectores $v \in N(\tilde{F}_i^{d_i})$ que no pertenezcan a los subespacios generados por el conjunto de vectores anteriores y que satisfagan $\tilde{F}_i^{r-1}(v) = 0$ pero $\tilde{F}_i^{r}(v) = 0$, con $r \leq d_i$, que generarán otros bloques de Jordan de dimensión $r \leq d_i$ con el mismo autovalor. La dimension total de $N(F - \lambda_i I)^{d_i}$ será así la multiplicidad $m_i \geq d_i$ de λ_i en el polinomio característico.
Si $d_i = 1$ los bloques son de dimensión 1 y los vectores correspondientes *autovectores* de F con autovalor λ_i. Este es el caso donde F es *diagonalizable* en el subespacio asociado a λ_i, es decir, donde la dimensión de $N(F - \lambda_i I)$ coincide con la multiplicidad de λ_i como raíz del polinomio característico.
Por lo tanto, si F es diagonalizable, el polinomio minimal es $P_m(x) = (x - \lambda_1) \ldots (x - \lambda_k)$.

17. Formas lineales, bilineales y cuadráticas

17.1 Formas lineales

Estudiaremos ahora funciones escalares lineales de argumento vectorial. Sea V un espacio vectorial sobre un cuerpo K. Una forma lineal es una función $F: V \to K$ que satisface las condiciones

$$F(\alpha v) = \alpha F(v) \;\; \forall v \in V,\; \alpha \in K \tag{1}$$
$$F(v_1 + v_2) = F(v_1) + F(v_2) \;\; \forall v_1, v_2 \in V \tag{2}$$

Una forma lineal puede ser considerada como un caso particular de transformación lineal si se considera el cuerpo K como un espacio vectorial de dimensión 1 sobre el mismo K. Nótese que se satisface $F(0) = 0$.
Ejemplos (se dejan las comprobaciones para el lector):
1) Si $K = \mathbb{R}$ y $V = \mathbb{R}^2$, $F(x, y) = x + y$ es claramente una forma lineal, mientras que $G(x, y) = x + y^2$ y $H(x, y) = 1 + x$ no son formas lineales.
2) Si $V = \mathbb{R}^{n\times n}$, la traza de una matriz $A \in V$, $\mathrm{Tr}[A] = \sum_{i=1}^{n} A_{ii}$, es una forma lineal.
3) Si $K = \mathbb{R}$ y $V = C_{[a,b]}$ (espacio de funciones continuas $f : [a, b] \to \mathbb{R}$),

$$T(f) = \int_a^b f(x)dx$$

es una forma lineal, y también lo es (para $\rho \in C_{[a,b]}$)

$$T_\rho(f) = \int_a^b f(x)\rho(x)dx\,.$$

4) En el mismo espacio anterior, y para $a < 0 < b$, $T(f) = f(0)$ es también una forma lineal. Nótese sin embargo que en este caso no existe $\rho(x)$ continua tal que $T(f) = \int_a^b f(x)\rho(x)dx$.
5) Si $V = \mathbb{R}^n$ y w es un vector fijo de $\mathbb{R}^n$,

$$F_w(v) = w \cdot v$$

(producto escalar usual) es una forma lineal. Por ej. el primer caso de 1), $F(x, y) = x + y$, puede ser escrito como $F(x, y) = (1, 1) \cdot (x, y)$. *Toda* forma lineal en $\mathbb{R}^n$ puede ser escrita de esta manera en términos de un único vector $w \in V$, como se verá a continuación.

Si $\dim V = n$ y F no es la forma lineal nula $\Rightarrow \dim I(F) = 1$, por lo que $\dim N(F) = n - 1$. Ejemplo: Hallar el núcleo de la forma lineal del ejemplo 2.

En un espacio vectorial V de dimensión finita n, la forma lineal queda completamente determinada por los valores que asigna a los elementos de una base: Si $B = (b_1, \ldots, b_n)$ es una base de V y $v = \sum_{i=1}^{n} \alpha_i b_i \Rightarrow$

$$F(v) = F(\sum_{i=1}^{n} \alpha_i b_i) = \sum_{i=1}^{n} \alpha_i F(b_i) = [F]_B[v]_B$$

donde

$$[F]_B = (F(b_1), \ldots, F(b_n))$$

es la matriz fila que representa a F en la base B y

$$[v]_B = \begin{pmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{pmatrix}$$

la matriz columna de coordenadas de v en dicha base. El producto $[F]_B[v]_B$ puede entonces visualizarse como el producto escalar usual de los vectores $[F]_B$ y $([v]_B)^t$ de K^n. En $V = \mathbb{R}^n$ toda forma lineal puede pues ser escrita en la forma $F(v) = w \cdot v$, con $w = [F]_e = (\beta_1, \ldots, \beta_n)$, siendo e la base canónica y $\beta_i = F(e_i) = w \cdot e_i$.
Frente a un cambio de base,

$$b'_i = \sum_{j=1}^{n} S_{ji} b_j, \quad i = 1, \ldots, n$$

con S una matriz no singular ($|S| \neq 0$) se obtiene $F(b'_i) = \sum_{j=1}^{n} S_{ji} F(b_j)$ y por lo tanto

$$[F]_{B'} = [F]_B S$$

con lo cual, dado que $[v]_B = S[v]_{B'}$, $F(v) = [F]_B[v]_B = [F]_B S[v]_{B'} = [F]_{B'}[v]_{B'}$.

Si $F : V \to K$ y $G : V \to K$ son dos formas lineales sobre V, la combinación lineal $\alpha F + \beta G$, definida por $(\alpha F+\beta G)(v) = \alpha F(v)+\beta G(v)$, es también una forma lineal $\forall\, \alpha, \beta \in K$, como es muy fácil comprobar. El conjunto de todas las formas lineales $F : \mathbb{V} \to K$ es pues un espacio vectorial denominado *espacio dual* V^*. Si V es de dimensión finita $\Rightarrow$

$$\dim V^* = \dim V$$

ya que existe un isomorfismo entre V^* y K^n (definido por $\mathcal{G}_B(F) = [F]_B \in K^n$, con $n = \dim V$) y por lo tanto entre V^* y V. Si $B = (b_1, \ldots , b_n)$ es una base ordenada de V, la base asociada de V^* es la base dual $B^* = \{F_1, \ldots , F_n\}$, donde $F_i : V \to K$ está definido por $F(\sum_i \alpha_i b_i) = \alpha_i$, es decir,

$$F_i(b_j) = \delta_{ij}\,.$$

F_i es pues representado por el vector fila $[F_i]_B = (0, \ldots , 1_i, \ldots , 0) = e_i$.

17.2 Formas bilineales

Una función escalar de dos variables vectoriales, $A : V \times V \to K$, es una forma bilineal si satisface

$$A(\alpha v, w) = \alpha A(v, w), \quad A(v_1 + v_2, w) = A(v_1, w) + A(v_2, w) \;\; \forall v, v_1, v_2, w \in V,\; \alpha \in K$$

$$A(v, \alpha w) = \alpha A(v, w), \quad A(v, w_1 + w_2) = A(v, w_1) + A(v, w_2) \;\; \forall w, w_1, w_2, v \in V,\; \alpha \in K$$

A es entonces una forma bilineal si es lineal con respecto a sus dos argumentos vectoriales.
Ejemplos (se dejan las comprobaciones como ejercicio):
1) Si $V = \mathbb{R}^2$ y $K = \mathbb{R}$, con $v = (x, y)$, $w = (z, t)$, las siguientes funciones son formas bilineales:

$$A(v, w) = v \cdot w = xz + yt \quad \text{(producto escalar)}$$

$$B(v, w) = xt - yz \;\; \text{(determinante de } (\begin{smallmatrix} x & y \\ z & t\end{smallmatrix}))$$

2) Si $V = C_{[a,b]}$ y $K = \mathbb{R}$, las siguientes funciones son formas bilineales:

$$A(f, g) = \int_a^b f(x)g(x)dx\,, \quad B(f, g) = \int_a^b f(x)g(x)\rho(x)dx\,, \quad C(f, g) = \int_a^b \int_a^b f(x)K(x, x')g(x')dxdx'$$

donde $\rho(x)$ y $K(x, x')$ son continuas.
3) En el mismo espacio anterior, para $a < 0 < b$, también son formas bilineales $T(f, g) = f(0)g(0)$ y $H(f, g) = f(0)g(c)$, con $c \in [a, b]$ fijo. Estas formas no pueden ser escritas en la forma integral del ejemplo anterior para ρ y K continuas.
4) En $V = \mathbb{R}^{n\times 1}$,

$$A(v, w) = v^t A w$$

donde $v^t = (\alpha_1, \ldots , \alpha_n)$, $w = \begin{pmatrix} \beta_1 \\ \ldots \\ \beta_n \end{pmatrix}$ y A es una matriz real de $n \times n$, es una forma bilineal. *Toda* forma bilineal en $\mathbb{R}^{n\times 1}$ (y por lo tanto $\mathbb{R}^n$) puede escribirse de esta manera, como se verá a continuación. Por ejemplo, las formas del ejemplo 1) pueden ser escritas como $A(v, w) = (x, y)(\begin{smallmatrix} 1 & 0 \\ 0 & 1\end{smallmatrix})(\begin{smallmatrix} z \\ t\end{smallmatrix})$, $B(v, w) = (x, y)(\begin{smallmatrix} 0 & 1 \\ -1 & 0\end{smallmatrix})(\begin{smallmatrix} z \\ t\end{smallmatrix})$.

Representación matricial. En un espacio vectorial V de dimensión finita n, la forma bilineal queda completamente determinada por los valores que asigna a pares ordenados de elementos de una base $B = (b_1, \ldots , b_n)$ de V. Si $v = \sum_{i=1}^n \alpha_i b_i$, $w = \sum_{j=1}^n \beta_j b_j \Rightarrow$

$$A(v, w) = A(\sum_{i=1}^n \alpha_i b_i, \sum_{j=1}^n \beta_j b_j) = \sum_{i=1}^n \alpha_i A(b_i, \sum_{j=1}^n \beta_j ebj) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i\beta_j A(b_i, b_j)$$

La igualdad anterior puede escribirse en forma compacta matricial como

$$A(v, w) = [v]_B^t [A]_B [w]_B$$

donde $[v]_B^t = (\alpha_1, \ldots, \alpha_n)$, $[w]_B = \begin{pmatrix} \beta_1 \\ \ldots \\ \beta_n \end{pmatrix}$ y

$$[A]_B = \begin{pmatrix} A(b_1, b_1) & \ldots & A(b_1, b_n) \\ & \ldots & \\ A(b_n, b_1) & \ldots & A(b_n, b_n) \end{pmatrix}$$

es la matriz de $n \times n$ que representa a la forma bilineal A en dicha base $[([A]_B)_{ij} = A(b_i, b_j)]$.
Por ej., si $V = \mathbb{R}^2$, $K = \mathbb{R}$ y e es la base canónica, obtenemos, para los casos del ejemplo 1) y $v = (x, y) = xe_1 + ye_2$, $w = (z, t) = ze_1 + te_2$,

$$A(v, w) = xz + yt = (x, y)[A]_e\binom{z}{t}, \quad [A]_e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$B(v, w) = xt - yz = (x, y)[B]_e\binom{z}{t}, \quad [B]_e = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Por otro lado, la matriz $[C]_e = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ determina la forma bilineal

$$C(v, w) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} z \\ t \end{pmatrix} = xz + 2xt + 3yz + 4yt$$

Una forma bilineal es *simétrica* si

$$A(v, w) = A(w, v) \ \forall v, w \in V$$

y *antisimétrica*

$$A(v, w) = -A(w, v) \ \forall v, w \in V$$

Toda forma bilineal puede escribirse como suma de una forma bilineal simétrica y otra antisimétrica:

$$A(v, w) = \frac{A(v, w) + A(w, v)}{2} + \frac{A(v, w) - A(w, v)}{2} = A_s(v, w) + A_a(v, w)$$

El conjunto de formas bilineales de $V \times V$ sobre K forma un espacio vectorial W (con las operaciones usuales de suma y multiplicación por escalar) y la anterior descomposición corresponde a la suma directa $W = W_s \oplus W_a$, con W_s, W_a los subespacios de formas bilineales simétricas y antisimétricas sobre K.
En espacios V de dimensión finita, las correspondientes matrices en cualquier base dada son simétricas y antisimétricas respectivamente:

$$A(v, w) = A(w, v) \Rightarrow [A]_B^t = [A]_B, \text{ pues } A(b_i, b_j) = A(b_j, b_i)$$

$$A(v, w) = -A(w, v) \Rightarrow [A]_B^t = -[A]_B, \text{ pues } A(b_i, b_j) = -A(b_j, b_i)$$

En los ejemplos anteriores, el primero (producto escalar) es una forma bilineal simétrica mientras que el segundo (determinante) es una forma antisimétrica.

Dada una forma bilineal A arbitraria, notemos que $A(v, 0) = A(0, w) = 0 \ \forall \ v, w \in V$, como el lector podrá fácilmente demostrar. Si además existe $w \neq 0$ tal que $A(v, w) = 0 \ \forall \ v \in V$, la forma bilineal se dice que es **singular**. En caso contrario se dice **no singular**.
En un espacio V de dimensión finita, *A es singular si y sólo si la matriz que la representa en una base cualquiera, $[A]_B$, es singular.*
Dem.: Si $[A]_B$ es singular, existe un vector columna $[w]_B$ no nulo tal que $[A]_B[w]_B = 0$ y por lo tanto, $A(v, w) = [v]_B^t[A]_B[w]_B = [v]_B^t 0 = 0 \ \forall \ v \in V$.
Por otro lado, si existe $w \neq 0$ tal que $A(v, w) = 0 \ \forall v \in V$, y B es una base cualquiera de $V \Rightarrow [v]_B^t[A]_B[w]_B = 0 \ \forall$ vector $[v]_B^t \in K^{n\times 1}$, lo que implica $[A]_B[w]_B = 0$. Como $[w]_B \neq 0$, la matriz $[A]_B$ es entonces singular.
En espacios de dimensión finita, si $\exists \ w \ / \ A(v, w) = 0 \ \forall \ v \in V \Rightarrow \exists \ u \in V \ / \ A(u, v) = 0 \ \forall \ v \in V$, pues si $[A]_B$ es singular $\Rightarrow [A]_B^t$ es también singular ($|[A]_B^t| = |[A]_B| = 0$).
Notemos también que si A es no singular y $A(v, w_1) = A(v, w_2) \ \forall \ v \in V \Rightarrow w_1 = w_2$, ya que en tal caso $A(v, w_1 - w_2) = 0 \ \forall \ v \in V$ y por lo tanto $w_1 - w_2 = 0$.

17.3 Cambio de base en formas bilineales

Consideremos una forma bilineal A. Frente a un cambio de base

$$b'_i = \sum_{j=1}^{n} S_{ji} b_j, \quad i = 1, \ldots, n$$

se tiene

$$A(b'_i, b'_k) = A(\sum_{j=1}^{n} S_{ji} b_j, \sum_{l=1}^{n} S_{lk} b_l) = \sum_{j=1}^{n} \sum_{l=1}^{n} S_{ji} A(b_j, b_l) S_{lk} = (S^t [A]_B S)_{ik}$$

Se obtiene entonces la ley de transformación

$$[A]_{B'} = S^t [A]_B S$$

donde $([A]_{B'})_{ij} = A(b'_i, b'_j)$ para $i, j = 1, \ldots, n$. De esta forma,

$$A(v, w) = [v]_B^t [A]_B [w]_B = (S[v]_{B'})^t [A]_B (S[w]_{B'}) = [v]_{B'}^t S^t [A]_B S [w]_{B'} = [v]_{B'}^t [A]_{B'} [w]_{B'}$$

Nótese la diferencia con la ley de transformación de matrices que representan operadores lineales $F : V \to V$ en una base, para las que $[F]_{B'} = S^{-1} [F]_B S$. Notemos también que ($| \ldots |$ denota el determinante)

$$|[A]_{B'}| = |S^t [A]_B S| = |S^t| |[A]_B| |S| = |S|^2 |[A]_B|$$

por lo que el signo del determinante no depende de la base (pues $|S| \neq 0$). Si A es singular, $|[A]_B| = 0$ y entonces $|[A]_{B'}| = 0$ en cualquier base.
Otra consecuencia es que como S es no singular ($|S| \neq 0$), el rango de $[A]_B$ (dimensión del espacio fila o columna de $[A]_B$) es también independiente de la base.
Podemos también corroborar que el carácter simétrico o antisimétrico es independiente de la base elegida:

$$[A]_{B'}^t = (S^t [A]_B S)^t = S^t [A]_B^t S$$

por lo que $[A]_{B'}^t = \pm [A]_{B'}$ si $[A]_B^t = \pm [A]_B$.

Ejemplo: Para el caso del producto escalar usual en $\mathbb{R}^n$, $[A]_e = I_n$ en la base canónica e y por lo tanto $[A]_{e'} = S^t [A]_e S = S^t S$ en una base arbitraria e', tal como se adelantó en el apunte 4 sobre cambio de base.
Ejemplo: Para el caso del determinante en $\mathbb{R}^2$, $[A]_e = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ en la base canónica y por lo tanto, en un base e' determinada por una matriz $S = [I]_e^{e'} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ no singular ($|S| \neq 0$),

$$[A]_{e'} = S^t [A]_e S = \left(\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right) \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = (ad - bc) \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) = |S| [A]_e$$

$[A]_{e'}$ es pues proporcional a $[A]_e$. Este resultado es obvio pues $[A]_{e'}$ debe ser antisimétrica y toda matriz antisimétrica de 2×2 debe ser proporcional a $[A]_e = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$.
Ejemplo: Si $[A]_B = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ y $b'_1 = (b_1 + b_2)$, $b'_2 = b_2 - b_1$, $S = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ y por lo tanto

$$[A]_{B'} = S^t [A]_B S = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

Así, si $v = x b_1 + y b_2 = x' b'_1 + y' b'_2$, $w = z b_1 + t b_2 = z' b'_1 + t' b'_2$,

$$A(v, w) = (x, y) [A]_B \left(\begin{smallmatrix} z \\ t \end{smallmatrix}\right) = (x', y') [A]_{B'} \left(\begin{smallmatrix} z' \\ t' \end{smallmatrix}\right)$$

o sea,

$$A(v, w) = xt + yz = 2(x'z' - y't')$$

lo que está de acuerdo con $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = S \left(\begin{smallmatrix} x' \\ y' \end{smallmatrix}\right) = \left(\begin{smallmatrix} x' - y' \\ x' + y' \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} z \\ t \end{smallmatrix}\right) = S \left(\begin{smallmatrix} z' \\ t' \end{smallmatrix}\right) = \left(\begin{smallmatrix} z' - t' \\ z' + t' \end{smallmatrix}\right)$

18 Formas cuadráticas

Si A es una forma bilineal de $V \times V$ en K, la función $\tilde{A} : V \to K$ dada por

$$\tilde{A}(v) = A(v, v)$$

se denomina *forma cuadrática*. Notemos que satisface $\tilde{A}(\alpha v) = \alpha^2 \tilde{A}(v)\ \forall\ \alpha \in K,\ v \in V$:
$A(\alpha v, \alpha v) = \alpha A(v, \alpha v) = \alpha^2 A(v, v)$.
Es importante notar que la forma cuadrática queda completamente determinada por la parte simétrica de la forma bilineal, ya que $A_a(v, v) = [A(v, v) - A(v, v)]/2 = 0$ y por lo tanto

$$A(v, v) = A_s(v, v)$$

Asimismo, la parte simétrica de una forma bilineal queda completamente determinada por la forma cuadrática respectiva, ya que

$$A_s(v + w, v + w) = A_s(v, v) + A_s(w, w) + 2A_s(v, w)$$

y por lo tanto

$$A_s(v, w) = [A_s(v + w, v + w) - A_s(v, v) - A_s(w, w)]/2$$

En un espacio vectorial V de dimensión finita n, podemos entonces escribir, para A *simétrica*,

$$\begin{aligned} A(v, v) &= [v]^t_B [A]_B [v]_B \\ &= \sum_{i,j=1}^{n} \alpha_i A(b_i, b_j)\alpha_j = \sum_{i=1}^{n} A(b_i, b_i)\alpha_i^2 + 2\sum_{i<j} A(b_i, b_j)\alpha_i\alpha_j \end{aligned} \tag{3}$$

Ejemplo: Si $V = \mathbb{R}^2$ y $K = \mathbb{R}$, la longitud al cuadrado de un vector $v = (x, y)$,

$$|v|^2 = x^2 + y^2$$

es una forma cuadrática y puede escribirse como

$$|v|^2 = (x, y)\binom{x}{y} = (x, y)[A]_e\binom{x}{y} = A(v, v)$$

con $[A]_e = I_2$, $A(v, w) = v \cdot w$ y e la base canónica.
También es una forma cuadrática

$$\tilde{B}(v) = 3x^2 + 5y^2 + 2xy = (x, y)[B]_e\binom{x}{y},\quad [B]_e = \begin{pmatrix} 3 & 1 \\ 1 & 5 \end{pmatrix}$$

Toda forma cuadrática en $V = \mathbb{R}^n$ o $V = \mathbb{R}^{n\times 1}$ puede escribirse como

$$\tilde{A}(v) = v^t A v = (\alpha_1, \ldots, \alpha_n) A \begin{pmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{pmatrix} = \sum_{i=1}^{n} a_{ii}\alpha_i^2 + 2\sum_{i<j} a_{ij}\alpha_i\alpha_j$$

con $a_{ij} = A_{ij} = a_{ji}$ los elementos de la matriz real *simétrica* A de $n \times n$ ($A^t = A$).
Ejemplo: Si $V = C_{[a,b]}$ y $K = \mathbb{R}$,

$$||f||^2 \equiv \int_a^b [f(x)]^2 dx = A(f, f)$$

es una forma cuadrática. También lo es $\tilde{C}(f) = \int_a^b \int_a^b K(x, x')f(x)f(x')dxdx'$.

18.1 Forma canónica de una forma cuadrática

Teorema: Sea V un espacio vectorial de dimensión finita n sobre un cuerpo K y sea $A : V \times V \to K$ una forma bilineal *simétrica*. Entonces existe una base B' en la que

$$A(b'_i, b'_j) = \begin{cases} 0 & i \neq j \\ a'_i & i = j \end{cases}$$

es decir, $A(b'_i, b'_j) = a'_i\delta_{ij}$. Esto implica, partiendo de una base arbitraria B, que existe una matriz de cambio de base S tal que

$$[A]_{B'} = S^t[A]_B S = \begin{pmatrix} a'_1 & 0 & \dots & 0 \\ 0 & a'_2 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & a'_n \end{pmatrix}$$

o sea, $([A]_{B'})_{ij} = a'_i\delta_{ij}$. En dicha base la forma bilineal toma entonces la forma diagonal o canónica

$$A(v, w) = \sum_{i=1}^{n} a'_i\alpha'_i\beta'_i$$

donde $v = \sum_{i=1}^n \alpha'_i b'_i$, $w = \sum_{i=1}^n \beta'_i b'_i$, y la correspondiente forma cuadrática toma la forma canónica

$$\tilde{A}(v) = A(v, v) = \sum_{i=1}^{n} a'_i{\alpha'_i}^2$$

Antes de proceder a la demostración, cabe destacar que ni los coeficientes a'_i, ni los vectores b'_i, son únicos. Por ejemplo, en la base B'' definida por $b''_i = \gamma_i b'_i$, $i = 1, \dots, n$, tenemos $A[b''_i, b''_j] = \gamma_i^2 a'_i\delta_{ij}$, y por lo tanto A toma también la forma canónica, con $a'_i \to a''_i = \gamma_i^2 a'_i$.

Notemos también que si la forma bilineal no es simétrica, no es posible encontrar una base en la que $[A]_{B'}$ sea diagonal: Si existiese, $[A]_{B'}$ sería simétrica y por lo tanto $[A]_{B''} = S^t[A]_{B'}S$ sería también simétrica en cualquier base B'' (y la forma bilineal sería entonces simétrica).

Demostración: En el caso de que $K = \mathbb{R}$, la demostración es inmediata si recordamos que toda matriz real simétrica A es siempre diagonalizable, que todos sus autovalores son reales y que sus autovectores pueden siempre elegirse ortogonales y de longitud 1 (véase apunte de autovalores).
Por lo tanto, existirá una matriz de cambio de base S formada por autovectores normalizados de $[A]_B$, con $|S| \neq 0$ y $S^{-1} = S^t$, tal que $S^{-1}[A]_B S = S^t[A]_B S$ es *diagonal*. Si B' es dicha base de autovectores, tendremos entonces

$$[A]_{B'} = S^t[A]_B S = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

con $a'_i = \lambda_i$ los autovalores de $[A]_B$.

No obstante, cabe destacar que diagonalizar $[A]_B$ no es el único procedimiento para llevar una forma cuadrática a una forma diagonal. Esto puede también lograrse utilizando la conocida y simple técnica de *completar cuadrados*, en la cual se basa la demostración del teorema para un cuerpo arbitrario K, que damos a continuación. En tales casos, los coeficientes diagonales a'_i *no son necesariamente iguales a los autovalores de* A.

Notemos primero que si encontramos una transformación lineal de coordenadas

$$\begin{pmatrix} \alpha_1 \\ \dots \\ \alpha_n \end{pmatrix} = S \begin{pmatrix} \alpha'_1 \\ \dots \\ \alpha'_n \end{pmatrix}$$

(o sea, $[v]_B = S[v]_{B'}$) con S una matriz de $n \times n$ no singular ($|S| \neq 0$), tal que

$$A(v, v) = \sum_{i,j=1}^{n} A(b_i, b_j)\alpha_i\alpha_j = \sum_{i,j,k,l} S_{ik}A(b_i, b_j)S_{jl}\alpha'_k\alpha'_l = \sum_{i=1}^{n} a'_i{\alpha'_i}^2$$

hemos entonces encontrado una base canónica para la forma bilineal, dada por

$$b'_i = \sum_{j=1}^{n} S_{ji}b_j \quad i = 1, \dots, n$$

ya que en tal caso $[v]_B = S[v]_{B'}$ y $([A]_{B'})_{ij} = (S^t[A]_B S)_{ij} = a'_i\delta_{ij}$. El problema se reduce pues al de encontrar variables α'_i relacionadas linealmente con las α_i por una transformación no singular, en las que la forma cuadrática sea diagonal.

Procederemos ahora por inducción sobre la dimensión n de V. Para $n=1$, toda forma cuadrática tiene trivialmente la forma canónica en cualquier base: Si $v \in V \to v = \alpha b_1$ y $\tilde{A}(v) = a'_1\alpha^2$, con $a'_1 = A(b_1, b_1)$. Para $n > 1$, supongamos que hemos demostrado que toda forma cuadrática en un espacio de dimensión $n-1$ puede escribirse en la forma canónica. Entonces,

$$A(v,v) = a_{nn}\alpha_n^2 + 2(a_{n1}\alpha_n\alpha_1 + \ldots + a_{n,n-1}\alpha_n\alpha_{n-1}) + g(\alpha_1, \ldots, \alpha_{n-1})$$

donde $v = \sum_{i=1}^n \alpha_i b_i$, $a_{ij} = A(b_i, b_j)$ y g representa una forma cuadrática de dimensión $n-1$. Si $a_{nn} \neq 0$ podemos escribir

$$A(v,v) = a_{nn}(\alpha_n^2 + 2\alpha_n \sum_{j=1}^{n-1} \alpha_j a_{nj}/a_{nn}) + g(\alpha_1, \ldots, \alpha_{n-1}) = a_{nn}(\alpha_n + \sum_{j=1}^{n-1} \alpha_j a_{nj}/a_{nn})^2 + h(\alpha_1, \ldots, \alpha_{n-1})$$

donde $h = g - a_{nn}(\sum_{j=1}^{n-1} \alpha_j a_{nj}/a_{nn})^2$. Por lo tanto

$$A(v,v) = a_{nn}{\alpha'_n}^2 + h(\alpha_1, \ldots, \alpha_{n-1}), \quad \alpha'_n = \alpha_n + \sum_{j=1}^{n-1} \alpha_j a_{nj}/a_{nn}$$

Y como h representa una forma cuadrática de dimensión $n-1$, podemos escribirla en forma canónica como $h = \sum_{i=1}^{n-1} a'_i{\alpha'_i}^2$, donde α'_i son combinaciones lineales de los α_j, $j = 1, \ldots, n-1$. Finalmente obtenemos la forma canónica

$$A(v,v) = a'_n{\alpha'_n}^2 + \sum_{i=1}^{n-1} a'_i{\alpha'_i}^2$$

donde $a'_n = a_{nn}$ y la matriz de transformación $T = S^{-1}$ es de la forma

$$T = \begin{pmatrix} T_{n-1} & 0 \\ t & 1 \end{pmatrix}$$

con T_{n-1} una matriz no singular de $(n-1) \times (n-1)$ y t el vector de $n-1$ componentes determinado por α'_n ($t_i = a_{ni}/a_{nn}$). T es por consiguiente no-singular y define una base B' determinada por $S = T^{-1}$ en la que A tiene la forma canónica.
Si $a_{nn} = 0$ pero $a_{ii} \neq 0$ para algún $i < n$, podemos proceder en forma similar realizando la correspondiente permutación $i \leftrightarrow n$. Finalmente, si todos los a_{ii} son nulos pero existe algún elemento $a_{in} \neq 0$ con $i \neq n$ (pues de lo contrario tendríamos una forma de dimensión $n-1$), podemos efectuar primero el cambio de variables $\alpha_n = \tilde{\alpha}_n + \tilde{\alpha}_i$, $\alpha_i = \tilde{\alpha}_n - \tilde{\alpha}_i$, con lo cual $2a_{ij}\alpha_i\alpha_j = 2a_{ij}(\tilde{\alpha}_n^2 - \tilde{\alpha}_i^2)$ y podemos entonces proceder como en los casos anteriores.

Ejemplo: para $V = \mathbb{R}^2$, $K = \mathbb{R}$ y $v = (x,y) = xe_1 + ye_2$, consideremos

$$A(v,v) = (x,y)[A]_e\binom{x}{y} = x^2 + y^2 + 4xy$$

que coresponde a

$$[A]_e = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Si optamos por el método (muy simple) de completar cuadrados, tenemos

$$x^2 + y^2 + 4xy = x^2 + (y+2x)^2 - 4x^2 = -3x^2 + (y+2x)^2$$

por lo que podemos escribir

$$A(v,v) = -3x'^2 + y'^2, \text{ con } y' = (2x+y),\ x' = x$$

Esto corresponde a la transformación

$$\binom{x'}{y'} = T\binom{x}{y}, \quad T = \left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$$

Por lo tanto
$$S = T^{-1} = \left(\begin{smallmatrix}1 & 0\\ -2 & 1\end{smallmatrix}\right)$$
y la base en la que A toma la forma canónica queda entonces determinada por las columnas de S:
$$e'_1 = e_1 - 2e_2 \;\; e'_2 = e_2$$
Se verifica entonces
$$[A]_{e'} = S^t [A]_e S = \left(\begin{smallmatrix}1 & -2\\ 0 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}1 & 2\\ 2 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}1 & 0\\ -2 & 1\end{smallmatrix}\right) = \left(\begin{smallmatrix}-3 & 0\\ 0 & 1\end{smallmatrix}\right)$$
es decir, $A(e'_1, e'_1) = -3$, $A(e'_2, e'_2) = 1$, $A(e'_1, e'_2) = 0$, como es posible corroborar directamente.

Podemos también optar por el método basado en la diagonalización de A. Tenemos $|[A]_e - \lambda I_2| = (1-\lambda)^2 - 4 = 0$, de donde $\lambda = 1 \pm 2$, o sea, $\lambda_1 = 3$, $\lambda_2 = -1$.
Las componentes de los autovectores correspondientes normalizados son $[e''_1]_e = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1\\ 1\end{smallmatrix}\right)$, $[e''_2]_e = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix}-1\\ 1\end{smallmatrix}\right)$, o sea, $e''_1 = (e_1 + e_2)/\sqrt{2}$, $e''_2 = (-e_1 + e_2)/\sqrt{2}$, y la correspondiente matriz de cambio de base es
$$S = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1 & -1\\ 1 & 1\end{smallmatrix}\right), \;\; S^{-1} = S^t = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1 & 1\\ -1 & 1\end{smallmatrix}\right)$$
Se verifica entonces
$$[A]_{e''} = S^t [A]_e S = \left(\begin{smallmatrix}3 & 0\\ 0 & -1\end{smallmatrix}\right)$$
Es muy importante que los autovectores esten normalizados para que $S^{-1} = S^t$. Finalmente, se obtiene,
$$A(v,v) = (x'', y'')^t [A]_{e''} \left(\begin{smallmatrix}x''\\ y''\end{smallmatrix}\right) = 3x''^2 - y''^2$$
donde $\left(\begin{smallmatrix}x''\\ y''\end{smallmatrix}\right) = [v]_{e''} = S^{-1}[v]_e = S^{-1}\left(\begin{smallmatrix}x\\ y\end{smallmatrix}\right) = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix}x+y\\ x-y\end{smallmatrix}\right)$, o sea, $x'' = (x+y)/\sqrt{2}$, $y'' = (x-y)/\sqrt{2}$.
Notemos que tanto los coeficientes diagonales como las bases obtenidas con los dos procedimientos anteriores son *distintos*. La diagonalización puede llevar más tiempo pero posee la ventaja que automáticamente proporciona una base *ortogonal* en la que la forma cuadrática tiene la forma canónica, lo cual es muy importante en diversas aplicaciones físicas.
Notemos también que el número de coeficientes positivos y negativos en la forma canónica obtenidos en ambos procedimientos *es el mismo*. Esta conclusión es general y se demostrará en el siguiente teorema, de gran importancia.

18.2 Teorema de inercia de formas cuadráticas:
Sea $A(v,v): V \times V \to \mathbb{R}$ una forma cuadrática sobre $\mathbb{R}$. El número de coeficientes a'_i positivos, negativos y nulos en cualquier forma canónica de A es el mismo.
Dem.: Consideremos dos formas canónicas distintas, tal que
$$A(v,v) = \sum_{i=1}^{n} a_i \alpha_i^2 = \sum_{i=1}^{n} a'_i {\alpha'_i}^2$$
con $v = \sum_{i=1}^n \alpha_i e_i = \sum_{i=1}^n \alpha'_i e'_i$, $A(e_i, e_j) = a_i \delta_{ij}$, $A(e'_i, e'_j) = a'_i \delta_{ij}$, y
$$\begin{pmatrix}\alpha_1\\ \dots\\ \alpha_n\end{pmatrix} = S \begin{pmatrix}\alpha'_1\\ \dots\\ \alpha'_n\end{pmatrix}$$
o sea, $\alpha_i = \sum_{j=1}^n S_{ij}\alpha'_j$ para $i = 1, \dots, n$, con $|S| \neq 0$.
Supongamos ahora que $a_i \begin{cases} > 0 & i = 1, \dots, k \\ < 0 & i = k+1, \dots, m \\ 0 & i = m+1, \dots, n \end{cases}$, $a'_i \begin{cases} > 0 & i = 1, \dots, p \\ < 0 & i = p+1, \dots, q \\ 0 & i = q+1, \dots, n \end{cases}$. Por consiguiente,
$$A(v,v) = \sum_{i=1}^{k} |a_i|\alpha_i^2 - \sum_{i=k+1}^{m} |a_i|\alpha_i^2 = \sum_{i=1}^{p} |a'_i|{\alpha'_i}^2 - \sum_{i=p+1}^{q} |a'_i|{\alpha'_i}^2$$

Veremos ahora que si se supone $k < p$ se llega a un absurdo. Si $k < p$, podemos elegir $v \in V$, $v \neq 0$, tal que las primeras k componentes de v en la base e sean nulas ($\alpha_i = 0$ si $i \leq k$), y tal que sus últimas $n-p$ componentes en la base e' sean también nulas ($\alpha'_i = 0$ si $i > p$). En efecto, esto conduce al sistema de k ecuaciones homogéneas $0 = \sum_{j=1}^{p} S_{ij}\alpha'_j$ para $i = 1, \ldots, k$, con $p > k$ incógnitas α'_j, $j = 1, \ldots, p$, el cual posee entonces infinitas soluciones (y por lo tanto, soluciones no nulas). Para tal vector, tendríamos

$$A(v,v) = -\sum_{i=k+1}^{m} |a_i|\alpha_i^2 = \sum_{i=1}^{p} |a'_i|{\alpha'_i}^2$$

pero el segundo miembre es menor o igual a 0 y el tercero mayor que 0, lo que es imposible. Por lo tanto, no puede ser $k < p$. De la misma manera se prueba que no puede ser $p < k$. Por lo tanto, la única posibilidad es $k = p$, es decir, que el número de coeficientes positivos es el mismo.
De la misma forma (se dejan los detalles para el lector) se prueba que $m-k = q-p$ (el número de coeficientes negativos es el mismo).
Finalmente, los dos resultados anteriores implican $n - m = n - q$, es decir, que el número de coeficientes nulos es el mismo.

El número k (número de coeficientes positivos de la forma canónica) se denomina índice de inercia positivo y $m - k$ (número de coeficientes negativos) índice de inercia negativo.
El rango de una forma bilineal simétrica coincide con el rango de la matriz $[A]_e$ y es por lo tanto m (es decir, el número de coeficientes no nulos).
Si A es no singular $\Rightarrow m = n$ (el número de coeficientes nulos es 0).

Ejemplo: Consideremos, para $V = \mathbb{R}^2$, $\mathbb{R} = K$,

$$A(v,v) = x^2 + y^2 + 2xy$$

Completando cuadrados llegamos fácilmente a

$$A(v,v) = (x+y)^2 = 1{x'}^2 + 0{y'}^2$$

con $x' = (x+y)$, $y' = y$. Es decir, existe un coeficiente positivo ($a_1 = 1$) y uno nulo ($a_2 = 0$).
Si en cambio optamos por diagonalizar la matriz correspondiente ($[A]_e = \left(\begin{smallmatrix}1 & 1\\ 1 & 1\end{smallmatrix}\right)$), obtenemos $|A - \lambda I_2| = (1-\lambda)^2 - 1 = 0$ y por lo tanto $\lambda = 1 \pm 1$, o sea, $\lambda_1 = 2$, $\lambda_2 = 0$. Obtenemos entonces un autovalor positivo y uno nulo.

Ejemplo: Consideremos, para $V = \mathbb{R}^3$,

$$A(v,v) = x^2 + y^2 + z^2 + 2xy + 2xz$$

Completando cuadrados,

$$A(v,v) = (x+y+z)^2 - (y+z)^2 + y^2 + z^2 = (x+y+z)^2 - 2yz = {x'}^2 + 2{y'}^2 - 2{z'}^2$$

donde $x' = x+y+z$, $z' = (z+y)/2$, $y' = (y-z)/2$ (se reemplazó $y = z' + y'$, $z = z' - y'$).
Esto implica que $[A]_e = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ tendrá dos autovalores positivos y uno negativo. En efecto, $|A - \lambda I_3| = (1-\lambda)((1-\lambda)^2 - 2) = 0$ conduce a $\lambda_1 = 1 > 0$, $\lambda_2 = 1 + \sqrt{2} > 0$, $\lambda_3 = 1 - \sqrt{2} < 0$.
Obtendremos en la correspondiente base de autovectores normalizados la forma canónica

$$A(v,v) = x''^2 + (1+\sqrt{2})y''^2 + (1-\sqrt{2})z''^2$$

18.3 Formas cuadráticas positivas y aplicaciones

Una forma cuadrática sobre $K = \mathbb{R}$ se denomina definida positiva (o estríctamente positiva) si

$$A(v,v) > 0 \quad \forall\, v \neq 0$$

Es fácil ver que A **es definida positiva si y sólo si los coeficientes diagonales** a_i **de la forma canónica son todos positivos:** $a_i > 0$ para $i = 1, \ldots, n$ (es decir, $k = n$). En efecto, en tal caso

$$A(v, v) = \sum_{i=1}^{n} a_i \alpha_i^2 > 0 \ \ \forall \ \ v \neq 0$$

donde ahora hemos escrito $v = \sum_{i=1}^{n} \alpha_i b_i$, con $B = (b_1, \ldots, b_n)$ una base donde A toma la forma canónica ($A(b_i, b_j) = a_i \delta_{ij}$). Por otro lado, si $A(v, v) > 0 \; \forall \; v \neq 0$, entonces $a_i = A(b_i, b_i) > 0$.

Para una forma cuadrática definida positiva, *podemos siempre elegir una base en la que* $a_i = 1$ para $i = 1, \ldots, n$: En efecto, si $A(b_i, b_j) = a_i \delta_{ij}$, con $a_i > 0$, podemos definir la base de elementos $e_i = b_i/\sqrt{a_i}$ en la que $A(e_i, e_j) = A(b_i, b_j)/\sqrt{a_i a_j} = (a_i/\sqrt{a_i^2})\delta_{ij} = 1\delta_{ij}$.
Notemos también que el determinante de la matriz que representa una forma cuadrática positiva es positivo en cualquier base. En la base B en la que A toma la forma canónica,

$$|[A]_B| = a_1 a_2 \ldots a_n > 0$$

y en cualquier otra base B' de V,

$$|[A]_{B'}| = \begin{vmatrix} A(b'_1, b'_1) & \ldots & A(b'_1, b'_n) \\ & \ldots & \\ A(b'_n, b'_1) & \ldots & A(b'_n, b'_n) \end{vmatrix} = |S^t [A]_B S| = |S|^2 |[A]_B| > 0$$

Además notemos que A sigue siendo positiva en cualquier subespacio de V (pues $A(v, v) > 0 \; \forall \; v \neq 0$), por lo que el determinante de cualquier menor de $[A]_{B'}$ (obtenido al suprimir un número dado de columnas y las respectivas filas de $[A]_{B'}$) es también siempre positivo. Por ejemplo, si consideramos el subespacio generado por los primeros $m \leq n$ elementos de la base B', tendremos

$$|[A]_m| > 0$$

donde $[A]_m$ es la matriz de $m \times m$ de elementos $A(b'_i, b'_j)$, $i \leq m$, $j \leq m$, que representa a A en la base $(b'_1, \ldots, b'_m)$ del subespacio anterior.

Más aún, A **es definida positiva si y sólo si todos los determinantes principales en una base arbitraria** B' **de** V **son positivos**, es decir, si $|[A]_m| > 0$ para $m = 1, \ldots, n$.
Dem.: Por inducción: Para $n = 1$ es obviamente válido. Asumiendo ahora que es válido para $n-1$, entonces existe una base canónica $(e_1, \ldots, e_{n-1})$ del subespacio generado por los primeros $n-1$ vectores de la base original B', en la que $A(e_i, e_j) = \delta_{ij}$. Definiendo ahora

$$e_n = b'_n - \sum_{i=1}^{n-1} \alpha_i e_i$$

con $\alpha_i = A(e_i, b'_n)$, obtenemos $A(e_i, e_n) = A(e_i, b'_n) - \alpha_i = 0$ para $i = 1, \ldots, n-1$. Se obtiene así una base canónica $e = (e_1, \ldots, e_{n-1}, e_n)$ de V en la que $A(e_i, e_j) = \delta_{ij} A(e_i, e_i)$, con $A(e_i, e_i) = 1$ si $i \leq n-1$ y entonces $A(e_n, e_n) = |[A]_e| > 0$ (pues $[A]_e = S^{\rm tr}[A]_{B'} S$ y $|[A]_e| = |S|^2 |[A]_{B'}| > 0$). La forma cuadrática es pues definida positiva.

Aplicaciones:
1) Clasificación de puntos críticos:
Consideremos un campo escalar $G : \mathbb{R}^n \to \mathbb{R}$ derivable a segundo orden orden en un entorno de un punto crítico $\vec{r}_0$ donde $\frac{\partial G}{\partial x_i}|_{\vec{r}=\vec{r}_0} = 0$, $i = 1, \ldots, n$. El polinomio de Taylor de segundo orden de $\Delta G(\vec{r}) = G(\vec{r}) - G(\vec{r}_0)$ alrededor de $\vec{r}_0$ es una forma cuadrática en $\Delta \vec{r} = \vec{r} - \vec{r}_0 = (\Delta x_1, \ldots, \Delta x_n)$:

$$\Delta G = \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 G}{\partial x_i \partial x_j}|_{\vec{r}=\vec{r}_0} \Delta x_i \Delta x_j + R_3 = \frac{1}{2}(\Delta \vec{r}) H (\Delta \vec{r})^t + R_3$$

donde H es una matriz simétrica de $n \times n$, denominada matriz *Hessiana*, de elementos

$$H_{ij} = \frac{\partial^2 G}{\partial x_i \partial x_j}|_{\vec{r}=\vec{r}_0}$$

y R_3 es el resto ($\lim_{\vec{r}\to\vec{r}_0} R_3/|\vec{r}-\vec{r}_0|^2 = 0$). Llevando la forma cuadrática anterior a una forma canónica (ya sea completando cuadrados o diagonalizando la matriz H), obtenemos

$$\Delta G = \frac{1}{2}\sum_{i=1}^{n} a'_i(\Delta x'_i)^2 + R_3$$

Si $a'_i > 0$ para $i = 1, \ldots, n$, $\Delta G > 0$ para $|\Delta\vec{r}|$ suf. pequeño y el punto crítico es un mínimo local o relativo. Si $a'_i < 0$ para $i = 1, \ldots, n$, $\Delta G < 0$ para $|\Delta\vec{r}|$ suf. pequeño y el punto crítico es un máximo local o relativo. Y si existen a'_i positivos y negativos, se trata de un punto silla ("saddle point").
Finalmente, si algunos a'_i son nulos y $a'_i \geq 0$ para $i = 1, \ldots, n$ (o $a'_i \leq 0$ para $i = 1, \ldots, n$) el presente criterio no decide y es necesario un desarrollo a orden más alto (que puede también no ser concluyente) o bien un análisis alternativo.
Por lo tanto, podemos clasificar el punto crítico en forma inmediata conociendo los autovalores de la matriz H (de $n \times n$), o bien simplemente completando cuadrados y observando los signos de los coeficientes diagonales a_i. El último método es en general más sencillo (pues no requiere determinar raíces de ninguna ecuación) pero el primero tiene la ventaja de determinar a la vez (mediante los autovectores de H) n direcciones *ortogonales* en las que la forma cuadrática tiene la forma canónica (y por lo tanto conocer las direcciones ortogonales en las que ΔG es positivo ($a'_i > 0$) o negativo ($a_i < 0$)). (Ver práctica para más detalles).

Notemos también que si definimos $f_{\vec{r}_0} : \mathbb{R} \to \mathbb{R}$ como

$$f_{\vec{r}_0}(t) = G(\vec{r}_0 + t\Delta\vec{r})$$

entonces

$$f''_{\vec{r}_0}(0) = \sum_{i,j} \frac{\partial^2 G}{\partial x_i \partial x_j}|_{\vec{r}=\vec{r}_0} \Delta x_i \Delta x_j = (\Delta\vec{r})H(\Delta\vec{r})^t$$

lo cual es una forma cuadrática en $\Delta\vec{r}$ definida por la matriz simétrica H. Si H es definida positiva $\Rightarrow f_{\vec{r}_0}(t)$ es cóncava hacia arriba en $t = 0$ para *cualquier dirección* $\Delta\vec{r}$, mientras que si es definida negativa, $f_{\vec{r}_0}(t)$ será cóncava hacia abajo para cualquier dirección $\Delta\vec{r}$. En el caso general, la concavidad dependerá de la dirección de $\Delta\vec{r}$.

2) Clasificación de curvas de nivel de formas cuadráticas. Consideremos la ecuación

$$\sum_{i,j=1}^{n} x_i a_{ij} x_j = C$$

que puede reescribirse como

$$\vec{r}A(\vec{r})^t = C$$

con $\vec{r} = (x_1, \ldots, x_n)$ y A la matriz (real) de elementos a_{ij}, que puede suponerse siempre simétrica ($a_{ij} = a_{ji}$). Llevándola a una forma canónica obtenemos la ecuación equivalente

$$\sum_{i=1}^{n} a'_i {x'_i}^2 = C$$

con los x'_i relacionados linealmente con los x_i. Si todos los a'_i son positivos (y $C > 0$) la ecuación anterior determina un *elipsoide*, mientras que si los a'_i tienen signos distintos la ec. determina un hiperboloide. Si la forma canónica se obtiene diagonalizando la matriz A, los autovectores pueden elegirse normalizados y ortogonales, en cuyo caso las variables x'_i serán las coordenadas a lo largo de ejes ortogonales en los que la forma cuadrática toma la forma canónica (ejes principales). (véase práctica para más detalles).

Ejemplo: Consideremos

$$G(x, y) = x^2 + y^2 + 2\alpha xy$$

$(0,0)$ es un pto. crítico de G y la matriz H de derivadas segundas es $H=2(\begin{smallmatrix}1&\alpha\\\alpha&1\end{smallmatrix})$. Sus autovalores son

$$\lambda_\pm = 2(1\pm\alpha)$$

(obtenidos de la ec. $|H-\lambda I_2|=(2-\lambda)^2-4\alpha^2=0$). Por lo tanto, Si $|\alpha|<1$ ambos autovalores son positivos y $(0,0)$ es un mínimo de G (en este caso mínimo absoluto). En cambio, si $|\alpha|>1$, un autovalor es positivo y el otro negativo (por ej., si $\alpha>1$, $\lambda_+>0$, $\lambda_-<0$), por lo que $(0,0)$ es en este caso un punto silla. Las componentes de los autovectores normalizados (y por su puesto ortogonales) de H son $[v_\pm]_e=(\begin{smallmatrix}\pm1\\1\end{smallmatrix})/\sqrt{2}$, por lo que $S=(\begin{smallmatrix}1&-1\\1&1\end{smallmatrix})/\sqrt{2}$ y podemos escribir

$$G(x,y)=(1+\alpha)x'^2+(1-\alpha)y'^2$$

con $(\begin{smallmatrix}x'\\y'\end{smallmatrix})=S^{-1}(\begin{smallmatrix}x\\y\end{smallmatrix})=(\begin{smallmatrix}x+y\\x-y\end{smallmatrix})/\sqrt{2}$, como puede verificarse directamente.
Si G representa la energía potencial de un sistema físico dependiente de dos coordenadas x,y en las cercanías de un punto estacionario, vemos pues que el sistema será estable sólo si $|\alpha|<1$. Si $\alpha>0$, la estabilidad del sistema en la dirección de e'_2 disminuye al aumentar α, tornándose inestable para $\alpha>1$.

Cabe destacar, no obstante, que la misma conclusión puede obtenerse simplemente completando cuadrados, lo cual conduce a

$$G(x,y)=(x+\alpha y)^2+y^2(1-\alpha^2)$$

Vemos pues que el coeficiente de y^2 es positivo si $|\alpha|<1$ y negativo si $|\alpha|>1$, mientras que el primero es siempre positivo.

Si consideramos ahora la ec.

$$x^2+y^2+2\alpha xy=C$$

el mismo análisis conduce a que para $C>0$, la ec. anterior representa una elipse si $|\alpha|<1$, con ejes principales inclinados 45 grados respecto de los originales (y radios de longitud $1/\sqrt{1\pm\alpha}$ para $C=1$), mientras que si $|\alpha|>1$ la ec. representa una hipérbola.
Ejemplo: Consideremos

$$G(x,y,z)=x^2+y^2+z^2+2xy+2xz$$

que ya fue analizado. $(0,0,0)$ es claramente un punto crítico. Completando cuadrados, se obtiene

$$G(x,y,z)=(x+y+z)^2-(y+z)^2+y^2+z^2=(x+y+z)^2-2yz=x'^2+2y'^2-2z'^2$$

donde $x'=x+y+z$, $z'=(z+y)/2$, $y'=(y-z)/2$, lo que implica que $(0,0,0)$ es un punto silla. El mismo resultado se obtiene de los autovalores de la matriz $H=2\begin{pmatrix}1&1&1\\1&1&0\\1&0&1\end{pmatrix}$, que son $\lambda_1=2>0$, $\lambda_2=2+2\sqrt{2}>0$, $\lambda_3=2-2\sqrt{2}<0$.
La ecuación

$$x^2+y^2+z^2+2xy+2xz=C$$

corresponde, por lo tanto, a un hiperboloide (de una hoja para $C>0$).

Ejemplo: Consideremos la función $T:\mathbb{R}^n\to\mathbb{R}$ dada por $\quad (X=(x_1,\ldots,x_n)^t)$

$$T(X)=\sum_{i,j}x_iA_{ij}x_j+2\sum_i r_ix_i$$

con $A_{ij}=A_{ji}$. Asumiendo que la matriz de coeficientes $A\in\mathbb{R}^{n\times n}$ es invertible, podemos reescribir T como

$$T(X)=X^tAX+(R^tX+X^tR)=Y^TAY-R^tA^{-1}R$$

donde $X^t=(x_1,\ldots,x_n)$, $R^t=(r_1,\ldots,r_n)$ y $Y=X+C$, con $C=A^{-1}R$. Es decir, $T(X)$ es una forma cuadrática en $Y=X+A^{-1}R$ (o sea, $y_i=x_i+\sum_j A^{-1}_{ij}r_j$) más una constante $R^tA^{-1}R$.

Si A es singular, podemos econtrar C tal que $AC=R$ sólo si $R\in EC(A)$ (espacio columna de A). En tal caso $T=Y^tAY-C^tR$ sigue siendo una forma cuadrática en $Y=X+C$, a menos de una constante $-C^tR$.

19. Espacios Euclídeos

Un espacio vectorial V sobre el cuerpo de los reales $\mathbb{R}$ es *Euclídeo* si está equipado con una operación denominada *producto escalar* y denotada por (v, w), que asigna a todo par de vectores un escalar real que satisface

$$(v, w) = (w, v) \;\; \forall\, v, w \in V$$

$$(v, w_1 + w_2) = (v, w_1) + (v, w_2), \;\; \forall\, v, w_1, w_2 \in V$$

$$(\alpha v, w) = \alpha(v, w) \;\; \forall\, v, w \in V, \; \alpha \in \mathbb{R}$$

$$(v, v) > 0 \; \forall v \neq 0, \;\; (0, 0) = 0$$

El producto escalar en un espacio euclídeo es pues *una forma bilineal simétrica* de $V \times V$ sobre $\mathbb{R}$ tal que la correspondiente forma cuadrática es *definida positiva.* Cualquier forma bilineal de este tipo es apta para definir un producto escalar. Notemos que $(0, v) = (v, 0) = 0 \;\forall\; v \in V$.

En un espacio de dimensión finita generado por una base $B = (b_1, \ldots, b_n)$, se obtiene, eligiendo para el producto escalar una forma bilineal simétrica G asociada a una forma cuadrática definida positiva,

$$(v, w) = G(v, w) = [v]_B^t [G]_B [w]_B = \sum_{i,j=1}^{n} \alpha_i g_{ij} \beta_j, \qquad g_{ij} = ([G]_B)_{ij} = (b_i, b_j) = g_{ji}$$

donde $v = \sum_i \alpha_i b_i$, $w = \sum_i \beta_i b_i$ y $[v]_B^t = (\alpha_1, \ldots, \alpha_n)$, $[w]_B^t = (\beta_1, \ldots, \beta_n)$. Recordemos que para este tipo de formas bilineales es siempre posible elegir una base B donde $[G]_B$ es diagonal, es decir, $(b_i, b_j) = g_i \delta_{ij}$, en cuyo caso el producto escalar toma la forma

$$(v, w) = \sum_{i,j=1}^{n} \alpha_i g_i \beta_i, \qquad g_i = (b_i, b_i) > 0$$

Definiendo ahora $e_i = b_i / \sqrt{g_i}$, podemos obtener así una *base canónica* $e = (e_1, \ldots, e_n)$ en la que $(e_i, e_j) = \delta_{ij}$ y por lo tanto $[G]_e = I_n$ (matriz identidad). El producto escalar en esta base adopta entonces la forma usual

$$(v, w) = [v]_e^t [w]_e = \sum_{i=1}^{n} \alpha_i \beta_i$$

A una base de este tipo la denominaremos *base canónica* o *base ortonormal* del espacio euclídeo.
Ejemplo 1: Si $V = \mathbb{R}^n$ y $v = (x_1, \ldots, x_n)$, $v' = (x'_1, \ldots, x'_n)$, el producto escalar usual, dado por

$$(v, v') = v \cdot v' = \sum_{i=1}^{n} x_i x'_i$$

satisface las 4 condiciones requeridas. Los vectores de la base canónica $e_1 = (1, 0, \ldots, 0) \ldots e_n = (0, \ldots, 0, 1)$ satisfacen $(e_i, e_j) = \delta_{ij}$ y forman pues una base ortonormal para este producto escalar.
Ejemplo 2: Si V es el espacio $C_{[a,b]}$ de funciones reales continuas $f : [a, b] \to \mathbb{R}$ (de dimensión infinita), podemos equiparlo con el producto escalar definido por

$$(f, g) = \int_a^b f(x) g(x) dx$$

que satisface todas las condiciones requeridas (probar como ejercicio).
Ejemplo 3: Si $V = \mathbb{R}^{m \times n}$ es el espacio de matrices reales de $m \times n$, podemos definir el producto escalar de dos matrices $A, B \in V$ como

$$(A, B) = \mathrm{Tr}\,[A^t B] = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} B_{ij} = (B, A)$$

que satisface también todas las condiciones requeridas (probar como ejercicio).

19.1 Norma de un vector

La **norma** (o longitud) de un vector $v \in V$ se define como

$$||v|| = \sqrt{(v,v)}$$

y satisface $||v|| \geq 0 \ \forall \ v \in V$, con $||v|| = 0$ si y sólo si $v = 0$. Por ejemplo, utilizando los productos escalares anteriores, en $V = \mathbb{R}^n$ se obtiene

$$||v|| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

mientras que en $V = C_{[a,b]}$,

$$||f|| = \sqrt{\int_a^b f^2(x)dx}$$

y en $V = \mathbb{R}^{m\times n}$,

$$||A|| = \sqrt{\mathrm{Tr}\,[A^t A]} = \sqrt{\sum_{i,j} A_{ij}^2}$$

Todo vector en un espacio euclídeo posee pues una norma, que es positiva si $v \neq 0$ y 0 si $v = 0$. Notemos que $\forall \alpha \in \mathbb{R}$ se cumple

$$||\alpha v|| = \sqrt{(\alpha v, \alpha v)} = \sqrt{\alpha^2 (v,v)} = |\alpha| ||v||$$

de modo que la norma de αv es $|\alpha|$ veces la longitud de v.
Un vector de norma 1 se denomina vector unitario. Todo vector v no nulo puede ser *normalizado*, es decir, convertido en vector unitario mediante la multiplicación por un escalar: Si $||\alpha v|| = |\alpha| ||v|| = 1 \Rightarrow$ basta con elegir α tal que $|\alpha| = 1/||v||$, o sea, $\alpha = \pm 1/||v||$. El vector normalizado con el mismo sentido de v es pues

$$v_n = v/||v||$$

Un conjunto C de V se dice que es **acotado** si existe $m \in \mathbb{R}$ tal que $||v|| < m \ \forall \ v \in C$. El conjunto $\{v, \ ||v|| \leq 1\}$ se l lama bola unidad, mientras que el conjunto $\{v, \ ||v|| = 1\}$ esfera unidad. Estos conjuntos no son subespacios (como el lector podrá fácilmente mostrar).

19.2 Desigualdad de Cauchy-Schwarz y ángulo entre vectores

Dados dos vectores v, w de un espacio euclídeo V, se cumple siempre la **desigualdad de Cauchy-Schwartz**

$$|(v,w)| \leq ||v|| \, ||w|| \qquad (19.1)$$

donde la igualdad rige si y sólo si v y w son LD (Linealmente Dependientes).
Demostración: Si v, w son LD $\Rightarrow v = \alpha\omega$ (o $w = \gamma v$) en cuyo caso $|(v,w)| = |\alpha||(w,w)| = |\alpha| \, ||w||^2 = ||v|| \, ||w||$. Esto incluye en particular el caso en que v o w es nulo ($\alpha = 0$ o $\gamma = 0$).
Si $v \neq 0$ y $w \neq 0$, se obtiene, para los correspondientes vectores normalizados $v_n = v/||v||$, $w_n = w/||w||$,

$$0 \leq ||v_n \pm w_n||^2 = (v_n \pm w_n, v_n \pm w_n) = (v_n, v_n) + (w_n, w_n) \pm 2(v_n, w_n) = 2(1 \pm (v_n, w_n))$$

lo que implica $-1 \leq (v_n, w_n) \leq 1$, o sea,

$$|(v_n, w_n)| \leq 1$$

de donde $|(v,w)| \leq ||v|| \, ||w||$, como se quería demostrar. Vemos también que la igualdad ($|(v_n, w_n)| = 1$) implica $||v_n \pm w_n||^2 = 0$ y por lo tanto $v_n \pm w_n = 0$, en cuyo caso v y w son LD.

Por ejemplo, en $\mathbb{R}^n$, la desigualdad de Cauchy-Schwarz implica, para $v = \sum_{i=1}^n x_i e_i$, $w = \sum_{i=1}^n y_i e_i$,

$$|\sum_{i=1}^{n} x_i x'_i| \leq \sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} x'^2_i}$$

y en $C_{[a,b]}$,

$$|\int_a^b f(x)g(x)dx| \leq \sqrt{\int_a^b f^2(x)dx}\sqrt{\int_a^b g^2(x)dx}$$

mientras que en $V = \mathbb{R}^{m\times n}$,

$$\mathrm{Tr}[A^tB] \leq \sqrt{\mathrm{Tr}[A^tA]}\sqrt{\mathrm{Tr}[B^tB]}$$

El **ángulo** θ entre dos vectores v, w no nulos se define como

$$\cos\theta = \frac{(v,w)}{||v||\,||w||} = (v_n, w_n)$$

donde $v_n = v/||v||$, $w_n = w/||w||$ son los vectores normalizados. La desigualdad de Cauchy-Schwartz asegura que $-1 \leq (v_n, w_n) \leq 1$, por lo que el ángulo θ está correctamente definido. Notemos que si $v = \alpha w$ entonces entonces $\theta = 0$ ($\alpha > 0$) o π ($\alpha < 0$).
Ejercicio: Determinar el ángulo entre los vectores $(1, 1, \ldots, 1)$ y $(-1, 1, \ldots, 1)$ pertenecientes a $\mathbb{R}^n$.

19.3 Desigualdad triangular y distancia entre vectores

La desigualdad de Cauchy-Schwarz permite demostrar en forma inmediata la desigualdad triangular

$$|||v|| - ||w||| \leq ||v + w|| \leq ||v|| + ||w||$$

ya que $||v + w||^2 = (v + w, v + w) = (v, v) + (w, w) + 2(v, w)$, y por lo tanto,

$$||v + w||^2 \leq ||v||^2 + 2||v||\,||w|| + ||w||^2 = (||v|| + ||w||)^2$$

$$||v + w||^2 \geq |v||^2 - 2||v||\,||w|| + ||w||^2 = (||v|| - ||w||)^2$$

Notemos que se cumple también (dado que $|| - w|| = ||w||$) $|||v|| - ||w||| \leq ||v - w|| \leq ||v|| + ||w||$.
La **distancia** entre dos vectores $d(v, w)$ se define como

$$d(v, w) = ||v - w||$$

y satisface las propiedades

$$d(v, w) \geq 0,\ \text{con } d(v, w) = 0 \text{ sii } v = w$$

$$d(v, w) = d(w, v)$$

$$d(v, w) \leq d(v, u) + d(u, w)$$

donde la última es consecuencia de la desigualdad triangular: $||v-w|| = ||v-u-(w-u)|| \leq ||v-u||+||w-u||$.

19.4 Ortogonalidad y bases ortonormales

Dos vectores $v, w \in V$ se dicen **ortogonales** si $(v, w) = 0$. En tal caso, si $v \neq 0$, $w \neq 0$, $\cos\theta = 0$ y por lo tanto, $\theta = \pi/2$.

Un conjunto de m vectores v_i son **ortogonales** si $(v_i, v_j) = 0\ \forall\ i \neq j$, es decir, si son mutuamente ortogonales de a pares. Y se dicen que son **ortonormales** si además tienen norma no nula e igual a 1: $(v_i, v_i) = 1$, $i = 1, \ldots, n$. Una base ortonormal es una base compuesta por vectores ortonormales. La base canónica en la que $(e_i, e_j) = \delta_{ij}$ es pues una base *ortonormal*.

Independencia lineal de vectores ortogonales: Si $v_1, v_2, \ldots, v_m$ son mutuamente ortogonales ($(v_i, v_j) = 0$ si $i \neq j$) y no nulos ($||v_i||^2 = (v_i, v_i) > 0$) $\Rightarrow$ *son linealmente independientes.*
Demostración: Si

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$$

multiplicando escalarmente por v_i, con $1 \leq i \leq m$, y teniendo en cuenta que $(v_i, v_j) = 0$ si $i \neq j$, se obtiene

$$(v_i, \alpha_1 v_1 + \ldots + \alpha_n v_n) = (v_i, \alpha_i v_i) = \alpha(v_i, v_i) = (v_i, 0) = 0$$

lo que implica $\alpha_i = 0$ pues $(v_i, v_i) = ||v_i||^2 > 0$. Esto muestra que son LI. La prop. recíproca no es, obviamente, válida.

Por lo tanto, si dim $V = n \Rightarrow$ cualquier conjunto de n vectores ortogonales no nulos forma una base de V.
Generalización del teorema de Pitágoras: Si v_1, v_2 son ortogonales ($(v_1, v_2) = 0) \Rightarrow$

$$||v_1 + v_2||^2 = (v_1 + v_2, v_1 + v_2) = (v_1, v_1) + (v_2, v_2) + 2(v_1, v_2) = ||v_1||^2 + ||v_2||^2$$

Y si $(v_1, v_2, \ldots, v_m)$ son mutuamente ortogonales ($(v_i, v_j) = 0$ si $i \neq j$), $\Rightarrow$

$$||\sum_{i=1}^{m} v_i||^2 = (\sum_{i=1}^{m} v_i, \sum_{j=1}^{m} v_j) = \sum_{i,j=1}^{m} (v_i, v_j) = \sum_{i=1}^{m} (v_i, v_i) = \sum_{i=1}^{m} ||v_i||^2$$

Expansión en una base ortonormal: Si escribimos, para un vector $v \in V$,

$$v = \sum_{i=1}^{n} x_i e_i$$

donde $(e_1, \ldots, e_n)$ es una base ortonormal de V, entonces

$$x_i = (e_i, v), \quad i = 1, \ldots, n$$

ya que $(e_i, v) = (e_i, \sum_{j=1}^{n} x_j e_j) = \sum_{j=1}^{n} x_j (e_i, e_j) = x_i$ por ortonormalidad de los e_i. Las coordenadas x_i de v en la base canónica se obtienen pues simplemente efectuando el producto escalar (e_i, v), no siendo necesario resolver explícitamente un sistema de ecuaciones lineales para su obtención. Además, por la generalización del teorema de Pitágoras anterior,

$$||v||^2 = \sum_{i=1}^{n} ||x_i e_i||^2 = \sum_{i=1}^{n} x_i^2$$

Los ángulos que forma v con e_i están determinados por

$$\cos(\theta_i) = \frac{(e_i, v)}{||e_i|| \, ||v||} = x_i / ||v||$$

(ángulos directores) y satisfacen

$$\sum_{i=1}^{n} \cos^2(\theta_i) = \sum_{i=1}^{n} x_i^2 / ||v||^2 = 1$$

Se cumple entonces $x_i = ||v|| \cos \theta_i$ para $i = 1, \ldots, n$.

Si una base e' es ortogonal pero no necesariamente ortonormal, entonces

$$x_i = (e'_i, v) / ||e'_i||^2$$

con $||v||^2 = \sum_{i=1}^{n} ||x_i e'_i||^2 = \sum_{i=1}^{n} x_i^2 ||e'_i||^2$ y $\cos(\theta_i) = \frac{(e'_i, v)}{||e'_i|| \, ||v||} = x_i ||e'_i|| / ||v||$. Se sigue cumpliendo que $\sum_{i=1}^{n} \cos^2(\theta_i) = 1$.

Notemos también que si $F : V \to W$ es una transformación lineal entre espacios euclídeos V y W de dimensiones n y m respectivamente, y e, $\tilde{e}$ son bases ortonormales de V y W, entonces los elementos F_{ij} de la matriz $[F]_{\tilde{e}}^{e} \in \mathbb{R}^{m \times n}$ que representa a F en estas bases están dados por el producto escalar

$$F_{ij} = (\tilde{e}_i, F(e_j))$$

dado que por definición, $F(e_j) = \sum_{i=1}^{m} F_{ij} \tilde{e}_i$.

Relación entre bases ortonormales.
Si e es una base ortonormal de V y e' es otra base de V, tenemos

$$e'_j = \sum_{i=1}^{n} S_{ij} e_i, \quad j = 1, \ldots, n$$

con $S_{ij} = (e_i, e'_j)$ y

$$(e'_j, e'_k) = \sum_{i=1}^{n} S_{ij} S_{ik} = (S^t S)_{jk}$$

Vemos que e' será una base ortonormal ($(e'_j, e'_k) = \delta_{jk}$) *si y sólo si* la matriz de cambio de base S satisface

$$S^t S = I_n$$

o sea, $S^{-1} = S^t$. Las matrices reales que satisfacen esta relación se denominan *ortonormales* (o a veces ortogonales). Dado que $(S^t S)_{ij}$ es el producto escalar de la columna i por la columna j de S, las columnas de estas matr ices son *ortonormales* ($(S^t S)_{ij} = \delta_{ij}$) formando entonces una base ortonormal de $\mathbb{R}^n$. Como la ec. anterior implica asimismo $SS^t = I_n$, las filas de S *son también ortonormales* y forman asimismo una base ortonormal de $\mathbb{R}^n$ (se prueba de la misma manera).
Notemos además que $|S| \equiv \text{Det}S = \pm 1$, pues $|S^t S| = |S|^2 = 1$.
Resumiendo, la base e' será ortonormal sii la matriz de cambio de base S es una matriz ortonormal.

Para un vector arbitrario $v = \sum_{i=1}^n x_i e_i = \sum_{i=1}^n x'_i e'_i$, tenemos entonces

$$x'_i = (e'_i, v) = \sum_{j=1}^n x_j (e'_i, e_j) = \sum_{j=1}^n S^t_{ij} x_j$$

es decir,

$$[v]_{e'} = S^t [v]_e$$

lo que esta de acuerdo con la relación general $[v]_{e'} = S^{-1}[v]_e$.

19.5 Teorema de ortogonalización de Gram-Schmidt

Las propiedades anteriores muestran claramente la ventaja de trabajar con bases y conjuntos ortonormales. Daremos ahora un método general para construir bases ortogonales de espacios y subespacios.
Sean $v_1, \ldots, v_m$ m vectores LI $\in V$, que generan un subespacio $S \subset V$ de dimensión $m \leq n = \dim V$. Entonces existen m vectores *ortogonales* no nulos $w_1, \ldots, w_m$ que general el mismo espacio (y que son, por lo tanto, combinaciones lineales de los $v_1, \ldots, v_m$).
La demostración es directamente constructiva. Comencemos con $w_1 = v_1$. Definimos luego

$$w_2 = v_2 - \alpha w_1$$

y exigimos que $0 = (w_1, w_2) = (w_1, v_2) - \alpha(w_1, w_1)$. Por lo tanto $\alpha = (w_1, v_2)/||w_1||^2$ y

$$w_2 = v_2 - \frac{(w_1, v_2)}{||w_1||^2} w_1$$

Análogamente, definimos

$$w_3 = v_3 - \alpha_2 w_2 - \alpha_1 w_1$$

Las condiciones $0 = (w_2, w_3) = (w_2, v_3) - \alpha_2 ||w_2||^2$, $0 = (w_1, w_3) = (w_1, v_3) - \alpha_1 ||w_1||^2$ (donde hemos utilizado la ortogonalidad $(w_1, w_2) = (w_2, w_1) = 0$) implican $\alpha_i = (w_i, v_i)/||w_i||^2$ para $i = 1, 2$, y por tanto

$$w_3 = v_3 - \frac{(w_2, v_3)}{||w_2||^2} w_2 - \frac{(w_1, v_3)}{||w_1||^2} w_1$$

En general, definiendo para $i = 2, \ldots, m$,

$$w_i = v_i - \sum_{j=1}^{i-1} \alpha_j w_j,$$

las $i - 1$ condiciones $(w_j, w_i) = 0$ para $j = 1, \ldots, i - 1$ implican $\alpha_j = \frac{(w_j, v_i)}{||w_j||^2}$, teniendo en cuenta la ortogonalidad $(w_j, w_k) = 0$ si $j < k < i$.

Por lo tanto,

$$w_1 = v_1, \quad w_i = v_i - \sum_{j=1}^{i-1} \frac{(w_j, v_i)}{||w_j||^2} w_j, \quad i = 2, \ldots, m$$

Los m vectores w_i así definidos son *no nulos*: si $w_i = 0 \Rightarrow v_i = \sum_{j=1}^{i-1} \alpha_j w_j$, lo que implicaría, dado que los w_j son combinaciones lineales de los v_j, que los vectores originales son LD, contradiciendo la hipótesis.

Los m vectores w_i así construidos son entonces mutuamente ortogonales por construcción ($(w_i, w_j) = 0$ si $i \neq j$) y no nulos, por lo que son LI, conformando entonces una base de S. Si $m = n$, se obtiene así un método para construir una base *ortogonal* del espacio completo V. Notemos que

$$||w_i||^2 = (w_i, w_i) = (w_i, v_i) = ||v_i||^2 - \sum_{j=1}^{i-1} (w_j, v_i)^2/||\omega_j||^2$$

Para obtener un conjunto ortonormal, se puede normalizar al final del procedimiento ($w_i \to w_i' = w_i/||w_i||$) o en cada paso. En este último caso, el método se resume en

$$w_1' = v_1/||v_1||, \quad w_i' = [v_i - \sum_{j=1}^{i-1} (w_j', v_i) w_j']/\,[||v_i||^2 - \sum_{j=1}^{i=1} (w_j', v_i)^2]^{1/2}, \ \ i = 2, \ldots, m$$

Ejemplo: Sean $v_1 = (1, 1, 1)$, $v_2 = (1, 1, -1)$ vectores de $\mathbb{R}^3$, no ortogonales ($(v_1, v_2) = 1$, con $(v_1, v_1) = (v_2, v_2) = 3$). Aplicando el método de Gram-Schmidt, se obtiene

$$w_1 = (1, 1, 1), \quad w_2 = (1, 1, -1) - \frac{1}{3}(1, 1, 1) = (2, 2, -4)/3$$

que son claramente ortogonales.
Para formar una base ortogonal de $\mathbb{R}^3$ que contenga a w_1 y w_2, podemos considerar un vector cualquiera v_3 tal que (w_1, w_2, v_3) sean LI. Por ejemplo, $v_3 = (1, 0, 0)$. Se obtiene entonces el resultado esperado

$$w_3 = v_3 - \frac{(w_1, v_3)}{||w_1||^2} w_1 - \frac{(w_2, v_3)}{||w_2||^2} w_2 = (1, 0, 0) - \frac{1}{3}(1, 1, 1) - \frac{1}{3}\frac{2/3}{24/9}(2, 2, -4) = \frac{1}{2}(1, -1, 0)$$

Ejemplo: Sean $p_1(t) = 1$, $p_2(t) = t$, $p_3 = t^2$ vectores de P_2 (polinomios de grado ≤ 2). Determinar una base ortogonal de P_2 para el producto escalar $(p, q) = \int_{-1}^{1} p(t)q(t)dt$.
Aplicando el método anterior, obtenemos, notando que $(p_1, p_1) = 2$, $(p_2, p_2) = 2/3$, $(p_3, p_3) = 2/5$, $(p_1, p_2) = 0 = (p_2, p_3)$, $(p_1, p_3) = 2/3$,

$$w_1(t) = 1, \ \ w_2(t) = t, \ \ w_3 = t^2 - \frac{2/3}{2} = t^2 - 1/3$$

Si exigimos que $w_i(1) = 1$ y extendemos $P_2 \to P_\infty$ se obtienen de esta manera los polinomios de Legendre: $P_1(t) = 1$, $P_2(t) = t$, $P_3(t) = (3t^2 - 1)/2$, etc.
De la misma manera, para productos escalares del tipo $(p, q) = \int_a^b p(t)q(t)\rho(t)dt$, donde $\rho(t) > 0$ para $t \in (a, b)$, se obtienen otras familias de polinomios ortogonales.

19.6 Proyección ortogonal

Sea w un vector no nulo $\in V$ y sea $v \in V$. Podemos descomponer v como una suma de un vector v_w paralelo a w y un vector $v - v_w$ *ortogonal* a w:

$$v = v_w + (v - v_w)$$

donde exigimos $(w, v - v_w) = 0$. Esta condición determina v_w. Escribiendo $v_w = \alpha w$, obtenemos

$$(w, v - \alpha w) = (w, v) - \alpha(w, w) = 0$$

por lo que $\alpha = (w, v)/||w||^2$ y

$$v_w = \frac{(w, v)}{||w||^2} w$$

El vector v_w es la *proyección ortogonal* de v sobre w y su significado geométrico es muy claro (recordar dibujo): Si trazamos la perpendicular desde el extremo de v a la recta generada por w, obtenemos un triángulo rectángulo formado por v, v_w y $v - v_w$, siendo $v_w \perp v - v_w$. Así, $v_w = 0$ si $v \perp w$, y $v_w = v$ si $v || w$.

El vector v_w puede también interpretarse como el vector paralelo a w cuya distancia a v es **mínima**. En efecto, si $u_w = \alpha w$,

$$d^2(v, u_w) = ||v - u_w||^2 = ||v - v_w + (v_w - u_w)||^2 = ||v - v_w||^2 + ||v_w - u_w||^2 + 2(v - v_w, v_w - u_w)$$

Pero el último término es nulo pues $v - v_w$ es $\perp$ a w y por tanto a $v_w - u_w$, por lo que

$$||v - u_w||^2 = ||v - v_w||^2 + ||v_w - u_w||^2 \geq ||v - v_w||^2$$

La distancia mínima se obtiene pues para $u_w = v_w$.

Operador de Proyección: El operador de proyección sobre w queda definido por

$$P_w(v) = v_w$$

y es un operador lineal que satisface $P_w^2 = P_w$. En una base canónica de V, $(w, v) = [w]_e^t [v]_e$, $||w||^2 = [w]_e^t [w]_e$ y entonces

$$[v_w]_e = \frac{[w]_e^t [v]_e}{[w]_e^t [w]_e} [w]_e = \frac{[w]_e [w]_e^t}{[w]_e^t [w]_e} [v]_e$$

La matriz $[P]_e$ que representa a P en una base canónica ($[v_w]_e = [P]_e [v]_e$) está entonces dada por

$$[P]_e = \frac{[w]_e [w]_e^t}{[w]_e^t [w]_e}$$

Ejemplo: Proyectar el vector $v = (1, 1, 1)$ sobre $w = (1, 1, -1)$.
Tenemos, como $(v, w) = 1$ y $||w||^2 = 3$,

$$v_w = \frac{1}{3}(1, 1, -1)$$

El operador de proyección correspondiente queda definido, para $v = (x, y, z) = xe_1 + ye_2 + ze_3$ (base canónica), por

$$v_w = P_w(v) = \frac{(w, v)}{||w||^2} w = \frac{x + y - z}{3}(1, 1, -1)$$

y la correspondiente matriz es entonces

$$[P_w]_e = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} (1, 1, -1) = \frac{1}{3} \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

de forma que $[v_w]_e = [P_w]_e [v]_e$.
Gram-Schmidt en términos de proyectores:
El proceso de ortogonalización de Gram-Schmidt puede ahora escribirse como

$$w_1 = v_1, \quad w_i = v_i - \sum_{j=1}^{i-1} P_{w_j}(v_i), \quad i = 2, \ldots, m$$

El significado es muy claro: w_i se construye a partir de v_i quitándole a este último las proyecciones sobre cada uno de los vectores anteriores w_j, $j < i$. De esta forma w_i sólo conserva la parte de v_i ortogonal al espacio generado por los w_j.

La expansión de un vector en una base ortonormal puede entonces verse también como la suma de proyecciones ortogonales: Tenemos, para $v \in V$ y $(e_1, \ldots, e_n)$ una base ortonormal,

$$v = \sum_{i=1}^{n} x_i e_i = \sum_{i=1}^{n} P_{e_i}(v)$$

ya que $x_i e_i = (e_i, v) e_i = P_{e_i}(v)$.

19.7 Subespacios ortogonales
El conjunto de vectores ortogonales a un cierto vector v es un subespacio de V: Si $(v,w_1)=0$, $(v,w_2)=0$ $\Rightarrow (v,w_1+w_2)=(v,w_1)+(v,w_2)=0$ y $(v,\alpha w_1)=\alpha(v,w_1)=0$. Además es no vacío pues $(v,0)=0$.
El conjunto de vectores ortogonales a todos los vectores de un cierto subespacio $S\subset V$ es también un subespacio (se prueba de la misma forma), denominado *complemento ortogonal* de S o $S_\perp$.

Mostraremos a continuación que $V=S\oplus S_\perp$.
Demostración: Sea $v\in V$ y v_s un vector $\in S$. Mostraremos que es siempre posible escribir

$$v=v_s+(v-v_s)$$

con $v_s\in S$ y $v-v_s\in S_\perp$. Si $(w_1,\ldots,w_m)$ es una base de S, que podemos escogerla ortogonal utilizando el método de Gram-Schmidt, entonces

$$v_s=\sum_{i=1}^m \alpha_i w_i,\quad \alpha_i=(w_i,v_s)/||w_i||^2$$

La condición $v-v_s\in S_\perp$ implica entonces

$$0=(w_i,v-v_s)=(w_i,v)-(w_i,v_s)=(w_i,v)-\alpha_i||w_i||^2,\ \ i=1,\ldots,m$$

o sea, $\alpha_i=(w_i,v)/||w_i||^2$. En tal caso, $(v-v_s)$ será también ortogonal a cualquier vector de S (pues estos serán combinaciones lineales de los w_i), por lo que $v-v_w\in S_\perp$. Además $S\cap S_\perp=\{0\}$, pues si $u\in S$ y $u\in S_\perp\Rightarrow (u,u)=0$ y por lo tanto $u=0$. Queda probado entonces que $V=S\oplus S_\perp$. Si V es de dimensión n y S de dimensión $m\Rightarrow$ dim $S_\perp=n-m$.
El vector v_s así construido es *la proyección ortogonal* de v sobre el subespacio S, y puede escribirse como

$$v_s=\sum_{i=1}^m P_{w_i}(v)=P_S(v)$$

donde $P_{w_i}(v)=\frac{(w_i,v)}{||w_i||^2}w_i$ es el proyector sobre w_i y

$$P_S=\sum_{i=1}^m P_{w_i}$$

el proyector otrogonal sobre S. En esta expresión los w_i deben formar una base *ortogonal* de S.

El vector v_s es el vector $\in S$ que posee distancia **mínima** a v: Si $u_s\in S$,

$$||v-u_s||^2=||v-v_s+(v_s-u_s)||^2=||v-v_s||^2+||v_s-u_s||^2+2(v-v_s,v_s-u_s)=||v-v_s||^2+||v_s-u_s||^2\geq||v-v_s||^2$$

Esta distancia mínima define la distancia de v a S:

$$d_{\min}(v,S)=||v-v_s||=||v-P_S(v)||$$

Al disponer de una métrica, en un espacio euclídeo podemos pues no sólo determinar si un vector v pertence al subespacio S generado por un conjunto de vectores $\{w_1,\ldots,w_m\}$, sino también determinar que tan lejos está v de este subespacio, a través de la distancia $d_{\min}(v,S)$.

El método de **Gram-Schmidt** puede entonces expresarse en forma aún más concisa como

$$w_1=v_1,\ \ w_i=v_i-P_{\overline{\{w_1,\ldots,w_{i-1}\}}}(v_i),\ \ i=2,\ldots,m$$

donde $P_{\overline{\{w_1,\ldots,w_{i-1}\}}}=P_{w_1}+\ldots+P_{w_{i-1}}$ es el proyector ortogonal sobre el subespacio generado por los $i-1$ vectores anteriores.

Ejemplo 1: El espacio nulo de una matriz A de $m\times n$, $N(A)=\{X|AX=0\}$ con X vectores de $n\times 1$, es el *complemento ortogonal* de las filas de A, es decir, del espacio fila de A ($EF(A)$), ya que $(AX)_i=A_iX$ es el producto escalar de la fila i de A por X.

Se cumple por lo tanto dim $EF(A)$+dim $N(A) = n$.

Ejemplo 2: Encontrar $S_\perp$ si S es el espacio generado por los vectores $(1,1,1),(1,1,-1)$.

Una manera es resolver el sistema homogéneo $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, que da como resultado el conjunto $\{x(1,-1,0),\ x\in\mathbb{R}\}$. $S_\perp$ es entonces el espacio generado por $(1,-1,0)$.
Se cumple dim S+dim $S_\perp$=2+1=3.

Ejemplo 3: a) Proyectar el vector $v = (1,2,3)$ sobre el plano generado por los vectores ortogonales $w_1 = (1,0,1)$ y $w_2 = (0,1,0)$.
Tenemos $(v,w_1) = 4$, $(v,w_2) = 2$, y

$$v_s = P_S(v) = P_{w_1}(v) + P_{w_2}(v) = \frac{4}{2}(1,0,1) + \frac{2}{1}(0,1,0) = (2,2,2)$$

b) Hallar la distancia mínima de v a S.
Tenemos $v - v_s = (-1,0,1)$ y $d_{\min} = ||v - v_s|| = \sqrt{2}$. Además, el ángulo entre v y S puede obtenerse a partir de $\cos\theta = ||v_s||/||v|| = 2\sqrt{3}/\sqrt{14}$.
c) Hallar la matriz que representa el proyector ortogonal sobre S en la base canónica de $\mathbb{R}^3$.

$$[P_S]_e = [P_{w_1}]_e + [P_{w_2}]_e = \frac{1}{2}[w_1]_e[w_1]_e^t + [w_2]_e[w_2]_e^t = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{pmatrix}$$

Se verifica $[v_s]_e = [P_S]_e[v]_e$.

19.8 Representación general del operador de proyección

Es posible dar la expresión general de la matriz que representa al proyector ortogonal sobre un subespacio S generado por un conjunto LI de m vectores $\{w_1,\ldots,w_m\}$ no necesariamente ortogonales. Definamos la matriz

$$R = ([w_1]_e,\ldots,[w_m]_e)$$

de $n\times m$, con $m \le n$, que contiene las coordenadas de los vectores en una base canónica e (con $n =$ dim V). Como $v_s \in S$ podemos escribir $v_s = \sum_{i=1}^m \alpha_i w_i$ y por lo tanto

$$[v_s]_e = \sum_{i=1}^m \alpha_i [w_i]_e = R\alpha$$

donde $\alpha = (\alpha_1,\ldots,\alpha_m)^t$. La condición $(w_i, v - v_s) = 0$ para $i = 1,\ldots,m$ implica

$$0 = R^t([v]_e - [v_s]_e) = R^t[v]_e - R^t[v_s]_e = R^t[v]_e - R^tR\alpha$$

de donde

$$\alpha = (R^tR)^{-1}R^t[v]_e$$

Por lo tanto, $[v_s]_e = R\alpha$ estará dado por

$$[v_s]_e = R(R^tR)^{-1}R^t[v]_e$$

La matriz que representa al proyector sobre S es entonces

$$[P_S]_e = R(R^tR)^{-1}R^t$$

Notar que $[P_S]_e^2 = [P_S]_e$, y que la expresión anterior no se puede simplificar, pues R no es cuadrada.
Discutiremos luego las propiedades de la matriz R^tR.
Ejemplo: Proyectar el vector $v = (1,2,3)$ sobre el plano generado por los vectores $w_1 = (1,1,1)$ y $w_2 = (2,1,2)$. Utilizando el método anterior, tenemos en este caso

$$R = \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$$

con $R^tR = \binom{3\ 5}{5\ 9}$, $R^{-1} = \binom{9\ -5}{-5\ 3}/2$ y

$$[P_S]_e = R(R^tR)^{-1}R^t = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{pmatrix}$$

que coincide con el resultado del último ejercicio. La razón es que el espacio generado por $(1,1,1)$ y $(2,1,2)$ coincide con el generado por los vectores ortogonales $(1,0,1)$ y $(0,1,0)$ ($(1,1,1)=(1,0,1)+(0,1,0)$, $(2,1,2)=2(1,0,1)+(0,1,0)$). Una forma general de obtener el resultado anterior es precisamente ver si los proyectores sobre el espacio generado son idénticos.

Solución de cuadrados mínimos
Consideremos nuevamente el sistema de m ecuaciones lineales con n incógnitas

$$AX = b$$

donde $A \in \mathbb{R}^{m\times n}$, $X \in \mathbb{R}^{n\times 1}$ $b \in \mathbb{R}^{m\times 1}$. Si A representa un monomorfismo $\Rightarrow$ rango$(A) = n \leq m$ y la matriz $A^tA \in \mathbb{R}^{n\times n}$ es no singular. De poseer solución, el sistema tiene entonces una solución única dada por

$$X = (A^tA)^{-1}A^t\,b$$

donde $(A^tA)^{-1}A^t$ es una inversa a izquierda de A. Esta solución se obtiene al multiplicar ambos miembros de $AX = B$ por $(A^tA)^{-1}A^t$, y es válida cuando b pertenece al espacio columna de A (EC(A)), es decir, cuando el sistema es compatible.

Cabe destacar, no obstante, que la expresión anterior para X tiene sentido aún si el sistema no tiene solución: En tal caso

$$AX = A(A^tA)^{-1}A^tb$$

es la **proyección ortogonal** de b sobre el espacio generado por las columnas de A, es decir, $AX = P_{EC(A)}(b)$, de modo que AX es el vector de EC(A) **más cercano a** b. En otras palabras, es el X que minimiza la distancia $||AX-b|| = \sqrt{\sum_{i=1}^m (AX-b)_i^2}$.

Ejemplo: Dado un conjunto de m puntos (x_i, y_i), $i = 1,\ldots,m$, con $x_i \neq x_j$ si $i \neq j$, hallar el polinomio de grado $n-1$ $p(x) = \sum_{j=0}^{n-1} c_j x^j$ tal que $\sum_{i=1}^m (p(x_i) - y_i)^2$ es mínimo. Considerar el caso $m \geq n$.

Tenemos un sistema de m ecuaciones $p(x_i) = y_i$, $i = 1,\ldots,m$, con n incógnitas c_j, $j = 0,\ldots,n-1$:

$$\begin{pmatrix} 1 & x_1 & \ldots & x_1^{n-1} \\ \ldots & & & \\ 1 & x_m & \ldots & x_m^{n-1} \end{pmatrix} \begin{pmatrix} c_0 \\ \ldots \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ \ldots \\ y_m \end{pmatrix}$$

Este sistema es en general incompatible si $m \geq n$. No obstante, el objetivo es buscar la solución que minimiza la distancia $||p(X) - Y||$ o equivalentemente $||p(X)-Y||^2$, donde $Y = (y_1,\ldots,y_m)^t$, $X = (x_1,\ldots,x_m)^t$ y $p(X) = AC$, con A la matriz de $m \times n$ de elementos $A_{ij} = x_i^{j-1}$ y C el vector columna de coeficientes c_i. Tal solución estará dada entonces por $C = (A^tA)^{-1}A^tY$, tal que $AC = A(A^tA)^{-1}A^tY$ es la proyección ortogonal de Y sobre $EC(A)$.

19.9 Matriz de Gram

Dado un conjunto de vectores $v_1,\ldots,v_m$ pertenecientes a un espacio vectorial euclídeo V de dimensión finita, la matriz simétrica G de $m \times m$ productos escalares, de elementos

$$G_{ij} = (v_i, v_j) = G_{ji}$$

se denomina matriz de Gram y posee importantes propiedades.
Notemos que en términos de la matriz $R = ([v_1]_e,\ldots,[v_m]_e)$ de $n \times m$, con e una base canónica,

$$G = R^tR$$

1) El producto escalar (w,u) de combinaciones lineales $w=\sum_{i=1}^n \alpha_i v_i$, $u=\sum_{i=1}^n \beta_i v_i$, con $[w]_e=R\alpha$, $[u]_e=R\beta$, y $\alpha=(\alpha_1,\ldots,\alpha_m)^t$, $\beta=(\beta_1,\ldots,\beta_m)^t$, puede expresarse como

$$(w,u)=[v]_e^t[u]_e=(R\alpha)^t(R\beta)=\alpha^tR^tR\beta=\alpha^tG\beta$$

2) La matriz G es no singular sii los vectores v_i son LI:
Si G es singular, existe un vector columna no nulo β de $m\times 1$ tal que $G\beta=0$ y por lo tanto, si $u=\sum_{i=1}^n \beta_i v_i$,

$$(u,u)=\beta^tG\beta=\beta^t0=0$$

por lo que necesariamente $u=0$. Por lo tanto, existe una combinación lineal nula u con coeficientes no todos nulos. Esto implica que los v_i son LD.
Análogamente, si existe una combinación lineal nula $u=\sum_{i=1}^n \beta_i v_i=0$, con los β_i no todos nulos, entonces $0=(v_j,u)=\sum_{i=1}^n G_{ji}\beta_i$ para cualquier j por lo que $G\beta=0$ y por lo tanto G es necesariamente singular.
Un método sencillo de determinar si los m vectores v_i son LI es pues evaluar el determinante $|G|=|R^tR|$: $\{v_1,\ldots,v_m\}$ es LI sii $|G|\neq 0$.
Para $m=n$, R es de $n\times n$ y $|G|=|R|^2$, por lo que se reobtiene la condición conocida $|R|\neq 0$ para n vectores en $\mathbb{R}^n$.
3) Si $w_i=\sum_{j=1}^n S_{ji}v_j$, $i=1,\ldots,m$, entonces $G'_{ij}\equiv(w_i,w_j)=\sum_{k,l}S_{ki}S_{lj}G_{kl}$, o sea,

$$G'=S^tGS$$

con $|G'|=|S|^2|G|$. En particular, si los v_i son LI, podemos ortogonalizarlos con el método de Gram-Schmidt, generando vectores ortogonales w_i. La correspondiente matriz S cumple, por construcción, $|S|=1$ y por lo tanto $|G|=|G'|=\prod_{i=1}^m ||w_i||^2$.
Este último producto representa *el cuadrado del volumen m dimensional del paralelepípedo formado por los vectores* $w_1,\ldots,w_m$, y por lo tanto, por $v_1,\ldots,v_m$. El volumen generado por estos m vectores es pues

$$Vol_{v_1,\ldots,v_m}=\sqrt{|G|}$$

Si $m=n$, $|G|=|R^tR|=|R|^2$, y $Vol_{v_1,\ldots,v_m}=|\mathrm{Det}(R)|$.
4) La matriz G es diagonalizable, por ser real y simétrica, y los autovalores de G son *positivos o nulos*. Los autovectores asociados a autovalores no nulos corresponden a vectores *ortogonales*, y los correspondientes a autovalores nulos a combinaciones lineales *nulas* de los vectores v_i.
En efecto, si $G\alpha=\lambda_\alpha\alpha$, con $\alpha\neq 0$, para $w=\sum_{i=1}^m \alpha_i v_i$ se obtiene

$$0\leq(w,w)=\alpha^tG\alpha=\lambda_\alpha\alpha^t\alpha$$

Como $\alpha^t\alpha>0$ entonces $\lambda_\alpha\geq 0$. Si $\lambda_\alpha>0\Rightarrow w$ es no nulo, mientras que si $\lambda_\alpha=0\Rightarrow w=0$, siendo pues una combinación lineal *nula* de los v_i. Además, si $G\beta=\lambda_\beta\beta$, con $\beta\neq 0$, tenemos, para $u=\sum_{i=1}^m \beta_i v_i$,

$$(w,u)=\alpha^tG\beta=\lambda_\beta\alpha^t\beta=0 \;\text{ si }\; \lambda_\alpha\neq\lambda_\beta$$

por ser α,β autovectores de una matriz simétrica. La diagonalización de G proporciona pues un método directo de extraer un conjunto ortogonal de k vectores LI de los m vectores w_i, que son los determinados por los autovectores asociados a los autovalores no nulos.
El número k de autovalores no nulos de G es precisamente el *rango* de G y determina entonces la *dimensión* del subespacio generado por los m vectores v_i: $k=r(G)=\dim\overline{\{v_1,\ldots,v_m\}}$.

Ejemplo: Consideremos los vectores $v_1=(1,1,1,1)$, $v_2=(1,1,-1,1)$, $v_3=(0,0,1,0)$ de $\mathbb{R}^4$. Tenemos

$$R=\begin{pmatrix}1&1&0\\1&1&0\\1&-1&1\\1&1&0\end{pmatrix},\quad G=R^tR=\begin{pmatrix}4&2&1\\2&4&-1\\1&-1&1\end{pmatrix}$$

Como $|G|=0$ los vectores son LD. Además, los autovalores de G son $\lambda=6,3,0$, con autovectores $(1,1,0)$, $(1,-1,1)$, $(-1,1,2)$.
Por lo tanto, los vectores $w'_1=v_1+v_2=(2,2,0,2)$, $w_2=v_1-v_2+v_3=(0,0,3,0)$, son ortogonales y $w'_3=-w_1+w_2+2w_3=(0,0,0,0)$ es la combinación lineal nula.
Pueden obtenerse resultados similares utilizando Gram-Schmidt. El determinante del primer menor de O, $16-4=12$, representa el cuadrado del área del paralelogramo determinado por w_1 y w_2.

19.10 Operadores adjuntos y autoadjuntos en espacios euclídeos

Sea $F : V \to V$ un operador lineal en un espacio euclídeo V. El operador *adjunto* $F^\dagger$ se define por

$$(v, F(w)) = (F^\dagger(v), w)$$

$\forall\ v, w \in V$. Si V es de dimensión finita y e denota una base canónica de V, la definición anterior implica

$$[v]_e^t[F]_e[w]_e = ([F^\dagger]_e[v]_e)^t[w]_e = [v]_e^t[F^\dagger]_e^t[w]_e$$

por lo que la matriz $[F^\dagger]_e \equiv [F^\dagger]_e^e$ que representa a $F^\dagger$ en dicha base es la traspuesta de la matriz que representa a F:

$$[F^\dagger]_e = [F]_e^t$$

Esto también muestra que $(F^\dagger)^\dagger = F$ (pues $[(F^\dagger)^\dagger]_e = ([F]_e^t)^t = [F]_e$) y que si $G : V \to V$ es otro operador lineal, $(FG)^\dagger = G^\dagger F^\dagger$ (pues $[(FG)^\dagger]_e = [FG]_e^t = [G]_e^t[F]_e^t$). Estas dos últimas propiedades pueden también demostrarse a partir de la defición de operador adjunto (se deja como ejercicio).
Notemos que $(F(v), w) = (v, F^\dagger(w))$.
Operador autoadjunto: Si $F^\dagger = F$ el operador se dice *autoadjunto*. En este caso debe cumplirse

$$[F]_e^t = [F]_e$$

por lo que F será autoadjunto si y sólo si es representado por una matriz *simétirca* en una base canónica.

Notemos que en una base arbitraria B, no necesariamente ortogonal, con $(b_i, b_j) = g_{ij} = g_{ji}$, tendríamos $(v, F(w)) = [v]_B G[F]_B[w]_B$, $(F^\dagger(v), w) = [v]_B^t[F^\dagger]_B^t G[w]_B$ y por lo tanto, $[F^\dagger]_B^t G = G[F]_B$, por lo que

$$[F^\dagger]_B = G^{-1}[F]_B^t G$$

Si F es autoadjunto $\Rightarrow [F]_B = G^{-1}[F]_B^t G$.

En general, si $F : V \to W$ es una transformación lineal entre espacios euclídeos V, W, podemos definir $F^\dagger : W \to V$ de la misma forma: $(w, F(v)) = (F^\dagger(w), v)\ \forall\ v,\ w$. Para espacios V, W de dimensión finita n y m respectivamente, esto implica $[F^\dagger]_e^{\tilde{e}} = ([F]_{\tilde{e}}^e)^{\rm tr}$ en bases ortonormales e y $\tilde{e}$ de V y W. De esta forma, $F_{ij} = (\tilde{e}_i, F(e_j)) = (F^\dagger(\tilde{e}_i), e_j) = (e_j, F^\dagger(\tilde{e}_i)) = F^\dagger_{ji}$.

Diagonalización de operadores autoadjuntos

Si $F : V \to V$ es un operador lineal autoadjunto en un espacio V de dimensión finita, demostraremos que *existe siempre una base canónica* e' *en la que* $[F]_{e'}$ *es diagonal* (Ya habíamos demostrado que los autovalores de matrices reales simétricas son todos reales y que los autovectores corresp. a autovalores distintos son ortogonales). Un resultado aún más general será demostrado luego para espacios complejos.
Para $n = \dim\ V = 1$, el resultado es trivial. Asumimos ahora que es válido para $\dim\ V = n - 1$. Si e'_1 es un autovector normalizado de F con autovalor λ_1, tal que $F(e'_1) = \lambda_1 e'_1$, $(e'_1, e'_1) = 1$, se puede construir, por Gram-Schmidt, una base ortonormal $\tilde{e}$ de V tal que $\tilde{e}_1 = e'_1$, definida por una matriz de cambio de base S ortonormal ($S^t = S$). En tal caso $[F]_{\tilde{e}} = S^{-1}[F]_e S = S^t[F]_e S$ será también *simétrica*: $[F]_{\tilde{e}}^t = S^t[F]_e S = [F]_{\tilde{e}}$. Pero como e'_1 es autovector, $[F]_{\tilde{e}}$ tendrá entonces la forma $\left(\begin{smallmatrix}\lambda_1 & 0\\ 0 & \tilde{F}\end{smallmatrix}\right)$, con $\tilde{F}$ una matriz *simétrica* de $(n-1)\times(n-1)$ que representa a un operador autoadjunto en un subespacio de dimensión $n-1$. Por hipótesis inductiva, existe una base ortonormal de este subespacio en la que el operador será representado por una matriz diagonal F'. Por lo tanto, agregando a esta base el autovector e'_1, tendremos una base ortonormal e' de V en la que $[F]_{e'} = \left(\begin{smallmatrix}\lambda_1 & 0\\ 0 & F'\end{smallmatrix}\right)$ será también diagonal.
Resumiendo, dado F autoadjunto ($[F]_e$ simétrica en una base canónica e) existe una base ortonormal definida por una matriz de cambio de base S, con $S^{-1} = S^t$, tal que

$$[F]_{e'} = S^t[F]_e S$$

es diagonal

19.11 Isometrías

Las isometrías son operadores $U : V \to V$ que conservan el producto escalar. Ejemplos comunes en $V = \mathbb{R}^n$ son rotaciones y reflexiones. Si U es una isometría,

$$(U(v), U(w)) = (v, w) \;\; \forall \;\; v, w \in V$$

Por lo tanto, si e es una base canónica, $(U(v), U(w)) = [v]_e^t [U]_e^t [U]_e [w]_e = [v]_e^t [w]_e \; \forall \; v, w \in V$, por lo que

$$[U]_e^t [U]_e = I_n$$

con I_n la matriz identidad, es decir, $[U]_e^{-1} = [U]_e^t$. Esto implica a su vez $[U]_e [U]_e^t = I_n$. Las matríces $[U]_e$ que representan a una isometría en una base canónica e son pues matrices *ortonormales*, y tanto las filas como las columnas de $[U]_e$ serán por lo tanto *ortonormales*, como se vió anteriormente: Si $U_{ij} = ([U]_e)_{ij}$,

$$\sum_{j=1}^{n} U_{ji} U_{jk} = \delta_{ik}, \;\; \sum_{j=1}^{n} U_{ij} U_{kj} = \delta_{ik}$$

En términos de operadores adjuntos, $(U(v), U(w)) = (v, U^\dagger U(w))$, por lo que U será una isometría si y sólo si

$$U^{-1} = U^\dagger$$

Demostraremos luego que toda isometría puede ser descompuesta en rotaciones y/o reflexiones.

Las isometrías transforman bases ortogonales en bases ortogonales. En efecto, al conservar todos los productos escalares, si $e_i' = U(e_i) = \sum_{j=1}^n U_{ji} e_j$, entonces

$$(e_i', e_j') = (U(e_i), U(e_j)) = (e_i, e_j) = \delta_{ij}$$

La recíproca es obviamente también válida: Cualquier par de bases canónicas e, e' de V estarán relacionadas por una isometría $e_i' = U(e_i)$. Cualquier matriz de cambio de base S que represente una isometría debe pues satisfacer $S^t S = I_n$, como se vió anteriormente.
Ejemplo: Si

$$[U]_e = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

entonces U es una isometría ya que $[U]_e^t [U]_e = I_3$. Tanto las filas como las columnas de $[U]_e$ son ortonormales (ortogonales y de longitud 1). Esta matriz representa una rotación de ángulo α antihoraria en el plano xy, compuesta con una reflexión respecto a este plano:

$$[U]_e = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Isomorfismo Euclídeo

Dados dos espacios euclídeos V, V' de la misma dimensión, podemos siempre elegir bases canónicas $e = (e_1, \ldots, e_n)$ en V y $e' = (e_1', \ldots, e_n')$ en V' tal que tales que $(e_i, e_j) = \delta_{ij}$, $(e_i', e_j') = \delta_{ij}$. Definiendo un isomorfismo $Q : V \to V'$ tal que $Q(e_i) = e_i'$, $i = 1, \ldots, n$, se tiene

$$(e_i', e_j') = (Q(e_i), Q(e_j)) = (e_i, e_j) = \delta_{ij}$$

Por lo tanto, si $v' = \sum_{i=1}^n \alpha_i e_i'$, $w' = \sum_{i=1}^n \beta_i e_i' \Rightarrow v' = Q(v)$, $w' = Q(w)$, con $v = \sum_{i=1}^n \alpha_i e_i$, $w = \sum_{i=1}^n \beta_i e_i$ y

$$(v', w') = (Q(v), Q(w)) = (v, w) = \sum_{i=1}^{n} \alpha_i \beta_i$$

Un isomorfismo $Q : V \to V'$ de este tipo (que conserva todos los productos escalares) se lo denomina isomorfismo euclídeo. La existencia de Q muestra que todas las propiedades geométricas de $\mathbb{R}^n$ pueden extenderse directamente a cualquier espacio euclídeo V' de dimensión n.

20 Descomposición en valores singulares (DVS)

Consideremos una matriz real A de $m \times n$. Podemos formar la matriz de $n \times n$

$$A^t A$$

la cual es simétrica ($(A^tA)^t = A^tA$) y tiene la mismas propiedades que la matriz de Gram. Por lo tanto, tiene un conjunto de n autovectores $v_i \in \mathbb{R}^{n\times 1}$ *ortonormales* asociados a autovalores λ_i*positivos o nulos*:

$$A^tAv_i = \lambda_i v_i, \quad i = 1, \ldots, n \ \text{ con } \ v_j^t v_i = \delta_{ij}, \ \ \lambda_i \geq 0$$

Sean $\lambda_1, \ldots, \lambda_k$, $k \leq n$, los autovalores no nulos de O. Podemos definir los k vectores de $m \times 1$

$$u_i = \frac{1}{\sqrt{\lambda_i}} A v_i, \quad i = 1, \ldots, k\,, \ \ \lambda_i \neq 0$$

que son ortonormales:

$$u_j^t u_i = \frac{1}{\sqrt{\lambda_j \lambda_i}} v_j^t A^t A v_i = \frac{\lambda_i v_j^t v_i}{\sqrt{\lambda_i \lambda_j}} = \delta_{ij}$$

Si $k < m$, podemos completar estos k vectores con $m-k$ vectores obtenidos por el método de Gram-Schmidt, tal que $(u_1, \ldots, u_m)$ forme un conjunto ortonormal (base de $\mathbb{R}^{m\times 1}$). Además, para $i = k+1, \ldots n$ se cumple $A^tAv_i = 0$ y entonces $(Av_i)^t(Av_i) = v_i^t A^t A v_i = 0$, es decir $||Av_i|| = 0$, lo que implica $Av_i = 0$. Tenemos entonces

$$A(v_1, \ldots, v_n) = (\sqrt{\lambda_1} u_1, \ldots, \sqrt{\lambda_k} u_k, 0 \ldots, 0) = (u_1, \ldots, u_m) A'$$

donde A' es una matriz "diagonal" de $m \times n$ de la forma

$$A' = \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 & \ldots & 0 \\ & & \ldots & & & \\ 0 & \ldots & \sigma_k & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & \ldots & 0 \\ & & \ldots & & & \\ 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}, \qquad \sigma_i = \sqrt{\lambda_i}, \ \ i = 1, \ldots, k$$

Por consiguiente, definiendo las matrices ortonormales $V = (v_1, \ldots, v_n)$, $U = (u_1, \ldots, u_m)$ (que satisfacen $V^tV = I_n$, $U^tU = I_m$), se tiene $AV = UA'$ y por lo tanto

$$A = UA'V^t$$

Esta representación de A se denomina *descomposición en valores singulares* (del inglés *singular value decomposition*) y los elementos σ_i de A' *los valores singulares* de A, que son las raíces de los autovalores no nulos de A^tA (necesariamente positivos). Vemos así que rango(A) =rango$(A') = k$, por lo que $k \leq \text{Min}[m, n]$. Además, por construcción, los primeros k vectores u_j, $j = 1, \ldots, k$ forman una base del espacio columna de A y los últimos $n - k$ vectores $v_{k+1}, \ldots, v_n$ una base del espacio nulo de A (el subespacio ortogonal al espacio fila de A).

Notemos también que si $A = UA'V^t$, con A' "diagonal" de $m \times n$ con elementos positivos o nulos y U, V matrices ortonormales, entonces necesariamente los elementos diagonales no nulos de A' son los valores singulares, pues

$$A^tA = VA'^tU^tUA'V^t = V(A'^tA')V^t$$

con A'^tA' diagonal de $n \times n$. Esto implica $V^tA^tAV = A'^tA'$, lo que muestra que V es necesariamente una matriz ortonormal de autovectores de A^tA y A'^tA' la correspondiente matriz diagonal de autovalores.

Desde el punto de vista operacional, A puede considerarse como la representación $[F]_{\tilde{e}}^{e}$ de una transformación lineal $F : V \to W$ entre espacios euclídeos V y W de dimensión n y m respectivamente, en bases canónicas $e = (e_1, \ldots, e_n)$ y $\tilde{e} = (e_1, \ldots, e_m)$ de V y W, siendo A^tA *la matriz de Gram* del conjunto de imágenes $\{F(e_1), \ldots, F(e_n)\}$: $(A^tA)_{ij} = (F(e_i), F(e_j))$.

La descomposición anterior muestra que es siempre posible encontrar *bases ortonormales* e' y $\tilde{e}'$ de V y W en la que F tiene una representación "diagonal", con elementos diagonales reales *positivos* o nulos, es decir

$$[F]_{\tilde{e}'}^{e'} = U^t[F]_{\tilde{e}}^{e}V = A'$$

con $V=[I]_e^{e'}$, $U=[I]_{\tilde{e}}^{\tilde{e}'}$ y $F(e'_i)=\sigma_i\tilde{e}'_i$, $i=1,\ldots,k$, con $F(e'_i)=0$ si $i>k$. Los primeros k vectores de $\tilde{e}'$ forman pues una base ortonormal de $Im(F)=F(V)$, y los últimos $n-k$ vectores de e' una base ortonormal de $N(F)$. Notemos que los valores singulares son independientes de las bases canónicas elegidas: Si $B=R^tAS$, con $R^tR=I_m$, $S^tS=I_n \Rightarrow B^tB=S^tA^tRR^tAS=S^tA^tAS$, y los autovalores de B^tB son entonces idénticos a los de A^tA.

Otro comentario importante es que si $A=UA'V^t \Rightarrow$

$$A^t=VA'^tU^t$$

que es necesariamente la descomposición singular de A^t. Esto muestra que los valores singulares son también las raíces de los autovalores no nulos de AA^t (matriz real simétrica de $m\times m$) y U una matriz ortonormal de autovectores de AA^t. Para la obtención de los valores singulares se puede pues diagonalizar la menor de las matrices A^tA y AA^t.

Se ve también que si A es de $n\times n$ y no singular,

$$A^{-1}=VA'^{-1}U^t$$

lo que muestra que los valores singulares de A^{-1} son los inversos de los valores singulares de A (y que si A es no singular estos son necesariamente no nulos). Notemos que para A de $n\times n$, $|A|=|U||A'||V^t|=\pm|A'|$, donde $|U|=\pm 1$, $|V|=\pm 1$, por lo que $|\mathrm{Det}[A]|=\mathrm{Det}[A']$.

Si A representa un monomorfismo $\rightarrow$ rango$(A)=n$, por lo que $k=n\leq m$. En tal caso, conociendo la descomposición singular de A, una **inversa a izquierda** $\tilde{A}$ (de $n\times m$) puede obtenerse como

$$\tilde{A}=V\tilde{A}'U^t$$

con $\tilde{A}'$ una matriz "diagonal" de $n\times m$ de elementos $\tilde{\sigma}_i=1/\sigma_i$, $i=1,\ldots,n$, ya que se verifica $\tilde{A}'A'=I_n$ y por tanto $\tilde{A}A=V\tilde{A}'A'V^t=I_n$. Esto muestra asimismo que los valores singulares de $\tilde{A}$ son los inversos de los de A. En forma análoga, si A representa un epimorfismo, rango$(A)=m$, por lo que $k=m\leq n$ y una **inversa a derecha** de A estará dada por $\tilde{A}=V\tilde{A}'U^t$, pues en este caso $A'\tilde{A}'=I_m$ y $A\tilde{A}=UA'\tilde{A}'U^t=I_m$.

Una última observación general muy importante es que la descomposición singular de A permite expandir a esta como

$$A=\sum_{i=1}^{k}\sigma_i u_i v_i^t$$

lo que constituye la generalización de la expansión de una matriz simétrica A de $n\times n$ en teŕminos de autovalores y autovectores ortonormales (ver siguiente comentario). En el caso de matrices de grandes dimensiones, un método general de compresión de información (utilizado en la compresión de imágenes digitales) consiste precisamente en conservar de la expansión anterior los términos con σ_i mayor a cierto valor inferior umbral.

En el caso especial de que A sea de $n\times n$ y *simétrica* ($A^t=A$) $\Rightarrow A^tA=A^2$, por lo que $\lambda_i=(\lambda_i^A)^2$, con λ_i^A los autovalores de A. Se obtiene entonces

$$\sigma_i=|\lambda_i^A|,\;\; i=1,\ldots,k$$

es decir, los valores singulares son los *valores absolutos* de los autovalores no nulos de A. La matriz V puede entonces elegirse como la matriz de autovectores de A y U como la matriz $U=(s_1v_1,\ldots,s_nv_n)$, con s_i el signo de λ_i. En este caso la expansión anterior se reduce a

$$A=\sum_{i=1}^{n}\lambda_i v_i v_i^t$$

con $v_iv_i^t$ la representación matricial del *proyector ortogonal* sobre el espacio generado por v_i.

Ejemplo : Consideremos

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

Tenemos

$$A^tA=\begin{pmatrix}2&1\\1&2\end{pmatrix}$$

Los autovalores de A^tA son entonces $\lambda_\pm = 2 \pm 1$ por lo que los valores singulares son $\sigma_1 = \sqrt{3}$, $\sigma_2 = 1$. Se obtiene $v_1 = (1,1)^t/\sqrt{2}$, $v_2 = (-1,1)^t/\sqrt{2}$, y $u_1 = Av_1/\sigma_1 = (1,2,1)^t/\sqrt{6}$, $u_2 = Av_2/\sigma_2 = (-1,0,1)/\sqrt{2}$. u_3 puede elegirse, utilizando GS a partir de u_1, u_2 y $(1,0,0)$, como $(1,-1,1)/\sqrt{3}$. Se obtiene entonces

$$A = \begin{pmatrix} 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} /\sqrt{2}$$

Algunas aplicaciones

20.1 Norma inducida de una matriz

Primeramente, consideremos una forma cuadrática real $\tilde{B}(v) = X^tBX$, con B de $n \times n$ real simétrica y $X = (x_1, \ldots, x_n)^t = [v]_e$ de $n \times 1$. Diagonalizando B, tenemos $S^tBS = B'$, con B' diagonal ($B'_{ij} = \lambda_i\delta_{ij}$) y $S = (X_1, \ldots, X_n)$ una matriz ortonormal de autovectores ($S^tS = I_n$). Por lo tanto, definiendo $X' = S^tX = (x'_1, \ldots, x'_n)$, tal que $X = SX'$, se obtiene

$$\tilde{B}(v) = X^tBX = X'^tS^tBSX' = X'^tB'X' = \sum_{i=1}^{n} \lambda_i x'^2_i$$

Como $||v||^2 = X^tX = X'^tS^tSX' = X'^tX'$, se obtiene, para $v \neq 0$,

$$Q(v) \equiv \frac{\tilde{B}(v)}{||v||^2} = \frac{X^tBX}{X^tX} = \frac{X'^tB'X'}{X'^tX'} = \frac{\sum_{i=1}^n \lambda_i x'^2_i}{\sum_{i=1}^n x'^2_i}$$

Si $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$, vemos entonces que

$$\lambda_1 \leq \frac{X^tBX}{X^tX} \leq \lambda_n$$

con el valor máximo λ_n alcanzado si $X = X_n$, con $BX_n = \lambda_n X_n$ y el mínimo λ_1 si $X = X_1$, con $BX_1 = \lambda_1 X_1$. Hemos pues demostrado que el valor máximo (mínimo) que toma la forma cuadrática X^tBX en la esfera unidad ($X^tX = 1$) es el máximo (mínimo) autovalor de B.

El cociente $Q(v)$ se denomina en contextos físicos cociente de Rayleigh y proporciona un *método variacional* para la determinación del autovalor máximo y mínimo de una matriz simétrica B:

$$\lambda_1 = \text{Min}_{v\neq 0} Q(v), \quad \lambda_n = \text{Max}_{v\neq 0} Q(v)$$

Consideremos ahora una transformación $F : \mathbb{R}^n \to \mathbb{R}^m$, representada en las bases canónicas por una matriz A de $m \times n$. Tenemos, para un vector no nulo $v \in \mathbb{R}^n$ tal que $[v]_e = X$,

$$\frac{||F(v)||^2}{||v||^2} = \frac{||AX||^2}{||X||^2} = \frac{(AX)^tAX}{X^tX} = \frac{X^tA^tAX}{X^tX}$$

y por lo tanto, utilizando el resultado anterior,

$$\sigma_m^2 \leq \frac{||AX||^2}{||X||^2} \leq \sigma_M^2$$

donde σ_M^2 y σ_m^2 denotan aquí el máximo y mínimo autovalor de A^tA (σ_M y σ_m son entonces los valores singulares extremos si son no nulos). Por lo tanto,

$$\sigma_m \leq \frac{||F(v)||}{||v||} \leq \sigma_M$$

Los valores σ_M y σ_m indican pues la máxima y mínima "dilatación" que puede experimentar un vector v al ser transformado por F. Si $m < n$ necesariamente $\sigma_n = 0$.

La **norma** de una matriz A de $m \times n$ (o de la transformación asociada F) **inducida por la norma del vector** se define como

$$||A|| = \text{Max}_{\{X, X\neq 0\}} \frac{||AX||}{||X||} = \text{Max}_{\{X, ||X||=1\}} ||AX||$$

El resultado anterior implica entonces

$$||A|| = \sigma_M$$

es decir, la norma es el mayor valor singular de A. Este resultado se denomina en realidad norma 2 de la matriz, pues está derivado de la norma $||X|| \equiv \sqrt{X^t X} = \sqrt{x_1^2 + \ldots + x_n^2}$.
Una consecuencia inmediata pero importante de esta norma es que se cumple

$$||AX|| \leq ||A||\,||X|| \ \ \forall X \in \mathbb{R}^n$$

Esta norma satisface las cuatro propiedades básicas siguientes:
1) $||A|| > 0$, con $||A|| = 0$ si y sólo si $A = 0$
2) $||\alpha A|| = |\alpha|||A||$
3) $||A + B|| \leq ||A|| + ||B||$
(pues $||A + B|| = ||(A + B)X_M||/||X_M|| \leq (||AX_M|| + ||BX_M||)/||X_M|| \leq ||A| + ||B||$).
4) $||AB|| \leq ||A||\,||B||$ ($B \in \mathbb{R}^{m\times n}$, $A \in \mathbb{R}^{p\times m}$)
(pues $||ABX|| = ||A(BX)|| = ||A||\,||BX|| \leq ||A||\,||B||\,||X|| \ \forall\ X \in \mathbb{R}^{n\times 1}$.

20.2 Imagen de la esfera unidad
Consideremos ahora la imagen por $F : \mathbb{R}^n \to \mathbb{R}^m$ de la esfera unidad C de $\mathbb{R}^n$, es decir $F(C) = \{F(v)|\,||v|| = 1\}$. La descomposición en valores singulares permite encontrar bases canónicas e' y $\tilde{e}'$ de $\mathbb{R}^n$ y $\mathbb{R}^m$ en las que la matriz A' que representa a F es "diagonal", con elementos diagonales $\sigma_i \geq 0$. Si $[v]_{e'} = X' = (x'_1, \ldots, x'_n)^t$, con $X'^t X' = 1 \Rightarrow Y' = [F(v)]_{\tilde{e}'} = A'X' = (\sigma_1 x'_1, \ldots, \sigma_k x'_k, 0, \ldots, 0)^t$. Por lo tanto, si $k = n \leq m$ las k componentes no nulas $y'_i = x'_i \sigma_i$ de Y' satisfacen

$$\sum_{i=1}^{n} y'^2_i/\sigma_i^2 = 1$$

lo que indica que la imagen en la base $\tilde{e}'$ es la superficie de un elipsoide de dimensión $k = n$ con ejes principales en la dirección de los $\tilde{e}'_i$ y radios de longitud σ_i. Si $k < n \Rightarrow$ al menos uno de los radios es nulo y la superficie del elipsoide degenera en el interior y borde de un elipsoide de dimensión $k < n$ (en este caso $\sum_{i=1}^{k} y'^2_i/\sigma_i^2 \leq 1$). En resumen, los valores singulares determinan los radios del elipsoide obtenido como imagen por F de la esfera unidad.

20.3 Número de condición de una matriz
Consideremos un sistema de n ecuaciones con n incógnitas representado por la ecuación matricial

$$AX = Y$$

con A de $n \times n$, y X, Y de $n \times 1$. Si A es no singular la única solución está dada por $X = A^{-1}Y$. Estudiemos ahora la estabilidad de esta solución frente a variaciones δY de Y. Tenemos $\delta X = A^{-1}\delta Y$ y por lo tanto

$$\frac{||\delta X||}{||X||} = \frac{||A^{-1}\delta Y||}{||X||} \leq \frac{||A^{-1}||||\delta Y||}{||X||} \leq ||A^{-1}||\,||A||\frac{||\delta Y||}{||Y||}$$

donde en la última expresión hemos utilizado la desigualdad $||Y|| = ||AX|| \leq ||A||\,||X||$. El número de condición de una matriz se define entonces como

$$n_c(A) = ||A||\,||A^{-1}||$$

y acota la inestabilidad de la solución del sistema asociado frente a variaciones en la inhomogeneidad Y:

$$\frac{||\delta X||}{||X||} \leq n_c(A)\frac{||\delta Y||}{||Y||}$$

En virtud del resultado previo, se tiene, utilizando la norma 2, $||A|| = \sigma_M$, $||A^{-1}|| = 1/\sigma_m$, con σ_M y σ_m el máximo y mínimo valor singular, y por lo tanto

$$n_c(A) = \sigma_M/\sigma_m \geq 1$$

El número de condición es entonces adimensional y queda determinado por el cociente entre los valores singulares extremos. Para matrices reales simétricas, $\sigma_M = |\lambda_M|$, $\sigma_m = |\lambda_m|$, con λ_M y λ_m los autovalores de mayor y menor valor absoluto respectivamente. Nótese que si la matriz A es singular, $\sigma_m = 0$ y en tal caso $n_c(A) = \infty$. Números de condición grandes indican matrices "cuasi singulares" (o mal condicionadas), para las que no se puede asegurar estabilidad en la solución del sistema asociado.

Es importante destacar que la estabilidad frente a variaciones en la matriz A queda también determinada por el mismo número de condición. Si $AX = Y$ y $(A + \delta A)(X + \delta X) = Y$, entonces, a primer orden en δX y δA, se obtiene $(\delta A)X + A\delta X = 0$ y

$$\delta X = -A^{-1}(\delta A)X$$

Por lo tanto

$$||\delta X|| = ||A^{-1}(\delta A)X|| \leq ||A^{-1}||\,||\delta A||\,||X||$$

de donde

$$\frac{||\delta X||}{||X||} \leq ||A^{-1}||\,||\delta A|| = n_c(A)\frac{||\delta A||}{||A||}$$

Ejemplo: Si

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

entonces

$$A^t A = \begin{pmatrix} \varepsilon^2 & 0 \\ 0 & 1 \end{pmatrix}$$

por lo que los valores singulares son $|\varepsilon|$ y 1 y el número de condición es

$$n_c(A) = 1/|\varepsilon|$$

si $|\varepsilon| \leq 1$. Notemos que $\mathrm{Det}[A] = -\varepsilon$ y que $n_c(A) \to \infty$ si $\varepsilon \to 0$. La solución al sistema $AX = Y$ es $X = (y_2/\varepsilon, y_1)^t$ con $\delta X = (\delta y_2/\varepsilon, \delta y_1)^t$ y $||\delta X||^2/||X||^2 = (\delta y_2^2/\varepsilon^2 + \delta y_1^2)/(y_2^2/\varepsilon^2 + y_1^2)$. Si por ejemplo $y_2 = 0$, $y_1 = 1$ y $\delta y_1 = 0 \Rightarrow ||\delta X||/||X|| = |\delta y_2|/|\varepsilon| = n_c(A)||\delta Y||/||Y||$, por lo que $||\delta X||/||X||$ puede ser mucho mayor que $||\delta Y||/||Y||$ cuando ε es suf. pequeño.

Notemos en cambio que la matriz

$$B = \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}$$

tiene número de condición 1 a pesar de que $\mathrm{Det}[B] = \varepsilon^2 \ll 1$ para $|\varepsilon| \ll 1$.

20.4 Pseudoinversa

Sea $A \in \mathbb{R}^{m\times n}$ con $A = UA'V^t = \sum_{i=1}^k \sigma_i u_i v_i^t$ su DVS. La pseudoinversa de A (denominada también pseudoinversa de Moore-Penrose) es una matriz $\tilde{A} \in \mathbb{R}^{n\times m}$ definida como

$$\tilde{A} = V\tilde{A}'U^{\mathrm{tr}} = \sum_{i=1}^{k} \frac{1}{\sigma_i} v_i u_i^t$$

con $\tilde{A}'$ una matriz de $n \times m$ de elementos diagonales $1/\sigma_i$ ($A'_{ij} = \delta_{ij}/\sigma_i$ si $i \leq k$ y 0 en caso contrario). Dado que $u_i^t u_j = \delta_{ij}$, $v_i^t v_j = \delta_{ij}$, se verifica que $A\tilde{A} = \sum_{i=1}^k u_i u_i^t$ es el **proyector ortogonal sobre el espacio columna de la matriz**, mientras que $\tilde{A}A = \sum_{i=1}^k v_i v_i^t$ es el proyector ortogonal sobre el espacio fila (es decir, sobre el espacio columna de A^t). Se verifica entonces

$$\tilde{A}A\tilde{A} = \tilde{A}\,, \quad A\tilde{A}A = A$$

Es facil ver que si rango(A)$= n \Rightarrow \tilde{A} = (A^tA)^{-1}A^t$, coincidiendo con una inversa a izquierda de A, mientras que si rango(A)$= m \Rightarrow \tilde{A} = A^t(AA^t)^{-1}$, coincidiendo con una inversa a derecha de A. Si rango(A)$= n = m$ $\Rightarrow \tilde{A} = A^{-1}$ es la inversa de A.

Consideremos ahora el sistema de ecuaciones lineales de $m \times n$

$$AX = b$$

donde $X \in \mathbb{R}^{n\times 1}$ y $b \in \mathbb{R}^{m\times 1}$. Si el sistema es compatible, $b = A\tilde{A}b$ (pues $b \in EC(A)$) y entonces una solución particular del sistema es

$$X = \tilde{A}b$$

pues $AX = A\tilde{A}b = b$. Si no existe solución ($b \not\in EC(A)$) entonces $X = \tilde{A}b$ es el vector que minimiza la diferencia $||AX - b||$, pues $A\tilde{A}b$ es la proyección ortogonal de b sobre $EC(A)$.

En el caso compatible, la solución general del sistema $AX = b$ puede expresarse como

$$X = \tilde{A}b + (I_n - \tilde{A}A)v$$

con v un vector arbitrario de $\mathbb{R}^n$. El segundo término es un vector general del núcleo de A, pues $I_n - \tilde{A}A$ es el proyector ortogonal sobre $\mathrm{Nu}(A)$ ($A(I - \tilde{A}A) = (A - A) = 0$), y representa una solución general del sistema homogéneo $AX = 0$. El primer término $\tilde{A}b$ es una solución particular de $AX = b$, y es **la solución particular de norma mínima**, pues es ortogonal a $(I - \tilde{A}A)w\ \forall\ w$ (ya que pertence al espacio fila de A).

En el caso general no necesariamente compatible, $X = \tilde{A}b$ es el vector de norma mínima que minimiza $||AX - b||$.

21. Espacios semieuclídeos y pseudoeuclídeos

Resumen. Para $\mathrm{dim}V = 2$ estos espacios quedan definidos por una forma bilineal $(v, w)_G = [v]_e^t[G]_e[w]_e$, con

$$[G]_e = \left(\begin{smallmatrix}0 & 0\\ 0 & 1\end{smallmatrix}\right)$$

en el caso semieuclídeo, tal que $(v, w)_G = yy'$, $(v, v)_G = y^2$ si $[v]_e = \left(\begin{smallmatrix}x\\ y\end{smallmatrix}\right)$, $[w]_e = \left(\begin{smallmatrix}x'\\ y'\end{smallmatrix}\right)$, y

$$[G]_e = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$$

en el pseudoeuclídeo, tal que $(v, w)_G = xx' - yy'$, $(v, v)_G = x^2 - y^2$. En estos casos $(v, v)_G$ puede ser 0 aun si $v \neq 0$, y en el caso pseudoeuclídeo puede ser también negativo.
Se demostró en clase que las transformaciones reales $\left(\begin{smallmatrix}x\\ y\end{smallmatrix}\right) = S\left(\begin{smallmatrix}x'\\ y'\end{smallmatrix}\right)$ que preservan estas formas bilineales (tales que $[G]_{e'} = S^t[G]_eS = [G]_e$) corresponden en el caso semieuclídeo a

$$S = \left(\begin{smallmatrix}a & b\\ 0 & d\end{smallmatrix}\right)$$

con $d = \pm 1$, y a, b arbitrarios, $a \neq 0$, y en el caso pseudoeuclídeo a

$$S = \left(\begin{smallmatrix}s\cosh(z) & s'\sinh(z)\\ s\sinh(z) & s'\cosh(z)\end{smallmatrix}\right)$$

con $s = \pm 1$, $s' = \pm 1$ y z arbitrario.
En particular, estas transformaciones comprenden las transformaciones de Galileo

$$\left(\begin{smallmatrix}x\\ t\end{smallmatrix}\right) = \left(\begin{smallmatrix}1 & v\\ 0 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}x'\\ t'\end{smallmatrix}\right)$$

en el caso semieuclídeo ($a = d = 1$, $b = v$) y las transformaciones de Lorentz

$$\left(\begin{smallmatrix}x\\ ct\end{smallmatrix}\right) = \left(\begin{smallmatrix}\cosh z & \sinh z\\ \sinh z & \cosh z\end{smallmatrix}\right)\left(\begin{smallmatrix}x'\\ ct'\end{smallmatrix}\right)$$

en el caso pseudoeuclídeo, con $\tanh(z) = v/c$, $s = s' = 1$, tal que $\cosh z = \frac{1}{\sqrt{1-v^2/c^2}}$, $\sinh z = \frac{v/c}{\sqrt{1-v^2/c^2}}$.
Para $v/c \to 0$, las transformaciones de Lorentz en las variables (x, t) se reducen a las de Galileo:

$$\left(\begin{smallmatrix}x\\ t\end{smallmatrix}\right) = \left(\begin{smallmatrix}\cosh z & c\sinh z\\ \frac{1}{c}\sinh z & \cosh z\end{smallmatrix}\right)\left(\begin{smallmatrix}x'\\ t'\end{smallmatrix}\right) \underset{v/c\to 0}{\longrightarrow} \left(\begin{smallmatrix}1 & v\\ 0 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}x'\\ t'\end{smallmatrix}\right)$$

Recordemos que para $n = 2$, las transformaciones que dejan invariante el producto escalar euclídeo son de la forma

$$S = \left(\begin{smallmatrix}s\cos\theta & -s'\sin\theta\\ s\sin\theta & s'\cos\theta\end{smallmatrix}\right)$$

con $s = \pm 1$, $s' = \pm 1$, que representan rotaciones (si $|S| = ss' = 1$) o reflexiones ($ss' = -1$).

22 Formas bilineales complejas

Sea V un espacio vectorial sobre el cuerpo de los complejos $\mathbb{C}$. Una función $A : V \times V \to \mathbb{C}$ se dice que es una forma bilineal *hermítica* si

$$A(v_1 + v_2, w) = A(v_1, w) + A(v_2, w), \quad A(v, w_1 + w_2) = A(v, w_1) + A(v, w_2)$$

$$A(v, \alpha w) = \alpha A(v, w), \quad A(\alpha v, w) = \alpha^* A(v, w)$$

$\forall\ v, v_1, v_2, w, w_1, w_2 \in V$ y $\alpha \in \mathbb{C}$. Nótese que α sale como conjugado cuando está en el primer miembro. Si V es de dimensión finita n y $e = (e_1, \ldots, e_n)$ es una base de V, escribiendo $v = \sum_{1=1}^{n} \alpha_i e_i$, $w = \sum_{j=1}^{n} \beta_j e_j$, con $[v]_e = (\alpha_1, \ldots, \alpha_n)^t$, $[w]_e = (\beta_1, \ldots, \beta_n)^t$, se obtiene

$$A(v, w) = \sum_{i,j=1}^{n} \alpha_i^* \beta_j A(e_i, e_j) = [v]_e^\dagger [A]_e [w]_e$$

donde el símbolo $\dagger$ denota traspuesto conjugado ($[v]_e^\dagger \equiv ([v]_e^t)^*$) y $[A]_e$ es la matriz de $n \times n$ de elementos

$$([A]_e)_{ij} = A(e_i, e_j)$$

Ejemplo: La siguiente es una forma bilineal de $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$:

$$\begin{aligned} A(v, w) &= \alpha_1^* \beta_1 + (1 + i)\alpha_1^* \beta_2 + (1 - i)\alpha_2^* \beta_1 + 2\alpha_2^* \beta_2 \\ &= (\alpha_1^*, \alpha_2^*) \begin{pmatrix} 1 & 1 + i \\ 1 - i & 2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \end{aligned}$$

donde hemos escrito en la base canónica $v = (\alpha_1, \alpha_2) = \alpha_1 e_1 + \alpha_2 e_2$, $w = (\beta_1, \beta_2) = \beta_1 e_1 + \beta_2 e_2$. A queda entonces representada en esta base por la matriz

$$[A]_e = \begin{pmatrix} 1 & 1 + i \\ 1 - i & 2 \end{pmatrix}$$

con $A(e_1, e_1) = 1$, $A(e_1, e_2) = 1 + i$, $A(e_2, e_1) = 1 - i$, $A(e_2, e_2) = 2$.

Si $A(v, w) = A(w, v)^*\ \forall v, w \Rightarrow$ la forma bilineal se dice que es hermíticamente simétrica y si $A(v, w) = -A(w, v)^*$, hermíticamente antisimétrica En el primer caso, la matriz que la representa es *hermítica*: $[A]_e^\dagger = [A]_e$, ya que $A(e_i, e_j) = A(e_j, e_i)^*$, y en el segundo caso *antihermítica*: $[A]_e^\dagger = -[A]_e$. Análogamente, si $[A]_e^\dagger = \pm[A]_e$, A es herm. simétrica $(+)$ o antisimétrica $(-)$. El ejemplo anterior corresponde a una forma bilineal herm. simétrica.

Notemos que una forma bilineal compleja que satisface las 4 condiciones no puede ser simplemente simétrica o antisimétrica a no ser que sea nula: Si $A(v, w) = \pm A(w, v)\ \forall\ v, w, \Rightarrow A(\alpha v, w) = \alpha^* A(v, w) = \pm A(w, \alpha v) = \pm\alpha A(w, v) = \alpha A(v, w)\ \forall\ \alpha,\ v, w$, lo que implica $(\alpha - \alpha^*)A(v, w) = 0$, es decir $A(v, w) = 0\ \forall\ v, w$.

Notemos también que toda forma bilineal puede expresarse como suma de una forma bilineal herm. simétrica y una forma bilineal herm. antisimétrica:

$$A(v, w) = \frac{1}{2}[A(v, w) + A(w, v)^*] + \frac{1}{2}[A(v, w) - A(w, v)^*]$$

22.1 Formas cuadráticas complejas

En forma análoga al caso real, la función $Q : V \to \mathbb{C}$ definida por

$$Q(v) = A(v, v)$$

se denomina *forma cuadrática* y satisface

$$Q(\alpha v) = A(\alpha v, \alpha v) = \alpha^* \alpha A(v, v) = |\alpha|^2 Q(v)$$

Una diferencia importante con las formas bilineales reales es que ahora la forma cuadrática determina completamente la forma bilineal (y no solamente la parte simétrica, como en el caso real). En efecto, podemos expandir $Q(v + w) = A(v + w, v + w)$ y $Q(v + iw) = A(v + iw, v + iw)$ como

$$Q(v + w) = Q(v) + Q(w) + A(v, w) + A(w, v),$$

$$Q(v+iw) = Q(v)+Q(w)+i[A(v,w)-A(w,v)]$$

de donde

$$\begin{aligned} A(v,w) &= Q(v+w)-iQ(v+iw)-(1-i)(Q(v)+Q(w)) \\ A(w,v) &= Q(v+w)+iQ(v+iw)-(1+i)(Q(v)+QA(w)) \end{aligned}$$

De aquí se deduce también una propiedad fundamental:
$Q(v)$ es *real* $\forall\ v$ si y sólo si $A(v,w)=[A(w,v)]^*\ \forall\ v,w$, es decir, sii la forma bilineal asociada es hermíticamente simétrica.
En efecto, de las expresiones anteriores se ve que si $Q(v)\in\mathbb{R}\ \forall\ v\in V \Rightarrow A(w,v)=[A(v,w)]^*\ \forall\ v,w\in V$. Y si $A(w,v)=[A(v,w)]^*\ \forall\ v,w \Rightarrow Q(v)=A(v,v)=[A(v,v)]^*$ es real. Formas cuadráticas *reales* determinan pues *formas bilineales hermíticamente simétricas y viceversa.*
Por otro lado, si $A(v,w)$ es hermíticamente simétrica, $A'(v,w)=iA(v,w)$ es hermíticamente antisimétrica. La forma cuadrática asociada $Q'(v)=A'(v,v)$ es obviamente imaginaria.
Ejemplo: La forma bilineal del ej. anterior origina la forma cuadrática

$$\begin{aligned} Q(v)=A(v,v)=(\alpha_1^*,\alpha_2^*)\begin{pmatrix}1 & 1+i\\ 1-i & 2\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} &= \alpha_1^*\alpha_1+2\alpha_2^*\alpha_2+(1+i)\alpha_1^*\alpha_2+(1-i)\alpha_2^*\alpha_1 \\ &= |\alpha_1|^2+2|\alpha_2|^2+2\mathrm{Re}[(1+i)\alpha_1^*\alpha_2] \end{aligned}$$

que es obviamente *real.*

22.2 Cambio de base

Si efectuamos un cambio de base $e'_i=\sum_{j=1}^n S_{ji}e_j$, con $|S|\neq 0$,

$$A(e'_i,e'_j)=A(\sum_{k=1}^n S_{ki}e_k,\sum_{l=1}^n S_{lj}e_l)=\sum_{k,l} S^*_{ki}S_{lj}A(e_k,e_l)$$

por lo que

$$[A]_{e'}=S^\dagger[A]_eS$$

donde $\dagger$ denota por su puesto la operación de traspuesto+conjugado.
Notemos que $\mathrm{Det}([A]_{e'})=|\mathrm{Det}(S)|^2\mathrm{Det}([A]_e)$, por lo que la fase del determinante es la misma en cualquier base. Obtenemos entonces

$$A(v,w)=[v]_e^\dagger[A]_e[w]_e=[v]_{e'}^\dagger[A]_{e'}[w]_{e'}$$

donde $[w]_{e'}=R[w]_e$, $[v]_{e'}^\dagger=[v]_eR^\dagger$ y $R=S^{-1}$.

22.3 Base canónica: Si A es herm. simétrica, *existe una base* e' (base canónica) donde $[A]'_e$ es diagonal:

$$[A]_{e'}=S^\dagger[A]_eS=\begin{pmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ & & \dots & \\ 0 & 0 & \dots & \lambda_n\end{pmatrix}$$

En esta base, si $v=\sum_{i=1}^n\alpha'_ie'_i$, $w=\sum_{i=1}^n\beta'_ie'_i$, tenemos

$$A(v,w)=\sum_{i=1}^n\lambda_i{\alpha'_i}^*\beta'_i$$

La demostración de la existencia de esta base puede efectuarse en forma similar al caso real, completando ahora módulos cuadrados, y se deja comos ejercicio. Sug.: Llamando $a_{ij}=([A]_e)_{ij}$ (con $a_{ji}=a^*_{ij}$) y asumiendo $a_{nn}\neq 0$, escribir la parte que contiene α_n y α_n^* en $A(v,v)$ como

$$a_{nn}\alpha_n^*\alpha_n+\sum_{j=1}^{n-1}(a_{nj}\alpha_n^*\alpha_j+a^*_{nj}\alpha_j^*\alpha_n)=a_{nn}{\alpha'_n}^*\alpha'_n-(\sum_{j=1}^{n-1}a^*_{nj}\alpha^*_j)(\sum_{j=1}^{n-1}a_{nj}\alpha_j)/a_{nn}$$

con $\alpha'_n=\alpha_n+\frac{1}{a_{nn}}\sum_{j=1}^{n-1}a_{nj}\alpha_j$, y proceder luego por inducción. Si $a_{nn}=0$ se comienza con una variable α_i tal que $a_{ii}\neq 0$, y si $a_{ii}=0\ \forall i$ se efectúa un cambio de variables simple para que a_{ii} sea no nulo para algún i (por ej., si $a_{ij}=a^*_{ji}\neq 0$, $a_{ij}\alpha^*_i\alpha_j+a_{ji}\alpha^*_j\alpha_i=2|a_{ij}|^2(|\alpha'_i|^2-|\alpha'_j|^2)$, con $\alpha_i=a_{ij}(\alpha'_i+\alpha_j)$, $\alpha'_j=\alpha_i-\alpha_j$. El cambio $\alpha'_i=\sum_{j=1}^n R_{ij}\alpha_j$ define una base $e'_i=\sum_{j=1}^n S_{ji}e_j$, con $S=R^{-1}$, en la que $[A]_{e'}=S^\dagger[A]S$ es diagonal.

Otra forma de demostrar la existencia es directamente diagonalizando la matriz $[A]_e$, que es en este caso hermítica y por lo tanto diagonalizable en una base ortonormal, tal que $S^{-1}=S^\dagger$ y $S^\dagger[A]_eS$ es diagonal. No obstante esto supone haber demostrado antes que tales matrices son diagonalizables, lo que nosotros realizaremos luego.

La base canónica no es única. Una base canónica puede obtenerse, al igual que en el caso real, completando módulos cuadrados o bien diagonalizando la matriz $[A]_e$.

Ejemplo: Hallar una base canónica para el ejemplo previo. Completando módulos cuadrados, obtenemos

$$Q(v)=(\alpha^*_1+(1-i)\alpha^*_2)(\alpha_1+(1+i)\alpha_2)+|\alpha_2|^2[2-(1+i)(1-i)]=|\alpha'_1|^2+0|\alpha'_2|^2$$

donde $\alpha'_1=\alpha_1+(1+i)\alpha_2$, $\alpha'_2=\alpha_2$, o sea $\binom{\alpha'_1}{\alpha'_2}=\left(\begin{smallmatrix}1&1+i\\0&1\end{smallmatrix}\right)\binom{\alpha_1}{\alpha_2}$. La matriz de cambio de base es entonces

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

y se verifica

$$[A]_{e'}=S^\dagger[A]_eS=\begin{pmatrix}1&0\\0&0\end{pmatrix}$$

Alternativamente, diagonalizando la matriz $[A]_e$ se obtienen los autovalores y autovectores

$$\lambda_1=1,\ v'_1=(1+i,2),\qquad \lambda_2=0,\ v_2=(-1-i,1)$$

Normalizando los autovectores, la matriz de cambio de base es entonces

$$S=\begin{pmatrix}(1+i)/\sqrt{6}&-(1+i)/\sqrt{3}\\2/\sqrt{6}&1/\sqrt{3}\end{pmatrix}$$

con $S^{-1}=S^\dagger$ (pues los autovectores en S están normalizados). Se obtiene así la reprentación diagonal

$$[A]_{e'}=S^\dagger[A]_eS=\begin{pmatrix}3&0\\0&0\end{pmatrix}$$

Vemos que el número de coeficientes diagonales positivos y nulos en las dos formas diagonales obtenidas es el mismo. Esta propiedad es general y constituye el

22.4 Teorema de Inercia para formas cuadráticas hermíticas: Si QA es una forma cuadrática herm. simétrica, el número de términos diagonales positivos, negativos, y nulos en una representación diagonal arbitraria es siempre el mismo. Se demuestra igual que en el caso real (Demostrar como ejercicio).
Es importante notar que el teorema de inercia no vale para formas cuadráticas comunes extendidas a los complejos: Si $Q(v)=\alpha_1^2+\alpha_2^2$, la transformación $\alpha'_1=i\alpha_1$, $\alpha'_2=\alpha_2$ la lleva a $-\alpha'^2_1+\alpha'^2_2$. Tal forma cuadrática no proviene de una forma bilineal hermítica, ya que no cumple $Q(\alpha v)=|\alpha|^2Q(v)$.

22.5 Formas cuadráticas positivas:
Una forma cuadrática se denomina *definida positiva* (o estríctamente positiva) si $Q(v)>0\ \forall\ v\neq 0$, y semipositiva (o no negativa) si $Q(v)\geq 0\ \forall\ v$ (obviamente, en cualquier caso, $Q(0)=0$). Por ser reales, estas formas cuadráticas están necesariamente asociadas a formas bilineales hermíticamente simétricas.
La forma cuadrática es pues definida positiva si y sólo si los coeficientes diagonales en una base canónica satisfacen $a_{ii}>0\ \forall\ i$ y semipositiva sii $a_{ii}\geq 0\ \forall\ i$. El teorema de inercia asegura que el número de coeficientes diagonales positivos y nulos para estas formas cuadráticas (pero no su valor particular) es siempre el mismo en cualquier base canónica.
En form análoga, una matriz $A\in\mathbb{C}^{n\times n}$ se dice definida positiva si $X^\dagger AX>0\ \forall\ X\neq 0$, $X\in\mathbb{C}^n$, en cuyo

caso podemos considerarla como la representación en la base canónica de $V=\mathbb{C}^n$ de una forma cuadrática definida positiva. Notemos que necesariamente A debe ser *hermítica* $(A^\dagger = A)$, para que $X^\dagger AX$ sea real.
Una matriz hermítica A es pues definida positiva si y sólo si todos sus autovalores son positivos.

Notemos que si $Q_A(v)$ es una forma cuadrática definida positiva $\Rightarrow$ *existe una base canónica e'' donde* $A(e''_i, e''_j) = \delta_{ij}$, es decir,

$$[A]_{e''} = I_n$$

(matriz identidad). En efecto, existirá una base canónica, obtenida completando módulos cuadrados o diagonalizando, en la que $A(e'_i, e'_j) = ([A]_{e'})_{ij} = \lambda_i \delta_{ij}$, con $\lambda_i > 0\ \forall\ i$. En la nueva base definida por $e''_i = e'_i/\sqrt{\lambda_i}$ tendremos $A(e''_i, e''_j) = A(e'_i, e'_j)/\sqrt{\lambda_i\lambda_j} = \lambda_i\delta_{ij}/\sqrt{\lambda_i\lambda_j} = \delta_{ij}$ para $i,j = 1,\ldots,n$.
Esto implica que existe una matriz $S = [I]_e^{e''}$ no singular tal que

$$[A]_{e''} = S^\dagger [A]_e S = I_n$$

con I_n la matriz identidad.
Esto implica a su vez que toda matriz $A \equiv [A]_e$ definida positiva puede escribirse como

$$A = R^\dagger R$$

con $R = S^{-1}$ no singular. La recíproca es también válida: La matriz $R^\dagger R$ es definida positiva $\forall\ R$ no singular (probar como ejercicio).

Para saber si una forma cuadrática es definida positiva o semipositiva se completan módulos cuadrados o se obtienen los autovalores de la matriz que la representa en alguna base, y se observan lo signos de los coeficientes diagonales resultantes. El determinante de una matriz asociada a una forma cuadrática definida positiva es obviamente positivo en cualquier base ($\mathrm{Det}[A] = \mathrm{Det}[R^\dagger R] = |\mathrm{Det}[R]|^2$), aunque esta condición no garantiza que A sea definida positiva.

Es válido no obstante el siguiente teorema: **Una matriz hermítica A de $n\times n$ es definida positiva si y sólo si todos sus determinantes principales son positivos ($\mathrm{Det}(A_m) > 0$, $m = 1,\ldots,n$).**
Dem.: Si A es definida positiva la forma cuadrática asociada debe ser positiva en cualquier subespacio de V y por lo tanto cualquier submatriz de A (obtenida quitando un conjunto de filas y las resp. columnas) es definida positiva. En particular, todas las submatrices principales ($A_m = [a_{ij}], i,j = 1,\ldots,m$, $m = 1,\ldots,n$) son definidas positivas y sus determinantes por ende positivos.
Recíprocamente, si todos los determinantes principales son positivos, procedemos por inducción sobre n. Para $n = 1$ se cumple trivialmente. Asumiendo la submatriz principal de $(n-1)\times(n-1)$ definida positiva, existirá una base $(e'_1,\ldots,e'_{n-1})$ del subespacio correspondiente en la que $A(e'_i, e'_j) = \delta_{ij}$. Definimos ahora $e'_n = e_n - \sum_{i=1}^{n-1}\alpha_i e'_i$, con $\alpha_i = A(e'_i, e_n)$, tal que $A(e'_i, e'_n) = A(e'_i, e_n) - \alpha_i = 0$ para $i = 1,\ldots,n-1$. En tal caso, $[A]_{e'}$ será diagonal y con todos sus elementos diagonales positivos, pues $A(e'_i, e'_i) = 1$ para $i < n$ y $A(e'_n, e'_n) = \mathrm{Det}([A]_{e'}) > 0$ por hipótesis (el signo del determinante no cambia al cambiar la base).

Recordemos también aqui los **círculos de Gershgorin:** Si $A \in \mathbb{C}^{n\times n}$ es una matriz cuadrada de elementos a_{ij}, entonces sus autovalores se encuentran en la unión de los círculos (en el plano complejo)

$$|\lambda - a_{ii}| \leq |\sum_{j\neq i}|a_{ij}|$$

En efecto, sea $v = (x_1,\ldots,x_n)^t \in \mathbb{C}^{n\times 1}$, $v \neq 0$, un autovector de A asociado al autovalor λ, tal que $Av = \lambda v$, es decir, $\sum_j a_{ij}x_j = \lambda x_i$. Si $|x_i| = \mathrm{Max}[|x_1|,\ldots,|x_n|] \neq 0$ es el módulo de la coordenada de v de módulo máximo, tenemos, dado que $(\lambda - a_{ii})x_i = \sum_{j\neq i} a_{ij}x_j$,

$$|\lambda - a_{ii}| = |\sum_{j\neq i} a_{ij}x_j/x_i| \leq \sum_{j\neq i}|a_{ij}|x_j/x_i| \leq \sum_{j\neq i}|a_{ij}|$$

Una desigualdad simular es válida para sumas sobre columnas, ya que los autovalores de A son idénticos a los de A^t.

En el caso de matrices hermíticas, tanto los elementos diagonales como los autovalores son todos reales. La cota anterior implica entonces la siguiente condición suficiente (aunque no necesaria) de positividad de una matriz hermítica A: Si $a_{ii} > 0\ \forall\ i$ y $\sum_{j\neq i}|a_{ij}| < a_{ii}\ \forall\ i$, los autovalores serán todos positivos y por ende A será definida positiva.

23 Espacios Unitarios (Espacios de Hilbert)

Un espacio vectorial V sobre $\mathbb{C}$ se denomina unitario o espacio de Hilbert si está equipado con una operación $V \times V \to \mathbb{C}$, denominada *producto interno* o producto escalar, y denotada por (v, w), que satisface

$$(v,w) = (w,v)^*, \quad (v, \alpha w) = \alpha(v,w), \quad (v, w_1 + w_2) = (v, w_1) + (v, w_2)$$

$$(v,v) > 0 \; \forall \; v \neq 0$$

Es decir, el producto interno no es otra cosa que una forma bilineal hermíticamente simétrica y *definida positiva.* En el caso de dimensión infinita, un espacio de Hilbert debe ser además completo: Si $\{u_n\}$ es una sucesión de vectores tal que $\sum_{n=0}^{\infty} ||u_n||$ es convergente entonces $\lim_{n\to\infty} u_n$ debe pertenecer al espacio.

En en el caso de dimensión finita, en una base arbitraria e tendremos, denotando con $[A]_e$ la matriz de elementos $a_{ij} = (e_i, e_j) = a_{ji}^*$,

$$(v,w) = [v]_e^\dagger [A]_e [w]_e = \sum_{i,j=1}^{n} \alpha_i^* a_{ij} \beta_j$$

donde $v = \sum_{i=1}^n \alpha_i e_i$, $w = \sum_{i=1}^n \beta_i e_i$
Y si e denota ahora la base canónica en la que $(e_i, e_j) = \delta_{ij}$, obtenemos la forma corriente

$$(v,w) = [v]_e^\dagger [w]_e = \sum_{i=1}^{n} \alpha_i^* \beta_i$$

Esta base es una *base ortonormal* para el producto escalar $((e_i, e_j) = \delta_{ij})$. En esta base,

$$(v,v) = [v]_e^\dagger [v]_e = \sum_{i=1}^{n} |\alpha_i|^2$$

Ejemplo: En $\mathbb{C}^n$, el producto interno usual en la base canónica está dado por

$$(v,w) = \sum_{i=1}^{n} x_i^* y_i, \quad (w,v) = \sum_{i=1}^{n} y_i^* x_i = (v,w)^*$$

para $v = (x_1, \ldots, x_n)$, $w = (y_1, \ldots, w_n)$, lo que implica $(v,v) = \sum_{i=1}^n |x_i|^2$.

Ejemplo: En el espacio de funciones complejas de parte real e imaginaria continua, $C_{[a,b]} = \{f : \mathbb{R} \to \mathbb{C}, \;\; f = f^r + i f^i, \; f^r, f^i \in R_{[a,b]}\}$, con $a < b$, el producto interno usual está dado por

$$(f,g) = \int_a^b f^*(x) g(x) dx = (g,f)^*$$

con $(f,f) = \int_a^b f^*(x) f(x) dx = \int_a^b |f(x)|^2 dx > 0$ si $f \neq 0$.
Ejemplo: En el caso de matrices complejas de $m \times n$, $V = \mathbb{C}^{m\times n}$, podemos definir el producto escalar

$$(A,B) = \mathrm{Tr}\,[A^\dagger\, B] = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}^* B_{ij}$$

con $(A,A) = \sum_{i,j} |A_{ij}|^2 > 0 \; \forall \; A \neq 0$.

En los espacios unitarios son válidas propiedades similares a las de espacios euclídeos. En particular:

La **norma** de un vector se define por

$$||v|| = \sqrt{(v,v)} \geq 0$$

con $||v|| = 0$ sii $v = 0$ y $||\alpha v|| = |\alpha|\, ||v||$. La distancia entre dos vectores es $d(v,w) = ||v - w||$.

La desigualdad de **Cauchy-Schwarz** también se verifica:

$$|(v,w)| \leq ||v||\, ||w||$$

donde la igualdad vale si y sólo si $\{v, w\}$ son LD.
Demostración: Si $v = 0$ o $w = 0$ la igualdad se cumple trivialmente: $0 = (v, w) = ||v||\,||w||$.
Idem si v y w son LD: En tal caso $w = \alpha v$ (o $v = \alpha w$) y por lo tanto $|(v, w)| = |\alpha(v, v)| = |\alpha|\,||v||^2 = ||v||\,||w||$.
Si $v \neq 0$ y $w \neq 0$, denotemos con $v_n = v/||v||$, $w_n = w/||w_n||$ los vectores normalizados ($||v_n|| = ||w_n|| = 1$), tal que $(v, w) = (v_n, w_n)||v||\,||w||$. Se obtiene, para s un número complejo arbitrario de módulo 1 ($|s| = 1$),

$$0 \leq (v_n - sw_n, v_n - sw_n) = ||v_n||^2 + |s|^2||w_n||^2 - s(v_n, w_n) - s^*(w_n, v_n) = 2 - 2\mathrm{Re}[s(v_n, w_n)] = 2(1 - \mathrm{Re}[s(v_n, w_n)])$$

Recordemos ahora que todo número complejo z puede escribirse como $z = |z|e^{i\phi}$, con $|z| = \sqrt{zz^*}$ (módulo) y ϕ reales. Por lo tanto, si $z = (v_n, w_n) = |(v_n, w_n)|e^{i\phi}$, eligiendo $s = e^{-i\phi}$ se obtiene

$$0 \leq 1 - |(v_n, w_n)|$$

de donde $|(v_n, w_n)| \leq 1$. Por lo tanto, $|(w, v)| = |(v, w)| \leq ||v||\,||w||$, q.e.d.
Además, si $|(w, v)| = 1 \Rightarrow |(w_n, v_n)| = 1$ y $(v_n - sw_n, v_n - sw_n) = 0$, por lo que $v_n - sw_n = 0$, es decir, $v = sw||v_n||/||w_n||$, lo que implica que v, w son L.D.
Las desigualdades triangulares *permanecen válidas* en espacios unitarios, por la vigencia de la desigualdad anterior: $|||v|| - ||w||| \leq ||v + w|| \leq ||v|| + ||w||$.
No obstante, no se pueden definir ahora ángulos entre vectores pues $(v, w)/(||v||\,||w||)$, si bien tiene módulo menor que 1, es en general complejo.
Ejemplo: Dados $v = (1 + i, i), w = (i, 1 + i) \in \mathbb{C}^2$, tenemos

$$(v, w) = (1 - i)i + (-i)(1 + i) = 2 \leq ||v||\,||w|| = \sqrt{|1 + i|^2 + 1}\sqrt{1 + |1 + i|^2} = \sqrt{3}\sqrt{3} = 3$$

Notación de Mecánica Cuántica:
La notación empleada en mecánica cuántica para los vectores de estado de un sistema (que pertenecen a un espacio de Hilbert) es $|v\rangle$, y para el producto interno $\langle w|v\rangle$. Es decir, $v \to |v\rangle$, $(w, v) \to \langle w|v\rangle$, con $\langle v|w\rangle = \langle w|v\rangle^*$.

23.1 Ortogonalidad y Método de ortogonalización Gram-Schmidt

Las propiedades de ortogonalidad son análogas al caso euclídeo. Dos vectores v, w de un espacio unitario son **ortogonales** si $(v, w) = 0$.
Al igual que en el caso euclídeo, dado un conjunto de m vectores v_i L.I., es posible construir con el método de Gram-Schmidt un conjunto ortogonal de vectores que genera el mismo espacio que los v_i, dados por:

$$w_1 = v_1\,, \quad w_i = v_i - \sum_{j=1}^{i-1} P_{w_j}(v_i)\,, \quad i = 2, \ldots, m$$

donde

$$P_{w_j}(v_i) = \frac{(w_j, v_i)}{||w_j||^2} w_j$$

es la *proyección ortogonal* de v_i sobre w_j. Notemos que en el caso complejo es necesario ser cuidadoso con el orden en el producto escalar, ya que $(w_j, v_i) \neq (v_i, w_j) = (w_j, v_i)^*$. Es fácil verificar que de esta forma, $(w_i, w_j) = 0$ si $i \neq j$, siendo los w_i no nulos si los vectores originales son L.I.
Dada una base arbitraria de V, es pues siempre posible por este método construir una base ortogonal de V, que puede convertirse en *ortonormal* normalizando los vectores resultantes.
Notemos que el cuadrado de la norma de los w_i está dado, para $i > 1$, por

$$||w_i||^2 = (w_i, w_i) = (v_i, w_i) = ||v_i||^2 - \sum_{j=1}^{i-1} \frac{|(w_j, v_i)|^2}{||w_j||^2} \leq ||v_i||^2$$

Notemos también que la matriz que representa al proyector sobre w_i en la base canónica es

$$[P_{w_i}]_e = \frac{[w_i]_e [w_i]_e^\dagger}{||w_i||^2}$$

Ejemplo 1 : Consideremos los vectores $v_1=(1+i,i,0)$, $v_2=(i,1+i,1)$. Tenemos

$$w_1=v_1=(1+i,i,0),\ \ w_2=v_2-\frac{(w_1,v_2)}{||w_1||^2}w_1=(i,1+i,1)-\frac{2}{3}(1+i,i,0)=(-2+i,3+i,3)/3$$

que verifican $(w_1,w_2)=0$.

Ejemplo 2: Las funciones $f_k(x)=e^{ikx}$, con k *entero*, son ortogonales con el producto interno $(f,g)=\int_{-\pi}^{\pi}f^*(x)g(x)dx$:

$$(f_{k'},f_k)=\int_{-\pi}^{\pi}e^{-ik'x}e^{ikx}dx=\int_{-\pi}^{\pi}e^{ix(k-k')}dx=\begin{cases}2\pi & k=k'\\ \frac{e^{ix(k-k')}}{i(k-k')}|_{-\pi}^{\pi}=0 & k\neq k'\end{cases}$$

Ejemplo 3 (Transformada de Fourier discreta): Sea $V=\mathbb{C}^n$ y sea $e=(e_1,\ldots,e_n)$ una base canónica $((e_i,e_j)=\delta_{ij})$. Los n vectores

$$\tilde{e}_k=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}e^{i2\pi kj/n}e_j$$

forman también una base ortonormal: $(\tilde{e}_k,\tilde{e}_l)=\delta_{kl}$.
En efecto, utilizando que $(e_i,e_j)=\delta_{ij}$ obtenemos, para $k,l=1,\ldots,n$,

$$(\tilde{e}_k,\tilde{e}_l)=\frac{1}{n}\sum_{j=1}^{n}e^{i2\pi j(l-k)/n}=\begin{cases}1 & k=l\\ \frac{1}{\sqrt{n}}\frac{1-e^{i2\pi(l-k)}}{1-e^{i2\pi(l-k)/n}}=0 & k\neq l\end{cases}$$

Ejemplo 4: Obtener una base ortonormal de $\mathbb{C}^{2\times 2}$ (con escalares complejos) para el producto escalar $(A,B)=\mathrm{Tr}\,A^\dagger B$, partiendo de $v_1=I_2=\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$.
Consideremos las matrices $v_1=I_2$, $v_2=\left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right)$, $v_3=\frac{1}{2}\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$, $v_4=\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)$, que forman una base no ortogonal de $\mathbb{C}^{2\times 2}$. Obtenemos, $w_1=v_1=I_2$,

$$w_2=v_2-\tfrac{1}{2}(w_1,v_2)w_1=v_2-\tfrac{1}{2}w_1=\tfrac{1}{2}\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right),\quad w_3=v_3-\tfrac{1}{2}(w_1,v_3)w_1-2(w_2,v_3)w_2=v_3=\tfrac{1}{2}\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right),$$
$$w_4=v_4-\tfrac{1}{2}(w_1,v_4)w_1-2(w_2,v_3)w_2-2(w_3,v_4)w_3=v_4-w_3=\tfrac{1}{2}\left(\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$$

Las matrices de Pauli se definen precisamente como

$$\sigma_0=I_2=\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right),\ \ \sigma_x=2w_3=\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right),\ \ \sigma_y=-2iw_4=\left(\begin{smallmatrix}0&-i\\i&0\end{smallmatrix}\right),\ \ \sigma_z=2w_2=\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)$$

y forman una base ortogonal y hermítica de $\mathbb{C}^{2\times 2}$: $\sigma_\mu^\dagger=\sigma_\mu$, $(\sigma_\mu,\sigma_\nu)=\mathrm{Tr}\,\sigma_\mu\sigma_\nu=2\delta_{\mu\nu}$.

Considerando ahora $\mathbb{C}^{2\times 2}$ sobre escalares reales, estas 4 matrices forman también una base del subespacio de matrices hermíticas de 2×2. Las matrices que representan las componentes del espín $\mathbf{s}=(s_x,s_y,s_z)$ en la base estándar de autoestados de s_z son precisamente

$$s_\mu=\frac{1}{2}\hbar\sigma_\mu,\ \ \mu=x,y,z$$

23.2 Expansión en una base ortonormal

Si $e=(e_1,\ldots,e_n)$ es una base ortonormal de V $((e_i,e_j)=\delta_{ij})$ y

$$v=\sum_{i=1}^{n}x_ie_i$$

entonces $(e_i,v)=\sum_{j=1}^{n}x_j(e_i,e_j)=x_j$. Por lo tanto,

$$x_i=(e_i,v)$$

Se cumple entonces

$$v=\sum_{i=1}^{n}P_{e_i}(v)$$

Se verifica también, por la ortogonalidad de los e_i, la generalización del teorema de Pitágoras,

$$||v||=(v,v)=\sum_{i=1}^{n}||x_ie_i||^2=\sum_{i=1}^{n}|x_i|^2$$

Notemos que en la notación de mecánica cuántica, $e_i\to|i\rangle$ y $|v\rangle=\sum_i\alpha_i|i\rangle$, con $\alpha_i=\langle i|v\rangle$.

23.3 Proyectores ortogonales y matriz de Gram

Dado un subespacio $S \subset V$, es posible construir el complemento ortogonal $S_\perp = \{v \in V | (w, v) = 0 \; \forall \; w \in S\}$, cumpliéndose que $V = S \oplus S_\perp$ y por lo tanto, dim $S+$ dim $S_\perp = n$
Si $v \in V$, podemos escribir

$$v = v_s + (v - v_s)$$

con $v_s \in S$ y $v - v_s \in S_\perp$. Si $(w_1, \ldots, w_m)$ es una base *ortogonal* de S, escribiendo $v_s = \sum_{i=1}^m \alpha_i w_i$, la condición $(w_i, v - v_s) = 0$ para $i = 1, \ldots, m$ implica $\alpha_i = (w_i, v)/||w_i||^2$ y por lo tanto

$$v_s = \sum_{i=1}^m P_{w_i}(v) = P_S(v)\,, \qquad P_S = \sum_{i=1}^m P_{w_i}$$

El vector v_s es el vector de S con distancia mínima a v: Si $w_s \in S$,

$$||v - w_s||^2 = ||(v - v_s) + (v_s - w_s)||^2 = ||v - v_s||^2 + ||w_s - v_s||^2 \geq ||v - v_s||^2$$

En general, para una base arbitraria $(w_i, \ldots, w_m)$ de S no necesariamente ortogonal,

$$[P_S]_e = R(R^\dagger R)^{-1} R^\dagger$$

donde $R = ([w_1]_e, \ldots, [w_m]_e)$ es la matriz de $n \times m$ donde cada columna son las coordenadas de los m vectores w_i de la base de S en una base canónica de V ($(e_i, e_j) = \delta_{ij}$). Las fórmulas del caso euclídeo se generalizan pues directamente al caso unitario reemplazando t (traspuesta) por $\dagger$ (traspuesto conjugado).
Recordemos que

$$P_S + P_{S_\perp} = I$$

Notemos también que la matriz de Gram

$$G = R^\dagger R$$

de $m \times m$, con $G_{ij} = (w_i, w_j) = G^*_{ji}$, es ahora una matriz *hermítica*, que posee las mismas propiedades anteriores: $|G| \neq 0$ sii los m vectores w_i son LI, los autovalores λ_i de G son reales y no negativos, los autovectores X_i ($GX_i = \lambda_i X_i$) correspondientes a autovalores no nulos determinan vectores u_i de componentes $[u_i]_e = RX_i$, que son ortogonales ($(u_i, u_j) = 0$ si $\lambda_i \neq \lambda_j$), y aquellos correspondientes a autovalores nulos dan las combinaciones lineales nulas de los w_i (Demostraciones totalmente similares al caso euclídeo).

Ejemplo: Proyectar el vector $v = (1, i, 1+i) \in \mathbb{C}^3$ sobre el espacio generado por los vectores $v_1 = (1+i, i, 0)$, $v_2 = (i, 1+i, 1)$. Aplicando la representación general, tenemos

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

con $R^\dagger R = \left(\begin{smallmatrix} 3 & 2 \\ 2 & 4 \end{smallmatrix}\right)$, $(R^\dagger R)^{-1} = \left(\begin{smallmatrix} 4 & -2 \\ -2 & 3 \end{smallmatrix}\right)/8$ y

$$[P_S]_e = R(R^\dagger R)^{-1} R^\dagger = \begin{pmatrix} 7 & 1-i & -2+i \\ 1+i & 6 & 3+i \\ -2-i & 3-i & 3 \end{pmatrix} /8$$

Podemos arribar a este mismo resultado considerando también la base ortogonal de S obtenida previamente al ortogonalizar v_1 y v_2 por Gram-Schmidt, dada por $w_1 = v_1$, $w_2 = (-2+i, 3+i, 3)/3$:

$$[P_S]_e = [P_{w_1}]_e + [P_{w_2}]_e = \frac{[w_1]_e [w_1]_e^\dagger}{||w_1||^2} + \frac{[w_2]_e [w_2]_e^\dagger}{||w_2||^2}$$

$$= \frac{1}{3} \begin{pmatrix} 1+i \\ i \\ 0 \end{pmatrix} (1-i, -i, 0) + \frac{1}{24} \begin{pmatrix} -2+i \\ 3+i \\ 3 \end{pmatrix} (-2-i, 3-i, 3) = \begin{pmatrix} 7 & 1-i & -2+i \\ 1+i & 6 & 3+i \\ -2-i & 3-i & 3 \end{pmatrix} /8$$

Se obtiene finalmente

$$[P_S(v)]_e = [P_S]_e [v]_e = \begin{pmatrix} 5 \\ 3+11i \\ 2+5i \end{pmatrix} /8$$

La distancia mínima al plano es $||v - v_s|| = 3/\sqrt{8}$.

23.4 Operadores adjuntos y autoadjuntos en espacios unitarios

Sea $F: V \to V$ un operador lineal. El operador adjunto $F^\dagger$ se define por la relación

$$(v, F(w)) = (F^\dagger(v), w)$$

$\forall\ v, w \in V$. Considerando una base canónica e de V ($(e_i, e_j) = \delta_{ij}$), y teniendo en cuenta que $[F(v)]_e = [F]_e[v]_e$, y $(v, w) = [v]_e^\dagger[w]_e$, se obtiene $(v, F(w)) = [v]_e^\dagger[F]_e[w]_e$, $(F^\dagger(v), w) = [v]_e^\dagger[F^\dagger]_e^\dagger[w]_e$ y por lo tanto

$$[F^\dagger]_e = [F]_e^\dagger$$

La matriz que representa al operador adjunto de F en una base canónica es pues la traspuesta conjugada de la que representa a F en dicha base. Notemos que:
1) si $G = \alpha F$, con $\alpha \in \mathbb{C} \Rightarrow G^\dagger = \alpha^* F^\dagger$ (pues $(\alpha^* F^\dagger(v), w) = \alpha(F^\dagger(v), w) = \alpha(v, F(w)) = (v, \alpha F(w))$)
2) $(F^\dagger)^\dagger = F$ (pues $((F^\dagger)^\dagger(v), w) = (v, F^\dagger(w)) = (F(v), w)\ \forall\ v, w$)
3) $(FG)^\dagger = G^\dagger F^\dagger$ (pues $(v, FG(w)) = (F^\dagger(v), G(w)) = (G^\dagger F^\dagger(v), w)$).

Un operador F es *autoadjunto* si $F^\dagger = F$. En tal caso la matriz que lo representa en una base canónica es *hermítica*:

$$[F]_e^\dagger = [F]_e$$

Una propiedad importante de operadores adjuntos es que **si S es un subespacio invariante por $F \Rightarrow S_\perp$ es invariante por $F^\dagger$.**
Demostración: si $F(v) \in S\ \forall\ v \in S$, y $w \in S_\perp \Rightarrow (w, F(v)) = 0\ \forall\ w \in S_\perp$ y $v \in S$. Por lo tanto,

$$(F^\dagger(w), v) = (w, F(v)) = 0$$

$\forall\ w \in S_\perp$ y $v \in S$, de modo que $F^\dagger(w) \in S_\perp$
En particular, si F es *autoadjunto* y S es invariante por $F \Rightarrow S_\perp$ *es también invariante por* F.

Comentemos finalmente que en una base B general, donde $(v, w) = [v]_B^\dagger A[w]_B$ con A es una matriz hermítica definida positiva ($A^\dagger = A$, $X^\dagger AX > 0\ \forall\ X \neq 0$, $X \in \mathbb{C}^{n\times 1}$) la condición $(v, F(w)) = (F^\dagger(v), w)\ \forall\ v, w \in V$ implica $[F^\dagger]_B^\dagger A = A[F]_B$, y por lo tanto,

$$[F^\dagger]_B = A^{-1}[F]_B^\dagger A$$

La matriz que representa el operador adjunto $F^\dagger$ en una base arbitraria es pues semejante (pero no necesariamente igual) a $[F]_B^\dagger$.

23.5 Operadores Unitarios

Un operador lineal $U: V \to V$ que conserva el producto interno en un espacio unitario se denomina unitario:

$$(U(v), U(w)) = (v, w)$$

$\forall\ v, w \in V$. Como $(U(v), U(w)) = (U^\dagger U(v), w) \Rightarrow U^\dagger U = I$ (identidad), por lo que en una base canónica tenemos

$$[U]_e^\dagger[U]_e = I_n$$

y por lo tanto $[U]_e[U]_e^\dagger = I_n$. Las matrices que representan a un operador unitario en una base canónica se denominan unitarias y satisfacen $[U]_e^{-1} = [U]_e^\dagger$, lo que implica filas y columnas ortonormales:

$$\sum_{j=1}^n S_{ji}^* S_{jk} = \delta_{ik} \quad \sum_{j=1}^n S_{ij} S_{kj}^* = \delta_{ik}$$

donde aquí $S_{ij} = ([U]_e)_{ij}$. El determinante de un operador unitario tiene módulo 1:

$$1 = \mathrm{Det}[U^\dagger U] = \mathrm{Det}[U]^*\mathrm{Det}[U] = |\mathrm{Det}[U]|^2$$

por lo que

$$|\mathrm{Det}[U]| = 1$$

Podemos entonces escribir $\text{Det}[U] = e^{i\phi}$, con ϕ real.
Debe remarcarse que los operadores unitarios transforman bases ortonormales en bases ortonormales: si $e'_i = U(e_i) = \sum_{j=1}^m S_{ji} e_j$, $i = 1, \ldots, n \Rightarrow$

$$(e'_i, e'_j) = (U(e_i), U(e_j)) = (e_i, e_j) = \delta_{ij}$$

Análogamente, cualquier par de bases ortonormales e, e' de V están relacionadas por una transformación unitaria, es decir, por una matriz de cambio de base S que satisface $S^\dagger S = SS^\dagger = I_n$, como es fácil verificar: Si $e'_i = \sum_j S_{ji} e_j$ y $(e'_i, e'_j) = (e_i, e_j) = \delta_{ij}$ entonces

$$(e'_i, e'_j) = \sum_{k,l} (S_{ki} e_i, S_{lj} e_l) = \sum_{k,l} S^*_{ki} S_{lj} (e_k, el) = \sum_k S^\dagger_{ik} S_{kj} = (S^\dagger S)_{ij} = \delta_{ij}$$

Remarquemos también que el *producto* (pero no la suma) de operadores unitarios es *unitario*: Si U, W son unitarios $\Rightarrow (UW)^{-1} = W^{-1} U^{-1} = W^\dagger U^\dagger = (UW)^\dagger$, por lo que UW es unitario. Está propiedad es también obvia a partir de la definición.

24. Autovalores y Autovectores de operadores autoadjuntos

1) Si $F : V \to V$ es un operador lineal autoadjunto $\Rightarrow$ sus autovalores son todos reales y los autovectores correspondientes a autovalores distintos son ortogonales.
Demostración: Si $F(v) = \lambda v$,

$$(v, F(v)) = (v, \lambda v) = \lambda (v, v)$$

pero por ser F autoadjunto,

$$(v, F(v)) = (F(v), v) = (\lambda v, v) = \lambda^* (v, v)$$

por lo que

$$(\lambda - \lambda^*)(v, v) = 0$$

lo que implica, si $v \neq 0$, $\lambda - \lambda^* = 0$, es decir, λ *real*. Todos los autovalores de F serán pues *reales*.
Además, si $F(v) = \lambda v$ y $F(v') = \lambda' v'$, entonces

$$(v', F(v)) = \lambda (v', v) = (F(v'), v) = \lambda' (v', v)$$

por lo que

$$(v', v)(\lambda - \lambda') = 0$$

lo que implica

$$(v', v) = 0 \text{ si } \lambda \neq \lambda'$$

2) Si $F : V \to V$ es un operador lineal autoadjunto en un espacio V de dimensión finita, existe siempre una base *ortonormal* de V formada por autovectores de F: $\exists\ e' = (e'_1, \ldots, e'_n)$, tal que

$$F(e'_i) = \lambda_i e'_i, \ \ i = 1, \ldots, n, \ \ (e'_i, e'_j) = \delta_{ij}$$

Es decir, F *es siempre diagonalizable* y además lo es en una base ortonormal, la cual estará relacionada con la base canónica original por una transformación unitaria U:

$$[F]_{e'} = S^\dagger [F]_e S = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}, \quad S^\dagger S = SS^\dagger = I$$

con $S = [U]_e$ y $e'_i = U(e_i)$.
Demostración: Por inducción sobre n. Para $n = 1$ todo F es trivialmente diagonal en cualquier base. Considerando ahora $n > 1$, si los n autovalores de F son todos distintos, entonces esta propiedad es inmediata, ya que por 1) existirán n autovectores ortogonales entre si, que puede ser convertidos en ortonormales luego de normalización ($e'_i \to e'_i/||e'_i||$).

En general, supongamos que e'_1 es un autovector normalizado de F ($F(e'_1) = \lambda_1 e'_1$, $((e'_1, e'_1) = 1)$ y sea S_1 el subespacio de V generado por e'_1. En tal caso S_1 es invariante por F y por lo tanto, el complemento ortogonal $S_{1\perp}$, de dimensión $n-1$, **será también invariante por** $F^\dagger = F$. F restringido a $S_{1\perp}$ es obviamente también autoadjunto. Por lo tanto, por hipótesis inductiva, existe una base ortonormal de $S_{1\perp}$ en la que F es diagonal. F resulta así diagonal en la base ortonormal de V formada por e'_1 y la base anterior de $S_{1\perp}$. F será entonces diagonalizable $\forall\ n$ en una base ortonormal.

3) Si F y G son dos operadores autoadjuntos y $[F, G] = 0$ (o sea, $FG = GF$) $\Rightarrow$ *existe una base ortonormal común* e' en la que ambos operadores son *simultáneamente diagonales:*

$$F(e'_i) = \lambda_i^F e'_i\,,\quad G(e'_i) = \lambda_i^G e'_i,\quad i = 1, \ldots, n$$

Demostración: Como F es autoadjunto, existe una base ortonormal donde F es diagonal. Como $[G, F] = 0$ $\Rightarrow$ si $F(e'_i) = \lambda_i^F e'_i$, $FG(e'_i) = GF(e'_i) = \lambda_i^F G(e'_i)$, por lo que $G(e'_i) \in V_F(\lambda_i^F)$ (espacio propio). $V_F(\lambda_i^F)$ es pues también invariante por G. Pero G restringido a $V_F(\lambda_i^F)$ es asimismo autoadjunto, por lo que es siempre posible elegir una base ortonormal de $V_F(\lambda_i^F)$ en la que G será también diagonal, con autovalores λ_i^G. Los elementos de dicha base serán, por pertenecer a $V_F(\lambda_i^F)$, también autovectores de F. Repitiendo esto para todos los autovalores, vemos que existirá una base ortonormal de V en la que tanto F y G serán diagonales.

24.1 Operadores normales

Un operador lineal $A : V \to V$ se dice *normal* si $A^\dagger A = AA^\dagger$, es decir, si

$$[A, A^\dagger] = 0$$

donde $[A, B] = AB - BA$ denota el conmutador.
Así, los operadores autoadjuntos ($F^\dagger = F$) son obviamente normales, y también son normales los unitarios ($U^\dagger = U^{-1}$, con $UU^\dagger = U^\dagger U = I$). Otro caso de operador normal son las antiautadjuntos ($F^\dagger = -F$).

Teorema de diagonalización para operadores normales:
Si $A : V \to V$ es un operador normal, entonces existe una base *ortonormal* e' en la cual $[A]_{e'}$ es *diagonal*:

$$[A]_{e'} = S^\dagger [A]_e S = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix},\qquad S^\dagger S = SS^\dagger = I_n$$

Además si $A : V \to V$ es diagonal en una base ortonormal $\Rightarrow$ es normal.
Demostración: Hemos ya demostrado que para todo operador autoadjunto existe una base ortonormal donde es diagonal. La extensión para todo operador normal se basa en la descomposición

$$A = A^r + iA^i,\quad A^r = \frac{A + A^\dagger}{2},\quad A^i = \frac{A - A^\dagger}{2i}$$

válida para *cualquier* operador A, donde A_r y A_i son claramente operadores *autoadjuntos*: $(A^r)^\dagger = A^r$, $(A^i)^\dagger = A^i$. Esta descomposición del operador es similar a la de un número complejo $z = x + iy$ en parte real x e imaginaria iy (caso particular $n = 1$).
Si A es *normal* $\Rightarrow$

$$[A^r, A^i] = \frac{1}{4i}[A + A^\dagger, A - A^\dagger] = 0$$

y por lo tanto, *existe una base ortonormal común* e' *donde* A^r y A^i *son simultáneamente diagonales*. Los autovalores de A serán entonces de la forma

$$\lambda_j = \lambda_j^r + i\lambda_j^i$$

con λ_j^r y λ_j^i *reales* y autovalores de A^r y A^i respect., por lo que λ_j será en general complejo.
Si A es autoadjunto ($A^\dagger = A$) $\Rightarrow$ $A_i = 0$ y por lo tanto $\lambda_j^i = 0$. Los autovalores de A son entonces todos reales, como ya habíamos demostrado.
Si A es antiautoadjunto ($A^\dagger = -A$) $\Rightarrow$ $A_r = 0$ y por lo tanto $\lambda_j^r = 0$. Los autovalores de A son entonces

todos *imaginarios puros.*
Finalmente, si A es unitario, $[A]^\dagger_{e'}[A]_{e'} = I_n$, lo que implica $\lambda_j\lambda_j^* = |\lambda_j|^2 = 1$, es decir $|\lambda_j| = 1$. Esto implica

$$\lambda_j = e^{i\phi_j} = \cos\phi_j + i\sin\phi_j$$

con $\lambda_j^r = \cos\phi_j$, $\lambda_j^i = \sin\phi_j$.

Por otro lado, si A es diagonal en una base e' ortonormal $\Rightarrow A^\dagger$ es también diagonal en dicha base, con

$$[A^\dagger]_{e'} = [A]^\dagger_{e'} = \begin{pmatrix} \lambda_1^* & 0 & \dots & 0 \\ 0 & \lambda_2^* & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n^* \end{pmatrix}$$

Por lo tanto

$$[AA^\dagger - A^\dagger A]_{e'} = [A]_{e'}[A^\dagger]_{e'} - [A^\dagger]_{e'}[A]_{e'} = 0$$

lo que implica $AA^\dagger - A^\dagger A = 0$. A es entonces normal.

En resumen, el teorema implica que en un espacio unitario, *un operador tiene representación diagonal en una base ortonormal si y sólo si es un operador normal.* En términos matriciales, si A es una matriz de $n \times n$, entonces *existe una matriz unitaria S tal que $A' = S^\dagger AS$ es diagonal si y sólo si A es normal* ($[A^\dagger, A] = 0$*)*. Esto comprende en particular las matrices hermíticas ($A^\dagger = A$), antihermíticas ($A^\dagger = -A$) y unitarias ($A^\dagger = A^{-1}$). Destaquemos también que todo $v \in V$ puede expandirse en la base e' de autovectores de un operador normal A,

$$v = \sum_{i=1}^n \alpha_i e'_i = \sum_{i=1}^n P_{e'_i}(v)$$

donde $\alpha_i = (e'_i, v)$ y $P_{e'_i}(v) = (e'_i, v)e'_i = \alpha'_i e'_i$. Por lo tanto

$$A(v) = \sum_{i=1}^n A(\alpha_i e'_i) = \sum_{i=1}^n \alpha_i\lambda_i e'_i = \sum_{i=1}^n \lambda_i P_{e'_i}(v)$$

Como v es arbitrario, esto implica

$$A = \sum_{i=1}^n \lambda_i P_{e'_i}$$

Un operador normal puede pues expresarse como *combinación lineal de proyectores ortogonales sobre sus espacios propios.*

De lo anterior se desprende además que todo operador unitario U puede escribirse en la forma

$$U = \exp[iF]$$

con F autoadjunto: Como los autovalores de U son de la forma $e^{i\phi_j}$, podemos definir F como el operador autoadjunto que es también diagonal en la base ortonormal e' en que U es diagonal y que tiene autovalores reales ϕ_j. En tal caso, $[U]_{e'} = \exp[i[F]_{e'}] = [\exp[iF]]_{e'}$, lo que implica $[U]_e = [\exp(iF)]_e$ en cualquier base. Esto conduce a $U = \exp[iF]$.

Ejercicio: Utilizando la representación diagonal, mostrar que si $F : V \to V$ es autoadjunto, entonces $\forall$ $v \in V$, con $v \neq 0$, se tiene

$$\lambda_m \leq \frac{(v, F(v))}{(v, v)} \leq \lambda_M$$

donde λ_m y λ_M denotan resp. el menor y mayor autovalor de F.

24.2 Isometrías en espacios euclideos

Hemos visto que los autovalores de un operador unitario U son necesariamente de la forma $\lambda = e^{i\phi} = \cos(\phi) + i\sin(\phi)$, con ϕ real. Mediante el "embedding" de un espacio euclídeo en un espacio unitario discutido en clase, esto permite demostrar que las isometrías U en espacios euclídeos sólo pueden ser rotaciones ($\text{Det}U = 1$) o rotaciones seguidas o precedidas de una reflexión ($\text{Det}U = -1$).

En efecto, si $S \equiv [U]_e$ es una matriz real que representa una isometría U en una base ortonormal de un espacio euclideo ($S^t = S^{-1}$), considerada en un espacio complejo representa una transformación unitaria ($S^\dagger = S^{-1}$). Dado que S es real, los autovalores vendrán de a pares conjugados con autovectores conjugados:

$$SX = \lambda X, \quad SX^* = \lambda^* X^*$$

Escribiendo $\lambda = \lambda_r + i\lambda_i$, $X = X_r + iX_i$, con $\lambda_r = \cos\phi$, $\lambda_i = \sin\phi$ y X_r, X_i reales, esto implica

$$SX_r = \lambda_r X_r - \lambda_i X_i\,, \quad SX_i = \lambda_r X_i + \lambda_i X_r$$

Si λ no es real ($\lambda_i \neq 0$) $\Rightarrow$ $X_i \neq 0$ (pues S es real) y la ortogonalidad de los autovectores para autovalores distintos (válido para cualquier matriz normal $[S^\dagger, S] = 0$) implica $(X^*)^\dagger X = 0$, o sea,

$$(X_r + iX_i)^t (X_r + iX_i) = X_r^t X_r - X_i^t X_i + 2iX_i^t X_r = 0$$

de donde $X_r^t X_r = X_i^t X_i$ y $X_i^t X_r = 0$. Por lo tanto, vemos que en el subespacio generado por X^*, X, existe una base *real* y *ortonormal con el producto escalar euclideo*, formada por (X_i, X_r), en la que el bloque correspondiente de $[U]_{e'}$ tiene la forma

$$S'_\phi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

que representa una rotación de ángulo ϕ ($\mathrm{Det}[S'_\phi] = 1$). Y en el espacio euclideo completo, vemos entonces que existe una base ortonormal e' donde $S' \equiv [U]_{e'}$ tiene la forma

$$S' = \begin{pmatrix} {S'}_{\phi_1} & 0 & \ldots & 0 & 0 \\ 0 & {S'}_{\phi_2} & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & \pm 1 & 0 \\ 0 & 0 & \ldots & 0 & \pm 1 \end{pmatrix}$$

donde S'_{ϕ_i} son bloques de la forma anterior que representan rotaciones en subespacios de dimensión 2, y los elementos ± 1 representan los posibles autovalores reales. U representa pues rotaciones ($\mathrm{Det}[S'] = 1$) o rotaciones compuestas con reflexiones ($\mathrm{Det}[S'] = -1$). Por ej., en $\mathbb{R}^3$, las posibilidades son un bloque A_ϕ seguido de $+1$ (rotación) o -1 (rotación compuesta con reflexión).

24.3 Elementos de matriz de un operador lineal en una base ortonormal

Recordemos que si $F : V \to V$ es un operador lineal $\Rightarrow$ la matriz $T = [F]_e$ ($\equiv [F]_e^e$) que lo representa en una base e de V queda definida por

$$F(e_i) = \sum_{j=1}^{n} T_{ji} e_j$$

En un espacio unitario y en una base ortonormal e, los elementos de matriz $T_{ji} = ([F]_e)_{ji}$ pueden entonces obtenerse, por ortonormalidad de los e_i, como

$$T_{ji} = (e_j, F(e_i))$$

De esta forma,

$$v = \sum_{i=1}^{n} \alpha_i e_i, \quad \alpha_i = (e_i, v)$$

y

$$F = \sum_{i,j=1}^{n} T_{ji} E_{ji}, \quad T_{ji} = (e_j, F(e_i))$$

con E_{ji} el operador lineal definido por

$$E_{ji}(v) = (e_i, v) e_j$$

ya que $F(e_i) = \sum_{j=1}^{n} T_{ji} e_j = \sum_{j,k=1}^{n} T_{jk}(e_k, e_i) e_j = (\sum_{j,k=1}^{n} T_{jk} E_{jk})(e_i)$.
Notemos que $([E_{ji}]_e)_{kl} = \delta_{kj}\delta_{il}$.

Notación de Mecánica cuántica:

$$T_{ji} = (e_j, F(e_i)) \to \langle j|F|i\rangle, \quad E_{ji} \to |j\rangle\langle i|$$

donde $|i\rangle \equiv e_i$ y $\langle i| \equiv f^i$ (vector asociado del espacio dual). Por lo tanto

$$F = \sum_{i,j} F_{ij}|i\rangle\langle j|, \quad F_{ij} = \langle i|F|j\rangle$$

Por ej., el proyector ortogonal sobre e_i se escribe como $P_{e_i} = |i\rangle\langle i|$ (ya que $(e_j, P_{e_i}(e_i)) = (e_j, e_i) = \delta_{ji}$) mientras que el operador identidad es $I = \sum_i |i\rangle\langle i|$. En general, para todo operador normal $F : V \to V$ existe entonces una base ortonormal $\{|i\rangle\}$ de V formada de autovectores de F en la que $\langle i|F|j\rangle = \delta_{ij}\lambda_i$ y

$$F = \sum_i \lambda_i |i\rangle\langle i|$$

25 Descomposición en valores singulares (DVS)

Sea $F : V \to W$ una transformación lineal arbitraria entre espacios unitarios V y W de dimensiónes n y m respectivamente, y e_V, e_W bases ortonormales de V y W. Entonces existen bases ortonormales e'_V, e'_W en las que F queda representado por una matriz *diagonal* de elementos no negativos. En otras palabras, dada una matriz A de $m \times n$, existen matrices unitarias U de $m \times m$ y V de $n \times n$ tales que

$$A = UA'V^\dagger$$

con $U^\dagger U = I_m$, $V^\dagger V = I_n$ y A' de $m \times n$ *diagonal* de elementos $A'_{kj} = \sigma_j \delta_{kj}$, con $\sigma_j \geq 0$. Aquí $A = [F]^{e_V}_{e_W}$, $A' = [F]^{e'_V}_{e'_W}$, $U = [I]^{e'_W}_{e_W}$ y $V = [I]^{e'_V}_{e_V}$, con $V^\dagger = [I]^{e_V}_{e'_V}$. Los σ_j no nulos se denominan *valores singulares* y son las *raíces* de los autovalores no nulos de la matriz hermítica $A^\dagger A$, de $n \times n$ (que posee autovalores no negativos). V es la correspondiente matriz de autovectores normalizados (tal que $V^\dagger(A^\dagger A)V$ es diagonal).
La demostración es similar al caso de matrices reales (espacios euclideos) y se deja como ejercicio.
Recordemos que si k es el número de autovalores no nulos de $A^\dagger A$, las primeras k columnas de U son los vectores $u_i = Av_i/\sigma_i$, $i = 1, \ldots, k$, $\sigma_i \neq 0$, obteniéndose las restantes $m - k$ columnas ortonormales de U por el método de Gram-Schmidt complejo.

Ejercicios: Para A de $m \times n$ compleja general, demostrar (en forma similar al caso euclideo) que:
0) Las matrices $A^\dagger A$ y $AA^\dagger$ son ambas hermíticas.
1) Los autovalores de $A^\dagger A$ son todos no negativos.
2) El número de autovalores no nulos de $A^\dagger A$ es igual al rango de A.
3) Los autovalores no nulos de las matrices $A^\dagger A$ y $AA^\dagger$ son iguales.
4) $||A||_2 = \sigma_M$, siendo σ_M el máximo valor singular y $||A||_2 \equiv \mathrm{Max}_{v\neq 0}||Av||/||v||$, con $||v|| = \sqrt{v^\dagger v}$.
5) Si $m = n$ y λ es autovalor de $A \Rightarrow |\lambda| \leq \sigma_M$.
6) Si $m = n$ y A es invertible $\Rightarrow n_c(A) = \sigma_M/\sigma_m$, donde σ_m es el mínimo valor singular de A y $n_c(A)$ es el número de condición.

25.1 Forma polar de un operador lineal

En el caso de $V = W$, la DVS permite obtener en forma inmediata la denominada *forma polar* de un operador: Si $F : V \to V$ es un operador lineal en un espacio unitario V entonces F puede escribirse como

$$F = WM = \tilde{M}W$$

donde W es un operador *unitario* y M, $\tilde{M}$ operadores autoadjuntos *positivos.*
Dem.: Utilizando la DVS para la representación $A = [F]_e$ de F en una base ortonormal e de V, se tiene

$$\begin{aligned} A &= UA'V^\dagger \\ &= (UV^\dagger)(VA'V^\dagger) = WM, \quad W = UV^\dagger, \quad M = VA'V^\dagger = \sqrt{A^\dagger A} \\ &= (UA'U^\dagger)(UV^\dagger) = \tilde{M}W, \quad \tilde{M} = UA'U^\dagger = \sqrt{AA^\dagger} \end{aligned}$$

donde W es unitario ($W^\dagger = VU^\dagger = W^{-1}$) y $A^\dagger A = VA'^2V^\dagger$, $AA^\dagger = UA'^2U^\dagger$.
Ejercicio: Discutir la DVS y la descomposición polar de una matriz hermítica.

26 Desigualdad de Cauchy Schwarz y relaciones de incerteza

Consideremos dos operadores autoadjuntos F, G. El valor medio de un operador F en un estado normalizado $|\psi\rangle$ ($\langle\psi|\psi\rangle = 1$) es

$$\langle F\rangle_\psi = \langle\psi|F|\psi\rangle$$

(o sea $\langle F\rangle_\psi = (\psi, F(\psi))$ en notación de A.L.). Si F es autoadjunto, $\langle F\rangle_\psi$ es *real*, ya que $\langle\psi|F|\psi\rangle = \langle\psi|F^\dagger|\psi\rangle^* = \langle\psi|F|\psi\rangle^*$ (o sea, $(\psi, F(\psi)) = (F(\psi),\psi)^* = (\psi, F^\dagger(\psi))^* = (\psi, F(\psi))^*$).

La varianza de un operador F en el estado $|\psi\rangle$ se define como el valor medio del cuadrado de la diferencia entre F y $\langle F\rangle_\psi$ y es una medida de la dispersión alrededor de la media:

$$\Delta^2 F = \langle (F - \langle F\rangle_\psi I)^2\rangle_\psi = \langle\psi|(F-\langle F\rangle_\psi I)^2|\psi\rangle = \langle\psi|F^2|\psi\rangle - \langle\psi|F|\psi\rangle^2$$

La desviación estándar es la raíz de la varianza: $\Delta F = \sqrt{\Delta^2 F} = \sqrt{\langle (F-\langle F\rangle_\psi I)^2\rangle_\psi}$.

Definamos ahora

$$\tilde F = F - \langle F\rangle_\psi I,\;\; \tilde G = G - \langle G\rangle_\psi I$$

tal que $\Delta F = \sqrt{\langle\psi|\tilde F^2|\psi\rangle}$, $\Delta G = \sqrt{\langle\psi|\tilde G^2|\psi\rangle}$, y consideremos el producto escalar $(\tilde F(\psi), \tilde G(\psi)) = (\psi, \tilde F\tilde G(\psi))$, es decir $\langle\psi|\tilde F\tilde G|\psi\rangle$ en notación cuántica. La desigualdad de Cauchy-Schwarz implica $|(\tilde F(\psi), \tilde G(\psi))| \leq ||\tilde F(\psi)||\,||\tilde G(\psi)||$, con $||\tilde F(\psi)||^2 = (\tilde F(\psi), \tilde F(\psi)) = (\psi, \tilde F^2(\psi))$, o sea,

$$|\langle\psi|\tilde F\tilde G|\psi\rangle| \leq \sqrt{\langle\psi|\tilde F^2|\psi\rangle}\sqrt{\langle\psi|\tilde G^2|\psi\rangle} = (\Delta F)(\Delta G)$$

Por otro lado, si $[F,G] = FG - GF$ denota el conmutador de F y G, entonces

$$\langle\psi|[F,G]|\psi\rangle = \langle\psi|[\tilde F,\tilde G]|\psi\rangle = \langle\psi|\tilde F\tilde G|\psi\rangle - \langle\psi|\tilde G\tilde F|\psi\rangle = 2i\,\mathrm{Im}[\langle|\psi|\tilde F\tilde G|\psi\rangle]$$

donde Im denota la parte imaginaria, ya que $\langle\psi|\tilde G\tilde F|\psi\rangle = \langle\psi|(\tilde G\tilde F)^\dagger|\psi\rangle^* = \langle\psi|\tilde F\tilde G|\psi\rangle^*$. Por lo tanto

$$\tfrac{1}{2}|\langle\psi|[F,G]|\psi\rangle| \leq |\langle\psi|\tilde F\tilde G|\psi\rangle| \leq (\Delta F)(\Delta G)$$

es decir,

$$(\Delta F)(\Delta G) \geq \tfrac{1}{2}|\langle [F,G]\rangle_\psi|$$

Esta es la denominada relación de incerteza entre dos operadores: Si el conmutador es no nulo entonces el producto de sus "incertezas" $(\Delta F)(\Delta G)$ en un estado $|\psi\rangle$ no puede ser menor que el módulo del valor medio del conmutador en dicho estado.

Como ejemplo fundamental, consideremos el espacio L^2 de funciones $\psi(x)$ de $\mathbb{R}\to\mathbb{C}$ de norma finita ($||\psi||^2 = \int_{-\infty}^{\infty}|\psi(x)|^2 dx < \infty$) y que tienden a 0 para $x\to\pm\infty$, tal que el producto escalar

$$(\psi,\phi) = \int_{-\infty}^{\infty}\psi^*(x)\phi(x)dx$$

esté bien definido. Los operatores X y $P = -i\hbar\partial_x = -i\hbar\frac{\partial}{\partial x}$, donde $\hbar = h/(2\pi)$, con $\hbar$ la constante de Planck, son autoadjuntos en este espacio: $(\psi, X\phi) = \int_{-\infty}^{\infty}\psi^*(x)x\phi(x)dx = (X\psi,\phi)$, y

$$(\psi, P\phi) = -i\hbar\int_{-\infty}^{\infty}\psi^*(x)\phi'(x)dx = -i\hbar[\,\psi^*(x)\phi(x)\,|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\psi^{*\prime}(x)\phi(x)dx] = \int_{-\infty}^{\infty}[-i\hbar\psi'(x)]*\phi(x)dx = (P(\psi),\phi)$$

Dado que $[X,P]\psi(x) = -i\hbar(x\psi'(x) - (x\psi(x))') = i\hbar\psi(x)\;\forall\;\psi$, es decir, $[X,P] = i\hbar I$, obtenemos $|\langle [X,P]\rangle_\psi| = \hbar\;\forall\;\psi$ y el resultado anterior implica entonces

$$(\Delta P)(\Delta X) \geq \frac{\hbar}{2}$$

El operador P representa en Mecánica Cuántica el operador impulso de una partícula (en una dimensión). Por lo tanto, en cualquier estado cuántico el producto de las desviaciones estándar de X y P es no nulo y mayor que $\hbar/2$.

27 Tensores (Resumen)

27. 1 Notación tensorial

Mediante la convención de Einstein para sumas, el cambio de base $e'_i=\sum_{j=1}^n S_{ji}e_j$, con $S=[I]_e^{e'}$ una matriz de $n\times n$ no singular, se escribe

$$e'_i=S_i^j e_j$$

donde $S_i^j e_j\equiv\sum_{j=1}^n S_i^j e_j$ y n es la dimensión del espacio. El índice superior en S denota fila y el inferior columna. En forma matricial, la relación anterior equivale pues a

$$(e'_1,\ldots,e'_n)=(e_1,\ldots,e_n)S$$

Por otro lado, la transformación $x'^i=\sum_{j=1}^n S^{-1}_{ij}x^j$ de las componentes de un vector $v=\sum_{i=1}^n x^i e_i=\sum_{i=1}^n x'^i e'_i$, se escribe en la forma

$$x'^i=R^i_j x^j,\qquad R=S^{-1}$$

donde $R^i_j x^j\equiv\sum_{j=1}^n R^i_j x^j$. En forma matricial, la relación previa equivale pues a

$$\begin{pmatrix}x'^1\\ \ldots\\ x'^n\end{pmatrix}=R\begin{pmatrix}x^1\\ \ldots\\ x^n\end{pmatrix}$$

lo que está también de acuerdo con el supraíndice como índice de fila. Notemos que

$$R^i_j S^j_k=S^i_j R^j_k=\delta^i_k$$

que es la expresión tensorial de la relación matricial $RS=SR=I$. El vector v se escribe entonces como

$$v=x^i e_i=x'^i e'_i$$

Como verificación, reemplazando $x'^i=R^i_j x^j$, $e'_i=S^k_i e_k$, se tiene $x'^i e'_i=R^i_j S^k_i x^j e_k=\delta^k_j x^j e_k=x^j e_j$.

En general, n componentes a_i que se transforman como

$$a'_i=S^j_i a_j$$

se denominan *covariantes*, mientras que n componentes b^i que se transforman como

$$b'^i=R^i_j b^j$$

con $R^i_j S^j_k=\delta^i_k$ (o sea, $R=S^{-1}$) se denominan *contravariantes*. En tal caso, el producto

$$b'^i a'_i=b^i a_i$$

(donde la suma sobre i está implícita) permanece invariante frente a cambios de base.

Notemos finalmente que las relaciones inversas están dadas por

$$a_i=R^j_i a'_j,\quad b^i=S^i_j b'^j$$

Transformación de las derivadas parciales:
Dado el cambio de variables lineal $x'^i=R^i_j x^j$ y su relación inversa $x^j=S^j_i x'^i$, con $S=R^{-1}$, y R,S *independientes de las coordenadas*, tenemos

$$S^j_i=\frac{\partial x^j}{\partial x'^i},\quad R^i_j=\frac{\partial x'^i}{\partial x^j}$$

En virtud de la regla de la cadena, se obtiene entonces

$$\frac{\partial}{\partial x'^i}=\sum_{j=1}^n\frac{\partial x^j}{\partial x'^i}\frac{\partial}{\partial x^j}$$

o sea, en notación covariante,

$$\partial'_i=S^j_i\partial_j$$

donde $\partial'_i\equiv\frac{\partial}{\partial x'^i}$, $\partial_j\equiv\frac{\partial}{\partial x^j}$. Las derivadas respecto de componentes contravariantes se transforman pues de manera covariante.

27.2 Transformación de vectores del dual

Dada una base $e = (e_1, \ldots, e_n)$ de V, los elementos de la base dual $f = (f^1, \ldots, f^n)$ del espacio dual V^* (el conjunto de formas lineales de V en K) quedan definidos por

$$(f^i, e_j) = \delta^i_j$$

(utilizamos la notación $f^i(v) = (f^i, v)$. Esto implica la ley de transformación contravariante

$$f'^i = R^i_j f^j$$

de forma que

$$(f'^i, e'_j) = R^i_k S^l_j (f^k, e_l) = R^i_k S^l_j \delta^k_l = R^i_l S^l_j = \delta^i_j$$

donde $e'_j = S^i_j e_i$. Un elemento arbitrario $h \in V^*$ puede entonces ser escrito como

$$h = a_i f^i = a'_i f'^i$$

donde

$$a'_i = S^j_i a_j$$

Notemos que si $v = x^i e_i$, $h = a_i f^i$,

$$a_i = (h, e_i), \quad x^i = (f^i, v)$$

Finalmente, mencionemos que si $(e_1, \ldots, e_n)$, $(f'^1, \ldots, f'^n)$ son bases arbitrarias de V y V^* respect., con

$$R^i_j = (f'^i, e_j)$$

una matriz no singular, la base dual de V asociada a la base f' de V^* está dada por

$$e'_i = S^j_i e_j$$

con $S = R^{-1}$, ya que $(f'^k, e'_i) = (f'^k, e_j) S^j_i = R^k_j S^j_i = \delta^k_i$. Análogamente, la base dual de V^* asociada a la base e de V está formada por

$$f^i = S^i_j f'^j$$

ya que $(f^i, e_k) = S^i_j (f'^j, e_k) = S^i_j R^j_k = \delta^i_k$.
Ejemplo: Sea $V = \mathbb{R}^2$ y sean (f'^1, f'^2) las formas lineales definidas por
$f'^1(x, y) = 2x - y$, $f'^2(x, y) = 3x + y$
donde hemos escrito $v = (x, y) = xe_1 + ye_2$, con (e_1, e_2) la base canónica de $\mathbb{R}^2$.
Hallar la base dual de V asociada a f'^1, f'^2.
Podemos escribir $f'^i = R^i_k f^k$, con $R = \left(\begin{smallmatrix} 2 & -1 \\ 3 & 1 \end{smallmatrix}\right)$ y (f^1, f^2) la base dual asociada a (e_1, e_2) ($f^1(x, y) = x$, $f^2(x, y) = y$).
Es claro que (f'^1, f'^2) es base de V^* pues $|R| = 5 \neq 0$. La base dual asociada de V está entonces dada por $e'_i = S^j_i e_j$, con $S = R^{-1} = \left(\begin{smallmatrix} 1 & 1 \\ -3 & 2 \end{smallmatrix}\right)/5$:

$$e'_1 = \frac{1}{5}(e_1 - 3e_2), \; e'_2 = \frac{1}{5}(e_1 + 2e_1)$$

verificándose que $f'^1(e'_1) = f'^2(e'_2) = 1$, $f'^1(e'_2) = f'^2(e'_1) = 0$.

27.3 Tensor métrico

Dado un espacio euclideo V de dimensión finita, con el producto escalar denotado por (v, w), y dada una base arbitraria $e = (e_1, \ldots, e_n)$ de V, el tensor métrico se define como

$$g_{ij} = (e_i, e_j)$$

Es una matriz simétrica ($g_{ij} = g_{ji}$) no singular ($|g| \neq 0$). En tal caso, la norma al cuadrado de un vector $v = x^i e_i$ (es decir, la distancia al cuadrado del extremo del vector al origen) está dada por

$$||v||^2 = (x^i e_i, x^j e_j) = x^i (e_i, e_j) x^j = x^i g_{ij} x^j$$

Podemos escribir lo anterior también en la forma

$$||v||^2 = x^i x_i, \quad x_i \equiv g_{ij} x^j$$

Frente a un cambio de base, el tensor métrico se transforma como

$$g'_{ij} = (e'_i, e'_j) = S_i^k S_j^l (e_k, e_l) = S_i^k S_j^l g_{kl}$$

que corresponde a un tensor de rango $(2, 0)$ (dos veces covariante), como se verá en breve.
Las componentes x_i se transforman pues en forma covariante:

$$x'_i = g'_{ij} x'^j = S_i^k S_j^l R_m^j g_{kl} x^m = S_i^k g_{kl} x^l = S_i^k x_k$$

En espacios euclideos V de dimensión finita, podemos identificar con cada elemento h del dual V^* uno y sólo un vector $w_h \in V$ tal que

$$(h, v) = (w_h, v)$$

$\forall\, v \in V$, donde el segundo paréntesis denota producto escalar: Si $h = a_i f^i$ y $w_h = a^i e_i$, con $(f^i, e_j) = \delta_j^i$,

$$(h, e_j) = a_j = (w_h, e_j) = a^i (e_i, e_j) = a^i g_{ij}$$

de modo que $a^i g_{ij} = a_j$. Por lo tanto,

$$a^i = g^{ji} a_j$$

donde g^{ij} denota los elementos de la *matriz inversa* de la matriz de elementos g_{ij}:

$$g^{ik} g_{kj} = \delta_j^i$$

En lo sucesivo denotaremos a w_h directamente como h. Por consiguiente, podemos escribir los elementos de la base dual como combinación lineal de los e_i. En notación tensorial,

$$f^i = g^{ik} e_k$$

con

$$(f^i, e_j) = g^{ik} (e_k, e_j) = g^{ik} g_{kj} = \delta_j^i$$

Notemos también que

$$(f^i, f^j) = g^{jk} (f^i, e_k) = g^{ji}$$

por lo que g^{ji} es el tensor métrico en la base dual. Un vector v puede pues escribirse en las formas

$$v = x^i e_i = x_i f^i$$

donde $x_i = g_{ij} x^j$, $f^i = g^{ik} e_k$, ya que $x_i f^i = g^{ik} g_{ij} x^j e_k = \delta_j^k x^j e_k = x^j e_j$. Para el producto escalar de dos vectores $v = x^i e_i$, $w = y^j e_j$ se tienen pues las expresiones

$$(v, w) = x^i g_{ij} y^j = x^i y_i = x_i y^i = x_i g^{ij} y_j$$

27.4 Tensores

Un tensor general de p índices covariantes y q indices contravariantes (que denotaremos aquí como tensor $\binom{q}{p}$) en un espacio de dimensión n, es un conjunto de n^{p+q} números $T_{i_1 \ldots i_p}^{j_1 \ldots j_q}$ dependientes de una base ordenada $B = (e_1, \ldots, e_n)$ de un espacio vectorial V, que se transforman frente a cambios de base $e'_i = S_i^j e_j$ en la forma

$$T'^{j'_1, \ldots j'_q}_{i'_1 \ldots i'_p} = S_{i'_1}^{i_1} \ldots S_{i'_p}^{i_p} R_{j_1}^{j'_1} \ldots R_{j_q}^{j'_q} T_{i_1 \ldots i_p}^{j_1 \ldots j_q}$$

con $R = S^{-1}$. Por ejemplo, para un tensor $\binom{1}{1}$, $T'^l_k = R_j^l S_k^i T_i^j$, que involucra una suma sobre i y j.

Una posible realización de un tensor $\binom{q}{p}$ es una forma *multilineal* $T : V^p \times (V^*)^q \to K$ de p vectores de V y q vectores del espacio dual V^* (una función es multilineal si es lineal en cada uno de sus argumentos:

$T(\alpha_1 v_1+\alpha'_1 v'_1, v_2,\ldots,v_p,w^1,\ldots,w^q) = \alpha_1 T(v_1,v_2,\ldots,v_p,w^1,\ldots,w^q)+\alpha'_1 T(v'_1,v_2,\ldots,v_p,w^1,\ldots,w^q)$, y similar para los restantes argumentos). En tal caso, si $v_i = x_i^j e_j$ y $w^i = a^i_j f^j$,

$$T(v_1,\ldots,v_p,w^1,\ldots,w^q) = x_1^{i_1}\ldots x_p^{i_p} a^1_{j_1}\ldots a^q_{j_q} T(e_{i_1},\ldots,e_{i_p},f^{j_1},\ldots,f^{j_q})$$

Si los f^i son los vectores de la base dual ($(f^i,e_j)=\delta^i_j$), los elementos

$$T^{j_1\ldots j_q}_{i_1\ldots i_p} \equiv T(e_{i_1},\ldots,e_{i_p},f^{j_1},\ldots,f^{j_q})$$

se transforman como un tensor $\binom{q}{p}$ frente a cambios de base: Si $e'_i = S^j_i e_j$, entonces $f'^i = R^i_j f^j$ y

$$\begin{aligned} T'^{j'_1\ldots j'_q}_{i'_1\ldots i'_p} &= T(e'_{i'_1},\ldots,e'_{i'_p},f'^{j'_1},\ldots,f'^{j'_q}) = T(S^{i_1}_{i'_1}e_{i_1},\ldots,S^{i_p}_{i'_p}e_{i_p},R^{j'_1}_{j_1}f^{j_1},\ldots,R^{j'_q}_{j_q}f^{j_q}) \\ &= S^{i_1}_{i'_1}\ldots S^{i_p}_{i'_p}R^{j'_1}_{j_1}\ldots R^{j'_q}_{j_q}T^{j_1\ldots j_q}_{i_1\ldots i_p} \end{aligned}$$

Otra posibilidad es considerar a $T^{j_1\ldots j_q}_{i_1\ldots i_p}$ como las coordenadas de un vector T perteneciente al **producto tensorial** de espacios $\underbrace{V\otimes\ldots\otimes V}_{\text{q veces}}\otimes\underbrace{V^*\otimes\ldots\otimes V^*}_{\text{p veces}}$ en una base $B=\{e_{j_1}\otimes\ldots\otimes e_{j_q}\otimes f^{i_1}\otimes\ldots\otimes f^{i_p}\}$, donde nuevamente $(f^i,e_j)=\delta^i_j$:

$$T = T^{j_1\ldots j_q}_{i_1\ldots i_p} e_{j_1}\otimes\ldots\otimes e_{j_q}\otimes f^{i_1}\otimes\ldots\otimes f^{i_p}$$

Si $e_i = R^j_i e'_j$ y $f^i = S^i_j f'^j$ (tal que $e'_i = S^j_i e_j$, $f'^i = R^i_j f^j$, con $R=S^{-1}$), tenemos

$$\begin{aligned} T &= T^{j_1\ldots j_q}_{i_1\ldots i_p} R^{j'_1}_{j_1}\ldots R^{j'_q}_{j_q} S^{i_1}_{i'_1}\ldots S^{i_p}_{i'_p} e'_{j'_1}\otimes\ldots\otimes e'_{j'_q}\otimes f'^{i'_1}\otimes\ldots\otimes f'^{i'_p} \\ &= T'^{j'_1\ldots j'_q}_{i'_1\ldots i'_p} e'_{j'_1}\otimes\ldots\otimes e'_{j'_q}\otimes f'^{i'_1}\otimes\ldots\otimes f'^{i'_p} \end{aligned}$$

por lo que

$$T'^{j'_1\ldots j'_q}_{i'_1\ldots i'_p} = S^{i_1}_{i'_1}\ldots S^{i_p}_{i'_p}R^{j'_1}_{j_1}\ldots R^{j'_q}_{j_q}T^{j_1\ldots j_q}_{i_1\ldots i_p}$$

Un tensor $\binom{0}{0}$ es un *escalar*. Permanece invariante frente a cambios de base:

$$T' = T$$

Un tensor $\binom{1}{0}$ representa el conjunto de coordenadas *contravariantes* de un vector. Se transforman como

$$T'^i = R^i_j T^j$$

En forma matricial esto corresponde a $T' = RT$, con T un vector columna.
Por ejemplo, las coordendas x^i de un vector $v = x^i e_i \in V$ se transforman como $x'^i = R^i_j x^j$.

Un tensor $\binom{0}{1}$ representa el conjunto de coordenadas *covariantes* de un vector. Se transforman como

$$T'_i = S^j_i T_j$$

En forma matricial esto corresponde a $T' = TS$, con T un vector fila.
Por ejemplo, las coordenadas a_i de un vector $h = a_i f^i \in V^*$ se transforman como $a'_i = S^j_i a_j$.

Un tensor $\binom{1}{1}$ se transforma como

$$T'^j_i = R^j_l S^k_i T^l_k$$

En forma matricial, esto corresponde a $T^j_i = (RTS)^j_i$, es decir, $T' = RTS$, con $R = S^{-1}$. Un ejemplo son pues las matrices que representan operadores lineales $F : V \to V$. Estos pueden expresarse como $F = F^j_i e_j f^i$, de forma que $F(e_k) = F^j_i e_j (f^i, e_k) = F^j_k e_j$, siendo $F^j_i = [F(e_i)]^j = ([F]^e_e)^j_i$ la matriz que lo representa en la base e. Recordemos que esta matriz se transforma precisamente como $F' = RFS$ con $R = S^{-1}$, o sea, $F'^j_i = R^j_l S^k_i F^l_k$.

Un tensor $\binom{0}{2}$ se transforma como

$$T'_{ij} = S^k_i S^l_j T_{kl}$$

En forma matricial, esto equivale a $T_{ij} = (S^t T S)_{ij}$, es decir, $T' = S^t T S$. Un ejemplo son pues las matrices que representan formas cuadráticas (funciones de $V \times V \to K$), de elementos $A_{ij} = A(e_i, e_j)$, las que se transforman como $A' = S^t A S$, es decir, $A'_{ij} = S^k_i A_{kl} S^l_j$. En forma análoga se ve el caso de un tensor $\binom{2}{0}$ (funciones de $V^* \times V^*$ en K).

27.5 Producto Tensorial de Espacios Vectoriales. Recordemos aquí que el **producto tensorial** $V \otimes W$ de dos espacios vectoriales V, W sobre el mismo cuerpo K, de dimensiones n y m respectivamente, es el espacio generado por los productos $\{e_i \otimes \tilde{e}_j\}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, donde $\{e_1, \ldots, e_n\}$ es una base de V y $\{\tilde{e}_1, \ldots, \tilde{e}_m\}$ una base de W. Se verifica, $\forall\, v \in V$, $w \in W$ y $\alpha \in K$,

$$\alpha(v \otimes w) = (\alpha v) \otimes w = v \otimes (\alpha w)$$

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w, \quad v \otimes (w_1 + w_2) = v \otimes w_2 + v \otimes w_2$$

$$0 \otimes w = v \otimes 0 = 0$$

Si $u \in V \otimes W \Rightarrow$

$$u = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij}\, e_i \otimes \tilde{e}_j, \quad c_{ij} \in K$$

Destaquemos que esto incluye vectores producto $u = v \otimes w$, con $v \in V$ y $w \in W$, como así también vectores que son combinaciones lineales de productos pero que no pueden ser escritos como un único producto. La dimensión de $V \times W$ es $n \times m$ (y no $n + m$, como sucede con $V \times W$).

En mecánica cuántica, el espacio de estados de un sistema compuesto por dos subistemas distinguibles es justamente el producto tensorial de los espacios de estados de cada subsistema, siendo estos últimos espacios de Hilbert ($K = \mathbb{C}$). Para $e_i \otimes \tilde{e}_j$ se emplea la notación $|i\rangle \otimes |\tilde{j}\rangle$ o directamente $|i\rangle|\tilde{j}\rangle$ o $|i\tilde{j}\rangle$.
Los estados producto $|u\rangle = |v\rangle \otimes |w\rangle$ se denominan estados **separables**, mientras que los estados que no pueden ser escritos como producto se denominan correlacionados o **entrelazados**.

27.6 Producto y Suma de tensores

Sea T un tensor $\binom{q}{p}$ y U un tensor $\binom{q'}{p'}$ sobre el mismo espacio. Su producto es un tensor $\binom{p+p'}{q+q'}$ dado por

$$(TU)^{j_1 \ldots j_{q+q'}}_{i_1 \ldots i_{p+p'}} = T^{j_1 \ldots j_q}_{i_1 \ldots i_p}\, U^{j_{q+1} \ldots j_{q+q'}}_{i_{p+1} \ldots i_{p+p'}}$$

La suma está definida para tensores del mismo rango $\binom{p}{q}$: $(T + U)^{j_1 \ldots j_q}_{i_1 \ldots i_p} = T^{j_1 \ldots j_q}_{i_1 \ldots i_p} + U^{j_1 \ldots j_q}_{i_1 \ldots i_p}$.

27.7 Producto tensorial de operadores. Si $F : V \to V$ y $G : W \to W$ son operadores lineales en espacios V, W, entonces $F \otimes G : V \otimes W \to V \otimes W$ es un operador lineal en el espacio producto tensorial $V \otimes W$, definido por

$$(F \otimes G)(v \otimes w) = F(v) \otimes G(w)$$

Si $F(v_i) = \lambda^F_i v_i$, $G(w_j) = \lambda^G w_j$, entonces

$$(F \otimes G)(v_i \otimes w_j) = \lambda^F_i \lambda^G_j v_i \otimes w_j$$

por lo que si F y G son diagonalizables, $(F \otimes G)$ también lo es, con $n \times m$ autovalores $\lambda^F_i \lambda^G_j$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. Además, $\mathrm{Det}(F \otimes G) = \mathrm{Det}(F)^m \mathrm{Det}(G)^n$. Notemos finalmente que $(F \otimes G)^k = F^k \otimes G^k$, válido para $k \in \mathbb{N}$ y también $k \in \mathbb{Z}$ si F y G son invertibles.

Si $F = F^i_j e_i f^k$, $G = G^k_l \tilde{e}_k \tilde{f}^l$, $\Rightarrow$ $F \otimes G = F^i_j G^k_l (e_i \otimes \tilde{e}_k)(f^j \otimes \tilde{f}^l)$, por lo que $(F \otimes G)^{ik}_{jl} = F^i_k G^j_l$. Esto corresponde pues al producto tensorial de las matrices que representan a F y G, denominado también producto Kronecker: Ordenando la base en la forma $b = (e_i \otimes \tilde{e}_1, e_1 \otimes \tilde{e}_2, \ldots, e_n \otimes \tilde{e}_m)$, la matriz de $nm \times nm$ que representa a $F \otimes G$ en esta base es

$$[F \otimes G]_b = [F]_e \otimes [G]_{\tilde{e}} = \begin{pmatrix} F^1_1 [G]_{\tilde{e}} & \ldots & F^1_n [G]_{\tilde{e}} \\ \vdots & \ddots & \vdots \\ F^n_1 [G]_{\tilde{e}} & \ldots & F^n_n [G]_{\tilde{e}} \end{pmatrix}$$

En notación de Mecánica Cuántica, $e_i \to |i\rangle$, $f^j \to \langle j|$ y $F \to \sum_{i,j} F_{ij} |i\rangle\langle j|$, $G = \sum_{k,l} G_{kl} |\tilde{k}\rangle\langle\tilde{l}|$, con $F \otimes G = \sum_{i,j,k,l} F_{ij} G_{kl} |i\tilde{j}\rangle\langle k\tilde{l}|$.

27.8 Contracción de tensores

La contracción de un tensor $\binom{p}{q}$, con $p \geq 1$, $q \geq 1$, queda definida por una suma de la forma

$$T^{j_1\ldots k\ldots j_q}_{i_1\ldots\ldots k\ldots i_p}$$

(donde la suma es sobre el índice repetido k), la cual se transforma como un tensor $\binom{p-1}{q-1}$, pues $S^i_k R^k_j = \delta^i_j$. Por ejemplo, si

$$U^j_i = T^{kj}_{ik}$$

entonces

$$U'^{j'}_{i'} = T'^{k'j'}_{i'k'} = S^i_{i'} S^l_{k'} R^{k'}_k R^{j'}_j T^{kj}_{il} = S^i_{i'} \delta^l_k R^{j'}_j T^{kj}_{il} = S^i_{i'} R^{j'}_j T^{kj}_{ik} = S^i_{i'} R^{j'}_j U^j_i$$

donde hemos utilizado $S^l_{k'} R^{k'}_k = \delta^l_k$. Vemos pues que se transforma como un tensor $\binom{1}{1}$.

Así, dado un tensor T^{ij}_{kl} (tensor $\binom{2}{2}$) son posibles las 4 contracciones

$$T^{kj}_{ki}\ ,\ T^{jk}_{ik}\ ,\ T^{kj}_{ik}\ ,\ T^{jk}_{ki}$$

que originan 4 tensores $\binom{1}{1}$ (en general distintos). Por otro lado, las dos posibles contracciones dobles que dan lugar a un escalar (tensor $\binom{0}{0}$) son

$$T^{kj}_{kj}\ ,\ T^{jk}_{kj}$$

Por ejemplo, dado el tensor T^j_i, la única contracción posible es el escalar T^i_i. Este representa la *traza* de la matriz T:

$$\mathrm{Tr}\, T = T^i_i$$

Esta es, como hemos visto, invariante frente a cambios de base.

Dado el tensor producto $T^{jk}_{il} = F^j_i G^k_l$, el escalar $T^{jk}_{jk} = F^j_j G^k_k$ representa, matricialmente, el producto de trazas: $(\mathrm{Tr}F)(\mathrm{Tr}G) = F^i_i G^k_k$, mientras que el escalar $T^{jk}_{kj} = F^j_k G^k_j$ representa la traza del producto: $\mathrm{Tr}(FG) = F^j_k G^k_j$.
Además, la contracción $T^{jk}_{ki} = F^j_k G^k_i$ es un tensor $\binom{1}{1}$, que representa el producto matricial FG.

Un tensor es simétrico respecto a dos índices del mismo tipo si $T^{\ldots i\ldots j\ldots}_{\ldots} = T^{\ldots j\ldots i\ldots}_{\ldots}$, y es antisimétrico si $T^{\ldots i\ldots j\ldots}_{\ldots} = -T^{\ldots j\ldots i\ldots}_{\ldots}$ (Definición similar respecto de índices inferiores). Esta propiedad es independiente de la base: Por ejemplo, si $T^{ij}_{kl} = T^{ji}_{kl}$,

$$T'^{i'j'}_{k'l'} = R^{i'}_i R^{j'}_j S^k_{k'} S^j_{j'} T^{ij}_{kl} = R^{i'}_i R^{j'}_j S^k_{k'} S^j_{j'} T^{ji}_{kl} = T'^{j'i'}_{k'l'}$$

Un tensor es completamente simétrico (antisimétrico) si es simétrico (antisimétrico) respecto de todo par de índices del mismo tipo.

27.9 Determinante: Consideremos una forma multilineal completamente antisimétrica de $V^n \to K$. En tal caso, si $v_i = x^j_i e_j$,

$$F(v_1, \ldots, v_n) = x^{i_1}_1 \ldots x^{i_n}_n F_{i_1,\ldots,i_n}$$

donde $F_{i_1,\ldots,i_n} = F(e_{i_1}, \ldots, e_{i_n})$. Se tiene $F_{\ldots,i,\ldots,j,\ldots} = -F_{\ldots,j,\ldots,i,\ldots}$ para cualquier par de índices i, j. Es claro entonces que $F_{\ldots,i,\ldots,j,\ldots} = 0$ si $i = j$, es decir, si dos (o más) índices coinciden, y que si los índices son todos distintos, $F_{i_1,\ldots,i_n} = (-1)^{n_{i_1,\ldots,i_n}} F_{1,2\ldots,n}$, donde $n_{i_1,\ldots,i_n}$ es el número de permutaciones necesarias para llevar $(i_1, \ldots, i_n)$ al orden normal $(1, 2, \ldots, n)$. Podemos pues escribir

$$F_{i_1,\ldots,i_n} = \lambda \epsilon_{i_1,\ldots,i_n}$$

donde $\lambda = F_{1,2,\ldots,n}$ y $\epsilon_{i_1,\ldots,i_n}$ es el símbolo completamente antisimétrico que satisface $\epsilon_{1,2,\ldots,n} = 1$ (símbolo de Levi-Civita). Por lo tanto,

$$F(v_1, \ldots, v_n) = \lambda x^{i_1}_1 \ldots x^{i_n}_n \epsilon_{i_1\ldots i_n} = \lambda\, \mathrm{Det}[X]$$

donde

$$\mathrm{Det}[X] = x^{i_1}_1 \ldots x^{i_p}_n \epsilon_{i_1\ldots i_n}$$

es el determinante de la matriz de elementos x^i_j (la cual es una función multilineal completamente antisimétrica de las columnas de la matriz, que vale 1 para la matriz identidad). Por ejemplo, para $n=2$, $\mathrm{Det}[X] = x^i_1 x^j_2 \epsilon_{ij} = x^1_1 x^2_2 \epsilon_{12} + x^2_1 x^1_2 \epsilon_{21} = x^1_1 x^2_2 - x^2_1 x^1_2$, mientras que para $n=3$,
$\mathrm{Det}[X] = x^i_1 x^j_2 x^k_3 \epsilon_{ijk} = x^1_1 x^2_2 x^3_3 \epsilon_{123} + x^1_1 x^3_2 x^2_3 \epsilon_{132} + x^2_1 x^3_2 x^1_3 \epsilon_{231} + x^2_1 x^1_2 x^3_3 \epsilon_{213} + x^3_1 x^1_2 x^2_3 \epsilon_{312} + x^3_1 x^2_2 x^1_3 \epsilon_{321}$
$= x^1_1 x^2_2 x^3_3 - x^1_1 x^3_2 x^2_3 + x^2_1 x^3_2 x^1_3 - x^2_1 x^1_2 x^3_3 + x^3_1 x^1_2 x^2_3 - x^3_1 x^2_2 x^1_3$.
Notemos también que $x^i_1 x^j_2 \epsilon_{ij} = x^1_1 x^2_2 - x^2_1 x^1_2 = x^1_1 x^2_2 - x^1_2 x^2_1 = x^1_i x^2_j \epsilon^{ij}$, donde $\epsilon^{ij} = \epsilon_{ij}$, y en general, $\mathrm{Det}[X] = x^{j_1}_1 \ldots x^{j_n}_n \epsilon_{j_1 \ldots j_n} = \frac{1}{n!} x^{j_1}_{i_1} \ldots x^{j_n}_{i_n} \epsilon_{j_1 \ldots j_n} \epsilon^{i_1 \ldots i_n} = x^1_{i_1} \ldots x^n_{i_n} \epsilon^{i_1 \ldots i_n}$,
donde $\epsilon^{i_1 \ldots i_n} = \epsilon_{i_1 \ldots i_n}$.

Observemos que frente a un cambio de base general, $F_{i_1,\ldots,i_n} = F(e_{i_1}, \ldots, e_{i_n})$ transforma como

$$F'_{i'_1 \ldots i'_n} = S^{i_1}_{i'_1} \ldots S^{i_n}_{i'_n} F_{i_1 \ldots i_n} = \lambda S^{i_1}_{i'_1} \ldots S^{i_n}_{i'_n} \epsilon_{i_1 \ldots i_n} = \lambda \mathrm{Det}(S) \epsilon_{i'_1 \ldots i'_n} = \mathrm{Det}(S) F_{i'_1,\ldots,i'_p}$$

Subida y bajada de índices y tensores cartesianos. En un espacio euclideo, es posible bajar o subir índices de un tensor mediante el tensor métrico$g_{ij} = (e_i, e_j)$, y su inversa $g^{ij} = (f^i, f^j)$, que son tensores simétricos de tipo $\binom{0}{2}$ y $\binom{2}{0}$ respectivamente:

$$\begin{aligned} T^{j_1,\ldots,j_q}_{i_1 \ldots, i_p} &= T(e_{i_1}, \ldots, e_{i_p}, f^{j_1}, \ldots, f^{j_q}) = T(e_{i_1}, \ldots, e_{i_p}, g^{j_1 j'_1} e_{j'_1}, \ldots, g^{j_q j'_q} e_{j'_q}) \\ &= g^{j_1 j'_1} \ldots g^{j_q j'_q} T(e_{i_1}, \ldots, e_{i_p}, e_{j_1}, \ldots, e_{jq}) = g^{j_1 j'_1} \ldots g^{j_q j'_q} T_{i_1 \ldots, i_p, j'_1, \ldots, j'_q} \end{aligned}$$

Por ejemplo, si T^j_i es un tensor $\binom{1}{1}$, $T^{ji} = g^{ki} T^j_k$ es un tensor $\binom{2}{0}$ y $T_{ji} = g_{jk} T^k_i$ es un tensor $\binom{0}{2}$. *Tensores cartesianos*: En un espacio euclideo V, si nos restringimos a transformaciones isométricas entre bases ortonormales, entonces $g_{ij} = (e_i, e_j) = \delta_{ij}$, $g^{ij} = \delta^{ij}$ y $f^i = g^{ij} e_j = e_i$. En tal caso no se puede distinguir entre índices covariantes y contravariantes y se tiene $T^i = T_i$, $T^i_j = T^{ij} = T_{ij}$, $T^{ij}_{kl} = T_{ijkl}$, etc.

Notemos precisamente que para transformaciones entre bases ortonormales (isometrías) $R = S^{-1} = S^t$, es decir, $R^i_j = S^j_i$. En tal caso, $T'^j = R^j_i T^i = \sum_i S^i_j T^i$, verificándose que T^j se transforma igual que T_j.

Pseudotensores cartesianos: Si frente a un cambio de base isométrico en un espacio euclideo se tiene

$$T'_{i'_1 \ldots i'_p} = \mathrm{Det}(S) S^{i_1}_{i'_1} \ldots S^{i_p}_{i'_p} T_{i_1 \ldots i_p}$$

se dice que T es un pseudotensor cartesiano de rango p. Se comporta como un tensor de rango p frente a cambios de base que satisfacen $\mathrm{Det}[S] = +1$ (rotaciones) pero exhibe un cambio de signo adicional si $\mathrm{Det}[S] = -1$ (reflexiones).

Por ejemplo, frente a isométrías, el tensor completamente antisimétrico $F_{i_1,\ldots i_n} = F(e_1, \ldots, e_n)$ es un pseudoescalar, mientras que $(\mathbf{a} \times \mathbf{b})_k = a^i b^j \epsilon_{ijk}$ es un pseudovector ($a'^{i'} b'^{j'} \epsilon_{i'j'k} = R^{i'}_i R^{j'}_j a^i a^j \epsilon_{i'j'k} = R^{i'}_i R^{j'}_j R^{k'}_l S^l_k a^i a^j \epsilon_{i'j'k'} = \mathrm{Det}(R) S^l_k a^i a^j \epsilon_{ijl} = \mathrm{Det}(S) S^l_k (\mathbf{a} \times \mathbf{b})_l$).

28 Campos tensoriales, símbolos de Christoffel y derivada covariante

Consideremos un cambio general de coordenadas $x'^i(x^1, \ldots, x^n)$ en $V = \mathbb{R}^n$. Tenemos

$$dx'^i = R^i_j dx^j, \quad R^i_j = \frac{\partial x'^i}{\partial x^j} = \partial_j x'^i$$

La matriz inversa es

$$S^i_j = \frac{\partial x^i}{\partial x'^j} = \partial'_j x^i$$

y satisface

$$S^i_j R^j_k = R^i_j S^j_k = \delta^i_k$$

Tanto S como R *dependen ahora de las coordenadas.* Podemos considerar en c/punto la base definida por

$$e'_i = S^j_i e_j$$

siendo aquí $e = (e_1, \ldots, e_n)$ una base de V *independiente* de las coordenadas, y $e' = (e'_1, \ldots, e'_n)$ *dependiente* de las coordenadas.

Si e es la base canónica, el tensor métrico original es $g_{ij} = (e_i, e_j) = \delta_{ij}$ mientras que en la nueva base, $g'_{ij} = (e'_i, e'_j) = S_i^k S_j^l g_{kl} = S_i^k S_j^l \delta_{kl}$, es decir, $g' = S^T S$ en notación matricial. Se obtiene entonces

$$ds^2 \equiv dx_i dx^i = dx'_i dx'^i = dx'^i dx'^j g'_{ij}$$

Un campo vectorial v dependiente de las coordenadas puede pues escribirse como

$$v = v^i(x^1, \ldots, x^n) e_i = v'^i(x'^1, \ldots, x'^n) e'_i, \quad v'^i = R_j^i v^j$$

Generalizando, si $D \subset V$, un campo tensorial real $\binom{q}{p}$ es una función $T : D \to \underbrace{V \otimes \ldots \otimes V}_{\text{q veces}} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{\text{p veces}}$:

$$T = T^{j_1, \ldots, j_q}_{i_1, \ldots, i_p}(x^1, \ldots, x^n) e_{j_1} \otimes \ldots \otimes e_{j_q} \otimes f^{i_1} \otimes f^{i_p}$$

Frente a un cambio general de coordenadas, se obtiene

$$T = T'^{j'_1, \ldots, j'_q}_{i'_1, \ldots, i'_p}(x'^1, \ldots, x'^n) e'_{j'_1} \otimes \ldots \otimes e'_{j'_q} \otimes f'^{i_1} \otimes f'^{i_p}$$

con

$$T'^{j'_1, \ldots, j'_q}_{i'_1, \ldots, i'_p}(x'^1, \ldots, x'^n) = S^{i_1}_{i'_1} \ldots S^{i_p}_{i'_p} R^{j'_1}_{j_1} \ldots R^{j'_q}_{j_q} T^{j_1, \ldots, j_q}_{i_1, \ldots, i_p}(x^1, \ldots, x^n)$$

Por ejemplo, un campo vectorial es un campo tensorial $\binom{1}{0}$.

Consideremos ahora la derivada de un campo tensorial $\binom{1}{0}$,

$$\partial'_j v = \partial'_j (v'^i e'_i) = (\partial'_j v'^i) e'_i + v'^i (\partial'_j e'_i)$$

El segundo término da cuenta de la dependencia de la base de las coordenadas. Dado que $e'_i = S_i^k e_k$, se tiene $\partial'_j e'_i = (\partial'_j S_i^l) e_l = (\partial'_j S_i^l) R_l^k e'_k$ y por lo tanto

$$\partial'_j e'_i = \Gamma^k_{ij} e'_k$$

donde $\Gamma^k_{ij} = (\partial'_j S_i^l) R_l^k = -S_i^l \partial'_j R_l^k$ son los *símbolos de Christoffel*, que dan cuenta de la variación de los elementos de la base. Como $S_j^i = \partial'_j x^i \Rightarrow \Gamma^k_{ij} = \Gamma^k_{ji}$, pues $\partial'_j S_i^l = \partial'_j \partial'_i x^l = \partial'_i \partial'_j x^l = \partial'_i S_j^l$.

Se obtiene entonces

$$\partial'_j v = [(\partial'_j v'^k) + v'^i \Gamma^k_{ij}] e'_k$$

La expresión

$$v'^k_{;j} \equiv v'^k_{,j} + v'^i \Gamma^k_{ij}$$

donde $v'^k_{,j} \equiv \partial'_j v'^k$, se denomina *derivada covariante* de las componentes contravariantes, y satisface las reglas correctas de transformación. Tenemos pues

$$\partial'_j v = v'^k_{;j} e'_k$$

En el caso de que la base sea independiente de las coordenadas, $\Gamma^k_{ij} = 0$ y la derivada covariante se reduce a la usual ($v^i_{;j} = v^i_{,j}$).

Por ejemplo, la divergencia de un campo vectorial $v = v^i e_i = v'^i e'_i$ puede entonces expresarse en la forma (demostrar como ejercicio)

$$\partial_i v^i = v^i_{,i} = v'^i_{;i} = (\partial'_i v'^i) + v'^i \Gamma^j_{ij}$$

Para componentes covariantes, tenemos $v = v_i f^i = v'_i f'^i$, con $f'^i = R_k^i f^k$, y f^k independiente de las coordenadas. Por lo tanto,

$$\partial'_j v = (\partial'_j v'_i) f'^i + v'^i (\partial'_j f'^i)$$

Pero $\partial'_j f'^i = (\partial'_j R_l^i) f^l = S_k^l (\partial'_j R_l^i) f'^k = -\Gamma^i_{kj}$ por lo que

$$\partial'_j v = [(\partial'_j v'_k) - v'_i \Gamma^i_{kj}] f'^k$$

La derivada covariante de componentes covariantes debe pues definirse como

$$v'_{k;j} = v'_{k,j} - v'_i \Gamma^i_{kj}$$

para que

$$\partial'_j v = v'_{k;j} f'^k$$

En forma análoga se definen las derivadas covariantes de tensores arbitrarios de rango $\binom{p}{q}$

Dado que $g'_{ik} = S^l_i S^m_k g_{lm}$, tenemos, para g_{lm} independiente de las coordenadas, $\partial'_j g'_{ik} = (\partial'_j S^l_i) S^m_k g_{lm} + S^l_i (\partial'_j S^m_k) g_{lm} = (\partial'_j S^l_i) R^r_l S^s_r S^m_k g_{sm} + (\partial'_j S^m_k) R^r_m S^s_r S^l_i g_{ls} = \Gamma^r_{ij} g'_{rk} + \Gamma^r_{kj} g'_{ir}$, por lo que

$$g'_{ik;j} = g'_{ik,j} - g'_{lk}\Gamma^l_{ij} - g'_{il}\Gamma^l_{kj} = 0$$

De esta forma, si $v'_i = g'_{ik} v'^k$ se verifica que $v'_{i;j} = g'_{ik} v'^k_{;j}$. La última ecuación permite también escribir los símbolos de Christoffel directamente en términos de derivadas del tensor métrico:

$$\Gamma^i_{kl} = \frac{1}{2} g^{im} (g_{mk,l} + g_{ml,k} - g_{kl,m})$$

Ejemplo: Para $V = \mathbb{R}^2$ y coordenadas polares, definidas por

$$x = r\cos\theta\,, \quad y = r\sin\theta$$

se obtiene $dx = dr\cos\theta - r\sin\theta d\theta$, $dy = dr\sin\theta + r\cos\theta d\theta$, de forma que

$$S = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}, \quad R = \frac{1}{r}\begin{pmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \quad g' = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$$

con $dr = dx\cos\theta + dy\sin\theta$, $d\theta = (-dx\sin\theta + dy\cos\theta)/r$,
$e_r = e_x\cos\theta + e_y\sin\theta$, $e_\theta = r(-e_x\sin\theta + e_y\cos\theta)$, y e_x, e_y la base canónica. Obtenemos entonces

$$ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\theta^2$$

En este caso, los únicos símbolos de Christoffel no nulos son $\Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = 1/r$, $\Gamma^r_{\theta\theta} = -r$.

La divergencia de un campo vectorial

$$v = v^x e_x + v^y e_y = v^r e_r + v^\theta e_\theta$$

es entonces

$$\partial_x v^x + \partial_y v^y = \partial_r v^r + \partial_\theta v^\theta + v^r \Gamma^\theta_{r\theta} = \partial_r v^r + \partial_\theta v^\theta + v^r/r$$

El gradiente de un campo escalar ϕ puede escribirse en la forma $(\partial^i\phi)e_i = (\partial'^i\phi)e'_i$, donde $\partial'^i = g'^{ij}\partial'_j$. Por lo tanto,

$$\frac{\partial\phi}{\partial x}e_x + \frac{\partial\phi}{\partial y}e_y = \frac{\partial\phi}{\partial r}e_r + \frac{1}{r^2}\frac{\partial\phi}{\partial\theta}e_\theta$$

Finalmente, el Laplaciano de un campo escalar ϕ (la divergencia del gradiente de ϕ) puede expresarse como

$$\partial_i\partial^i\phi = \partial'_i\partial'^i\phi + \Gamma^i_{ji}\partial'^j\phi = \frac{\partial^2\phi}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2\phi}{\partial\theta^2} + \frac{1}{r}\frac{\partial\phi}{\partial r}$$