FACULTAD DE MATEMÁTICAS
PONTIFICIA UNIVERSIDAD
CATÓLICA DE CHILE

Pontificada Universidad Católica

MAT1226

Álgebra Lineal

Autor:
Sebastián Lepe V.

2 de julio de 2025

Índice

Introducción y Motivación

Comentario: Hay que arreglar la relación entre un operador lineal y su matriz.

Hay que especificar que $\|x\|_V^2 = (x, x)_V$.

Hay que estudiar la relación entre $(\cdot, \cdot)_V$ y $(\cdot, \cdot)_W$ donde V, W tienen distitnos cuerpos.

1. Determinantes

Sabemos que el determinante de una matriz 2×2 es,

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

Si a la primera fila le cambiamos el signo, obtenemos,

$$\det \begin{bmatrix} -a & -b \\ c & d \end{bmatrix} = -(ad - bc)$$

Y si intercambiamos las filas obtenemos,

$$\det \begin{bmatrix} c & d \\ a & b \end{bmatrix} = -(ad - bc)$$

Pero nos interesa determinar el valor del determinante de una matriz de 3×3 o de 4×4. Por lo que necesitamos estructurar el determinante formalmente.

El determinante es una función $\det(\cdot)$ que toma una matriz A y entrega un valor (real, complejo, otros).

Axiomas: *Consideremos la función determinante* $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ *(también se puede definir en los complejos). Los axiomas del determinante son:*

(1) $\det(I) = 1$, *donde I es la matriz identidad.*

(2) Sea A una matriz $n \times n$ con filas $f_1, \cdots, f_n$. Entonces,

$$\det \begin{bmatrix} \vdots \\ f_i \\ \vdots \\ f_j \\ \vdots \end{bmatrix} = -\det \begin{bmatrix} \vdots \\ f_j \\ \vdots \\ f_i \\ \vdots \end{bmatrix}$$

para todo $i, j = 1, \cdots, n$ distintos. Es decir, al intercambiar dos filas, sale un factor (-1).

(3) Sean λ, μ constantes, entonces,

$$\det \begin{bmatrix} \vdots \\ \lambda a + \mu b \\ \vdots \end{bmatrix} = \lambda \det \begin{bmatrix} \vdots \\ a \\ \vdots \end{bmatrix} + \mu \det \begin{bmatrix} \vdots \\ b \\ \vdots \end{bmatrix}$$

Es decir, es lineal por filas.

Proposición: *Consideremos la función determinante y sea A una matriz $n \times n$. Entonces se cumple que:*

(1) Si la matriz A tiene dos filas iguales, entonces $\det(A) = 0$.

(2) Se puede sumar múltiplos de una fila a otra fila.

(3) Si la matriz A tiene una fila de 0*, entonces* $\det(A) = 0$.

(4) Si A es una matriz triangular (superior o inferior) con diagonal $\lambda_1, \cdots, \lambda_n$. Entonces $\det(A) = \lambda_1 \cdots \lambda_n$.

Demostración: Sea A una matriz $n \times n$ con filas $f_1, f_2, \cdots, f_n$. Entonces,

(1) Si las filas i, j son iguales, entonces por el axioma dos se tiene que,

$$\det \begin{bmatrix} \vdots \\ f_i \\ \vdots \\ f_i \\ \vdots \end{bmatrix} = -\det \begin{bmatrix} \vdots \\ f_i \\ \vdots \\ f_i \\ \vdots \end{bmatrix} \Longleftrightarrow 2\det(A) = 0$$

Finalmente se tiene que $\det(A) = 0$.

(2) Consideremos las filas f_i, f_j, entonces por linealidad se tiene que,

$$\det \begin{bmatrix} \vdots \\ f_i \\ \vdots \\ f_j + \lambda f_i \\ \vdots \end{bmatrix} = \det \begin{bmatrix} \vdots \\ f_i \\ \vdots \\ f_j \\ \vdots \end{bmatrix} + \lambda \det \begin{bmatrix} \vdots \\ f_i \\ \vdots f_i \\ \vdots \end{bmatrix}$$

$$= \det \begin{bmatrix} \vdots \\ f_i \\ \vdots \\ f_j \\ \vdots \end{bmatrix}$$

(3) Por linealidad se tiene que,

$$\det \begin{bmatrix} \vdots \\ 0 \cdots 0 \\ \vdots \end{bmatrix} = \det \begin{bmatrix} \vdots \\ 0 \cdots 0 \\ \vdots \end{bmatrix} + \det \begin{bmatrix} \vdots \\ 0 \cdots 0 \\ \vdots \end{bmatrix} \Longleftrightarrow \det(A) = 0$$

(4) Supongamos que A es una matriz superior. Por eliminación de Gauss se cumple que,

$$\det \begin{bmatrix} \lambda_1 & \cdot & \cdots \\ 0 & \lambda_2 & \cdots \\ \vdots & \vdots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = \det \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \lambda_1 \cdots \lambda_n \det(I)$$

Por tanto $\det(A) = \lambda_1 \cdots \lambda_n$.

■

Ejemplo:

$$\det \begin{bmatrix} 1 & 2 & 4 \\ -2 & 2 & -2 \\ 1 & 0 & 5 \end{bmatrix} = \det \begin{bmatrix} 1 & 2 & 4 \\ 0 & 6 & 6 \\ 0 & -2 & 1 \end{bmatrix} = \det \begin{bmatrix} 1 & 2 & 4 \\ 0 & 6 & 6 \\ 0 & 0 & 3 \end{bmatrix} = 18$$

Teorema: *Los axiomas de determinantes fijan univocamente el valor del determinante.*

Nota: Más adelante le daremos sentido al teorema.

Definamos formalmente los determinantes.

Definición: *Sea $F : V_1 \times V_2 \times \cdots \times V_n \to \mathbb{R}$ una función donde $V_1, \cdots, V_n$ son espacios vectoriales de dimensión n. Entonces,*

- *F es **alternante** si para todo $v_i \in V_i$ fijos con $i = 1, \cdots, n$, se tiene que $F(v_1, \cdots, v_n) = 0$ cuando existen $i \neq j$ distintos tales que $v_i = v_j$*
- *F es **multilineal** si para todo $v_i = \lambda w + \mu z \in V_i$ fijos con $i = 1, \cdots, n$ (toda fila) se cumple que,*

$$\begin{aligned} F(v_1, \cdots, v_{i-1}, \lambda w + \mu z, v_{i+1}, \cdots, v_n) &= \lambda F(v_1, \cdots, v_{i-1}, w, v_{i+1}, \cdots, v_n) \\ &+ \mu F(v_1, \cdots, v_{i-1}, z, v_{i+1}, \cdots, v_n) \end{aligned}$$

Observación: Si consideramos F multilineal, entonces la definición de alternante es equivalente a la condición de intercambio de vectores (intercambio de filas).

Definición: *Definimos la función determinante por la función* $\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ *tal que es multilineal y alternante tal que* $\det(e_1, \cdots, e_n) = 1$.

Ejemplo:

$$\det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = -\det(I) = -1$$

Ejemplo:

$$\det \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \det \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix} = 2\det(I) = 2$$

Observación: Pedimos que $\det(e_1, \cdots, e_n) = 1$ ya que gracias a la multilinealidad y la alternancia, se puede descomponer la matriz hasta obtener la identidad.

Teorema: *La función determinante es única.*

Idea de la demostración: Consideremos la función determinante $\det(\cdot)$. Mediante la multilinealidad y la alternancia veremos que podemos representar el determinante mediante una fórmula.

Consideremos una matriz 2×2 de la forma $A = (a_{ij})_{i,j=1,2}$. Entonces,

$$\begin{aligned}\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} &= \det \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \\ &= a_{11} \det \begin{bmatrix} 1 & 0 \\ a_{21} & a_{22} \end{bmatrix} + a_{12} \det \begin{bmatrix} 0 & 1 \\ a_{21} & a_{22} \end{bmatrix}\end{aligned}$$

Hacemos lo mismo para la segunda fila, de forma que,

$$\begin{aligned}a_{11} \det \begin{bmatrix} 1 & 0 \\ a_{21} & a_{22} \end{bmatrix} &= a_{11}a_{21} \det \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + a_{11}a_{22} \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = a_{11}a_{22} \\ a_{12} \det \begin{bmatrix} 0 & 1 \\ a_{21} & a_{22} \end{bmatrix} &= a_{12}a_{21} \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + a_{12}a_{22} \det \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} = -a_{12}a_{22}\end{aligned}$$

Por lo tanto,

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Consideremos ahora $A = (a_{ij})_{i,j=1,2,3}$. Entonces,

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \det \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \det \begin{bmatrix} 0 & 0 & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Ahora podemos replicar el proceso de sacar la constante y luego trabajar con los determinantes que se formaron. En particular, se puede observar que,

$$\begin{aligned}\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} &= a_{11}a_{22}a_{33} \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + a_{11}a_{23}a_{32} \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \\ &+ a_{12}a_{21}a_{33} \det \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} + a_{12}a_{23}a_{31} \det \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \\ &+ a_{13}a_{21}a_{32} \det \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + a_{13}a_{22}a_{31} \det \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\end{aligned}$$

Notemos que los coeficientes son de la forma $a_{1i_1}a_{2i_2}a_{3i_3}$ donde $i_1, i_2, i_3 = 1, 2, 3$ no pueden tomar al menos dos valores iguales. Y además que los coeficientes determina la forma de la matriz acompañante, en particular $a_{1i_1}a_{2i_2}a_{3i_3}$ está asociado a la matriz con 1 en la fila 1 y columna i_1, en la fila 2 y columna i_2 y en la fila 3 y columna i_3.

Finalmente obtenemos que,

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}$$

Observación: Al replicar esto en una matriz $n \times n$, solo sobreviven $n!$ términos.

Ahora vamos a generalizar. Definimos$[n] := \{1, \cdots, n\}$, definimos la permutación $P : [n] \to [n]$ por una función biyectiva de $[n]$ a $[n]$ y definimos S_n la colección de permutaciones de $[n]$, es decir,

$$S_n := \{P : [n] \to [n] : P \,\text{es permutación}\}$$

Observación: Una permutación P se puede interpretar como una matriz llamada matriz de permutación, por ejemplo, la permutación $P : [3] \to [3]$ tal que $P(1) = 2, P(2) = 3, P(3) = 1$ tiene forma de matriz de permutación por,

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

Puesto que,

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

Definimos $\tau(P)$:='el número de transposiciones para enviar P a la identidad'. Así por ejemplo,

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

Y por último definimos el signo de una matriz de permutación P por,

$$\operatorname{sgn}(P) := (-1)^{\tau(P)} = \det(P)$$

De forma que,

$$\operatorname{sgn} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \det \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = -1$$

Ahora sea $A = (a_{ij})_{i,j=1}^n$ una matriz $n \times n$. Por axiomas de determinantes, se puede demostrar que,

$$\det(A) = \sum_{P \in S_n} a_{1P(1)} a_{2P(2)} \cdots a_{nP(n)} \operatorname{sgn}(P)$$

Es decir, que el valor del determinante es único. ■

Proposición: *Consideremos una matriz cuadradra A y su transpuesta A^T. Entonces se cumple que* $\det(A) = \det(A^T)$.

Demostración: Sea $A = (a_{ij})_{i,j=1}^n$ una matriz $n \times n$. Por la demostración anterior, se tiene que,

$$\det(A) = \sum_{P \in S_n} \prod_{j=1}^{n} a_{jP(j)} \operatorname{sgn}(P)$$

Por otro lado sabemos que $(A^T)_{ij} = (A)_{ji}$ y que,

$$\det(A^T) = \sum_{Q \in S_n} \prod_{j=1}^{n} (a_{jQ(j)})^T \operatorname{sgn}(Q) = \sum_{Q \in S_n} \prod_{j=1}^{n} a_{Q(j)j} \operatorname{sgn}(Q)$$

Estudiemos sumando por sumando. Sea $Q \in S_n$, luego tenemos el término,

$$\prod_{j=1}^{n} a_{Q(j)j} \operatorname{sgn}(Q)$$

Si Q es una permutación, entonces podemos arreglar los $a_{Q(j)j}$ de tal forma que se cumple,

$$\prod_{j=1}^{n} a_{Q(j)j} = \prod_{k=1}^{n} a_{kQ^{-1}(k)}$$

Donde Q^{-1} es una permutación. Además, se cumple $\operatorname{sgn}(Q^{-1}) = \operatorname{sgn}(Q)$ debido a que $\tau(Q) = \tau(Q^{-1})$ (se necesita las mismas cantindad transformaciones elementales para que Q, Q^{-1} lleguen a la identidad). Finalmente tenemos que,

$$\det(A^T) = \sum_{Q \in S_n} \prod_{k=1}^{n} a_{kQ^{-1}(k)} \operatorname{sgn}(Q^{-1}) = \sum_{P \in S_n} \prod_{k=1}^{n} a_{kP(k)} \operatorname{sgn}(P) = \det(A)$$

■.

Proposición: *Sea A una matriz cuadrada. Se tiene que A es singular si y sólo si* $\det(A) = 0$.

Demostración: Supongamos que A es singular (no invertible), entonces las filas no linealmente independiente, es decir, si las filas son $A_1, \cdots, A_n$, entonces existen $x_1, \cdots, x_n$ no todos nulos tales que $A_1x_1 + \cdots + A_nx_n = 0$, luego,

$$\det(A) = \det \begin{bmatrix} A_1x_1 + \cdots + A_nx_n \\ A_2 \\ \vdots \\ A_n \end{bmatrix} = 0$$

Supongamos ahora que $\det(A) = 0$. Si A fuera invertible, entonces por eliminación de Gauss se obtiene una matriz triangular superior (o inferior), con diagonal no nula, pero entonces $\det(A) \neq 0$ siendo imposible. Por lo tanto A es singular. ■

Proposición: *Sean A, B matrices cuadradas $n \times n$. Entonces $\det(AB) = \det(A)\det(B)$.*

Demostración: Si $\det(B) = 0$, entonces B es singular, lo que implica que AB es singular, por lo que $\det(AB) = 0$.

Supongamos que B no son singular fijo. Consideremos la función $f(A) := \det(AB)/\det(B)$.

Afirmación: *La función $f(\cdot)$ sastisface los axioma de determinantes.*

Demostración: Debemos probar los tres axiomas de determinante.

- $f(I) = \det(B)/\det(B) = 1$.
- Sean $A_1, \cdots, A_n$ las filas de A y sean $B_1, \cdots, B_n$ las columnas de B. Entonces,

$$AB = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} \begin{bmatrix} B_1 & B_2 & \cdots & B_n \end{bmatrix} = \begin{bmatrix} A_1B_1 & A_1B_2 & \cdots & A_1B_n \\ A_2B_1 & A_2B_2 & \cdots & A_2B_n \\ \vdots & \vdots & \cdots & \vdots \\ A_nB_1 & A_nB_2 & \cdots & A_nB_n \end{bmatrix}$$

  (considerando el producto punto). Si intercambiamos las filas A_i, A_j, entonces las filas de AB también se intercambian en el mismo orden, por lo que,

$$f(A) = \frac{\det(AB)}{\det(B)} = \frac{-\det(A'B)}{\det(A)} = f(A')$$

  con A' la matriz A con las filas i, j intercambiadas. De forma que f es alternante.
- Supongamos que $A_i = \lambda A_i' + \mu A_i''$, luego,

$$\begin{aligned} [A_iB_1 \; A_iB_2 \; \cdots \; A_iB_n] &= [\lambda A_i'B_1 + \mu A_i''B_1 \; \cdots \; \lambda A_i'B_n + \mu A_i''B_n] \\ &= \lambda[A_i'B_1 \; A_i'B_2 \; \cdots \; A_i'B_n] + \mu[A_i''B_1 \; A_i''B_2 \; \cdots \; A_i''B_n] \end{aligned}$$

  Luego se tiene que $f(A) = \lambda f(A') + \mu f(A'')$, donde A', A'' son matriz A pero en la fila i tiene A_i', A_i'' respectivamente. De aquí se puede concluir que es multilineal.

Finalmente se tiene que $f(\cdot)$ es una función determinante, por tanto $f(A) = \det(A)$, o dicho de otra forma $\det(AB) = \det(A)\det(B)$. ■

1.1. Determinante de un operador lineal

Sea $f : V \to V$ un operador lineal con V un espacio vectorial de dimensión n con base $\{v_1, \cdots, v_n\}$. Queremos definir el determinante del operador f.

Dado que f es operador lineal, entonces existe una matriz A que representa al operador que está asociado a la base $\{v_1, \cdots, v_n\}$, es decir, dado $a \in V$ con coordenadas,

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

en la base $\{v_1, \cdots, v_n\}$, entonces $f(a)$ tiene coordenadas,

$$A \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Por lo que definimos el determinante del operador f con respecto a la base $\{v_1, \cdots, v_n\}$ por $\det(f) := \det(A)$.

Afirmación: *El determinante del operador f es invariante de las bases, es decir, no cambia de valor si cambiamos de base.*

Demostracíón: Consideremos otra base $\{w_1, \cdots, w_n\}$ de V, entonces,

$$w_j = c_{1j}v_1 + c_{2j}v_2 + \cdots + c_{nj}v_n = \sum_{i=1}^{n} c_{ij}v_i$$

Los c_{ij} forman una matriz C conocida como cambio de base. Así, si $a \in V$ tiene coordenadas,

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

respecto a $\{w_1, \cdots, w_n\}$, entonces,

$$a = x_1w_1 + \cdots + x_nw_n = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} x_jc_{ij} \right) v_i$$

Luego,

$$y_i = \sum_{j=1}^{n} c_{ij}x_i$$

De esta forma,

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = C \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Sea B la matriz de f con respecto a la base respecto a la base $\{w_1, \cdots, w_n\}$. Luego se cumple,

$$\underset{a \text{ en base } \{w_j\}}{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}} \quad \overset{C}{\to} \quad \underset{a \text{ en base } \{v_j\}}{\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}} \quad \overset{A}{\mapsto} \quad \underset{f(a) \text{ en base } \{v_j\}}{A \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}} \quad \overset{C^{-1}}{\to} \quad \underset{f(a) \text{ en base } \{w_j\}}{C^{-1}A \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}} = C^{-1}AC \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

De esta forma si B es la matriz de f, entonces,

$$B = C^{-1}AC$$

Aplicando determinante se obtiene que,

$$\det(B) = \det(C^{-1}AC) = \det(C^{-1})\det(A)\det(C) = \det(A)$$

Por tanto, el determinante del operador f es invariante de las bases. ■

Observación: Por lo tanto podemos definir el determinante de l operador f por el determinante de su matriz asociada a una base respectiva.

Ejemplo: Sea V es espacio vectorial de los polinomios de grado 2. Consideremos la transformación linea f derivada, por lo que,

$$f(p(x))) = p'(x) \in V$$

Una base de V es $\{x^2, x, 1\}$. Consideremos $p(x) = ax^2 + bx + c$, entonces p tiene coordenadas,

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Aplicando la transformación f se obtiene,

$$f(p(x)) = 2ax + b$$

El cual tiene coordenadas,

$$\begin{bmatrix} 0 \\ 2a \\ b \end{bmatrix}$$

Observemos que la matriz asociada a f en base $\{x^2, x, 1\}$ es,

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

Por tanto, el determinante del operador lineal es $\det(f) = 0$, es decir, f no es invertible.

1.2. Cofactores

Consideremos la matriz $A = (a_{ij})_{i,j=1}^{n}$. Sabemos que podemos descomponer el determinante de la matriz A de la siguiente forma,

$$\det\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix} + \det\begin{bmatrix} 0 & a_{12} & \cdots & 0 \\ a_{21} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & 0 & \cdots & a_{nn} \end{bmatrix} + \cdots + \det\begin{bmatrix} 0 & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{bmatrix}$$

Sea M_{ij} la matriz generada a partir de quitar la fila i y columna j de A. Y sea $C_{ij} := (-1)^{i+j}\det(M_{ij})$ Luego se satisface que,

$$\det\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix} = a_{11}C_{11}$$

Puesto que,

$$det\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix} = a_{11}\sum_{P\in\widetilde{S}_n}\prod_{i=1}^{n-1} a_{(i+1)P(i+1)}\operatorname{sgn}(P)$$

donde $\widetilde{S}_n$ son las permutaciones de $[n]$ a $[n]$ tales que $P(1) = 1$. (esto debido a que tenemos en la primera fila un elemento fijo a_{11}). Por otro lado,

$$\det(M_{1j}) = \sum_{P\in S_{n-1}}\prod_{k=1}^{n-1}(M_{1j})_{kP(k)}\operatorname{sgn}(P)$$

donde $(M_{1j})_{kP(k)} = a_{(k+1)P(k+1)}$. Por lo tanto,

$$\det\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix} = a_{11}\det(M_{1j})$$

Ahora, en general tenemos que,

$$\begin{aligned}\det\begin{bmatrix} 0 & \cdots & a_{1j} & \cdots & 0 \\ a_{21} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{bmatrix} &= (-1)^{1+j}\det\begin{bmatrix} a_{1j} & 0 & \cdots & 0 \\ 0 & a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n1} & \cdots & a_{nn} \end{bmatrix} \\ &= (-1)^{1+j}a_{1j}\det(M_{1j}) = a_{1j}C_{1j}\end{aligned}$$

Estudiando M_{1j} se verifica la igualdad. Finalmente se tiene que,

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$$

Definición: *Sea A una matriz $n \times n$. Se define la matriz de cofactores de A por la matriz $n \times n$,*

$$Cof(A) := \begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \cdots & C_{nn} \end{bmatrix}$$

Observación: Se cumple que,

$$\det(A) = \sum_{j=1}^{n} a_{1j}C_{1j} = \sum_{j=1}^{n} a_{2j}C_{2j} = \cdots = \sum_{j=1}^{n} a_{nj}C_{nj}$$

Para comprobarlo basta usar el mismo método para probar con el caso de la primera fila, además, como el determinante es invariante sobre la transpuesta, se puede concluir con respecto a las columnas,

$$\det(A) = \sum_{i=1}^{n} a_{i1}C_{i1} = \sum_{i=1}^{n} a_{i2}C_{i2} = \cdots = \sum_{i=1}^{n} a_{in}C_{in}$$

Proposición: *Sea A una matriz $n \times n$ tal que* $\det(A) \neq 0$. *Entonces A es invertible y se cumple que,*

$$\frac{(Cof(A))^T}{\det(A)} = A^{-1}$$

Demostración: Si $\det(A)$, entonces Sea $A_1, \cdots, A_n$ las filas de A y sea $C_1, \cdots, C_n$ las filas de $\mathrm{Cof}(A)$, entonces,

$$A(\mathrm{Cof}(A))^T = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} \begin{bmatrix} C_1 & C_2 & \cdots & C_n \end{bmatrix} = \begin{bmatrix} A_1C_1 & A_1C_2 & \cdots & A_1C_n \\ A_2C_1 & A_2C_2 & \cdots & A_2C_n \\ \vdots & \vdots & \cdots & \vdots \\ A_nC_1 & A_nC_2 & \cdots & A_nC_n \end{bmatrix}$$

Por definición, para todo $i = 1, \cdots, n$ se tiene que,

$$A_iC_i = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} C_{i1} \\ C_{i2} \\ \vdots \\ C_{in} \end{bmatrix} = \sum_{j=1}^{n} a_{ij}C_{ij} = \det(A)$$

Por otro lado, para todo $i \neq j$ se tiene que,

$$A_iC_j = \sum_{k=1}^{n} a_{ik}C_{jk} = \det \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} = 0$$

Es decir, A_iC_j es el determinante de una matriz con dos filas iguales (las filas i y j dado que $i \neq j$). Por tanto,

$$A(\mathrm{Cof}(A))^T = \det(A)I \Leftrightarrow A^{-1} = \frac{\mathrm{Cof}(A)^T}{\det(A)}$$

■

Regla de Cromer: Consideremos el sistema linea $Ax = b$ con A invertible. Luego $x = A^{-1}b$, por lo que,

$$x = \frac{(\mathrm{Cof}(A))^T b}{\det(A)}$$

Luego si $A_1, \cdots, A_n$ son las columnas de A se tiene que,

$$x_j = \frac{1}{\det(A)} \sum_{j=1}^{n} C_{ij}b_i = \frac{1}{\det(A)} \det \begin{bmatrix} a_{11} & \cdots & a_{1(j-1)} & b_1 & a_{1(j+1)} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2(j-1)} & b_2 & a_{2(j+1)} & \cdots & a_{2n} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{n(j-1)} & b_n & a_{n(j+1)} & \cdots & a_{nn} \end{bmatrix}$$

Por lo tanto, la coordenada j es de la forma,

$$x_j = \frac{\det(B_j)}{\det(A)}$$

donde B_j es la matriz A pero en la columna j se reemplaza por el vector b.

Ejemplo: Consideremos la matriz,

$$A = \begin{bmatrix} 1 & 0 & 12 & 0 & 0 \\ 0 & 1 & -3 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 17 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Determinemos la entrada $(A^{-1})_{3,4}$. Notemos que $\det(A) = 4$, por lo que es invertible y además,

$$A^{-1} = \frac{(\mathrm{Cof}(A))^T}{\det(A)}$$

Por tanto,

$$\begin{aligned}(A^{-1})_{3,4} &= \frac{(\mathrm{Cof}(A))^T_{3,4}}{\det(A)} = \frac{\mathrm{Cof}(A)_{4,3}}{\det(A)} = \frac{C_{4,3}}{\det(A)} \\ &= -\frac{1}{\det(A)} \det \begin{bmatrix} 1 & 0 & 12 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 17 \\ 0 & 0 & 0 & 1 \end{bmatrix} = 0\end{aligned}$$

Ejemplo: Consideremos el sistema lineal,

$$\begin{bmatrix} 1 & 5 & -4 & 3 & 2 \\ 0 & 1 & 2 & 4 & 5 \\ 0 & 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 1 & 12 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 0 \\ 0 \end{bmatrix}$$

Queremos determinar x_3. Sabemos que el determinante de la matriz es 1, por lo que la matriz es invertible y que,

$$\det \begin{bmatrix} 1 & 5 & 1 & 3 & 2 \\ 0 & 1 & 2 & 4 & 5 \\ 0 & 0 & 4 & 2 & 7 \\ 0 & 0 & 0 & 1 & 12 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = 4$$

Por lo tanto $x_3 = 4$.

2. Geometría de los Operadores Lineales

Definición: *Sea $f : V \to V$ un operador lineal. Si $f(a) = \lambda a$ con $a \neq 0$ y con $\lambda \in \mathbb{R}$ (o en otro cuerpo), decimos que a es un **vector propio** de f y λ es su **valor propio** asociado.*

Consideremos un operador lineal A con valor propio λ y vector propio v. Entonces se cumple que,

$$\begin{aligned} Av = \lambda v &\Leftrightarrow Av - \lambda v = 0 \\ &\Leftrightarrow (A - \lambda I)v = 0 \\ &\Leftrightarrow A - \lambda I \text{ es singular} \end{aligned}$$

Esto último se cumple puesto que $v \neq 0$. Luego,

$$Av = \lambda v \Leftrightarrow \det(A - \lambda I) = 0$$

Concluimos que si x es tal que $\det(A - xI) = 0$, entonces x es un valor propio de la matriz A.

Definición: *Sea A una matriz $n \times n$. Definimos ewl polinomio característico de A por,*

$$p_A(x) := \det(A - xI)$$

Ejemplo: Consideremos la matriz,

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

Entonces,

$$P_A(x) = \det \begin{bmatrix} 1-x & 2 \\ 2 & 1-x \end{bmatrix} = x^2 - 2x - 3 = (x+1)(x-3)$$

Por tanto, $x = -1, 3$ son valores propios reales de la matriz A. Determinemos vectores propios. Para $x = -1$ se tiene que,

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Entonces $x_1 + x_2 = 0$, es decir, los vectores propios de -1 son de la forma,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

con x_1 un escalar.

Observación: Los vectores propios de $x = -1$ del ejemplo anterior, forman un subespacio vectorial.

Ejemplo: Consideremos la matriz,

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

Entonces $p_A(x) = x^2 + 1$, con soluciones complejas $x = \pm i$. Es decir, A no tiene valores propios reales pero si complejos. Por lo que no toda matriz tiene valor propio.

Definición: *Sea A una matriz $n \times n$. Sea λ un valor propio de A. Definimos el* ***espacio propio*** *del valor propio λ por el subespacio vectorial de los vectores propios de λ.*

Ejemplo: Consideremos la matriz,

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

Luego $p_A(x) = (1-x)^2(2-x)$, de forma que los valores propios son $x = 1, 2$. Determinemos vectores propios. Sea el valor propios $x = 1$, entonces,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 1 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Entonces se obtiene el sistema de ecuaciones,

$$\begin{aligned} x_1 &= x_1 \\ x_2 &= x_2 \\ 2x_3 &= x_3 \end{aligned}$$

Por lo que x_1, x_2 puede tomar cualquier valor y $x_3 = 0$. Si pensamos en $V = \mathbb{R}^n$ con escalar también en $\mathbb{R}$. Se tiene que el espacio propio de $x = 1$ es,

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} : x_1, x_2 \in \mathbb{R} \right\}$$

Por otro lado, el espacio propio para $x = 2$ es,

$$\left\{ \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} : x_3 \in \mathbb{R} \right\}$$

Proposición: *Sea A una matriz $n \times n$. Supongamos que $\lambda_1, \cdots, \lambda_k$ son k valores propios distintos con $k \leq n$. Entonces los vectores propios $v_1, \cdots, v_k$ respectivos son linealmente independientes.*

Demostración: Demostraremos por inducción. En el caso base se tiene $\alpha_1 v_1 = 0$, luego $\alpha_1 = 0$.

Supongamos que para $s < k$ se tiene que toda colección de vectores $\{v_{i_1}, \cdots, v_{i_s}\}$ con $i_s = 1, \cdots, n$, son independientes. Sin pérdida de generalidad consideremos los vectores $v_1, \cdots, v_{s+1}$ y sin pérdida de generalidad supongamos que $\lambda_1 \neq 0$. Sean $\alpha_1, \cdots, \alpha_{s+1}$ escalares tales que,

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{s+1} v_{s+1} = 0$$

Aplicando A se obtiene,

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \cdots + \alpha_{k+1} \lambda_{k+1} v_{k+1} = 0 \Leftrightarrow \alpha_1 v_1 + \alpha_2 \frac{\lambda_2}{\lambda_1} v_2 + \cdots + \alpha_{s+1} \frac{\lambda_{s+1}}{\lambda_1} v_{s+1} = 0$$

Restando la última ecuación con la primera, se obtiene que,

$$\alpha_2 \left(\frac{\lambda_2}{\lambda_1} - 1 \right) v_2 + \cdots + \alpha_{k+1} \left(\frac{\lambda_{s+1}}{\lambda_1} - 1 \right) v_{s+1} = 0$$

Dado que la colección de k vectores $\{v_2, \cdots, v_{k+1}\}$ es independiente, entonces se tiene que,

$$\alpha_2 \left(\frac{\lambda_2}{\lambda_1} - 1 \right) = \cdots = \alpha_{k+1} \left(\frac{\lambda_{k+1}}{\lambda_1} - 1 \right) = 0$$

Como los valores propios son distintos, se tiene que $\alpha_2 = \cdots = \alpha_{s+1} = 0$. Finalmente en la ecuación inicial se obtiene $\alpha_1 v_1 = 0$, lo que implica que $\alpha_1 = 0$. Por lo tanto $v_1, \ldots, v_{s+1}$ son independientes y por por inducción se tiene que para toda colección de s vectores con $s \leq k$. ■

Corolario: *Si A es una matriz $n \times n$ con respecto a un espacio vectorial V de dimensión n y A tiene n valores propios distintos. Entonces los vectores propios forman una base del espacio vectorial.*

Demostración: Por la proposición anterior, se tiene que la colección de vectores propios $\{v_1, \ldots, v_n\}$ es linealmente independiente. Como el espacio vectorial es de dimensión n, entonces es evidente que $\{v_1, \ldots, v_n\}$ es una base. ■

Proposición: *Sean A, B matrices $n \times n$. Se tiene que $AB = BA$ si A, B tienen una base común de vectores propios.*

Demostración: Sea $\{v_1, \cdots, v_n\}$ base de vectores propios de A y B. Supongamos que $Av_i = \lambda_i v_i$ y $Bv_i = \mu_i v_i$ donde λ_i, μ_i son los valores propios de A y B respectivamente para todo $i = 1, \cdots, n$. Sea $v \in V$, entonces existen $\alpha_1, \cdots, \alpha_n$ tales que,

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

Entonces,

$$\begin{aligned} ABv &= AB(\alpha_1 v_1 + \cdots + \alpha_n v_n) \\ &= \alpha_1 ABv_1 + \cdots + \alpha_n ABv_n \\ &= \alpha_1 A(\mu_1 v_1) + \cdots + \alpha_n A(\mu_n v_n) \\ &= \alpha_1 \lambda_1 \mu_1 v_1 + \cdots + \alpha_n \lambda_n \mu_n v_n \\ &= \alpha_1 \mu_1 \lambda_1 v_1 + \cdots + \alpha_n \mu_n \lambda_n v_n \\ &= \alpha_1 B(\lambda_1 v_1) + \cdots + \alpha_n B(\lambda_n v_n) \\ &= \alpha_1 BAv_1 + \cdots + \alpha_n BAv_n \\ &= BA(\alpha_1 v_1 + \cdots + \alpha_n v_n) \\ &= BAv \end{aligned}$$

Por lo tanto,

$$AB = BA$$

Como queriamos demostrar. ■

Observación: Si A es invertible y tiene valor propio λ, entonces,

$$Av = \lambda v \Leftrightarrow A^{-1} v = \frac{1}{\lambda} v$$

es decir, $1/\lambda$ es valor propio de A^{-1}.

Proposición: *Sea $f : V \to V$ un operador lineal con V de dimensión n. Si $\{v_1, \cdots, v_n\}$ es una base de vectores propios de V. Entonces la matriz de f está representado por una matriz diagonal en la misma base de vectores propios.*

Demostración: Consideremos la base de vectores propios $\{v_1, \ldots, v_n\}$. Sea A la matriz asociada a f con respecto a la base de vectores propios. Entonces, respecto a esta base se cumple que

$$A \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \lambda_i \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

donde el 1 está en la fila i. Esto implica que la columna i de A es de la forma,

$$\begin{bmatrix} 0 \\ \vdots \\ \lambda_i \\ \vdots \\ 0 \end{bmatrix}$$

donde λ_i está en la fila i el vector. Por lo tanto $A = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$ con respecto a la base donde estamos trabajando. ■

Observación: Si $f : V \to V$ es un operador lineal donde V tiene una base de vectores propios, entonces la matriz asociada a f con esta base, es una matriz diagonal con elementos los valores propios.

Definición: *Un operador (o matriz) es* ***diagonalizable*** *si existe una base del espacio vectorial donde la matriz asociada al operador es una matriz diagonal.*

Observación: Consideremos A la matriz asociada al espacio vectorial V. Se tiene que A es diagonalizable si y sólo si existe una matriz invertible C tal que $C^{-1}AC$ es diagonal (C es cambio de coordenada).

Supongamos que,

$$C^{-1}AC = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

Entonces si $C = \begin{bmatrix} C_1 & \cdots & C_n \end{bmatrix}$ se tiene que,

$$AC = A[C_1 \cdots C_n] = [AC_1 \cdots AC_n] = [\lambda_1 C_1 \cdots \lambda_n C_n]$$

Es decir, $AC_k = \lambda_k C_k$. Por lo tanto, si una matriz es diagonalizable, entonces la matriz diagonal está compuesto por los valores propios y las columnas de la matriz C son los vectores propios. Además, se puede concluir que la matriz C no es única.

Proposición: *El polinomio característico de una matriz no depende de la base elegida.*

Demostración: Sea A una matriz con respecto a la base $\{v_1, \cdots, v_n\}$. Consideremos otra base $\{w_1, \cdots, w_n\}$ con cambio matriz cambio de coordenada C, por lo que A en la segunda base es de la forma $C^{-1}AC$. Ahora notemos que,

$$\begin{aligned}
p_{C^{-1}AC}(x) &= \det(C^{-1}AC - xI) = \det(C^{-1}AC - xC^{-1}C) \\
&= \det(C^{-1}(AC - xC)) = \det(C^{-1}(AC - xI)C) \\
&= \det(C^{-1}C)\det(A - xI) \\
&= \det(A - xI) = p_A(x)
\end{aligned}$$

Por lo tanto, el polinomio característico es independiente de la base. ■

Observación: El valor propio de una matriz no cambia al hacer un cambio de base, es decir, el valor propio siempre será valor propio si cambiamos las base del espacio vectorial.

Definición: *Sea A una matriz $n \times n$. Sea λ un valor propio de A. Definimos:*

- *La* ***multiplicidad algebráica*** *de λ por la multiplcidad como raíz de $p_A(x)$ y lo denotaremos por $ma(\lambda)$.*
- *La* ***multiplicidad geométrica*** *de λ como la dimensión del espacio propio de λ y lo denotaremos por $mg(\lambda)$.*

Ejemplo: Consideremos la matriz,

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

Claramente $p_A(x) = (1-x)^2(2-x)$. Se puede ver que $x = 1$ tiene multiplicidad algebráica y geométrica 2.

Ejemplo: Consideremos la matriz,

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

Proposición: *Consideremos un valor propio λ, entonces siempre se cumple que,*

$$ma(\lambda) \geq mg(\lambda)$$

Demostración: Sea E_λ el espacio propio del valor propio de λ. Sea $f : V \to V$ el operador lineal que tiene valor propio λ, entonces se tiene que,

$$f\big|_{E_\lambda} = \lambda \mathrm{Id}_{E_\lambda}$$

Observemos que $f\big|_{E_\lambda}$ en una base es de la forma,

$$\lambda I$$

donde I es la matriz identidad de $\text{mg}(\lambda) \times \text{mg}(\lambda)$. Entonces existe una base tal que la matriz de f es de la forma,

$$A = \begin{bmatrix} \lambda I & * \\ 0 & B \end{bmatrix}$$

Donde I es la matriz identidad de $\text{mg}(\lambda) \times \text{mg}$, B es una matriz $(n - \text{mg}(\lambda)) \times (n - \text{mg})$ no necesariamente conocido, $*$ es para simplificar y 0 es para representar que hay solo 0 debajo de λI. Esto se puede hacer debido a que podemos tomar bases ideales. Obteniendo una abreviación de la matriz A.

Afirmación: *Consideremos la matriz A de $(n+m) \times (n+m)$ abreviada,*

$$A = \begin{bmatrix} A_1 & * \\ 0 & A_2 \end{bmatrix}$$

donde A_1, A_2 son matrices de $n \times n$ y $m \times m$ respectivamente. Entonces,

$$\det(A) = \det(A_1)\det(A_2)$$

Demostración: Consideremos la fórmula general del determinante. Tenemos que,

$$\det(A) = \sum_{P \in S_{n+m}} \prod_{i=1}^{n+m} a_{iP(i)} \text{sgn}(P)$$

Notemos lo siguiente, sea $\widetilde{P}$ una permutación tal que $\widetilde{P}(1) = n+1$, entonces existe un j donde $n < j$ tal que $\widetilde{P}(j) \leq n$, esto implica que $a_{jP(j)} = 0$ y entonces el sumando asociada a la permutación $\widetilde{P}$ cumple que,

$$\prod_{i=1}^{n+m} a_{i\widetilde{P}(i)} \text{sgn}(\widetilde{P}) = 0$$

En general, si $\widetilde{P}$ es una permutación tal que $P(l) > n$ para algún $l = 1, \ldots, n$, entonces existe un $j = n+1, \ldots, n+m$ tal que $P(j) \leq n$, esto implica que $a_{jP(i)} = 0$ y por tanto,

$$\prod_{i=1}^{n+m} a_{i\widetilde{P}(i)} \text{sgn}(\widetilde{P}) = 0$$

Por lo tanto, solo sobreviven la suma de permutaciones tales que $P(l) \leq n$ para todo $l = 1, \ldots, n$. Que son las mismas permutaciones tales que $P(l) > n$ para todo $l = n+1, \ldots, n+m$. Obteniendo,

$$\det(A) = \sum_{\substack{P \in S_{n+m} \\ P(j) \leq n : \forall j = 1, \ldots, n}} \prod_{i=1}^{n+m} a_{iP(i)} \text{sgn}(P)$$

Es decir, en la parte $*$ no importa su contenido y solo nos interesa los coeficientes en A_1, A_2. Probemos por casos el enunciado.

- A_2 **no invertible:** Claramente se cumple igualdad puesto que mediante combinaciones lineales con respecto a filas mayores a n, se obtiene una fila de 0. Por lo que,

$$\det \begin{bmatrix} A_1 & * \\ 0 & A_2 \end{bmatrix} = 0 = \det(A_1)\det(A_2)$$

Y por tanto,

$$\det \begin{bmatrix} A_1 & * \\ 0 & A_2 \end{bmatrix} = 0 = \det(A_1)\det(A_2)$$

- A_2 **es invertible y** A_1 **no:** Como A_2 es invertible, entonces mediante combinaciones lineales podemos eliminar $*$, de forma que,

$$\det \begin{bmatrix} A_1 & * \\ 0 & A_2 \end{bmatrix} = \det \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

Si A_1 no es invertible basta estudiar el caso anterior concluyendo que,

$$\det \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} = 0 = \det(A_1)\det(A_2)$$

- A_2, A_1 **invertibles:** Usando el caso anterior se obtiene que,

$$\det \begin{bmatrix} A_1 & * \\ 0 & A_2 \end{bmatrix} = \det \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

Ahora, como A_1, A_2 son invertibles se tiene que,

$$\det(A_1) = (-1)^{k_1} \det \begin{bmatrix} \widetilde{a}_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \widetilde{a}_{nn} \end{bmatrix} = (-1)^{k_1} \prod_{i=1}^{n} \widetilde{a}_{ii}$$

$$\det(A_2) = (-1)^{k_2} \det \begin{bmatrix} \widetilde{b}_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \widetilde{b}_{mm} \end{bmatrix} = (-1)^{k_2} \prod_{i=1}^{n} \widetilde{b}_{ii}$$

con $\widetilde{a}_{ii}, \widetilde{b}_{ii} \neq 0$ para todo $i = 1, \ldots, n$ y donde k_1, k_2 es la cantidad de intercambios de filas para que A_1, A_2 respectivamente sean una matriz diagonal. Finalmente,

$$\begin{aligned} \det \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} &= (-1)^{k_1+k_2} \det \begin{bmatrix} \widetilde{a}_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \widetilde{b}_{mm} \end{bmatrix} \\ &= \left((-1)^{k_1} \prod_{i=1}^{n} \widetilde{a}_{ii} \right) \left((-1)^{k_2} \prod_{i=1}^{n} \widetilde{b}_{ii} \right) = \det(A_1)\det(A_2) \end{aligned}$$

Por lo tanto,

$$\det(A) = \det(A_1)\det(A_2)$$

Demostrando la afirmación. ■

Consideremos el polinomio característico de A, entonces,

$$\begin{aligned}\det(A - xI) &= \det\begin{bmatrix}\lambda I - xI & * \\ 0 & B - xI\end{bmatrix} = \det((\lambda - x)I)\det(B - xI) \\ &= (\lambda - x)^{\dim E_\lambda} p_B(x)\end{aligned}$$

donde $\dim E_\lambda = \mathrm{mg}(\lambda)$. Como el polinomio característico es independiente de la base, se cumple que el factor $(\lambda - x)^{\dim E_\lambda}$ divide a $p_A(x)$ de forma entera y por tanto $\mathrm{mg}(\lambda) = \dim E_\lambda \leq \mathrm{ma}(\lambda)$ como queriamos probar. ■

Suma directa: Sea V un espacio vectorial y sean $W_1, \ldots, W_k$ subespacios vectoriales. Denotamos la suma de los k subespacios por:

$$W = W_1 + \cdots + W_k := \{w_1 + \cdots + w_k : w_1 \in W_1, \cdots, w_k \in W_k\}$$

Decimos que la suma es directa si cumple una de las tres definiciones:

(1) Para todo $w \in W$ existe un único $w_1 \in W, \cdots, w_k \in W_k$ tal que $w = w_1 + \cdots + w_k$.

(2) Si $w_1 \in W_1, \cdots, w_k \in W_k$ tales que $0 = w_1 + \cdots + w_k$, entonces $w_1 = \cdots = w_k = 0$

(3) Si juntamos una base de cada subespacio $W_1, \cdots, W_k$, se obtiene una base de W

Si $W_1, \cdots, W_k$ están en suma directa, entonces lo denotamos por,

$$W_1 \oplus \cdots \oplus W_k$$

Afirmación: *Las tres definiciones de suma directa son equivalentes.*

Demostración:

- **(1) implica (2):** Sean $w_1 \in W_1, \cdots, w_k \in W_k$ tales que $0 = w_1 + \cdots + w_k$. Si $0 \in W$, entonces existen únicos $\widetilde{w}_1 \in W_1, \cdots \widetilde{w}_k \in W_k$ tales que,

  $$0 = \widetilde{w}_1 + \cdots + \widetilde{w}_k$$

  Restando ambas igualdes, obtenemos que,

  $$0 = (\widetilde{w}_1 - w_1) + \cdots + (\widetilde{w}_k - w_k)$$

  donde $\widetilde{w}_i - w_i \in W_i$ para todo $i = 1, \cdots, k$. Pero sabemos que 0 tiene escritura única, de forma que $\widetilde{w}_i - w_i = \widetilde{w}_i$ para todo $i = 1, \cdot, k$, por lo que $w_1 = w_2 = \cdots = w_k = 0$.

- **(2) implica (3):** Consideremos $E = W_1 + W_2$, probemos que $W_1 \cap W_2 = \{0\}$. Sea $w \in W_1 \cap W_2$, entonces,

$$w = w_1 + w_2 \Leftrightarrow (w - w_1) + (-w_2) = 0 \Leftrightarrow (-w_1) + (w - w_2) = 0$$

con $w_1 \in W_1, w_2 \in W_2$. Por (2) se cumple que $w_1 = w_2 = 0$, es decir, $w = 0$. Por lo que $W_1 \cap W_2 = \{0\}$. Ahora, si se cumple (2) para k, entonces se cumple para todo $i \leq k$.

Sean $\{v_1, \cdots, v_n\}$ y $\{w_1, \cdots, w_m\}$ las bases de W_1, W_2 respectivamente, claramente,

$$\{v_1, \cdots, v_n, w_1, \cdots, w_m\}$$

es un generador de E. Probemos que es una base de E, notemos que si W_1, W_2 son de dimensión n, m respectivamente, entonces E es de dimensión a lo más $n + m$. Sean $\alpha_1, \cdots, \alpha_{n+m}$ tales que,

$$\alpha_1 v_1 + \cdots \alpha_{n+m} v_{n+m} = 0 \Leftrightarrow \alpha_1 v_1 + \cdots + \alpha_n v_n = -\alpha_{n+1} w_1 - \cdots - \alpha_{n+m} w_m$$

Por tanto se tiene que,

$$-\alpha_{n+1} w_1 - \cdots - \alpha_{n+m} w_m \in W_1$$

Es decir,

$$-\alpha_{n+1} w_1 - \cdots - \alpha_{n+m} w_m = 0$$

Y por tanto $\alpha_{n+1} = \cdots = \alpha_{n+m} = 0$. Realizando lo mismo pero con la combinación lineal de $v_1, \cdots, v_n$, se concluye que,

$$\alpha_1 = \cdots = \alpha_{n+m} = 0$$

Es decir, E tiene dimensión $n + m$ y $\{v_1, \cdots, v_n, w_1, \cdots, w_m\}$ es base.

- **(3) implica (1):** Sea $w \in W$, entonces existen $w_1 \in W_1, \cdots, w_k \in W_k$ tales que,

$$w = w_1 + \cdots + w_k$$

Supongamos que,

$$w_i = \alpha_{i1} v_{i1} + \cdots + \alpha_{in_i} v_{in_i}$$

donde $\{v_{i1}, \cdots, v_{in_i}\}$ es base de W_i para todo $i = 1, \cdots, k$. Entonces, $\{v_{11}, \cdots, v_{kn_k}\}$ es base de W y luego,

$$w = \alpha_{11} v_{11} + \cdots + \alpha_{kn_k} v_{kn_k}$$

Como las coordenadas son únicas, se tiene que los coeficientes α_{ij} son únicos y por tanto, los w_i son únicos para todo $i = 1, \cdots, n$.

Probando de esta forma que las tres definiciones son equivalentes. ■

Proposición: *Si $\lambda_1, \cdots, \lambda_k$ son valores propios distintos, entonces sus espacios propios están en suma directa.*

Demostración: Sea E_{λ_i} el espacio propio del valor propio λ_i para $i = 1, \ldots, n$.

$$E := E_{\lambda_1} + \cdots + E_{\lambda_k}$$

Sean $v_1 \in E_{\lambda_1}, \cdots, v_k \in E_{\lambda_k}$ tales que $0 = v_1 + \cdots + v_k$. Si $\{v_1, \cdots, v_k\}$ es linealmente independiente, entonces 0 no puede ser representado por una combinación lineal de los vectores $v_1, \ldots, v_k$, por lo que necesariamente la colección $\{v_1, \cdots, v_k\}$ no puede ser linealmente independiente, sin pérdida de generalidad supongamos que existen $\alpha_1, \cdots, \alpha_k$ no todos nulos tales que,

$$\alpha_1 v_1 + \cdots + \alpha_{k-1} v_{k-1} = v_k$$

(es decir, estamos suponiendo que v_k depende linealmente de $v_1, \cdots, v_{k-1}$). Por lo que,

$$v_1 + \cdots + v_{k-1} + v_k = (\alpha_1 + 1)v_1 + \cdots + (\alpha_{k-1} + 1)v_{k-1} = 0$$

Aquí aplicamos el mismo argumento anterior, es decir, se tiene que $\{v_1, \cdots, v_{k-1}\}$ no es linealmente independiente, de forma que existen $\beta_1, \cdots, \beta_{k-2}$ no todos nulos tales que,

$$\beta_1 v_1 + \cdots + \beta_{k-2} v_{k-2} = v_{k-1}$$

(sin pérdida de generalidad suponemos que v_{k-1} depende de $v_1, \cdots, v_{k-2}$). Por lo que,

$$\begin{aligned} v_1 + \cdots + v_k &= (\alpha_1 + 1)v_1 + \cdots + (\alpha_{k-2})v_{k-2} + (\alpha_{k-1} + 1)(\beta_1 v_1 + \cdots + \beta_{k-2} v_{k-2}) \\ &= \gamma_1 v_1 + \cdots + \gamma_{k-2} v_{k-2} = 0 \end{aligned}$$

Sin pérdida de generalidad suponemos que v_s depende linealmente de $v_1, \cdots, v_{s-1}$ para todo $s = 2, \cdots, k$. Por lo que de forma recursiva se concluye que $v_1 = 0$ y por tanto $v_1 = v_2 = \cdots = v_k = 0$. Probando que los espacios propios están en suma directa. ■

Proposición: *Sea $f : V \to V$ un operador lineal y supongamos que* $\dim V = n$. *Entonces son equivalentes:*

(1) f es diagonalizable.

(2) V es suma directa de los espacios propios de f.

(3) La suma de la multiplicidad geométrica de los valores propios de f es igual a n.

Demostración: Supongamos que f tiene k valores propios distintos, entonces sus espacios propios están en suma directa. Consideremos,

$$E = E_{\lambda_1} + \cdots + E_{\lambda_k}$$

- **(1) implica (2):** Si f es diagonalizable, entonces existe una base de V tal que f tiene matriz diagonal, y esto quiere decir que, hay una base de vectores propios, es decir, si juntamos las bases de los espacios propios, obtenemos una base de V y por tanto $E = V$.

- **(2) implica (3):** Si $V = E$, entonces se tiene que la unión de las bases de E_{λ_i} con $i = 1, \cdots, k$ es igual a una base de V, aquí se tiene que la suma de la multiplicidad geométrica de los valores propios es n.
- **(3) implica (1)**: Si la suma de las multiplicidades geométricas de los valores propios de f es igual a n, entonces hay una base de vectores propios y esto implica que f tiene forma de diagonal en esta base, es decir, f es diagonalizable.

Demostrando la proposición. ■

Corolario: *Si $f : V \to V$ es un operador lineal donde $\dim V = n$, y tiene n valores propios distintos, entonces es diagonalizable.*

Demostración: Si f tiene n valores propios distintos, entonces podemos formar una base de n vectores propios, luego la matriz asociada de f a esta base es una matriz diagonal. ■

Ejemplo: Consideremos la matriz,

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

con espacio vectorial $V = \mathbb{R}^2$ y cuerpo $\mathbb{R}$. Entonces λ es un valor propio con multiplicidad algebráica 2. Determinemos la multiplicidad geométrica. Se tiene que,

$$E_\lambda = \left\{ v = \begin{bmatrix} v_1 \\ 0 \end{bmatrix} : v_1 \in \mathbb{R} \right\}$$

Es decir, la multiplicidad geométrica es 1 y como λ es el único valor propio, se tiene que f no es diagonalizable.

2.1. Matrices Triangularizables

Definición: *Un operador es **triangularizable** si existe una base tal que el operador es una matriz triangular.*

Definición: *Una matriz A es triangularizable si existe C es invertible tal que $C^{-1}AC$ es triangular.*

Proposición: *Sea $f : V \to V$ un operador lineal tal que $\dim V = n$. Entonces f es triangularizable si y sólo si en V existe una cadena creciente maximal de subespacios f-invariantes.*

Sea V espacio vectorial de dimensión n y sea $f : V \to V$ operador lineal. Que sea una cadena creciente maximal de subespacios f-invariante, significa que existen subespacios V_i con $i = 0, 1, \cdots, n$, tal que,

$$\{0\} = V_0 \subset V_1 \subset \cdots \subset V_{n-1} \subset V_n = V$$

tales que $f(V_i) \subseteq V_i$ para todo $i = 0, \cdots, n$. Decimos que es maximal puesto que tiene largo maximal, ya que se cumple que,

$$\dim V_i < \dim V_{i+1}$$

para todo $i = 0, \cdots, n-1$, entonces V_i tiene dimensión i.

Demostración: Sea $\{v_1, \cdots, v_n\}$ base de V tal que f es triangularizable, es decir, es de la forma,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Donde A es la matriz asociada a f en la base $\xi = \{v_1, \cdots, v_n\}$. Notemos que,

$$[f(v_1)]_\xi = A \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = a_{11}[v_1]_\xi$$

donde $[\cdot]_\xi$ son las coordenadas con respecto a la base ξ, por lo que $f(v_1) = a_{11}v_1$. Evaluando por v_2 obtenemos,

$$[f(v_2)]_\xi = A \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ 0 \end{bmatrix} = a_{12}[v_1]_\xi + a_{22}[v_2]_\xi$$

Por lo que $f(v_2) = a_{12}v_1 + a_{22}v_2$. En general,

$$f(v_i) = a_{1i}v_1 + \cdots + a_{ii}v_i$$

Definimos,

$$V_i = \langle v_1, \cdots, v_i \rangle, \;\; V_0 = \{0\}, \;\; V_n = V$$

Entonces se satisface la cadena

$$V_0 \subset V_1 \subset \cdots \subset V_{n-1} \subset V_n$$

Y también se cumple que,

$$f(V_i) \subseteq V_i$$

Encontrando una cadena creciente maximal.

Supongamos ahora que existe una cadena creciente maximal f -invariante $V_0, \cdots, V_n$. Si se cumple que,

$$i = \dim V_i < \dim V_{i+1} = i + 1$$

para todo $i = 0, \cdots, n-1$. Entonces en V_{i+1} existe un vector v_{i+1} que no está en V_i tal que $\langle V_i, v_{i+1}\rangle = V_{i+1}$ (v_{i+1} es linealmente independiente de la base de V_i). Sin pérdida de generalidad supongamos que,

$$V = \langle v_1, v_2, \cdots, v_n\rangle$$

tal que,

$$V_i = \langle v_1, \cdots, v_i\rangle$$

para todo $i = 1, \cdots, n$.

Afirmamos que la base $\{v_1, \cdots, v_n\}$ triangulariza el operador lineal. Si $f(V_i) \subseteq V_i$, entonces existen $a_{11}, \cdots, a_{ii}$ tales que,

$$f(v_i) = a_{1i}v_1 + \cdots + a_{ii}v_i \in V_i$$

para todo $i = 1, \cdots, n$. Aquí concluimos que f en la base $\{v_1, \cdots, v_n\}$ tiene matriz asociada,

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Es decir, f es triangularizable.

■

Ejemplo: La matriz,

$$\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

No es diagonalizable pero si es triangularizable.

Ejemplo: Consideremos la matriz,

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

Está matriz es triangularizable, veamos como pasar de A a su forma triangular. Queremos construir una cadena creciente maximal,

$$\{0\} = V_0 \subset V_1 \subset V_2 \subset V_3 = V$$

tal que $AV_i \subseteq V_i$ y donde $V_i = \{v_1, \cdots, v_i\}$ con $i = 1, 2, 3$. Ya que al tomar $C = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$, vemos que es una matriz invertible tal que,

$$C^{-1}AC = \text{Matriz Triangular}$$

Notemos que $p_A(x) = (1-x)^3$, entonces $x = 1$ es un vector propio y podemos tomar como vector propio,

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

En particular, $Av_1 = v_1$, luego escogemos $V_1 := \langle v_1 \rangle$. Ahora necesitamos v_2 tal que $Av_2 = \alpha v_1 + \beta v_2$ y que sea linealmente independiente de v_1.

Tenemos que,

$$Av_2 = \alpha v_1 + \beta v_2 \Leftrightarrow (A - \beta I)v_2 = \alpha v_1$$

$$\Leftrightarrow \begin{bmatrix} 1-\beta & 0 & 1 \\ 0 & 1-\beta & 0 \\ 0 & 1 & 1-\beta \end{bmatrix} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{=v_2} = \begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix}$$

Obteniendo el sistema,

$$\begin{aligned} (1-\beta)x_1 + x_3 &= \alpha \\ (1-\beta)x_2 &= 0 \\ x_2 + (1-\beta)x_3 &= 0 \end{aligned}$$

Si $\beta = 1$, entonces $x_2 = 0, x_3 = \alpha$. Sin pérdida de generalidad podemos escoger $x_1 = 0$ y $\alpha = 1$, obteniendo que,

$$v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Y funciona, puesto que es linealmente independiente de v_1 y,

$$Av_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = v_1 + v_2$$

Luego tomamos $V_2 = \langle v_1, v_2 \rangle$. Finalmente debemos tomar v_3 tal que $Av_3 = \widetilde{\alpha} v_1 + \widetilde{\beta} v_2 + \widetilde{\gamma} v_3$ y que sea linealmente independiente de v_1, v_2, sin embargo, al estar en un espacio vectorial de dimensión 3, solo basta encontrar v_3 que sea linealmente independiente de v_1, v_2. Sea

$$v_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Y tomamos $V_3 = \langle v_1, v_2, v_3 \rangle = V$. Por lo tanto hemos construido una cadena creciente maximal A-invariante, entonces,

$$C = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Finalmente triangularizamos A,

$$\begin{aligned} C^{-1}AC &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Observación: Este método se puede replicar para cualquier dimensión. Se debe ir construyendo el espacio vectorial V con vectores v_i tales que,

$$Av_i = \alpha_1 v_1 + \cdots + \alpha_i v_i$$

y que sea linealmente independiente a los vectores $v_1, \cdots, v_{i-1}$. Luego la matriz $C = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ triangulariza la matriz.

Ejemplo: Queremos triangularizar usando otro método. Sea,

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

Notemos que el polinomio característico de A es $p_A(x) = (3-x)^3$ y notemos que,

$$A \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 3 \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{=:v_1}$$

Definimos $V_1 := \langle v_1 \rangle$. Sea W subespacio vectorial de V tal que,

$$V = V_1 \oplus W$$

En particular podemos tomar,

$$W = \left\langle \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\rangle$$

Consideremos la submatriz,

$$\widetilde{A} = \begin{bmatrix} 1 & 1 \\ -4 & 5 \end{bmatrix}$$

Vamos a triangularizar $\widetilde{A}$ para luego triangularizar A. Notemos que,

$$p_{\widetilde{A}}(x) = (3 - x)^2$$

Luego podemos tomar el vector propio,

$$\widetilde{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Como estamos trabajando con una matriz 2×2, basta completar la base, en particular escogemos,

$$\widetilde{v}_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Ahora volviendo a la matriz original, tomamos,

$$v_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Pero falta ver si es buena elección. Observamos que v_1, v_2, v_3 son linealmente independiente y que,

$$\begin{aligned} Av_1 &= 3v_1 \\ Av_2 &= 2v_1 + 3v_2 \end{aligned}$$

Finalmente definiendo $V_2 = \langle v_1, v_2 \rangle$ y $V_3 = \langle v_1, v_2, v_3 \rangle = V$ obtenemos una cadena creciente maximal A-invariante. Por tanto la matriz,

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

Triangulariza la matriz A. En particular,

$$\begin{aligned} C^{-1}AC &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ -4 & 5 & 0 \\ -6 & 4 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 2 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} \end{aligned}$$

Observación: El método anterior funciona bien porque la matriz posee una columna casi nula. Si no fuera tal el caso sería más complicado de determinar los vectores.

Ejemplo: Consideremos una matriz de rotación,

$$\begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

con $\theta \in [0, 2\pi]$. Pensemos en el espacio vectorial $V = \mathbb{R}^2$ con cuerpo $C = \mathbb{R}$. Si la matriz fuera triangularizable, entonces existen valores propios λ_1, λ_2 reales. Estudiando el polinomio característico,

$$\det \begin{bmatrix} \cos\theta - x & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta - x \end{bmatrix} = (\cos\theta - x)^2 + \operatorname{sen}^2\theta = 1 - 2x\cos\theta + x^2$$

Luego los valores propios son de la forma,

$$x_\pm = \frac{2\cos\theta \pm \sqrt{4\cos^2\theta - 4}}{2}$$

Observemos que $x_\pm$ es real si y sólo si $\cos\theta = \pm 1$, en este caso se tiene que la matriz tiene dos valores propios reales y luego es triangularizable. Por otro lado, si $\cos\theta \neq \pm 1$, entonces los valores propios son complejos, de forma que no es triangularizable por ninguna base.

Observación: Si en el ejemplo anterior extendemos el espacio vectorial V al cuerpo $C = \mathbb{C}$, entonces se tiene que la matriz si es triangularizable.

Proposición: *Sea V un espacio vectorial sobre $\mathbb{C}$. Entonces todo operador lineal $f : V \to V$ es triangularizable.*

Demostración: Vamos a demostrar por inducción sobre la dimensión de V. Si $\dim V = 1$, entonces es evidente puesto que toda matriz es triangular superior.

Supongamos que $\dim V = n$ y que toda matriz/operador $(n-1) \times (n-1)$ es triangularizable. Sea A una matriz $n \times n$ asociada a f, si p_A es un polinomio definido sobre $\mathbb{C}$, entonces todas sus raíces están en $\mathbb{C}$ y luego se puede descomponer. Esto implica que f tiene vectores propios. Sea $V_1 := \langle v_1 \rangle$ donde V_1 es vector propio, luego existe un subespacio vectorial $W = \langle w_2, \cdots, w_n \rangle$ tal que,

$$V = V_1 \oplus W$$

Definimos $\widetilde{f} : W \to W$ la restricción de f al espacio W. Se tiene que todo $v \in V$ tiene representación única $v = \widetilde{v}_1 + w$ donde $\widetilde{v}_1 \in V_1, w \in W$ y podemos definir proyecciones:

$$\begin{cases} P_{V_1} : V \to V_1, & P_{v_1}(v) = \widetilde{v}_1 \\ P_W : V \to W, & P_w(v) = w \end{cases}$$

Por lo que $\widetilde{f} : W \to W$ se define por $\widetilde{f}(w) = P_W(f(w))$. Lo importante es que W es de dimensión $n-1$, por lo que existe una cadena creciente de subespacios de W $\widetilde{f}$-invariantes,

$$\{0\} = W_0 \subset W_1 \subset \cdots W_{n-1} = W$$

donde $\dim W_i = i$ y $\widetilde{f}(W_i) \subseteq W_i$ para todo $i = 0, 1, \cdots, n-1$.

Ahora, si V_1 está en suma directa con W, entonces V_1 está en suma directa con cualquier subespacio de W. Con esto en cuenta definimos,

$$\begin{aligned} V_2 &:= V_1 \oplus W_1 \\ V_3 &:= V_1 \oplus W_2 \\ &\vdots \\ V_n &:= V_1 \oplus W_{n-1} = V \end{aligned}$$

Veamos que son f-invariantes. Sea $v \in V_i = V_1 \oplus W_{i-1}$, entonces,

$$v = \alpha v_1 + w_{i-1}$$

de forma única. Luego,

$$\begin{aligned} f(v) &= f(\alpha v_1 + w_{i-1}) \\ &= \alpha\lambda v_1 + f(w_{i-1}) \\ &= \alpha\lambda v_1 + P_{V_1}(f(w_{i-1})) + P_W(f(w_{i-1})) \\ &= \underbrace{\alpha\lambda + P_{V_1}(f(w_{i-1}))}_{\in V_1} + \underbrace{\widetilde{f}(w_{i-1})}_{\in W_{i-1}} \end{aligned}$$

Por lo tanto, $f(v) \in V_i$. Finalmente obtenemos una cadena creciente maximal,

$$V_0 \subset V_1 \subset \cdots V_n = V$$

Que es f-invariante, lo que implica que f es triangularizable (o bien que toda matriz n es triangularizable). ■

Proposición: *Sea $A = (a_{ij})_{i,j=1}^n$ una matriz. Luego se cumple,*

(1) $p_A(x) = \det(A - xI) = (-x)^n + Tr(A)(-x)^{n-1} + \cdots + \det(A)$. Donde $Tr(A) = a_{11} + \cdots + a_{nn}$.

(2) Para todo C invertible se tiene que $Tr(A) = Tr(C^{-1}AC)$

Demostración:

(1) Tenemos que,

$$p_A(x) = \det \begin{bmatrix} a_{11} - x & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} - x \end{bmatrix}$$

Por la fórmula de determinante, se cumple que,

$$\det \begin{bmatrix} a_{11} - x & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} - x \end{bmatrix} = \sum_{P \in S_n} \prod_{i=1}^{n} \widetilde{a}_{iP(i)} \operatorname{sgn}(P)$$

donde $\widetilde{a}_{iP(i)} = a_{ii} - x$ si $P(i) = i$. Notemos que $P = \text{Id}$, entonces,

$$\sum_{P \in S_n} \prod_{i=1}^{n} \widetilde{a}_{iP(i)} \text{sgn}(P) = (a_{11} - x)(\cdots)(a_{nn} - x) + \sum_{P \in S_n \setminus \{\text{Id}\}} \prod_{i=1}^{n} \widetilde{a}_{iP(i)} \text{sgn}(P)$$

En la sumatoria de derecha el mayor término posible es un x^{n-2} por como está determinado las permutaciones, por tanto, los coeficientes de los términos x^n, x^{n-1} se determinan solo por el producto,

$$\begin{aligned}(a_{11} - x)(\cdots)(a_{nn} - x) &= (-x)^n + (-x)^{n-1}(a_{11} + \cdots + a_{nn}) + \cdots \\ &= (-x)^n + \text{Tr}(A)(-x)^{n-1} + \cdots\end{aligned}$$

Ahora para determinar el coeficiente libre, basta tomar $x = 0$, de forma que,

$$p_A(0) = \det(A)$$

Por tanto,

$$p_A(x) = (-x)^n + \text{Tr}(A)(-x)^{n-1} + \cdots + \det(A)$$

(2) Sea C invertible. Como sabemos que el polinomio característico es invariante sobre cambio de base, entonces se tiene que,

$$p_A(x) = p_{C^{-1}AC}(x)$$

Por el índice anterior se tiene que,

$$(-x)^n + \text{Tr}(A)(-x)^{n-1} + \cdots + \det(A) = (-x)^n + \text{Tr}(C^{-1}AC)(-x)^{n-1} + \cdots + \det(A)$$

En particular,

$$(-x)^{n-1}(\text{Tr}(A) - \text{Tr}(C^{-1}AC)) + \cdots + (\det(A) - \det(C^{-1}AC)) = 0$$

Si $\{1, x, \cdots, x^{n-1}\}$ es una base de los polinomios de grado menor o igual a $n-1$, se cumple que,

$$\text{Tr}(A) = \text{Tr}(C^{-1}AC)$$

Demostrando la proposición. ■

Observación: Sea V un esapcio vectorial definido en $\mathbb{C}$. Entonces el polinomio característico se puede descomponer de la siguiente forma,

$$p_A(x) = (\lambda_1 - x)(\cdots)(\lambda_n - x) = (-x)^n + (\lambda_1 + \cdots + \lambda_n)(-x)^{n-1} + \cdots + \lambda_1 \cdots \lambda_n$$

Por la proposición anterior se concluye que,

$$\begin{aligned}\text{Tr}(A) &= \lambda_1 + \cdots + \lambda_n \\ \det(A) &= \lambda_1 \cdots \lambda_n\end{aligned}$$

Notación: Sea $p(x) = a_0 + a_1x + \cdots + a_nx^n$ un polinomio. Sea A una matriz y sea $f : X \to X$ una función, entonces denotamos,

$$\begin{aligned} p(A) &= a_0I + a_1A + \cdots + a_nA^n \\ p(f)(x) &= a_0\mathrm{Id}(x) + a_1f(x) + \cdots + a_nf^n(x) \end{aligned}$$

Proposición: *Sea A matriz $n \times n$ con valor propio λ. Entonces, si $q(x)$ es un polinomio, entonces $q(A)v = q(\lambda)v$.*

Demostración: Sea $q(x) = a_0 + a_1x + \cdots + a_nx^n$, entonces,

$$q(A) = a_0 + a_1A + \cdots + a_nA^n$$

Luego,

$$\begin{aligned} q(A)v &= (a_0 + a_1A + \cdots + a_nA^n)v \\ &= a_0v_0 + a_1Av + \cdots + a_nA^nv \\ &= a_0v + a_1\lambda v + \cdots + a_n\lambda^nv \\ &= (a_0 + a_1\lambda + \cdots + a_n\lambda^n)v \\ &= q(\lambda)v \end{aligned}$$

Por tanto $q(A)v = q(\lambda)v$. ■

Proposición: *Sea $p(x)$ un polinomio y sea $f : V \to V$ un operador lineal. Si λ es valor propio con vector propio v con respecto a f, entonces,*

$$p(f)(v) = p(\lambda)v$$

Demostración: Sabemos que $f(v) = \lambda v$, entonces por definición,

$$\begin{aligned} p(f)(v) &= a_0v + a_1f(v) + \cdots + a_nf^n(v) \\ &= a_0v + a_1\lambda v + \cdots a_n\lambda^nv \\ &= p(\lambda)v \end{aligned}$$

■

Teorema (Cayley-Hamilton): *Sea A una matriz $n\times n$ y sea $p_A(x)$ su polinomio característico. Entonces $p_A(A) = 0$.*

Ejemplo: Sea A una matriz $n \times n$ con polinomio característico $(-x)^n$. Es decir, A tiene un único valor propio de multiplicidad algebraica n. Además, por el teorema de Cayley-Hamilton se cumple que $A^n = 0$.

Definición: *Decimos que una matriz A es **nilpotente** si existe $k \in \mathbb{N}$ tal que $A^k = 0$. Y decimos que un operador $f : V \to V$ es nilpotente si existe $k \in \mathbb{N}$ tal que $f^k = 0$.*

Observación: Sea A tal que $Av = \lambda v$ con v no nulo y tal que A es nilpotente, entonces para algún $k \in \mathbb{N}$ se cumple que,

$$0 = A^k v = \lambda^k v$$

Entonces $\lambda = 0$.

Ejemplo: Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ operador lineal con matriz asociada en la base canónica,

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

Es decir, A solo tiene 1 en una diagonal que no es la principal. Se cumple que,

$$\begin{aligned} f(e_n) &= e_{n-1} \\ f(e_{n-1}) &= e_{n-2} \\ &\vdots \\ f(e_2) &= e_1 \\ f(e_1) &= 0 \end{aligned}$$

Por lo tanto, sea $v \in \mathbb{R}^n$ tal que,

$$v = x_1 e_1 + \cdots + x_n e_n$$

Entonces,

$$\begin{aligned} f^n(v) &= x_1 f^n(e_1) + \cdots + x_n f^n(e_n) \\ &= 0 \end{aligned}$$

Es decir, f es nilpotenten, en particular, n es el mínimo natural tal que $f^n = 0$.

Ejemplo: Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ operador lineal con matriz asociada en la base canónica A una matriz triangular superior con diagonal de ceros. Entonces se cumple que,

$$\begin{aligned} f(e_n) &\in \langle e_1, \cdots, e_{n-1} \rangle \\ f(e_{n-1}) &\in \langle e_1, \cdots, e_{n-2} \rangle \\ &\vdots \\ f(e_2) &\in \langle e_1 \rangle \\ f(e_1) &= 0 \end{aligned}$$

Por lo tanto $f^n = 0$, siendo nilpotente.

Proposición: *Sea A una matriz. Entonces A es nilpotente si y sólo si $p_A(x) = (-x)^n$.*

Demostración: Sea A nilpotente. Si A es triangularizable, entonces podemos definir n valores propios, digamos que $\lambda_1, \cdots, \lambda_k$ son todos sus valores propios con $k \leq n$ con vectores propios $v_1, \cdots, v_k$. Como A es nilpotente, se tiene que,

$$0 = A^m v_i = \lambda_i^m v_i$$

Es decir, $\lambda_i = 0$ para todo $i = 1, \cdots, k$. Por lo tanto,

$$p_A(x) = (-x)^n$$

Supongamos ahora que $p_A(x) = (-x)^n$, es decir, A solo tiene un valor propio nulo de multiplicidad algebraica n. Sea C matriz invertible que triangulariza a A, por lo que la diagonal de la matriz $C^{-1}AC$ son los valores propios de A, es decir, tiene una diagonal solo de 0. Por tanto,

$$(C^{-1}AC)^n = 0 \Leftrightarrow C^{-1}A^nC = 0$$

Lo que implica que $A^n = 0$, de forma que A es nilpotente. ■

Observación: Sea $q(x)$ un polinomio. Entonces,

$$\begin{aligned} q(C^{-1}AC) &= a_0 I + a_1 C^{-1}AC + \cdots + a_n C^{-1}A^nC \\ &= a_0 C^{-1}C + a_1 C^{-1}AC + \cdots + a_n (C^{-1}AC)^n \\ &= C^{-1}(a_0 AC + a_1 AC + \cdots + a_n A^n C) \\ &= C^{-1}(a_0 I + a_1 A + \cdots + a_n A^n) \\ &= C^{-1}q(A)C \end{aligned}$$

Demostración Cayley-Hamilton: Sin pérdida de generalidad supongamos que A es una matriz trinagular superior, entonces su diagonal son sus valores propios, es decir,

$$A = \begin{bmatrix} \lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

donde $\lambda_1, \cdots, \lambda_n$ son los valores propios de A. Sea $p_A(x)$ su polinomio característico, entonces,

$$p_A(x) = (x - \lambda_1)(\cdots)(x - \lambda_n)$$

Luego,

$$p_A(A) = (A - \lambda_1 I)(\cdots)(A - \lambda_n I)$$

Observemos que,

$$A - \lambda_n I = \begin{bmatrix} \lambda_1 - \lambda_n & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

Entonces,

$$(A - \lambda_n I)e_n \in \langle e_1, \cdots, e_{n-1}\rangle$$
$$(A - \lambda_n I)e_i \in \langle e_1, \cdots, e_i\rangle, \quad i = 1, \cdots, n-1$$

De aquí podemos concluir que en general se cumple,

$$(A - \lambda_j I)e_j \in \langle e_1, \cdots, e_{j-1}\rangle$$
$$(A - \lambda_j I)e_i \in \langle e_1, \cdots, e_i\rangle, \quad i = 1, \cdots, n; \;\; i \neq j$$

para todo $j = 2, \cdots, n$. Con respecto a $j = 1$ se tiene,

$$(A - \lambda_1 I)e_1 = 0$$
$$(A - \lambda_1 I)e_i \in \langle e_1, \cdots, e_i\rangle, \quad i = 2, \cdots, n$$

Sea $V_i = \langle e_1, \cdots, e_i\rangle$ para todo $i = 1, \cdots, n$. Entonces,

$$V = V_n \overset{(A-\lambda_n I)}{\longrightarrow} V_{n-1} \overset{(A-\lambda_{n-1} I)}{\longrightarrow} V_{n-2} \longrightarrow \cdots \longrightarrow V_1 \overset{(A-\lambda_1 I)}{\longrightarrow} \{0\}$$

Por lo tanto,

$$p_A(A) = 0$$

Probando que toda matriz triangular A satisface que $p_A(A) = 0$.

Supongamos ahora que A es una matriz $n \times n$ cualquiera. Sabemos que es triangularizable, por lo que existe una matriz C invertible tal que $C^{-1}AC$ es una matriz triangular superior. Luego,

$$0 = p_{C^{-1}AC}(C^{-1}AC) = p_A(C^{-1}AC) = C^{-1}p_A(A)C$$

Por tanto $p_A(A) = 0$, probando el teorema. ■

Observación: Si $q(x)$ es un polinomio tal que $q(A) = 0$, entonces no necesariamente $q(x)$ es múltiplo de $p_A(x)$, por ejemplo, si p_A es de grado 2 y tomamos $A = I$, entonces $p_A(x) = (1-x)^2$ y $q(x) = 1 - x$ es tal que $q(A) = I - A = 0$ pero q no es múltiplo de p_A.

Definición: *Sea A una matriz $n \times n$. El polinomio de grado mínimo $q(x)$ tal que $q(A) = 0$, se llama* ***polinomio mínimo*** *de A.*

2.2. Forma canónica de Jordan

Toda matriz A es triangularizable. Ahora vamos a estuidar un tipo de matrices triangulares, las formas canónica de Jordan. Veremos que toda matriz tiene una única forma canónica de J y que este está relacionado con el polinomio minimal.

Definición: *Sea A un matríz $n \times n$. Sea λ un valor propio. Definimos el* ***espacio propio generalizado*** *de λ por el conjunto,*

$$G_\lambda(A) := \{v \in V : \exists k \in \mathbb{N}, (A - \lambda I)^k v = 0\}$$

Observación:

- Si E_λ es espacio propio de λ, entonces se cumple que,

$$E_\lambda \subseteq G_\lambda(A)$$

- G_λ es invariante bajo $(A-\lambda I)$, es decir,

$$(A-\lambda I)G_\lambda(A) \subseteq G_\lambda$$

- G_λ es un subespacio vectorial de V.
- La transformación $(A-\lambda I) : G_\lambda \to G_\lambda$, es nilpotente, para ver esto tomemos $\{v_1, \cdots, v_k\}$ base de G_λ, entonces para cada vector existe n_i tal que,

$$(A-\lambda I)^{n_i} v_i = 0$$

Tomando $n := \max\{n_1, \cdots, n_k\}$, se obtiene que,

$$(A-\lambda I)^n v = 0$$

para todo $v \in G_\lambda$, es decir, $(A-\lambda I)^n = 0$. Luego se tiene que $(A-\lambda I)$ tiene polinomio característico $(-x)^k$ donde $k = \dim G_\lambda$. Y esto implica que $A\big|_{G_\lambda}$ tiene polinomio característico $(\lambda - x)^k$, ya que λ tiene multiplicidad algebraica k.

Proposición (Descomposición del espacio generalizado): *Sea A un matriz $n \times n$ y sea λ un valor propio con $mg(\lambda) = h$. Entonces el espacio propio generalizado se puede escribir como,*

$$G_\lambda(A) = Z_A(w_1) \oplus \cdots \oplus Z_A(w_h)$$

donde $Z_A(w) := \{q(A)w : \ q(x) \text{ es polinomio con coeficientes complejos}\}$..

Ejemplo: Sea $w \in V$, sea $q(x) = x^3 + (1+i)x^2 - 2$, entonces,

$$q(A)w = (A^3 + (1+i)A^2 - 2I)w \in Z_A(w)$$

Observación: Sea A una matriz $n \times n$ y sea λ un valor propio de A. Sea $v \in G_\lambda$. Supongamos que $Av = \lambda v$, entonces,

$$(A-\lambda I)v = 0$$

Sea $q(x)$ un polinomio, entonces $q(A)v = q(\lambda)v$, es decir,

$$\{q(A)v : \ q \text{ es un polinomio complejo}\} = \langle v \rangle$$

O bien $Z_A(v) = \langle v \rangle$.

Supongamos que $(A-\lambda I)^2 v = 0$ y que $(A-\lambda I)v \neq 0$, entonces,

$$(A-\lambda I)^2 v = A^2 v - 2\lambda A v + \lambda^2 v = 0 \Leftrightarrow A^2 v = 2\lambda A v - \lambda^2 v$$

Por lo que $A^2v \in \langle Av, v\rangle$, también notemos que,

$$A^3v = 2\lambda A^2 - \lambda^2 Av = 3\lambda^2 Av - 2\lambda^3 v$$

En general se concluye que $A^kv \in \langle Av, v\rangle$ y por tanto $Z_A(v) = \langle Av, v\rangle$.

Afirmación: *Si v es tal que $(A-\lambda I)^n v = 0$ donde $(A-\lambda I)^{n-1}v \neq 0$, entonces $Z_A(v) = \langle A^{n-1}v, \cdots, Av, v\rangle$.*

Demostración: Si $(A-\lambda I)^{n-1}v \neq 0$, entonces $(A-\lambda I)^i v \neq 0$ para todo $1 \leq 1 \leq n-1$. Notemos que,

$$(A-\lambda I)^n v = \sum_{i=0}^{n} \binom{n}{i} A^i (-\lambda)^{n-i} v = 0 \Leftrightarrow A^n v = -\sum_{i=0}^{n-1} \binom{n}{i} A^i (-\lambda)^{n-i} v$$

Aquí es evidente que $A^m v \in \langle A^{n-1}v, \cdots, Av, v\rangle$ para todo $m \geq n$. Por lo tanto $Z_A(v) = \langle A^{n-1}v, \cdots, Av, v\rangle$.

■

Afirmación: *Si v es tal que $(A-\lambda I)^n v = 0$ donde $(A-\lambda I)^{n-1}v \neq 0$. Entonces,*

$$Z_A(v) \cap E_\lambda = \langle (A-\lambda I)^{n-1} v\rangle$$

Demostración: Si $(A-\lambda I)^n v = 0$, entonces,

$$(A-\lambda I)(A-\lambda I)^{n-1} v = 0 \Leftrightarrow A(A-\lambda I)^{n-1}v = \lambda (A-\lambda I)^{n-1} v$$

Es decir, $(A-\lambda I)^{n-1}v \in E_\lambda$. Además, si $(A-\lambda I)^{n-1}$ es un polinomio formado por el generador $\{v, Av, \cdots, A^{n-1}v\}$, entonces $(A-\lambda I)^{n-1} \in Z_A(v)$. Por tanto $\langle (A-\lambda I)^{n-1}v\rangle \subseteq Z_A(v) \cap E_\lambda$.

Probemos la igualdad. Sea $w \in Z_A(v) \cap E_\lambda$, esto implica que,

$$\begin{aligned} w = p(A)v &= \alpha_0 v + \cdots + \alpha_{n-1}A^{n-1}v \\ (A-\lambda I)w &= 0 \end{aligned}$$

Observemos que $p(x)$ es un polinomio de grado a lo más $k-1$. Consideremos $B := A - \lambda I$ y sea $q(B) := p(B+\lambda I)$ otro polinomio de a lo más $k-1$ grado, de esta forma obtenemos,

$$\begin{aligned} Bq(B)v &= 0 \\ B^k v &= 0 \end{aligned}$$

Supongamos que $q(B) = \beta_0 I + \cdots + \beta_{n-1}B^{n-1}$, luego,

$$Bq(B)v = \beta_0 Bv + \cdots + \beta_{n-2}B^{n-1}v + \underbrace{\beta_{b-1}B^n v}_{=0} = \beta_0 Bv + \cdots + \beta_{n-2}B^{n-1}v$$

Aplicando nuevamente B, obtenemos,

$$B^2 q(B)v = \beta_0 B^2 v + \cdots + \beta_{n-3}B^{n-1}v$$

De forma general, multiplicando B^{n-1} a $q(B)v$, se obtiene,

$$B^{n-1}q(B)v = \beta_0 B^{n-1}v = 0$$

Si $B^{n-1}v \neq 0$, se obtiene que $\beta_0 = 0$. Luego,

$$q(B) = \beta_1 B + \cdots + \beta_{n-1}B^{n-1}v$$

Multiplicando por B^{n-2} a $q(B)v$, se obtiene,

$$B^{n-2}q(B)v = \beta_1 B^{n-1}v = 0$$

Luego $\beta_1 = 0$. De esta forma podemos concluir que $\beta_0 = \beta_1 = \cdots = \beta_{n-2} = 0$. Obteniendo que,

$$q(B) = \beta_{n-1}B^{n-1}$$

Por lo tanto,

$$\begin{aligned} w &= p(A)v = p(B + \lambda I) \\ &= q(B)v = \beta_{n-1}B^{n-1}v \\ &= \beta_{n-1}(A - \lambda I)^{n-1} \in \langle (A - \lambda I)^{n-1}v \rangle \end{aligned}$$

Por lo tanto,

$$Z_A(v) \cap E_\lambda = \langle (A - \lambda I)^{n-1}v \rangle$$

Demostrando la afirmación ■

Demostración (Descomposición del espacio generado): (Revisar en la segunda vuelta) Demostraremos la proposición paso por paso.

- Sea $v \in E_\lambda$ no nulo, tal que $(A - \lambda I)v = 0$ (v existe al tomar un espacio definido sobre los complejos). Definimos $v_1 := v$.
- Sea k el mayor entero tal que el sistema lineal,

  $$(A - \lambda I)^{k-1}w = v_1$$

  Tenga solución, es decir, existe un vector w tal que soluciona el sistema lineal anterior. Definimos $v_k := w$ y definimos,

  $$v_i := (A - \lambda I)^{k-i}v_k$$

  para todo $i = 2, \cdots, k-1$. En particular, $(A - \lambda I)^k v_k = 0$ y $(A - \lambda I)^{k-1}v_k \neq 0$.

  Afirmación: *Los vectores $v_1, \cdots, v_k$ son independientes.*

  Demostración: Sean $\alpha_1, \cdots, \alpha_k$ tales que,

  $$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$$

Aplicando $(A-\lambda I)^{k-1}$, se obtiene,

$$\underbrace{\alpha_1(A-\lambda I)^{k-1}v_1+\cdots}_{=0}+\alpha_k\underbrace{(A-\lambda I)^{k-1}v_k}_{=v_1}=\alpha_k v_1=0$$

Si $v_1\neq 0$, entonces $\alpha_k=0$. Obteniendo la combinación lineal,

$$\alpha_1 v_1+\cdots+\alpha_{k-1}v_{k-1}=0$$

Aplicando el método anterior para cada coeficiente de cada vector, concluimos que $\alpha_1=\cdots=\alpha_k=0$. Por tanto, $v_1,\cdots,v_k$ son vectores independientes. ■

La afirmación anterior implica que $k\leq n$, puesto que si $k>n$, entonces la colección $\{v_1,\cdots,v_k\}$ tiene más elementos que la base de V, siendo contradicción, puesto que $v_i\in V$ para todo $i=1,\cdots,k$.

- Si $(A-\lambda I)^k v_k=0$ y $(A-\lambda I)^{k-1}v_k=v\neq 0$, entonces se cumple que,

$$Z_A(v_k)=\langle v_k,Av_k,\cdots,A^{k-1}v_k\rangle$$

- Sea $w\in E_\lambda$ un vector independiente de v. Entonces w no es un elemento de $Z_A(v_k)$. Supongamos que lo es, es decir, $w\in Z_A(v_k)\cap E_\lambda$. Como hemos visto, se tiene que,

$$Z_A(v_k)\cap E_\lambda=\langle(A-\lambda I)^{k-1}v_k\rangle$$

Esto implica que existe una constante α tal que,

$$w=\alpha(A-\lambda I)^{k-1}v_k=\alpha v$$

Sin embargo obtenemos que w,v son dependientes, siendo contradicción. Por tanto $w\notin Z_A(v_k)$.

- Sean $v,w\in E_\lambda$. Sean $k_1,k_2\leq n$ los enteros mayores tales que los sistemas,

$$(A-\lambda I)^{k_1-1}x_1=v$$
$$(A-\lambda I)^{k_2-1}x_2=w$$

Afirmación: *$Z_A(v_{k_1}),Z_A(w_{k_2})$ están en suma directa.*

Demostración: Sabemos que,

$$Z_A(v_{k_1})=\langle v_{k_1},Av_{k_1},\cdots,A^{k_1-1}v_{k_1}\rangle$$
$$Z_A(w_{k_2})=\langle w_{k_2},Aw_{k_2},\cdots,A^{k_2-1}w_{k_2}\rangle$$

Probemos que $Z_A(v_{k_1})\cap Z_A(w_{k_2})=\{0\}$. Para ello veremos que,

$$A^i w_{k_2}\notin Z_A(v_{k_1})$$

para todo $i = 0, \cdots, k_2 - 1$. Estudiemos el caso cuando $i = 0$, es decir, veamos que $w_{k_2} \notin Z_A(v_{k_1})$. Supongamos que $w_{k_2} \in Z_A(v_{k_1})$, luego se cumple que,

$$w_{k_2} = \alpha_1 v_{k_1} + \cdots + \alpha_{k_1} A^{k_1 - 1} v_{k_1}$$

Aplicando $(A - \lambda I)^{k_2 - 1}$, obtenemos que,

$$\begin{aligned} w &= (A - \lambda I)^{k_2 - 1} w_{k_2} \\ &= \alpha_1 (A - \lambda I)^{k_2 - 1} v_{k_1} + \cdots + \alpha_{k_1} (A - \lambda I)^{k_2 - 1} A^{k_1 - 1} v_{k_1} \end{aligned}$$

Se cumple que $(A - \lambda I)^{k_2 - 1} A^i v_{k_1} \in Z_A(v_{k_1})$ para todo $i = 0, \cdots, k_1 - 1$ dado que $(A - \lambda I)^{k_1} v_{k_1} = 0$, por lo que las potencias $A^m \in \langle v_{k_1}, \cdots, A^{k_1 - 1} v_{k_1} \rangle$ para todo $m \geq k_1$. Lo importante es que,

$$w \in Z_A(v_{k_1})$$

Sin embargo por el punto anterior es imposible que esto ocurra, por lo tanto $w_{k_2} \notin Z_A(v_{k_1})$. Consideremos $A^i w_{k_2}$ con $i = 1, \cdots, k_2 - 1$. Supongamos que $A^i w_{k_2} \in Z_A(v_{k_1})$ y que $A^s w_{k_2} \notin Z_A(v_{k_2})$ para todo $s = 0, \cdots, i - 1$, entonces se tiene,

$$A^i w_{k_2} = \alpha_1 v_{k_1} + \cdots + \alpha_{k_1} A^{k_1 - 1} v_{k_1}$$

Notemos que,

$$A^{i+1} w_{k_2} = \alpha_1 A v_{k_1} + \cdots + \alpha_{k_1} A^{k_1} v_{k_1} \in Z_A(v_{k_1})$$

como ya hemos visto. Luego se obtiene que $A^j w_{k_2} \in Z_A(v_{k_1})$ para todo $i \leq j \leq k_2$. De esta forma,

$$\begin{aligned} (A - \lambda I)^{k_2} w_{k_2} = 0 &\Leftrightarrow \sum_{l=0}^{k_2} \binom{k_2}{l} A^l (-\lambda)^{k_2 - l} w_{k_2} = 0 \\ &\Leftrightarrow \sum_{l=i}^{k_2} \binom{k_2}{i} A^i (-\lambda)^{k_2 - i} w_{k_2} + \sum_{l=0}^{i-1} \binom{k_2}{l} A^l (-\lambda)^{k_2 - l} w_{k_2} = 0 \\ &\Leftrightarrow \underbrace{\sum_{l=i}^{k_2} \binom{k_2}{i} A^i (-\lambda)^{k_2 - i} w_{k_2}} \end{aligned}$$

- Supongamos que $\{v_1, \cdots, v_h\}$ es una base de E_λ, luego los espacios $Z_A(v_{k_i})$ están en suma directa para todo $i = 1, \cdots, h$. Sea,

  $$E := Z_A(v_{k_1}) \oplus \cdots \oplus Z_A(v_{k_h})$$

  Queremos proba que $E = G_\lambda(A)$. Si $x \in G_\lambda(A)$, entonces existe un l natural tal que $(A - \lambda I)^l x = 0$, es decir, $(A - \lambda I)^{l-1} x \in E_\lambda$, luego existen $\alpha_1, \cdots, \alpha_h$ tales que,

  $$(A - \lambda I)^{l-1} x = \alpha_1 v_1 + \cdots + \alpha_h v_h$$

  Si sabemos que $(A - \lambda I)^{k_i - 1} v_{k_i} = v_i$, entonces obtenemos que,

  $$(A - \lambda I)^{l-1} x = \alpha_1 (A - \lambda I)^{k_1 - 1} v_{k_1} + \cdots + \alpha_h (A - \lambda I)^{k_h - 1} v_{k_h}$$

Proposición: *Sea $A : V \to V$ un operador lineal con vectores propios distintos $\lambda_1, \cdots, \lambda_k$. Entonces,*

$$V = G_{\lambda_1}(A) \oplus \cdots \oplus G_{\lambda_k}(A)$$

Proposición: *Sea A una matriz $n \times n$ y sea λ un valor propio, sea $v \in V$ y $k \in \mathbb{N}$ tal que $(A - \lambda I)^k v = 0$ y $(A - \lambda I)^{k-1} v \neq 0$. Entonces,*

(1) $Z_A(v)$ es A-invariante.

(2) La matriz $A\big|_{Z_A(v)}$ respecto a la base $\{(A - \lambda I)^{k-1} v, \cdots, (A - \lambda I)v, v\}$ es de la forma,

$$\begin{bmatrix} \lambda & 1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

*A esta matriz le llamamos **bloque de Jordan**.*

Demostración:

(1) Sea $w \in Z_A(v)$, entonces existe un polinomio complejo q tal que $w = q(A)v$, luego,

$$Aw = Aq(A)v = Aq(\lambda)v$$

Tomando $p(x) = xq(\lambda)$ otro polinomio complejo, se obtiene que $Aw = p(A)v$, es decir, $A(Z_A(v)) \subseteq Z_A(v)$.

(2) Por el item anterior tenemos que la matriz,

$$A : Z_A(v) \to Z_A(v)$$

Está bien definida. Consideremos $B := A\big|_{Z_A(v)}$. Sea $v_i := (A - \lambda I)^{k-i} v$ donde $i = 1, \cdots, k$. Observemos que $v_i \in Z_A(v)$ para todo $i = 1, \cdots, k$. Notemos que,

$$v_{k-1} = (A - \lambda I)v_k = Av_k - \lambda v_k \Leftrightarrow Av_k = v_{k-1} - \lambda v_k$$

Restringiendo a $Z_A(v)$, obtenemos que,

$$Bv_k = v_{k-1} - \lambda v_k$$

Entonces, las coordenadas de Bv_k en la base es,

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \lambda \end{bmatrix}$$

Con v_{k-2} se cumple que,

$$v_{k-2} = (A - \lambda I)v_{k-1} = Av_{k-1} - \lambda v_{k-1} \Leftrightarrow Av_{k-1} = \lambda v_{k-1} + v_{k-2}$$

Restrigiendo obtenemos que Bv_{k-1} tiene coordenadas,

$$\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \lambda \\ 0 \end{bmatrix}$$

Realizamos para todo $2 \leq i \leq k$, obteniendo todas las columnas de B menos la primera. Para la primera notemos que $(A - \lambda I)v_1 = 0$, entonces,

$$Av_1 = \lambda v_1$$

Y por tanto Bv_1 tiene coordenadas,

$$\begin{bmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Finalmente B en la base es de la forma,

$$\begin{bmatrix} \lambda & 1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

Como se quería probar.

Demostrando la proposición. ■

Ejemplo: Las siguientes matrices son bloques de Jordan:

$$\begin{bmatrix} 5 & 1 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Sin embargo, las siguientes matrices no son bloques de Jordan:

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

Sea $A : V \to V$ un operador lineal con $\lambda_1, \cdots, \lambda_k$ valores propios distintos. Entonces se cumple que,

$$V = G_{\lambda_1}(A) \oplus \cdots \oplus G_{\lambda_k}(A)$$

Si cada término se puede descomponer de la siguiente forma:

$$G_{\lambda_i}(A) = Z_A(v_i^{(1)}) \oplus \cdots \oplus Z_A(v_i^{(\mathrm{mg}(\lambda_i))})$$

Obtenemos que,

$$V = Z_A(v_1^{(1)}) \oplus \cdots \oplus Z_A(v_k^{\mathrm{mg}(\lambda_k)})$$

Entonces para todo $i = 1, \cdots, k$ y $j = 1, \cdots, \mathrm{mg}(\lambda_i)$, se tiene que,

$$A\big|_{Z_A(v_i^{(j)})} : Z_A(v_i^{(j)}) \to Z_A(v_i^{(j)})$$

Es un bloque de Jordan. Es decir, existe una base de V tal que A tiene forma compuesta por bloques de Jordan de distintos tipos. El siguiente teorema resume todo.

Teorema (Jordan): *Sea $f : V \to V$ es un operador lineal donde V es un espacio vectorial de dimensión finita, entonces hay una base de V respecto a la cual f tiene matriz "diagonal por bloques" formada por bloques de Jordan.*

Demostración: Bastar usar la descomposición de V y de G_{λ_i}. Luego es evidente que existe una base tal que f tiene matriz con diagona de bloques donde cada elemento es un bloque de Jordan. ■

Teorema (Jordan II): *La forma de Jordan está determinada por la descomposición,*

$$V = G_{\lambda_1}(A) \oplus \cdots \oplus G_{\lambda_k}(A)$$

y por la descomposición,

$$G_{\lambda_j}(A) = Z_A(v_j^1) \oplus \cdots \oplus Z_A(v_j^{l_j})$$

donde $j = 1, \cdots, k$. Entonces es únicamente determinada módulo permutación de los bloques de Jordan.

Los dos teorema de Jordan nos dicen que toda matriz se puede expresar como una diagonal por bloques conformado por bloques de Jordan y que esta matriz es única salvo cambio en el orden de los bloques de Jordan.

Definición: *Sea $f : V \to V$ un operador lineal con V espacio vectorial de dimensión finita. Sea J la matriz asociada a f formada por la descomposición de V y de los subespacios generalizados. A J le llamamos forma canónica de Jordan del operador f.*

Por último presentamos el teorema de Jordan en el caso real, el cual no demostraremos.

Teorema (Jordan caso real): *Sea A una matriz con coeficientes reales. Sean $\lambda_1, \cdots, \lambda_m$ los valores propios reales de A, y sea $\mu_1, \overline{\mu}_1, \cdots, \mu_k, \overline{\mu}_k$ los valores propios complejos de A. Donde,*

$$\mu_j = a_j + ib_j$$
$$\overline{\mu}_j = a_j - ib_j$$

donde $b_j \neq 0$ para todo $j = 1, \cdots, k$.

(1) Los bloques asociados a $\overline{\mu}_j$ son los conjugados de aquellos asociados a μ_j.

(2) Si los bloques,

$$\begin{bmatrix} \mu_j & 1 & \cdots & 0 & 0 \\ 0 & \mu_j & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mu_j & 1 \\ 0 & 0 & \cdots & 0 & \mu_j \end{bmatrix}, \begin{bmatrix} \overline{\mu}_j & 1 & \cdots & 0 & 0 \\ 0 & \overline{\mu}_j & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \overline{\mu}_j & 1 \\ 0 & 0 & \cdots & 0 & \overline{\mu}_j \end{bmatrix}$$

Están en pareja como definimos en (1). Definimos,

$$M_j := \begin{bmatrix} a_j & -b_j \\ b_j & a_j \end{bmatrix}$$

Luego existe una matriz invertible C con coeficientes reales tal que $C^{-1}AC$ tiene los siguientes bloques diagonales:

(I) Bloques correspondientes a los $\lambda_1, \cdots, \lambda_m$ iguales a la forma de Jordan compleja.

(II) Cada pareja de bloques se reemplaza por un bloque de dimensión doble:

$$\begin{bmatrix} M_j & I & \cdots & 0 & 0 \\ 0 & M_j & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & M_j & I \\ 0 & 0 & \cdots & 0 & M_j \end{bmatrix}$$

donde tomamos I la matriz identidad 2×2.

2.3. Ejemplos y Aplicaciones

Las matrices diagonales son fáciles de determinar: Sea A una matriz diagonalizable, entonces existe una matriz invertible C tal que $C^{-1}AC$ es diagonal. Supongamos que,

$$C^{-1}AC = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

Entonces,

$$(C^{-1}AC)^k = C^{-1}A^nC = \begin{bmatrix} \lambda_1^k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^k \end{bmatrix}$$

Por lo que si conocemos A, podemos deducir C y por tanto determinar A^k, el cual es de la forma,

$$A^k = C \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} C^{-1}$$

Es decir, si A es diagonalizable, entonces es fácil de calcular las potencias de A.

Número de Fibonacci: Definimos la secuencia de Fibonacci por,

$$F_{n+2} = F_{n+1} + F_n$$

para todo $n \geq 0$ y donde $F_1 = F_0 = 1$. Notemos que,

$$\begin{bmatrix} F_{n+2} \\ F_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix}$$

Estudiemos la matriz,

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

Notemos que,

$$p_A(x) = x^2 - x - 1$$

Entonces los valores propios son,

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

Notemos que,

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_\pm \\ 1 \end{bmatrix} = \begin{bmatrix} \lambda_\pm + 1 \\ \lambda_\pm \end{bmatrix} = \begin{bmatrix} \lambda_\pm^2 \\ \lambda_\pm \end{bmatrix} = \lambda_\pm \begin{bmatrix} \lambda_\pm \\ 1 \end{bmatrix}$$

Luego los vectores propios forman una base en $\mathbb{R}^2$. Por lo tanto A es diagonalizable. En particular, existe C invertible tal que,

$$C^{-1}AC = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}$$

En particular,

$$C = \begin{bmatrix} \lambda_+ & \lambda_- \\ 1 & 1 \end{bmatrix}$$

Entonces,

$$\begin{aligned}
\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^2 \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} \\
&\vdots \\
&= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} \\
&= C \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{bmatrix} C^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}
\end{aligned}$$

Si,

$$C^{-1} = \frac{(\mathrm{Cof}(C))^T}{\det(A)} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ -1 & \frac{-1+\sqrt{5}}{2} \end{bmatrix}$$

Entonces,

$$\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{n+2} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+2} \\ \left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} \end{bmatrix}$$

Por lo tanto, para todo $n \geq 0$ se cumple que,

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} \right)$$

Determinando una fórmula para la secuencia de Fibonacci.

Gráfos: Un grafo consiste en vértices y aristas, donde los vértices se conectan mediante aristas, por ejemplo,

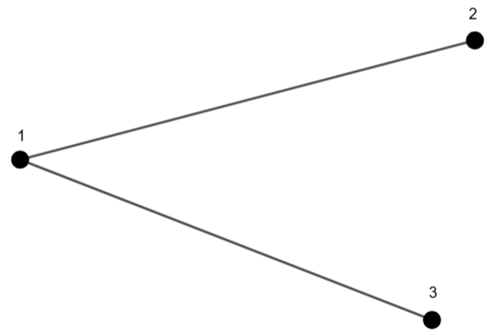


Figura 1: Grafo de tres vértices y dos aristas.

De forma formal, un grado es un conjunto $G = (V, L)$, donde V es el conjunto de vértices y L de aristas.

Consideremos nueve puntos en el espacio como muestra la figura de abajo,

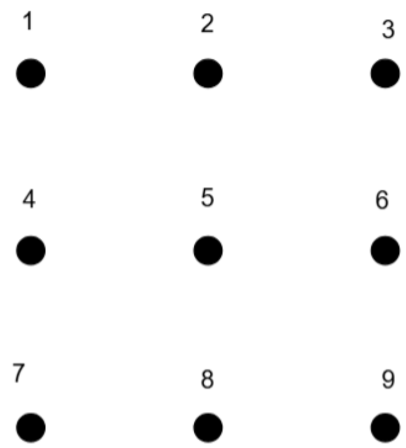


Figura 2: Grafo con 9 puntos

Queremos conectar al menos 5 puntos de solo con sus vecinos.

Sea $G = (V, L)$ un grado donde $|V| = n$, es decir, G tiene n vértices. Definimos la matriz de incidencia del grafo por la matriz A_G de $n \times n$ por,

$$(A_G)_{ij} = \begin{cases} 0, & \text{si no hay una arista entre } i, j \\ 1, & \text{si hay una arista entre } i, j \end{cases}$$

Ejemplo: Consideremos el grafo $G = (V, L)$ de la figura 1, por lo que $V = \{1, 2, 3\}$ y $L = \{(1, 2), \}$. Luego la matriz de incidencia del grafo es:

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

Notemos que $(A_G)_{12} = 1$, significa que hay una sola forma de llegar a 2 partiendo el punto 1 moviendonos un solo paso, este es el camino $1 \to 2$.

Supongamos que estamos en el punto 1 del grafo G.

$$v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Luego $A_G v$ representa las posiciones donde puedo acabar al iniciar en 1, de hecho,

$$A_G v = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Es decir, si partimos en 1, al avanzar un paso, terminamos en 2 o en 3. Por otro lado al estudiar la matriz A_G^2 obtenemos la matriz,

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

Notemos que $(A_G^2)_{11} = 2$, esto se puede interpretar que para llegar de 1 a 1 usando dos pasos, hay dos formas, la primera es $1 \to 2 \to 1$, y la otra es $1 \to 3 \to 1$. Si multiplicamos por v obtenemos,

$$A_G^2 v = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

Es decir, si partimos en 1, entonces al movernos dos pasos vamos a terminar en 1, y que $(A_G^2)v = 2$ nos dice que hay dos formas de hacer.

Volvamos al ejemplo principal. Si juntamos todos los puntos usandos las reglas, obtenemos la siguiente matriz,

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

Observemos que se cumple lo siguiente:

$$A^5 \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + \text{diag}A^5 = 2 \cdot \text{número de caminos de 5 pasos}$$

Revidsar.

Queremos diagonalizar a mano, pero resulta complicado, por lo que usamos computadores para obtener este tipos de resultados.

Hecho: Sea A una matriz de $n \times n$, sea C matriz $n \times n$ invertible y sea D matriz diagonal de $n \times n$.

- La A matriz se diagonaliza en a lo mpas $c_1 n^3$ pasos.
- La matriz AB se calcula en a lo más $c_1 n^3$.
- La matriz D^k se determina en kn pasos

Por lo tanto, obtenemos que,

$$A \xrightarrow{c_1 n^3} C^{-1}DC \xrightarrow{kn} D^k \xrightarrow{2c_1 n^3} C^{-1}D^kC$$

Es decir, desde A hasta $C^{-1}D^kC$, se realizó $3c_1n^3 + kn$. Por otro lado, los pasos desde A hasta A^k son de a lo más a c_2kn^3. Nos preguntamos si hay una relación en los pasos, es decir, algo del estilo,

$$3c_1n^3 + kn \overset{?}{\leftrightarrow} c_2kn^3$$

revidasr

- Si k es muy grande, entonces $n < cn^3$.
- Si k es pequeño, entonces no es tan claro.

En general, se mira más en detalle la estructura de las matrices.

La respuesta al problema original es que el número de combinaciones posibles de largo 5 es 12416.

Espacios de polinomios: Definimos $\mathcal{P}_d$ el espacio de los polinomios de grado a lo más d. Sabemos que $\{1, x, \cdots, x^d\}$ es una base, por lo que $\dim \mathcal{P}_d = d + 1$ y que todo polinomio tiene expresión única, es decir, dado $p(x) \in \mathcal{P}_d$, entonces existen únicos $c_0, c_1, \cdots, c_d$ tales que,

$$p(x) = c_0 + c_1x + \cdots + c_dx^d$$

Observación: Se cumple que,

$$\mathcal{P}_d = \{p(x) \in \mathcal{P}_{d+1} : \text{coeficiente } c_{d+1} = 0\}$$

Operador derivada: Definimos el operador $D : \mathcal{P}_d \to \mathcal{P}_d$ tal que $D(p(x)) = p'(x)$. Por ejemplo, si $p(x) = c_0 + c_1 x + \cdots + c_d x^d$, entonces,

$$D(p(x)) = c_1 + 2c_2 x + \cdots + dc_d x^{d-1}$$

Notemos que D es un operador lineal, entonces se le puede asociar a una matriz, es más, D es nilpotente puesto que,

$$D^{d+1}(p(x)) = 0$$

para todo $p(x) \in \mathcal{P}_d$. Notemos que,

$$D(x^i) = \begin{cases} 0, & i = 0 \\ ix^i, & i = 1, \cdots, d \end{cases}$$

Entonces la matriz derivada es,

$$D = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & d-1 & 0 \\ 0 & 0 & 0 & \ddots & 0 & d \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Una matriz $(d+1) \times (d+1)$.

Operador Translación: Sea a una constante. Sea $p(x)$ un polinomio, definimos el operador transación por a $T_a(p(x)) := p(x-a)$.

Ejemplo: Sea $\mathcal{P}_3 = \langle 1, x, x^2, x^3 \rangle$, entonces,

$$\begin{aligned} 1 &\overset{T_a}{\mapsto} 1 \\ x &\overset{T_a}{\mapsto} x - a \\ x^2 &\overset{T_a}{\mapsto} (x-a)^2 \\ x^3 &\overset{T_a}{\mapsto} (x-a)^3 \end{aligned}$$

Luego la matriz de translación en la base $\{1, x, x^2, x^3\}$ es de la forma,

$$A = \begin{bmatrix} 1 & -a & a^2 & a^3 \\ 0 & 1 & -2a & 3a^2 \\ 0 & 0 & 1 & -3a \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Ejemplo: Consideremos una matriz con polinomo característico $p_A(x) = x^2(1-x)^3(2-x)$. Entonces **terminar**

Ejemplo: Sea A una matriz con polinomio característico $p_A(x) = (1-x)^6$. Entonces A es una matriz de 6×6. Supongamos que el polinomio minimal es $p_A^{\text{mín}}(x) = 1 - x$, esto implica que,

$$p_A^{\text{mín}}(A) = I - A = 0$$

Es decir, $A = I$.

Supongamos ahora que el polinomio minimal es $p_A^{\text{mín}}(x) = (1-x)^2$. Claramente $A \neq I$, pero sabemos que la forma triangular de A tiene por diagonal solo valores 1. Sabemos que toda matriz tiene una forma canónica de Jordan, por lo tanto queremos saber la forma de fordan.

Observemos que $G_1(A) = V$ puesto que $(A-I)^2 = 0$. Por otro lado, los bloques de Jordan son de dimensión ≤ 2, porque las bases son de la forma a lo más $\{(A-I)^v, v\}$ para algún vector $v \in V$ y esto define el tamaño de los bloques de Jordan. Como candidatos tenemos,

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

La primera forma tiene **terminar**

Ejemplo: Consideremos la matriz,

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

Queremos triangularizar A. Determinemos el polinomio característico,

$$\begin{aligned} \det(A - xI) &= \det \begin{bmatrix} 1-x & 1 & 0 & 0 & 0 \\ 0 & 1-x & 0 & 0 & 0 \\ 0 & 0 & 1-x & 1 & 0 \\ 0 & 0 & -1 & 1-x & 1 \\ 1 & 0 & 0 & 1 & 1-x \end{bmatrix} \\ &= (1-x) \det \begin{bmatrix} 1-x & 0 & 0 & 0 \\ 0 & 1-x & 1 & 0 \\ 0 & -1 & 1-x & 1 \\ 0 & 0 & 1 & 1-x \end{bmatrix} \\ &= (1-x)^2 \det \begin{bmatrix} 1-x & 1 & 0 \\ -1 & 1-x & 1 \\ 0 & 1 & 1-x \end{bmatrix} \\ &= (1-x)^2 \left((1-x)\left((1-x)^2 - 1\right) + (1-x)\right) \\ &= (1-x)^3(-x(2-x)+1) \\ &= (1-x)^3(x^2 - 2x + 1) = (1-x)^5 \end{aligned}$$

Por tanto el polinomio característico de A es $p_A(x) = (1-x)^5$. Esto implica que su forma triangular tiene por diagonal los valores 1.

Ejemplo: Sea $f : V \to V$ un operador lineal donde V es de dimensión finita sobre $\mathbb{R}$. Supongamos que $f^4 = f^2$.

Afirmación: f *es triangularizable.*

Demostración: Supongamos que λ un valor propio complejo, por lo que existe un vector no nulo v tal que $f(v) = \lambda v$, luego por hipótesis se cumple que,

$$f^4(v) = f^2(v) \Leftrightarrow \lambda^4 v = \lambda^2 v$$

Es decir $\lambda^4 = \lambda^2$, por tanto todo valor propio de f cumple,

$$\lambda^2(\lambda+1)(\lambda-1=$$

Por tanto, lo únicos valores propios de λ son $0, 1, -1$, que son reales. Por tanto se tiene que f es triangularizable

terminar

Ejemplo: Consideremos la matriz,

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

Probemos que es nilpotente. Tenemos que,

$$\det(A - xI) = (-x)\det(M_{11}) - \det(M_{12})$$

Observemos que,

$$\det(M_{11}) = \det \begin{bmatrix} -x & 1 & 0 & 0 \\ -1 & -x & 0 & 0 \\ 0 & 0 & -x & 1 \\ 0 & 0 & 0 & -x \end{bmatrix} = \det \begin{bmatrix} -x & 1 \\ -1 & -x \end{bmatrix} \det \begin{bmatrix} -x & 1 \\ 0 & -x \end{bmatrix} = (x^2+1)(-x)^2$$

Y que,

$$\det(M_{12}) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -x & 0 & 0 \\ 0 & 0 & -x & 1 \\ 1 & 0 & 0 & -x \end{bmatrix} = \begin{bmatrix} -x & 0 & 0 \\ 0 & -x & 1 \\ 0 & 0 & -x \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & -x & 1 \\ 1 & 0 & -x \end{bmatrix} = (-x)^3$$

Por tanto,

$$\det(A - xI) = (-x)^3(x^2+1) - (-x)^3 = (-x)^5$$

Es decir, A es nilpotente ya que todos sus valores propios son 0 y por tanto, su forma triangular tiene diagonal nulo. Probemos que $I + A$ es invertibe. Veremos dos caminos para ver esto.

- **Primer camino:** Sea C invertible tal que $C^{-1}AC$ es matriz triangular con diagonal los valores propios de A, entonces,

$$C^{1}(I + A)C = I + C^{-1}AC$$

es una matriz triangular con diagonal de solo 1's, esto implica que,

$$\det(I + A) = \det(I + C^{-1}AC) = 1$$

Es decir, $I + A$ es invertible.

- **Segundo camino:** Consideremos el sistema lineal $(I + A)v = 0$. Supongamos que v es solución no trivial, entonces se cumple que,

$$Av = -v$$

Es decir, -1 es valor propio de A, sin embargo sabemos que esto es imposible, por lo tanto la única solución es que $v = 0$ y por tanto, $I + A$ es invertible.

Ejemplo (Flujo discreto): Sea $G = (V, L)$ un grafo donde $V = \{1, \cdots, 10\}$. Supongamos que está orientado en forma de árbol con raíz en 1.

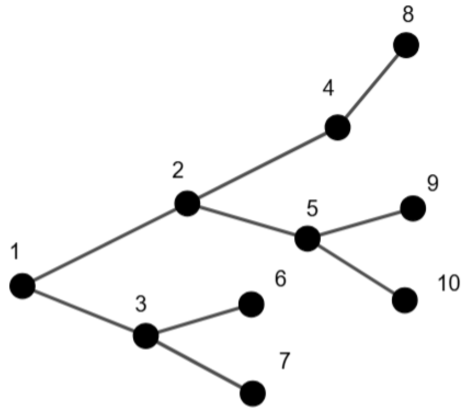


Figura 3: Ejemplo de grafo orientado en forma de árbol.

Hay un flujo a lo largo del árbol. Si en el vértice i hay una cantidad x_i y este se divide en k vértices siguientes, entonces al dar un paso en el árbol, el fluido x_i se reparte en los siguientes vértices. Si i no tiene sucesor, el fluido desaparece en el paso siguiente.

Sea A la matriz que representar el grafo, notemos que **termunar**

2.4. Matrices Positivas

Motivación: Google quiere ordenar las páginas web según su importancia. Para ello usaremos grafos.

Consideremos un grafo $G = (V, L)$, donde los vértices representa las páginas web y las aristas son flechas con sentido $a \to b$, que significa que hay enlaces en a que envían a b.

Definimos la probabilidad de llegar a a, a partir de b por P_{ab}. Entonces se cumple:

(1) Para todo $b \in V$ se tiene que,

$$\sum_{a \in V} P_{ab} = 1$$

(2) Si no hay camino de b hacía a, entonces $P_{ab} = 0$.

(3) Para todo $a, b \in V$ se cumple que P_{ab}.

Ahora supongamos que los usuarios parten con una distribución inicial, y vamos observando como se reparten, moviendose por apsos, según sus probabilidad $\{P_{ab}\}_{a,b}$.

Sea $P_a^{(0)}$ la probabilidad de encontrar el usuario en a en el primer paso. Luego definimos $P_a^{(k+1)}$ la probabilidad de encontrar el usuario en a en el paso $k+1$. Notemos que,

$$\begin{aligned}
P_a^{(k+1)} &= \sum_{b\in V} \text{Probabilidad de encontrarse en } V \text{ en el paso } k \text{ y que } b \to a. \\
&= \sum_{b\in V} \text{Probabilidad de encontrarse en } V \text{ en el paso } k\text{probabilidad de que } b \to a. \\
&= \sum_{b\in V} P_{ab}P_b^{(k)}
\end{aligned}$$

Supongamos que $V = \{1, \cdots, n\}$. Luego se tiene que,

$$\begin{aligned}
P^{(k+1)} := \begin{bmatrix} P_1^{(k+1)} \\ \vdots \\ P_n^{(k+1)} \end{bmatrix} &= \begin{bmatrix} \sum_{i=1}^n P_{1i}P_i^{(k)} \\ \vdots \\ \sum_{i=1}^n P_{ni}P_i^{(k)} \end{bmatrix} \\
&= P \begin{bmatrix} P_1^{(k)} \\ \vdots \\ P_n^{(k)} \end{bmatrix}
\end{aligned}$$

donde $P = (P_{ij})_{i,j=1}^n$. Por tanto, obtenemos una matriz asociada a vectores de probabilidad. De forma resumida se cumple $P^{(k+1)} = PP^{(k)}$.

Definición: *Consideremos la motivación anterior.*

(1) La sucesión de vectores de probabilidades $\{P^{(k)}\}_{k\in\mathbb{N}}$ lo llamaremos ***proceso aleatorio****.*

(2) Un proceso aleatorio $\{P^{(k)}\}_{k\in\mathbb{N}}$ tal que $P^{(k+1)}$ depende solo de los pasos k. Decimos que $\{P^{(k)}\}_{k\in\mathbb{N}}$ es un ***proceso de Markov****.*

(3) Consideremos un proceso de Markov. A P le llamamos ***matriz de transición*** *del proceso de Markov.*

Ejemplo: Sea $G = (V, L)$ un grado donde $V = \{A, B, C, D, E\}$. Consideremos un proceso de Markov sobre este grafo con matriz de transición:

$$P = \begin{bmatrix} 0 & 1/2 & 1/3 & 1 & 0 \\ 1 & 0 & 1/3 & 0 & 1/3 \\ 0 & 1/2 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 1/3 \\ 0 & 0 & 1/3 & 0 & 0 \end{bmatrix}$$

Notemos que si tomamos cualquier columna, y sumamos las filas, nos da 1, es decir, se cumple que para todo $b \in V$, se tiene que,

$$\sum_{a\in V} P_{ab} = 1$$

Si partimos inicialmente en C, entonces tenemos el vector,

$$P^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Luego,

$$\begin{aligned} P^{(1)} &= \begin{bmatrix} 0 & 1/2 & 1/3 & 1 & 0 \\ 1 & 0 & 1/3 & 0 & 1/3 \\ 0 & 1/2 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 1/3 \\ 0 & 0 & 1/3 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 1/3 \\ 1/3 \\ 0 \\ 0 \\ 1/3 \end{bmatrix} \end{aligned}$$

Nuestro interés es ver como se comportar $P_{(k+1)}$, que como sabemos se obtiene siguiendo la identidad $P^{(k+1)} = PP^{(k)} = \cdots P^{k+1}P^{(0)}$. Entonces, necesitamos estudiar P^{k+1}.

Proposición: *Sea P una matriz de transición de un proceso de Markov. Entonces 1 es un valor propio de P.*

Demostración: Sea $G = (V, L)$ un grafo que forma un proceso de Markov con matriz transposición P. Sabemos que para $b \in V$ se cumple que,

$$\sum_{a \in V} P_{ab} = 1 \Leftrightarrow \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \\ P_b \\ \\ \end{bmatrix} = 1$$

donde P_b es la columna b de P. Luego se tiene que,

$$\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} P = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \Leftrightarrow P^T \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Es decir, P^T tiene valor propio 1. Veamos que P^T, P tienen los mismos valores propios. Notemos que,

$$p_{P^T}(x) = \det(P^T - xI) = \det((P - xI)^T) = \det(P - xI) = p_P(x)$$

Por lo tanto, si P, P^T tienen los mismos valores propios, entonces P tiene valor propio 1. ■

Proposición (Característica de la matriz transición) : *Sea P una matriz de transición. Si λ es valor propio de P, entonces $|\lambda| \leq 1$. Además, si v es un vector propio de $\lambda = 1$, entonces las coordenadas son ≥ 0.*

Demostraremos la proposición más adelante.

Sea P una matriz de transición. Supongamos que,

(1) Existe un único valor propio λ tal que $|\lambda| = 1$.

(2) El valor propio $\lambda = 1$ tiene multiplicidad 1.

(3) P es diagonalizable.

Afirmación: *Existe un único vector π con coordenadas ≥ 0 con suma 1, tal que $P\pi = \pi$.*

Demostración: Por la proposición anterior existe π^* con coordenadas ≥ 0 asociado al valor propio $\lambda = 1$. Entonces se cumple que $P\pi^* = \pi^*$.

Si la suma de coordenadas de π^* es $M \geq 0$, entonces definimos el vector,

$$\pi := \frac{\pi^*}{M}$$

donde claramente cumple π tiene coordenadas ≥ 0, tiene suma de coordenadas 1 y que $P\pi = \pi$.
■

Afirmación: *Consideremos el vector π de la afirmación anterior. Para todo vector inicial $P^{(0)}$, se cumple que,*

$$\lim_{n\to\infty} P^n P^{(0)} = \pi$$

Demostración: Supongamos que P tiene n valores propios $\lambda_1, \cdots, \lambda_n$ tales que,

$$|\lambda_n| \leq \cdots \leq |\lambda_2| < |\lambda_1| = 1$$

Le desigualdad estricta se cumple dado que hay un único valor propio con módulo 1. Si P es diagonalizable, entonces existe una base de vectores propios $v_1, \cdots, v_n$ asociados a los valores propios $\lambda_1, \cdots, \lambda_n$. Luego se tiene que,

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

Sobre P sabemos que,

$$\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} P = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}$$

Supongamos que $u^t = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}$ Para todo $i = 2, \cdots, n$ se cumple que,

$$u^t v_i = u^t P v_i = \lambda_i u^t v_i$$

Entonces,

$$u^t v_i (1 - \lambda_i) = 0$$

Si $i \neq 1$, entonces necesariamente $u^t v_i = 0$, es decir,

$$\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} (v_i)_1 \\ \vdots \\ (v_i)_n \end{bmatrix} = (v_i)_1 + \cdots + (v_i)_n = 0$$

Notemos que,

$$\begin{aligned} 1 &= \sum_{j=1}^{n} P_j^{(0)} \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} c_i (v_i)_j \\ &= \sum_{i=1}^{n} c_i \underbrace{\sum_{j=1} (v_i)_j}_{=0 \text{ para } i \neq 1} \\ &= c_1 \sum_{j=1}^{n} (v_1)_j \end{aligned}$$

Podemos pensar en $v_1 = \pi$, por lo que la suma de las coordenadas es 1 y por tanto,

$$c_1 = 1$$

Para todo vector inicial $P^{(0)}$. Luego se tiene que,

$$P^m P^{(0)} = \sum_{j=1}^{n} c_j P^m v_j = \sum_{j=1} c_j \lambda_j^m v_j = c_1 v_1 + \sum_{j=2}^{n} \lambda_j^m c_j v_j = \pi + \sum_{j=2}^{n} \lambda_j^m c_j v_j$$

Si $|\lambda_j| < 1$ para todo $j = 2, \cdots, n$, entonces se cumple que,

$$\begin{aligned} \lim_{m \to \infty} P^m P^{(0)} &= \pi + \lim_{m \to \infty} \sum_{j=2}^{n} \lambda_j^m c_j v_j \\ &= \pi \end{aligned}$$

Probando la afirmación. ■

Conclusión: Si P es una matriz de transición tal que hay un único valor propio tal que $|\lambda| = 1$, el cual es $\lambda = 1$, que además tiene multiplicidad 1 y que P sea diagonalizable. Entonces P tiene un punto fijo, es decir, existe un vector π tal que $P\pi = \pi$ donde las coordenadas de π son ≥ 0.

Definición: *Sea P una matriz real $m \times n$.*

(1) Denotaremos $P \geq 0$ si todas las entradas son ≥ 0.

(2) Denotaremos $P > 0$ si $P \geq 0$ y $P \neq 0$.

(3) Denotaremos $P \gg 0$ si las entradas son > 0. En tal caso diremos que P es una ***matriz positiva****.*

Observación: La definición anterior también se aplica a los vectores.

Nota: Si P, Q son matrices de igual dimensiones, diremos que,

$$P \geq Q \Leftrightarrow P - Q \geq 0$$

Análogamente con los símbolos $>, \gg$.

Consideremos P matriz $n \times n$ tal que $P \geq 0$, consideremos $x > 0$ y el sistema de inecuaciones:

$$\star \begin{cases} Px \geq \lambda x \\ x > 0 \end{cases}$$

Definimos $\Lambda(P) := \{\lambda \in \mathbb{R} : \exists x > 0 \text{ tal que } Px \geq \lambda x\}$

Observación: Todo valor propio λ de P, satisface que $|\lambda| \in \Lambda(P)$.

Afirmación: *El conjunto $\Lambda(P)$ tiene supremo finito.*

Demostración: Definimos $M := \text{máx}_{i,j=1,\cdots,n} P_{ij}$. Sea $x > 0$ asociado a $\lambda \in \Lambda(P)$, entonces se cumple para todo $i = 1, \cdots, n$ que,

$$\lambda x_i \leq (Px)_i = \sum_{j=1}^{n} P_{ij} x_j \leq M \sum_{j=1}^{n} x_k$$

Luego,

$$\lambda \max_{i=1,\cdots,n} x_i \leq M \sum_{j=1}^{n} x_j \leq M \left(n \max_{i=1,\cdots,n} x_i \right)$$

Aquí concluimos que,

$$\lambda \leq Mn$$

Por lo tanto $\Lambda(P)$ es un conjunto acotado superiormente, por lo que su supremo es finito. $\blacksquare$.

Afirmación: *El supremo del conjunto $\Lambda(P)$ se alcanza.*

Demostración: Sea $\{(\lambda^{(j)}, x^{(j)})\}_{i \in \mathbb{N}}$ una secuencia tal que $\lambda^{(j)} \in \Lambda(P)$ y para tal valor se satisface con el vector $x^{(j)}$ tal que,

$$\lim_{j \to \infty} \lambda^{(j)} = \sup \Lambda(P) \leq \frac{M}{n}$$

Sin pérdida de generalidad supongamos que para todo $j \in \mathbb{N}$,

$$\sum_{i=1}^{n} x_i^{(j)} = 1$$

(Siempre se puede reescalar los vectores). Luego $x_i^{(j)} \in [0,1]$ para todo $i = 1, \cdots, n$ y para todo $j \in \mathbb{N}$. Por último supondremos que $0 \leq \lambda^{(j)} \leq Mn$ (notar que $\lambda = 0$ cumple la inecuación). Consideremos el siguiente resultado.

Lema: *Toda sucesión real acotada tiene una subsucesión convergente.*

Observemos que la secuencia de vectores $\{x^{(j)}\}_{j\in\mathbb{N}}$ define n suceiones reales acotadas de la forma $\{x_i^{(j)}\}_{j\in\mathbb{N}}$ donde $i = 1, \cdots, n$. Entonces para $i = 1$ tenemos la sucesión $\{x_1^{(j)}\}_{j\in\mathbb{N}}$ y por el lema anterior existe una subsucesión $\{x_1^{(j_k)}\}_{k\in\mathbb{N}}$ tal que converge en $\mathbb{R}$. Digamos que,

$$x_1^{(j_k)} \xrightarrow{k\to\infty} x_1^{(\infty)} \in [0,1]$$

Ahora de forma conveniente reemplazamos la sucesión $\{x_1^{(j)}\}_{j\in\mathbb{N}}$ por la sucesión convergente $\{x_1^{(j_k)}\}_{k\in\mathbb{N}}$.

Volvemos hacer esto para todo $i = 2, \cdots, n$. De forma que se obtiene una nueva sucesión $\{x^{(j)}\}_{j\in\mathbb{N}}$ que converge a un límite $x^{(\infty)}$ donde,

$$\sum_{i=1}^{n} x_i^{(\infty)} = 1$$

Y por tanto $x^{(\infty)} > 0$. (Los vectores con coordenadas que suman 1 conforman un espacio compacto). Finalmente obtenemos la secuencia $\{(x^{(j)}, \lambda^{(j)})\}_{j\in\mathbb{N}}$ convergente al límite $(x^{(\infty)}, \lambda^{(\infty)})$ donde $x^{(\infty)} > 0$ y $\lambda^{(\infty)} = \sup \Lambda(P)$.

Demostremos que el par $(x^{(\infty)}, \lambda^{(\infty)})$ satisface la inecuación. Observemos que,

$$\lim_{j\to\infty} Px^{(j)} \geq \lim_{j\to\infty} \lambda^{(j)} x^{(j)} \Leftrightarrow \lim_{j\to\infty} Px^{(j)} \geq \lambda^{(\infty)} x^{(\infty)}$$

Estudiemos el límite $\lim_{j\to\infty} Px^{(j)}$. Por definición se tiene que,

$$(Px^{(j)})_i = \sum_{k=1}^{n} P_{ik} x_k^{(j)} \xrightarrow{j\to\infty} \sum_{k=1}^{n} P_{ik} x_k^{(\infty)} = (Px^{(\infty)})_i$$

Entonces $\lim_{j\to\infty} Px^{(j)} = Px^{(\infty)}$ y por tanto,

$$Px^{(\infty)} \geq \lambda^{(\infty)} x^{(\infty)}$$

Finalmente concluimos que $\lambda^{(\infty)} \in \Lambda(P)$, es decir, el supremo se alcanza. ■

Afirmación: *Si* $P \gg 0$, *entonces* $\max \Lambda(P)$ *es estríctamente positivo.*

Demostración: Si $P \gg 0$, entonces todo elemento de la matriz, es positivo. Definimos,

$$m := \min_{i,j=1,\cdots,n} P_{ij}$$

Donde claramente $m > 0$. Entonces sigue que,

$$P \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \geq mn \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Es decir, $mn \in \Lambda(P)$, si $mn > 0$ entonces claramente $\max \Lambda(P) > 0$. ■

Nota: El sistema de inecuación $(\star)$ se puede considerar $P > 0$ o $P \gg 0$.

Observación: Sea $P \gg 0$ y $v > 0$, entonces se cumple que $Pv \gg 0$, puesto que,

$$(Pv)_i = \sum_{j=1}^{n} P_{ij} v_j > 0$$

para todo $i = 1, \cdots, n$.

Teorema (Perron-Frobenius terminar): *Sea $P > 0$ una matriz $n \times n$ y sea $\lambda_0 := \max \Lambda(P)$. Entonces,*

(a) Existe un $x_0 > 0$ tal que,

$$Px_0 = \lambda_0 x_0$$

(b) Si $\lambda \in \mathbb{C}$ es un valor propio de P, entonces $|\lambda| \leq \lambda_0$.

Demostración (Perron-Frobenius):

(a) Primero vamos a demostrar que si $P \gg 0$, entonces existe x_0 tal que $Px_0 = \lambda_0 x_0$, luego demostraremos para $P > 0$.

Supongamos que $P \gg 0$. Existe un $x_0 > 0$ tal que $Px_0 \geq \lambda_0 x_0$ donde $\lambda_0 = \max \Lambda(P)$. Supongamos que $Px_0 \neq \lambda_0 x_0$, luego sigue que,

$$Px_0 > \lambda_0 x_0 \Leftrightarrow Px_0 - \lambda_0 x_0 > 0$$

Aplicando P obtenemos,

$$P(Px_0) - P\lambda_0 x_0 \gg 0 \Leftrightarrow P(Px_0) \gg \lambda_0 (Px_0)$$

Definimos $y_0 := Px_0 \gg 0$, por lo que obtenemos,

$$Py_0 \gg \lambda_0 y_0$$

Observemos que podemos escoger $\varepsilon > 0$ tal que se satisface,

$$Py_0 - (\lambda_0 + \varepsilon) y_0 \geq 0$$

Sin embargo, esto implica que $\lambda_0 + \varepsilon$ satisface las condiciones de $\Lambda(P)$ y además, $\lambda_0 + \varepsilon > \lambda_0$, es decir, λ_0 no es máximo, siendo una contradicción, por lo tanto necesariamente,

$$Px_0 = \lambda_0 x_0$$

Observación: Si $Px_0 \gg 0$, entonces $\lambda_0 x_0$, y esto implica que $x_0 \gg 0$. Es decir, si $P \gg 0$, entonces existe $x_0 \gg 0$ tal que $Px_0 = \lambda_0 x_0$.

Ahora supongamos que $P > 0$, consideremos,

$$P_\varepsilon := P + \varepsilon \underbrace{\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}}_{=:A}$$

Donde $P_\varepsilon \gg 0$, esto implica que $\lambda_\varepsilon := \text{máx}\, \Lambda(P_\varepsilon)$ y que existe $x_1 \gg 0$ tales que,

$$P_\varepsilon x_1 = \lambda_\varepsilon x_1$$

Aplicando $\varepsilon \to 0$ se concluye que,

$$\lim_{\varepsilon \to 0} P_\varepsilon x_1 = \lim_{\varepsilon \to 0} \lambda_\varepsilon x_1 \Leftrightarrow P x_1 = \lim_{\varepsilon \to 0} \lambda_\varepsilon x_1$$

Afirmación: *Se cumple que,*

$$\lim_{\varepsilon \to 0} \lambda_\varepsilon = \lambda_0$$

Demostración: De la igualdad anterior se cumple que existe un $x_1 > 0$ tal que,

$$P x_1 \geq \lim_{\varepsilon \to 0} \lambda_\varepsilon x_1$$

Por lo que $\lim_{\varepsilon \to 0} \lambda_\varepsilon \in \Lambda(P)$. Por otro lado, existe un $x_0 > 0$ tal que,

$$P x_0 \geq \lambda_0 x_0$$

Entonces,

$$P_\varepsilon x_0 = P x_0 + \varepsilon A x_0 \geq \lambda_0 x_0 + \varepsilon A x_0 \geq \lambda_0 x_0$$

Por lo tanto $\lambda_0 \in \Lambda(P_\varepsilon)$. Finalmente se concluye que,

$$\lim_{\varepsilon \to 0} \lambda_\varepsilon \leq \lambda_0$$
$$\lambda_0 \leq \lim_{\varepsilon \to 0} \lambda_\varepsilon$$

Cumpliendo la igualdad y probando la afirmación. ■

Finalmente se concluye que dado $P > 0$ existe $x_0 > 0$ tal que $P x_0 = \lambda_0 x_0$.

(b) Sea $\lambda \in \mathbb{C}$ valor propio distinto a λ_0. Se cumple que $Pz = \lambda z$ donde z es un vector no nulo. Luego,

$$(Pz)_i = \sum_{j=1}^{n} P_{ij} z_j = \lambda z_i$$

para todo $i = 1, \cdots, n$. Entonces,

$$|\lambda z_i| = |\lambda||z_i| = \left|\sum_{j=1}^{n} P_{ij} z_j\right| \leq \sum_{j=1}^{n} P_{ij}|z_j|$$

para todo $i = 1, \cdots, n$. De aquí se concluye que,

$$P \begin{bmatrix} |z_1| \\ \vdots \\ |z_n| \end{bmatrix} \geq |\lambda| \begin{bmatrix} |z_1| \\ \vdots \\ |z_n| \end{bmatrix} \Leftrightarrow P\widetilde{z} \geq |\lambda|\widetilde{z}; \quad \widetilde{z} > 0$$

Es decir, $|\lambda| \in \Lambda(P)$ y por lo tanto $|\lambda| \leq \lambda_0$.

Demostrando el teorema. ■

Proposición (Característica de la matriz transición) : *Sea P una matriz de transición. Si λ es valor propio de P, entonces $|\lambda| \leq 1$. Además, si v es un vector propio de $\lambda = 1$, entonces las coordenadas son ≥ 0.*

Demostración (Característica de la matriz transición): Si P es una matriz de transición, entonces $P > 0$. Sea $\lambda_0 := \text{máx}\,\Lambda(P)$, demostraremos que $\lambda_0 = 1$.

Consideremos $Q = P^T$, entonces de antemano sabemos que se cumple,

$$\sum_{j=1}^{n} Q_{ij} = 1$$

para todo $i = 1, \cdots, n$. Observemos que $Q > 0$, luego por Perron-Frobenius, existe $x_0 > 0$ tal que,

$$Qx_0 = \lambda_0 x_0$$

Luego para todo $i = 1, \cdots, n$ se tiene que,

$$\begin{aligned} \lambda_0 (x_0)_i &= (Qx_0)_i \\ &= \sum_{j=1} nQ_{ij}(x_0)_j \\ &\leq \sum_{j=1}^{n} Q_{ij} \max_{k=1,\cdots,n} (x_0)_k = \max_{k=1,\cdots,n} (x_0)_k \end{aligned}$$

Por tanto,

$$\lambda \max_{k=1,\cdots,n} (x_0)_k \leq \max_{k=1,\cdots,n} (x_0)_k$$

Es decir $\lambda_0 \leq 1$. Por otro lado, si 1 es valor propio de P al ser matriz de transición, finalmente se concluye que $\lambda_0 = 1$. ■

3. Producto Interno

Consideremos el espacio vectorial $\mathbb{R}^n$. Podemos definir el producto punto entre dos vectores $x, y \in \mathbb{R}^n$ por:

$$(x,y) := x_1y_1 + \cdots + x_ny_n = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = y^T x$$

En particular,

$$\begin{aligned} (\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\ (x,y) &\mapsto y^T x \end{aligned}$$

El producto punto satisface que,

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

(2) Es bilineal, es decir, es lineal en cada argumento.

(3) Para todo $x \in \mathbb{R}^n$ se tiene que,

$$(x,x) = x_1^2 + \cdots + x_n^2 \geq 0$$

Podemos definir el producto punto con respecto al espacio $\mathbb{C}^n$, aunque tiene un comportamiento distinto. Sean $x, y \in \mathbb{C}^n$, luego se define el producto punto de x, y por,

$$(x,y) := x_1\overline{y_1} + \cdots + x_n\overline{y_n} = \overline{y}^T x$$

Observemos que $(x,y) = \overline{(y,x)}$ y que para todo $x \in \mathbb{C}^n$ se tiene que,

$$(x,x) = |x_1|^2 + \cdots + |x_n|^2 \geq 0$$

Sin embargo no tiene una propiedad bilineal pero si casi-bilineal, en particular, en la primera coordenada si es lineal pero en la segunda ocurre lo siguiente: Sean $x, y_1, y_2 \in \mathbb{C}^n$ y sea $c_1, c_2 \in \mathbb{C}$, entonces,

$$\begin{aligned} (x, c_1y_1 + c_2y_2) &= (\overline{c_1y_1 + c_2y_2})^T x \\ &= (\overline{c_1y_1} + \overline{c_2y_2})^T x \\ &= \overline{c_1}(\overline{y_1})^T x + \overline{c_2}(\overline{y_2})^T x \\ &= \overline{c_1}(x,y_1) + \overline{c_2}(x,y_2) \end{aligned}$$

Ejemplo: Sea $z = a + ib \in \mathbb{C}$, luego,

$$(z,z) = z\overline{z} = (a+ib)(a-ib) = a^2 + b^2 = |z|^2$$

Ejemplo: Sea $x_{\mathbb{C}} = \begin{bmatrix} x_1 + ix_2 & x_3 + ix_4 \end{bmatrix} \in \mathbb{C}^2$ y sea $x_{\mathbb{R}} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix} \in \mathbb{R}^4$. Sigue que,

$$(x_{\mathbb{C}}, x_{\mathbb{C}}) = x_1^2 + x_2^2 + x_3^2 + x_4^2 = (x_{\mathbb{R}}, x_{\mathbb{R}})$$

Definición: *Sea V un espacio vectorial sobre $\mathbb{C}$. Una función $(\cdot,\cdot) : V \times V \to \mathbb{C}$ es un* ***producto interno*** *si se satisface las siguientes condiciones:*

(1) ***Hermitiano:*** *$(x, y) = \overline{(y, x)}$ para todo $x, y \in V$.*

(2) ***Lineal para el primer argumento:*** *Para todo $x_1, x_2 \in V$ y para todo $c_1, c_2 \in \mathbb{C}$ se cumple que,*

$$(c_1x_1 + c_2x_2, y) = c_1(x_1, y) + c_2(x_2, y)$$

(3) Para todo $x \in V$ se tiene que,

$$(x, x) \geq 0$$

Con igualdad si y sólo si $x = 0$.

Observación: Sea V espacio vectorial sobre $\mathbb{C}$. Se puede definir el producto interno $(\cdot,\cdot) : V \times V \to \mathbb{R}$ con la única diferencia que el primer argumento es lineal con respecto a las constantes reales.

Observación: Veamos que pasa con el segundo argumento. Sean $y_1, y_2 \in V$ y sean $c_1, c_2 \in \mathbb{C}$, entonces,

$$(x, c_1y_1 + c_2y_2) = \overline{(c_1y_1 + c_2y_2, x)} = \overline{c_1(y_1, x)} + \overline{c_2(y_2, x)} = \overline{c}_1(x, y_1) + \overline{c}_2(x, y_2)$$

Por lo que,

$$(x, c_1y_1 + c_2y_2) = \overline{c}_1(x, y_1) + \overline{c}_2(x, y_2)$$

Nota: Sea V un espacio vectorial. Si este posee producto interno, entonces diremos que V es un EVPI.

Ejemplo: $\mathbb{R}^n$ es un EVPI con producto interno $(x, y) := x^T y$. En particular, $(\cdot,\cdot)$ es un producto interno real.

Ejemplo: Pensemos en $\mathbb{C}^n$ con función $(x, y) := \overline{y}^T x$, entonces $\mathbb{C}^n$ es un EVPI.

Notación: Para vectores y matrices denotamos,

$$x^* := \overline{x}^T$$

Ejemplo: Consideremos $V = \mathcal{P}(\mathbb{C})$ el espacio de polinomios complejos. Definimos la función,

$$(p(x), q(x)) := \int_0^1 p(t)\overline{q(t)}dt$$

Así por ejemplo,

$$\begin{aligned}
(x+ix^2, 2i+(1+i)x) &= \int_0^1 (t+it^2)(-2i+(1-i)t)dt \\
&= \int_0^1 (-2it+(1-i)t^2+2t^2+(i+1)t^3)dt \\
&= -i+\frac{1}{3}(3-i)+\frac{1}{4}(i+1)
\end{aligned}$$

Afirmación: $\mathcal{P}(\mathbb{C})$ *es un EVPI.*

Demostración: Debemos probar los axiomas de producto interno.

- Sean $p, q \in \mathcal{P}(\mathbb{C})$, supongamos que $p(t)\overline{q}(t) = a(t)+ib(t)$ donde a, b son polinomios reales. Entonces,

$$\begin{aligned}
(p,q) &= \int_0^1 p(t)\overline{q(t)}dt \\
&= \int_0^1 (a(t)+ib(t))dt \\
&= \int_0^1 a(t)dt + i\int_0^1 b(t)dt \\
&= \overline{\int_0^1 a(t)dt - i\int_0^1 b(t)dt} \\
&= \overline{\int_0^1 (a(t)-ib(t))dt} \\
&= \overline{\int_0^1 \overline{p(t)}q(t)dt} \\
&= \overline{(q,p)}
\end{aligned}$$

- Sean $p, r, q \in \mathcal{P}(\mathbb{C})$ y sean $c_1, c_2 \in \mathbb{C}$, entonces,

$$\begin{aligned}
(c_1p+c_2r, q) &= \int_0^1 (c_1p(t)+c_2r(t))\overline{q(t)}dt \\
&= c_1\int_0^1 p(t)\overline{q(t)}dt + c_2\int_0^1 r(t)\overline{q(t)}dt \\
&= c_1(p,q)+c_2(r,q)
\end{aligned}$$

- Sea $p \in \mathcal{P}(\mathbb{C})$, entonces,

$$(p,p) = \int_0^1 p(t)\overline{p(t)}dt$$

Supongamos que $p(t) = a(t) + ib(t)$ donde a, b son polinomios reales. Entonces,

$$p(t)\overline{p(t)} = a^2(t) + b^2(t) \geq 0$$

Por lo que,

$$(p,p) = \int_0^1 (a^2(t) + b^2(t))dt \geq 0$$

Por lo tanto, $\mathcal{P}(\mathbb{C})$ es un EVPI. ■

En general, si $V = \mathcal{F}(\Omega) := \{f : \Omega \to \mathbb{C}\}$ y consideramos,

$$(f,g) := \int_\Omega f(x)\overline{g(x)}dx$$

Entonces $\mathcal{F}(\mathbb{C})$ también es un EVPI.

Ejemplo: Sea $V = \mathbb{C}^{n\times n}$ el espacio de las matrices complejas. Sea,

$$(A,B) := \mathrm{Tr}(B^*A)$$

Afirmación: *$\mathbb{C}^{n\times n}$ es un EVPI.*

Demostración: Estudiemos algunas propiedades de la función $\mathrm{Tr}(\cdot)$. Sea $A \in \mathbb{C}^{n\times n}$, luego se cumple,

$$\begin{aligned} \mathrm{Tr}\left(\overline{A}\right) &= \overline{\mathrm{Tr}(A)} \\ \mathrm{Tr}(A) &= \mathrm{Tr}(A^T) \end{aligned}$$

La primera se cumple usando definición de traza y la segunda se cumple de forma directa puesto que la transpuesta no cambia los valores de la diagonal. Veamos que pasa con la suma de dos matrices. Sean $A, B \in \mathbb{C}^{n\times n}$, donde la diagonal de A es $(a_{11}, \cdots, a_{nn})$ y la diagonal de B es $(b_{11}, \cdots, b_{nn})$, entonces,

$$\begin{aligned} \mathrm{Tr}(A+B) &= (a_{11} + b_{11}) + \cdots + (a_{nn} + b_{nn}) \\ &= (a_{11} + \cdots + a_{nn}) + (b_{11} + \cdots + b_{nn}) \\ &= \mathrm{Tr}(A) + \mathrm{Tr}(B) \end{aligned}$$

Probemos los axiomas de producto interno.

- Luego se tiene que para todo $A, B \in \mathbb{C}^{n\times n}$ se cumple que,

$$\begin{aligned}
(A,B) &= \mathrm{Tr}(B^*A)\\
&= \overline{\mathrm{Tr}\left(\overline{B^*A}\right)}\\
&= \overline{\mathrm{Tr}\left(\overline{A^TB^*}\right)}\\
&= \overline{\mathrm{Tr}\left(\overline{A^T(B^*)^T}\right)}\\
&= \overline{\mathrm{Tr}\left(\overline{A}^TB\right)}\\
&= \overline{\mathrm{Tr}(A^*B)}\\
&= \overline{(B,A)}
\end{aligned}$$

- Sean $A, B, C \in \mathbb{C}^{n\times n}$ y sean $c_1, c_2 \in \mathbb{C}$, entonces,

$$\begin{aligned}
(c_1A + c_2C, B) &= \mathrm{Tr}(B^*(c_1A + c_2C))\\
&= \mathrm{Tr}(c_1B^*A + c_2B^*C)\\
&= c_1\mathrm{Tr}(B^*A) + c_2\mathrm{Tr}(B^*C)\\
&= c_1(A,B) + c_2(C,B)
\end{aligned}$$

- Sea $A \in \mathbb{C}^{n\times n}$ con diagonal $(a_{11}, \cdots, a_{nn})$, entonces,

$$\begin{aligned}
(A,A) &= \mathrm{Tr}(A^*A)\\
&= a_{11}\overline{a}_{11} + \cdots + a_{nn}\overline{a}_{nn}\\
&= |a_{11}|^2 + \cdots + |a_{nn}|^2 \geq 0
\end{aligned}$$

  Es decir, $(A,A) \geq 0$.

Por lo tanto, $\mathbb{C}^{n\times n}$ es un EVPI. ■

Lema: *Sea V un EVPI. Entonces $(x,y) = 0$ para todo $y \in V$ si y sólo $x = 0$.*

Demostración: Si $x = 0$, entonces $(x,y) = 0$ para todo $y \in V$. Supongamos que $(x,y) = 0$ para todo $y \in V$. Tomando $(x,x) = 0$, entonces se obtiene que $x = 0$. Probando lema.

Lema: *Sea V un EVPI. Sean $x_1, x_2 \in V$, entonces $x_1 = x_2$ si y sólo si $(x_1, y) = (x_2, y)$ para todo $y \in V$.*

Demostración: Notemos que,

$$(x_1, y) = (x_2, y) \Leftrightarrow (x_1 - x_2, y) = 0$$

Entonces $x_1 = x_2$ si y sólo si $(x_1 - x_2, y) = 0$ para todo $y \in V$. ■

Lema: *Sean $A, B : V \to W$ donde V, W son EVPI. Si para todo $x \in V, y \in W$ se cumple que,*

$$(Ax, y) = (Bx, y)$$

Entonces $A = B$.

Demostración: Tomemos $x \in V$ fijo, si para todo $y \in W$ se tiene que,

$$(Ax, y) = (Bx, y)$$

Entonces $Ax = Bx$. En particular, se tiene $Ax = Bx$ para todo $x \in V$, es decir, $A = B$. ■

Teorema (Desigualdad de Cauchy-Schwarz): *Sea V un EVPI. Entonces para todo $x, y \in V$ se cumple que,*

$$\overline{(x,y)}(x,y) \leq (x,x)(y,y)$$

Caso producto interno real: Supongamos que V es un espacio vectorial sobre $\mathbb{C}$ con producto interno definido sobre $\mathbb{R}$. Sea $t \in \mathbb{R}$, de forma que,

$$0 \leq \|x - ty\|^2 =: (x - ty, x - ty)$$

Por lo que,

$$0 \leq \|x - ty\|^2 = (x,x) - t(x,y) - 2t(y,x) + t^2(y,y) = \|x\|^2 - 2t(x,y) + t^2\|y\|^2$$

Obtenemos una parábola que está por encima de la recta real, es decir, el discriminante cumple que,

$$4(x,y)^2 - 4t^2\|x\|^2\|y\|^2 \leq 0 \Leftrightarrow (x,y)(x,y) \leq (x,x)(y,y)$$

Probando Cauchy-Schwarz cuando el producto interno toma solo valores reales. Probemos el caso genereal.

Demostración (Cauchy-Schwarz): Sea V un espacio vectorial sobre $\mathbb{C}$ con producto interno definido sobre $\mathbb{C}$. Sean $x, y \in V$ fijos, queremos estudiar $\|x - zy\|^2$ cuando $z \in \mathbb{C}$, en particular, queremos minimizar la expresión.

Si $z = v + iw$, notemos que,

$$\|x - zy\|^2 = (x - zy, x - zy) = (x,x) + (v^2 + w^2)(y,y) - 2v\mathrm{Re}(x,y) + 2w\mathrm{Im}(x,y)$$

Es equivalente a estudiar la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por,

$$f(v,w) = \|x\|^2 + (v^2 + w^2)\|y\|^2 - 2v\mathrm{Re}(x,y) + 2w\mathrm{Im}(x,y)$$

con x, y fijos. Que es una función diferenciable, en particular, nos interesa minimizar la función, notemos que,

$$\frac{df}{dv}(v,w) = \frac{df}{dw}(v,w) = 0 \Leftrightarrow 2v\|y\|^2 - 2\mathrm{Re}(x,y) = 2w\|y\|^2 + 2\mathrm{Im}(x,y) = 0$$

En particular,

$$v = \frac{\mathrm{Re}(x,y)}{\|y\|^2}, \quad w = -\frac{\mathrm{Im}(x,y)}{\|y\|^2}$$

donde $y \neq 0$ (si $y = 0$, entonces se cumple directamente Cauchy-Schwarz), por lo que la función f se minimiza cuando,

$$(v^*, w^*) = \frac{1}{\|y\|^2}(\mathrm{Re}(x,y), -\mathrm{Im}(x,y))$$

Por tanto,

$$\begin{aligned}\min_{(v,w)\in\mathbb{R}^2} f(v,w) = f(v^*,w^*) &= \|x\|^2 + \frac{\mathrm{Re}^2(x,y)+\mathrm{Im}^2(x,y)}{\|y\|^2} - 2\frac{\mathrm{Re}^2(x,y)}{\|y\|^2} - 2\frac{\mathrm{Im}^2(x,y)}{\|y\|^2}\\ &= \frac{\|x\|^2\|y\|^2 - (\mathrm{Re}^2(x,y)+\mathrm{Im}^2(x,y))}{\|y\|^2}\\ &= \frac{\|x\|^2\|y\|^2 - \overline{(x,y)}(x,y)}{\|y\|^2}\end{aligned}$$

Finalmente,

$$\min_{(v,w)\in\mathbb{R}^2} f(v,w) \geq 0 \Leftrightarrow \overline{(x,y)}(x,y) \leq \|x\|^2\|y\|^2$$

Demostrando la desigualdad de Cauchy-Schwarz. ■

Observación: La desigualdad de Cauchy-Schawrz permite definir ángulos. Sea V un EVPI, entonces sabemos que,

$$|(x,y)|^2 \leq \|x\|^2\|y\|^2 \Leftrightarrow -1 \leq \frac{(x,y)}{\|x\|\|y\|} \leq 1$$

Por lo tanto, definimos el ángulo de x, y por

$$\angle(x,y) = \arccos\frac{(x,y)}{\|x\|\|y\|}$$

Ejemplo: Sea $V = \mathbb{R}^2$ con el producto punto. Entonces,

$$x \cdot y = \|x\|\|y\|\cos\theta$$

donde $\theta \in [0, 2)$ es el ángulo entre x e y.

Proposición: *Sea V un EVPI con producto interno definido en $\mathbb{R}$. Sean $x, y, z \in V$, entonces se satisface que,*

$$\|y-z\|^2 = \|x-y\|^2 + \|x-z\|^2 - 2(x-y, x-z)$$

Demostración: Por definición,

$$\begin{aligned}\|y-z\|^2 = (y-z, y-z) &= (y-x+x-z, y-x+x-z)\\ &= (y-x,y-x)+(y-x,x-z)+(x-z,y-x)+(x-z,x-z)\\ &= \|x-y\|^2 + \|x-z\|^2 - 2(x-y,x-z)\end{aligned}$$

■

Lema (Polarización del producto interno): *Sea V un EVPI. Entonces, si el producto interno está definido,* *(a) Sobre* $\mathbb{R}$*, se cumple,*

$$(x, y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

(b) Sobre $\mathbb{C}$*, se cumple,*

$$(x, y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2)$$

Demostración ((Polarización del producto interno)):

(a) Por definición,

$$\begin{aligned}\|x + y\|^2 - \|x - y\|^2 &= (x + y, x + y) - (x - y, x - y) \\ &= (x, x) + (x, y) + (y, x) + (y, y) - (x, x) + (x, y) + (y, x) - (y, y) \\ &= 2(x, y) + 2(y, x) = 4(x, y)\end{aligned}$$

Por lo tanto,

$$\frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) = (x, y)$$

(b) Notemos que,

$$\begin{aligned}i\|x + iy\|^2 - i\|x - iy\|^2 &= i(x + iy, x + iy) - i(x - iy, x - iy) \\ &= i(x, iy) + i(iy, x) + i(x, iy) + i(iy, x) \\ &= 2(x, y) - 2(y, x)\end{aligned}$$

Por lo tanto,

$$\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2 = 4(x, y)$$

Por lo tanto,

$$\frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2) = (x, y)$$

Probando la proposición. ■

3.1. Normas

Hemos definido una "norma" con respecto a un producto interno. Sin embargo podemos definir de forma general una norma.

Definición: *Sea* V *un espacio vectorial. Una función* $\|\cdot\| : V \to \mathbb{R}$ *es una* ***norma*** *sobre* V *si:*

(a) *Para todo $v \in V, \alpha \in \mathbb{K}$ se tiene que,*

$$\|\alpha v\| = |\alpha|\|v\|$$

donde $\mathbb{K}$ puede ser $\mathbb{R}$ o $\mathbb{C}$.

(b) *Para todo $u, v \in V$ se tiene que,*

$$\|u + v\| \leq \|u\| + \|v\|$$

(c) *Para todo $v \in V$ se tiene que $\|v\| \geq 0$ y la igualdad se cumple si y sólo si $v = 0$.*

Ejemplo: Sea V un EVPI. Consideremos la función $\|x\| = \sqrt{(x, x)}$.

Afirmación: *La función $\| \cdot \|$ es una norma.*

Demostración: Probemos los tres axiomas de norma.

(a) Sea $x \in V$ y α una constante, luego,

$$\|\alpha x\|^2 = (\alpha x, \alpha x) = |\alpha|^2(x, x) = |\alpha|\|x\|$$

(b) Sean $x, y \in V$, usando la desigualdad de Cauchy-Schawrz se observa que,

$$\begin{aligned}
\|x + y\|^2 &= (x + y, x + y) \\
&= \|x\|^2 + \|y\|^2 + (x, y) + (y, x) \\
&\leq \|x\|^2 + \|y\|^2 + 2|(x, y)| \\
&\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \\
&= (\|x\| + \|y\|)^2
\end{aligned}$$

Finalmente se concluye que,

$$\|x + y\| \leq \|x\| + \|y\|$$

(c) Claramente $\|x\| \geq 0$ para todo $x \in V$. Si $x = 0$, entonces claramente $\|x\| = 0$. Supongamos ahora que $\|x\| = 0$, entonces para todo $y \in V$ se cumple que,

$$|(x, y)|^2 = \overline{(x, y)}(x, y) \leq (x, x)(y, y) = 0$$

Es decir, $(x, y) = 0$ para todo $y \in V$, por lo que necesariamente $x = 0$.

Demostrando que $\| \cdot \|$ es una norma. ■

Ejemplo: Consideremos $V = \mathbb{K}^n$ (donde $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$). Para $p \in [1, \infty)$ definimos la función $\| \cdot \|_p$ de la siguiente forma:

$$\|x\|^p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

Afirmación: *La función $\| \cdot \|_p$ es una norma y la llamamos **norma**-p.*

Demostración: Probemos los tres axiomas de norma.

(a) Sea $x \in V$ y sea α una constante, luego,

$$\begin{aligned}\|\alpha x\|_p &= \left(\sum_{i=1}^{n} |\alpha x_i|^p\right)^{1/p} \\ &= \left(\sum_{i=1}^{n} |\alpha|^p |x_i|^p\right)^{1/p} \\ &= |\alpha| \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} = |\alpha| \|x\|_p\end{aligned}$$

(b) Para todo $x, y \in V$ se cumple que,

$$\|x + y\|_p = \left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p}$$

terminar

(c) Claramente $\|x\|_p \geq 0$ para todo $x \in V$. Si $x = 0$, entonces $\|x\|_p = 0$. Supongamos ahora que $\|x\|_p = 0$, es decir,

$$\sum_{i=1}^{n} |x_i|^p = 0$$

Como tenemos sumandos no negativos, se tiene que $|x_i| = 0$ para todo $i = 1, \cdots, n$. Por lo tanto $x = 0$.

Demostrando que $\| \cdot \|_p$ es una norma. ■

Ejemplo: Consideremos $V = \mathbb{K}^n$ (donde $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$). Definimos la función dada por

$$\|x\|_\infty = \max_{i=1,\cdots,n} |x_i|$$

Afirmación: *La función $\| \cdot \|_\infty$ es una norma y la llamamos* ***norma infinita***.

para todo $x \in V$. Verifiquemos que es una norma, para ello debemos probar los axiomas de normal.

(a) Sea $x \in V$ y α una constante, luego,

$$\begin{aligned}\|\alpha x\|_\infty &= \max_{i=1,\cdots,n} |\alpha x_i| \\ &= |\alpha| \max_{i=1,\cdots,n} |x_i| = |\alpha| \|x\|_\infty\end{aligned}$$

(b) Para todo $x, y \in V$ se cumple que,

$$\begin{aligned}\|x + y\|_\infty &= \max_{i=1,\cdots,n} |x_i + y_i| \\ &\leq \max_{i=1,\cdots,n} |x_i| + \max_{i=1,\cdots,n} |y_i| = \|x\|_\infty + \|y\|_\infty\end{aligned}$$

(c) Claramente $\|x\|_\infty \geq 0$. Observemos que,

$$\|x\|_\infty = 0 \Leftrightarrow \max_{i=1,\cdots,n} |x_i| = 0 \Leftrightarrow |x_i| = 0; \quad i = 1, \cdots, n$$

Por lo tanto, $\|x\|_\infty = 0$ si y sólo si $x = 0$.

Probando que $\|\cdot\|_\infty$ es una norma.

Nos preguntamos si existe una relación entre la norma-p y la norma infinita. Veamos que pasa al estudiar el conjunto, **mejorar nombre**

$$\text{Cua}_p := \{...\}$$

Afirmación: *Sea V un espacio vectorial. Consideremos las normas $\|\cdot\|_p, \|\cdot\|_\infty$ con $p \geq 1$. Entonces para todo $x \in V$ se cumple*

$$\lim_{p\to\infty} \|x\|_p = \|x\|_\infty$$

Demostración: Sea $x \in V$ cualquiera. Vamos a demostrar la siguiente desigualdad:

$$\|x\|_\infty \overset{(1)}{\leq} \|x\|_p \overset{(2)}{\leq} n^{1/p}\|x\|_\infty$$

- **Primera desigualdad:** Por definición de norma infinita tenemos,

$$\begin{aligned}\|x\|_\infty &= \max_{i=1,\cdots,n} |x_i| \\ &= \left(\left(\max_{i=1,\cdots,n} |x_i|\right)^p\right)^{1/p} \\ &\leq \left(\sum_{i=1}^{n} |x_i|\right)^{1/p} = \|x\|_p\end{aligned}$$

  Esta última desigualdad se cumple puesto que en la sumatoria se suma el máximo de los módulos de las coordenadas. Probando la primera desigualdad.

- **Segunda desigualdad:** Notemos que,

$$\begin{aligned}\|x\|_p &= \left(\sum_{i=1}^{n} |x_i|\right)^{1/p} \\ &\leq \left(n\left(\max_{i=1,\cdots,n} |x_i|\right)^p\right)^{1/p} \\ &\leq n^{1/p} \max_{i=1,\cdots,n} |x_i| = n^{1/p}\|x\|_\infty\end{aligned}$$

  Probando la segunda desigualdad.

Tomando $p \to \infty$ obtenemos,

$$\|x\|_\infty = \lim_{p\to\infty} \|x\|_p = \|x\|_\infty$$

Por lo tanto, el límite $\|x\|_p$ existe para todo $x \in V$ y,

$$\lim_{p\to\infty} \|x\|_p = \|x\|_\infty$$

Demostrando la afirmación. ■

Teorema: *Sea V un espacio vectorial. Una norma de V es inducida por un producto interno ($\|x\| = \sqrt{(x,x)}$) si y sólo si,*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

para todo $x, y \in V$.

Demostración: Supongamos que una norma $\|\cdot\|$ es inducida por un producto interno, por lo que $\|x\| = \sqrt{(x,x)}$ para todo $x \in V$. Demostremos la ley del paralelogramo. Por definición, para todo $x, y \in V$ se tiene que,

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= (x+y, x+y) + (x-y, x-y)\\ &= 2\|x\|^2 + 2\|y\|^2 + (x,y) + (y,x) - (x,y) - (y,x)\\ &= 2(\|x\|^2 + \|x\|^2)\end{aligned}$$

Probando la ley del paralelogramo.

Supongamos ahora que $\|\cdot\|$ es una norma tal que satisface la ley del paralelogramo. Queremos demostrar que existe una norma $(\cdot,\cdot)$ en V tal que $\|x\| = \sqrt{(x,x)}$ para todo $x \in V$. Primero demostremos el caso real.

Caso real: Consideremos,

$$(x,y) := \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$$

Verifiquemos que es producto interno, para ello debemos probar los axiomas de producto interno real.

(1) Observemos que para todo $x, y \in V$ se tiene que,

$$(x,y) = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) = \frac{1}{4}(\|y+x\|^2 - \|y-x\|^2) = (y,x)$$

Por lo que $(x,y) = (y,x)$.

(2) Aqui probaremos dos cosas. Probaremos que el primer argumento es lineal y luego que es homogeneo.

Afirmación: *Para todo $x, y, z \in V$ se cumple que $(x+y, z) = (x,z) + (y,z)$. Es decir, $(\cdot,\cdot)$ es lineal en el primer argumento.*

Demostración: Sean $x, y, z \in V$. Por ley de paralelogramo tenemos que,

$$2\|x+z\|^2 + 2\|y\|^2 = \|x+y+z\|^2 + \|x-y+z\|^2$$

Luego,

$$\|x+y+z\|^2 = 2\|x+z\|^2 + 2\|y\|^2 - \|x-y-z\|^2 = 2\|y+z\|^2 + 2\|x\|^2 - \|y-x+z\|^2$$

Esto implica que,

$$\|x+y+z\|^2 = \|x\|^2 + \|y\|^2 + \|x+z\|^2 + \|y+z\|^2 - \frac{1}{2}\|x-y+z\|^2 - \frac{1}{2}\|y-x+z\|^2$$

Aplicando $z \mapsto -z$ obtenemos una nueva identidad,

$$\|x+y-z\|^2 = \|x\|^2 + \|y\|^2 + \|x-z\|^2 + \|y-z\|^2 - \frac{1}{2}\|x-y-z\|^2 - \frac{1}{2}\|y-x-z\|^2$$

Finalmente obtenemos que,

$$\begin{aligned}(x+y, z) &= \frac{1}{4}(\|x+y+z\|^2 - \|x+y-z\|^2)\\ &= \frac{1}{4}(\|x+z\|^2 - \|x-z\|^2) + \frac{1}{4}(\|y+z\|^2 - \|y-z\|^2)\\ &= (x,z) + (y,z)\end{aligned}$$

Por tanto $(\cdot, \cdot)$ es lineal en el primer argumento. ■

Afirmación: *Para todo $x, y \in V$ y para todo $\lambda \in \mathbb{R}$ se cumple que $(\lambda x, y) = \lambda(x, y)$. Es decir, $(\cdot, \cdot)$ es homogeneo en el primer argumento.*

Demostración: Observemos que para $n \in \mathbb{N}$ se cumple,

$$(nx, y) = \underbrace{(x,y) + \cdots + (x,y)}_{n\text{-veces}} = n(x,y)$$

Y que,

$$(-x, y) = \frac{1}{4}(\|x-y\|^2 - \|x+y\|^2) = -(x,y)$$

usando que $\|w\| = \|-w\|$ para todo $w \in V$. Esto implica que para todo $k \in \mathbb{Z}$ se tiene que,

$$(kx, y) = k(x,y)$$

Consideremos $\frac{p}{q} \in \mathbb{Q}$ donde $q \neq 0$ $(p, q \in \mathbb{Z})$, entonces,

$$q\left(\frac{p}{q}x, y\right) = (px, y) = p(x,y)$$

Por lo tanto,

$$\left(\frac{p}{q}x, y\right) = \frac{p}{q}(x, y)$$

Hemos probando que $(\lambda x, y) = \lambda(x, y)$ para todo $\lambda \in \mathbb{Q}$. Extendamos a los reales. Consideremos la función,

$$t \mapsto \frac{1}{t}(tx, y)$$

definida en $\mathbb{R} \setminus \{0\}$. Esta función es continua puesto que **falta demsotrar continuidad**.

Luego, sea $\lambda \in \mathbb{R} \setminus \{0\}$ y sea $\{\lambda_n\}_{n\in\mathbb{N}}$ una sucesión tal que $\lambda_n \in \mathbb{Q}$ para todo $n \in \mathbb{N}$ y que $\lambda_n \to \lambda$ cuando $n \to \infty$. Entonces por continuidad de la función se tiene que,

$$\begin{aligned}
(x, y) &= \lim_{n\to\infty} \frac{1}{\lambda_n}(\lambda_n x, y) \\
&= \frac{1}{\lim_{n\to\infty} \lambda_n}\left(\lim_{n\to\infty} \lambda_n x, y\right) \\
&= \frac{1}{\lambda}(\lambda x, y)
\end{aligned}$$

Es decir, $\lambda(x, y) = (\lambda x, y)$ para todo $\lambda \in \mathbb{R} \setminus \{0\}$. Para concluir el resultado basta ver que es evidente que se cumple cuando $\lambda = 0$, de forma que $\lambda(x, y) = (\lambda x, y)$ para todo $\lambda \in \mathbb{R}$. Demostrando la homogienidad. ■

Finalmente concluimos que $(\cdot, \cdot)$ es lineal en el primer argumento.

(3) Sea $x \in V$, entonces,

$$(x, x) = \|x\|^2 \geq 0$$

Y $(x, x) = 0$ si y sólo si $x = 0$.

Nota: De forma general sobre un V EVPI seguiremos trabajando con la norma $\|x\| = \sqrt{(x, x)}$.

Por lo tanto $(\cdot, \cdot)$ es un producto interno real que induce a la norma $\|\cdot\|$.

Caso complejo: Consideremos,

$$(x, y) := \frac{1}{4}\sum_{k=0}^{3} i^k\|x + i^k y\|^2$$

Observemos que,

$$\begin{aligned}
(ix, y) &= \frac{1}{4}\sum_{k=0}^{3} i^k \|ix + i^k y\|^2 \\
&= \frac{1}{4}\sum_{k=0}^{3} i^k \|ix + i^k y\|^2 \\
&= i\frac{1}{4}\sum_{k=0}^{3} i^{k-1} |i|^2 \|x + i^{k-1} y\|^2 \\
&= i\frac{1}{4}\sum_{k=0}^{3} i^k \|x + i^k y\|^2 \\
&= i(x, y)
\end{aligned}$$

falta segunda vuelta

3.2. Ortogonalidad

Definición: *Sea V un EVPI. Decimos que dos vectores u, v son* ***ortogonales o perpendiculares*** *si $(u, v) = 0$. En tal caso denotamos $u \perp v$.*

Observación: Consideremos la definición de ángulo, el cual satisface que,

$$(u, v) = \cos\theta \|u\|\|v\|$$

Si (u, v) son ortogonales, entonces,

$$(u, v) = 0 \Leftrightarrow \cos\theta\|u\|\|v\| = 0 \Leftrightarrow \cos\theta = 0$$

Por lo que $\theta \in \{k\pi + \pi/2 : k \in \mathbb{Z}\}$.

Lema: *Sea V un EVPI. Si $u \perp v$, entonces $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.*

Demostración: Por definición,

$$\begin{aligned}
\|u + v\|^2 &= (u + v, u + v) = (u, u) + (u, v) + (v, u) + (v, v) \\
&= \|u\|^2 + \|v\|^2
\end{aligned}$$

puesto que $(u, v) = (v, u) = 0$.

Definición: *Sea V un EVPI y sea $E \subseteq V$ subespacio de V. Sea $v \in V$, decimos que v es ortogonal a E si para todo $e \in E$ se tiene que $v \perp e$. En tal caso denotamos $v \perp E$.*

Lema: *Sea V EVPI. Sea $E = \langle v_1, \cdots, v_k\rangle$, entonces $v \perp E$ si y sólo si $v \perp v_1, \cdots v \perp v_k$.*

Demostración: Si $v \perp E$, entonces es claro que $v \perp v_1, \cdots, v \perp v_k$.

Supongamos que v es perpendicular a la cada elemento de la base. Si $w \in E$, entonces existen coeficientes $\alpha_1, \cdots, \alpha_k$ tales que,

$$w = \alpha_1 v_1 + \cdot + \alpha_k v_k$$

Luego,

$$(v,w) = (v, \alpha_1 v_1 + \cdots + \alpha_k v_k) = \overline{\alpha_1}(v,v_1) + \cdots + \overline{\alpha_k}(v,v_k) = 0$$

Por tanto $v \perp E$. ■

Definición: *Sea V un EVPI.*

(a) Una colección $\{v_1, \cdot, v_k\}$ son ***ortogonales*** *si para todo $i,j = 1,\cdots,k$ distintos, se tiene que $v_i \perp v_j$.*

(b) Supongamos que $\{v_1,\cdots,v_k\}$ es una colección ortogonal. Diremos que son ***ortonormales*** *si,*

$$(v_i,v_j) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

Lema: *Sea V un EVPI. Sea $\{v_1,\cdots,v_k\}$ una colección ortogonal. Entonces,*

$$\left\| \sum_{i=1}^{k} \alpha_i v_i \right\|^2 = \sum_{i=1}^{k} |\alpha_i|^2 \|v_i\|^2$$

Demostración: Por definición,

$$\left\| \sum_{i=1}^{k} \alpha_i v_i \right\|^2 = \left(\sum_{i=1}^{k} \alpha_i v_i, \sum_{i=1}^{k} \alpha_i v_i \right) = \sum_{i=1}^{k} \|\alpha_i v_i\|^2 + \underbrace{\sum_{1 \leq i < j \leq k} (\alpha_i v_i, \alpha_j v_j)}_{=0}$$

$$= \sum_{i=1}^{k} |\alpha_i|^2 \|v_i\|^2$$

■

Lema: *Sea V un EVPI. Si $\{v_1,\cdots,v_k\}$ es una colección ortogonales de elementos no nulos. Entonces es una colección independiente.*

Demostración: Consideremos $\alpha_1,\cdots,\alpha_k$ tales que,

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$$

Si $v_1,\cdots,v_k$ son ortogonales, entonces,

$$0 = \|0\|^2 = \left\| \sum_{i=1}^{k} \alpha_i v_i \right\|^2 = \sum_{i=1}^{k} |\alpha_i|^2 \|v_i\|^2$$

Si tenemos una suma de valores no negativos que da 0, entonces cada sumando vale cero, es decir, $\alpha_i = 0$ para todo $i = 1,\cdots,k$, por lo que la colección es independiente. ■

Definición: *Sea V un EVPI de dimensión n. Si $\{v_1, \cdots, v_n\}$ es una base y es una colección ortogornal, entonces decimos que es una* ***base ortogonal.*** *De forma análoga si es una colección ortonormal, decimos que es una* ***base ortonormal****.*

Proyectar para obtener coordendas de un vector respecto a una base ortogonal: Sea V un EVPI con base ortogonal $\{v_1, \cdots, v_n\}$. Sea $x = \alpha_1 v_1 + \cdots + \alpha_n v_n$, entonces para todo $j = 1, \cdots, n$ se tiene,

$$(x, v_j) = \sum_{k=1}^{n} \alpha_k (v_k, v_j) = \alpha_j \|v_j\|^2$$

Luego se tiene que,

$$\alpha_j = \frac{(x, v_j)}{\|v_j\|^2}$$

para todo $j = 1, \cdots, n$. Siendo esto las coordenadas de x en la base ortogonal. Por tanto podemos escribir,

$$x = \sum_{k=1}^{n} \frac{(x, v_k)}{\|v_k\|^2} v_k$$

Si $\{v_1, \cdots, v_n\}$ es base ortonormal, entonces escribimos,

$$x = \sum_{k=1}^{n} (x, v_k) v_k$$

Lema: *Sea V un EVPI con base ortonormal $\{v_1, \cdots, v_n\}$. Sean $x, y \in V$ con coordenadas $(x_1, \cdots, x_n), (y_1, \cdots, y_n)$ respectivamente. Entonces,*

$$(x, y) = \sum_{k=1}^{n} x_k \overline{y}_k = y^* x$$

Demostración: Sean $x, y \in V$ tales que tienen coordenadas $(x_1, \cdots, x_n), (y_1, \cdots, y_n)$ respectivamente, es decir,

$$\begin{aligned} x &= x_1 v_1 + \cdots + x_n v_n \\ y &= y_1 v_1 + \cdots + y_n v_n \end{aligned}$$

Luego,

$$\begin{aligned} (x, y) &= \left(\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_i v_i \right) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \overline{y}_i (v_i, v_j) \\ &= \sum_{i=1}^{n} x_i \overline{y}_i (v_i, v_i) = \sum_{i=1}^{n} x_i \overline{y}_i \end{aligned}$$

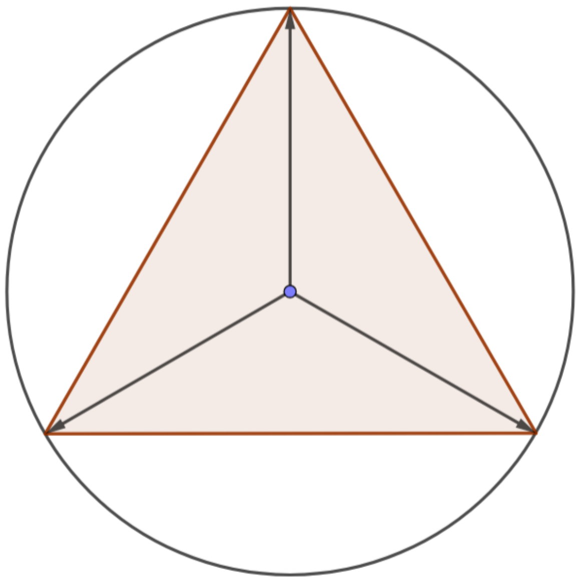
Demostrando el lema. ■

Tetraedro regular en una esfera (revsiar 20): En $\mathbb{R}^{n+1}$, definimos la n-esfera por el conjunto:

$$\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} : \|x\|^2 = 1\}$$

donde $\|x\| = \sqrt{(x, x)}$.

Ejemplo: En $\mathbb{R}^2$ consideremos una circuferencia de radio 1 y un triángulo equilatero inscrito en la circuferencia. Consideremos los vectores $x_1, x_2, x_3 \in \mathbb{R}^3$ del centro de la circuferencia a los vértice del triángulo.



Nos interesa estudiar (x_i, x_j) y $\|x_i - x_j\|$ donde $i, j = 1, 2, 3$. Observemos las siguientes propiedades:

- $\|x_i\| = 1$ para todo $i = 1, 2, 3$.
- El lado del triángulo mide $\|x_i - x_j\| =: L_2$ donde $i \neq j$ con L cosntante.
- $\sum_{i=1}^{3} x_i = 0$.

Luego para $i \neq j$ se cumple que,

$$\|x_i - x_j\|^2 = \|x_i\|^2 + \|x_j\|^2 - 2(x_i, x_j) \Leftrightarrow (x_i, x_j) = 1 - \frac{\|x_i - x_j\|^2}{2}$$

Entonces para todo $i = 1, 2, 3$ se tiene que,

$$0 = (x_i, x_1 + x_2 + x_3) = 3 - L_2^2$$

Por lo tanto el lado del triángulo mide,

$$L_2 = \sqrt{3}$$

Por otro lado, para todo $i \neq j$ se tiene que,

$$(x_i, x_j) = 1 - \frac{3}{2} = -\frac{1}{2}$$

Usando la definición de ángulo concluimos que los ángulos entre los vectores x_i, x_j es de,

$$\arccos\left(-\frac{1}{2}\right) = 120^{\circ}$$

Veamos que pasa en $\mathbb{R}^3$. Consideremos la esfera unitaria $\mathbb{S}^2$ y consideremos un tetraedro regular inscrito de la esfera. Sean $x_1, x_2, x_3, x_4 \in \mathbb{R}^3$ los vectores del centro de la esfera hasta el vértice del tetraedro. Al igual que en dos dimensiones, se cumple las siguientes propiedades:

- $\|x_i\| = 1$ para todo $i = 1, 2, 3, 4$.
- El lado del tetraedro mide $\|x_i - x_j\| = L_3$ donde $i \neq j$ con L constante.
- $\sum_{i=1}^{4} x_i = 0$.

Del caso anterior ya sabemos como determinar $\|x_i - x_j\|$ y (x_i, x_j), por lo que calculamos de forma directa. Entonces para todo $i = 1, 2, 3, 4$ se tiene que,

$$(x_i, x_1 + x_2 + x_3 + x_4) = 0 \Leftrightarrow L_3^2 = \frac{8}{3}$$

Por lo tanto, para todo $i \neq j$ se tiene que,

$$\|x_i - x_j\| = 2\sqrt{\frac{2}{3}} \ \text{ y } \ (x_i, x_j) = -\frac{1}{3}$$

Por definición de ángulos se tiene que los vectores x_i, x_j con $i \neq j$ forman un ángulo de,

$$\arccos\left(-\frac{1}{3}\right) \approx 109^{\circ}$$

Veamos por último que pasa en $\mathbb{R}^4$. Repitiendo el proceso de tomar una 3-esfera unitara, un tetraedro regular en $\mathbb{R}^4$ inscrito, tomamos los vectores desde el centro de la 3-esfera a los vértices del tetraedo $x_1, x_2, x_3, x_4, x_5 \in \mathbb{R}^4$. Se cumplen las tres propiedades anteriores y podemos concluir que,

$$\|x_i - x_j\| = \sqrt{\frac{5}{2}} \ \text{ y } \ (x_i, x_j) = -\frac{1}{4}$$

Es decir, el ángulo entre los vectores x_i, x_j es de,

$$\arccos\left(-\frac{1}{4}\right) \approx 104^{\circ}$$

Generalizemos para $\mathbb{R}^n$. Consideremos la n-esfera unitara y sea un tetraedo regular en $\mathbb{R}^n$ inscrito en la $n-1$-esfera. Sean $x_1, \cdots, x_n \in \mathbb{R}^n$ vectores definimos desde el centro de la $n-1$-esfera a los vértices del tetraedo. Se satisfacen las siguientes propiedades:

- $\|x_i\| = 1$ para todo $i = 1, \cdots, n$.
- El lado del tetraedro mide $\|x_i - x_j\| = L_n$ donde $i \neq j$ con L constante.
- $\sum_{i=1}^{n} x_i = 0$.

Luego para todo $i \neq j$ se satisface que,

$$\|x_i - x_j\|^2 = \|x_i\|^2 + \|x_j\|^2 - 2(x_i, x_j) \Leftrightarrow (x_i, x_j) = 1 - \frac{\|x_i - x_j\|^2}{2}$$

Entonces para todo $i = 1, \cdots, n$ se tiene que,

$$0 = \left(x_i, \sum_{i=1}^{n} x_j\right) = n - \frac{n-1}{2} L_n^2 \Leftrightarrow L_n = \sqrt{\frac{2n}{n-1}}$$

Determinando el lado del tetraedo regular. Y por otro lado, para todo $i \neq j$ se tiene que,

$$(x_i, x_j) = 1 - \frac{n}{n-1} = -\frac{1}{n-1}$$

Por tanto, los vectores x_i, x_j son "casi perpendiculares". Es más, si tomamos $n \to \infty$ generamos el espacio vectorial $\mathbb{R}^{\mathbb{N}}$ definido de la siguiente forma:

$$\mathbb{R}^{\mathbb{N}} := \{(x_1, x_2, \cdots) : x_i \in \mathbb{R},\ i \in \mathbb{N}\}$$

Y al tomar toda las piezas anteriores, obtenemos que un tetraedo regular inscrito en una esfera infinita unitaria tiene lado

$$\lim_{n \to \infty} L_n = \sqrt{2}$$

Donde los vectores x_i, x_j que estan formados desde el centro de la esfera hasta un vértice del tetraedro, para todo $i \neq j$ satisfacen que,

$$(x_i, x_j) = \lim_{n \to \infty} -\frac{1}{n-1} = 0$$

Por lo tanto podemos determinar una base infinita ortogonal de $\mathbb{R}^{\mathbb{N}}$ usando estos vectores.

Hay un hecho más fuerte de lo que hemos visto, y es que no es necesario tomar un tetraedo, sino que solo basta tomar vectores al azar. El siguiente hecho lo demuestra.

Hecho: *En $\mathbb{R}^n$, sea $\varepsilon > 0$ pequeño y sea $k \in \mathbb{N}$. Si se escogen k vectores en la esfera $\mathbb{S}^{n-1}$, la probabilidad de que todos los ángulos de los vectores sean entre $90^\circ - \varepsilon$ y $90^\circ + \varepsilon$ tiende a 1 cuando $n \to \infty$.*

De forma resumida: para todo $\varepsilon > 0$ y para todo $k \geq 2$ natural, escogemos k vectores cualesquiera tales que $\|x_i\|^2 = 1$ para todo $i = 1, \cdots, k$. Se cumple que,

$$\mathbb{P}\left(\arccos \frac{(x_i, x_j)}{\|x_i\|\|x_j\|} \in \left[\frac{\pi}{2} - \varepsilon, \frac{\pi}{2} + \varepsilon\right] : x_i, x_j \in \mathbb{R}^n,\ i \neq j \in \{1, \cdots, k\}\right) \overset{n \to \infty}{\longrightarrow} 1$$

3.3. Proyecciones Ortogonales

Definición: *Sea V un EVPI y sea $E \subseteq V$ subespacio. Sea $v \in V$, definimos la* ***proyección ortogonal de*** *v* ***en*** *E como el único vector $w \in V$ tal que,*

$$\begin{cases} w \in E \\ v - w \perp E \end{cases}$$

En tal caso denotamo $P_E v := w$.

Observación: Podemos definir la función proyección de un vector en V sobre el subespacio E por $P_E : V \to E$.

Al estudiar la proyección ortogonal, surgen algunas preguntas.

- ¿Cómo sabemos que existe este w?
- En caso de existir, ¿de verdad es única?
- ¿Cómo determinamos w?

Teorema: *Sea V un EVPI y $E \subseteq V$ subespacio. Sea $v \in V$ y supongamos que existe su proyección ortogonal sobre E donde $w = P_E v$, entonces se cumple,*

(a) $\|v - w\| \leq \|v - x\|$ *para todo $x \in E$.*

(b) $\|v - w\| = \|v - x\|$ *con $x \in E$, si y sólo si $x = w$.*

Demostración:

(a) Sea $x \in E$, entonces $v - x = v - w + w - x$, por definición se tiene que $v - w \perp E$ y $w - x \in E$, entonces,

$$\|v - x\|^2 = \|v - w\|^2 + \|w - x\|^2 \geq \|v - w\|^2$$

Es decir, $\|v - x\| \geq \|v - w\|$ para todo $x \in E$.

(b) Si $x = w$, entonces claramente se cumple la igualdad. Supongamos que dado $x \in E$ se tiene que $\|v - w\| = \|v - x\|$. Sabemos que se cumple que,

$$\|v - x\|^2 = \|v - w\|^2 + \|w - x\|^2 \Leftrightarrow 0 = \|w - x\|^2$$

Entonces $w = x$.

Probando el teorema. ■

Observación: El teorema anterior nos dice que la proyección ortogonal, es caso de existir, es única, puesto que si w_1, w_2 son proyecciones de v, entonces $\|v - w_1\| = \|v - w_2\|$, lo que implica que $w_1 = w_2$.

Nota: El teorema anterior nos dice que podemos pensar la proyección de forma intuitiva como estudiar la distancia entre un vector y el subespacio, en particular, la proyección ortogonal, es la distancia mínima de un vector v, a un subespacio E.

Proposición: *Sea V un EVPI y E subespacio. Sea $\{v_1, \cdots, v_r\}$ una base ortogonal de E. Entonces se cumple que,*

$$P_E v = \sum_{k=1}^{r} \frac{(v, v_k)}{\|v_k\|^2} v_k$$

Demostración: Definimos,

$$w := \sum_{k=1}^{r} \frac{(v, v_k)}{\|v_k\|^2} v_k$$

Donde claramente $w \in E$. Si demostramos que $v - w \perp E$, entonces necesariamente $w = P_E v$. Notemos que,

$$\begin{aligned}
(v - w, v_j) &= (v, v_j) - \left(\sum_{k=1}^{r} \frac{(v, v_k)}{\|v_k\|^2} v_k, v_j \right) \\
&= (v, v_j) - \sum_{k=1}^{r} \frac{(v, v_k)}{\|v_k\|^2} (v_k, v_j) \\
&= (v, v_j) - \frac{(v, v_j)}{\|v_j\|^2} (v_j, v_j) = 0
\end{aligned}$$

Por lo tanto,

$$P_E v = \sum_{k=1}^{r} \frac{(v, v_k)}{\|v_k\|^2} v_k$$

Demostrando la proposición. ■

Ejemplo: Consideremos $V = \mathbb{R}^3$ y el producto interno habitual. Sean,

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ x_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \ x_3 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

Sea $E_2 = \langle x_1, x_2 \rangle$, queremos determinar $P_{E_2} x_3$. Determinemos una base ortogonal de E_2. Sean $v_1 := x_1$ y $v_2 := x_2 - x_1$, entonces,

$$(v_1, v_2) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = 0$$

Luego $\{v_1, v_2\}$ es una base ortogonal de E_2. Luego por la proposición anterior se tiene que,

$$\begin{aligned}
P_{E_2} x_3 &= \frac{(x_3, v_1)}{\|v_1\|^2} v_1 + \frac{(x_3, v_2)}{\|v_2\|^2} v_2 \\
&= v_1 + \frac{1}{2} v_2 = \begin{bmatrix} 1/2 \\ 1 \\ 3/2 \end{bmatrix}
\end{aligned}$$

Corolario (Existencia de la proyección ortogonal): *Sea V un EVPI y sea E un subespacio. Entonces para todo $v \in V$, existe la proyección ortogonal $P_E v$.*

Para demostrar la existencia, necesitamos tener una base ortogonal del subespacio. Para ello vamos aplicar un método para obtener bases ortogonales.

3.4. Ortogalización de Gram-Schmidt

El método de Gram-Schmidt permite obtener a partir de una colección de vectores $\{x_1, \cdots, x_m\}$ independientes, una colección de vectores $\{v_1, \cdots, v_m\}$ ortogonales independientes con el mismo espacio generado

Para realizar este método necesitamos aplicar la proyección ortogonal de forma reiterativa.

Ortogonalización Gram-Schmidt: Consideremos la colección $\{x_1, \cdots, x_m\}$ de vectores independientes. Definimos $v_1 := x_1$. Notemos que existe la proyección $P_{\langle v_1 \rangle} v$ para todo $v \in V$ al tomar un solo elemento. Ahora tomamos,

$$v_2 := x_2 - P_{\langle v_1 \rangle} x_2 = x_2 - \frac{(v_1, x_2)}{\|v_1\|^2} v_1$$

Entonces claramente v_1, v_2 son ortogonales, puesto que,

$$(v_1, v_2) = (x_1, x_2) - \frac{(x_1, x_2)}{\|x_1\|^2}(x_1, x_1) = 0$$

Luego por la propisición anterior podemos tomar,

$$v_3 := x_3 - P_{\langle v_1, v_2 \rangle} x_3$$

Y este es perpendicular a v_1, v_2. De forma recursiva definimos,

$$v_{r+1} := x_{r+1} - P_{\langle v_1, \cdots, v_r \rangle} x_{r+1}$$

El cual es ortogonal a $v_1, \ldots, v_r$. De esta forma $\{v_1, \cdots, v_{r+1}\}$ es una colección ortogonal e independiente. Ahora notemos lo siguiente,

$$v_{r+1} \in \langle x_1, \cdots, x_{r+1} \rangle$$

Por tanto,

$$\langle v_1, \cdots, v_{r+1} \rangle = \langle x_1, \cdots, x_{r+1} \rangle$$

Finalmente obtenemos que $\{v_1, \cdots, v_m\}$ una colección de vectores ortogonales que tiene el mismo espacio generado por $\{x_1, \cdots, x_m\}$.

Ejemplo: Con $V = \mathbb{R}^3$ y el producto interno habitual. Sean los vectores,

$$x_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 0 \\ 5 \\ 4 \end{bmatrix}$$

Notemos que es una base de $\mathbb{R}^3$. Construyamos una base ortogonal. Por el método de Gram-Schmidt se tiene que,

$$\begin{aligned}
v_1 &= x_1 = \begin{bmatrix}1\\0\\1\end{bmatrix} \\
v_2 &= x_2 - P_{\langle v_1\rangle}x_2 = x_2 - \frac{4}{2}x_1 = \begin{bmatrix}2\\1\\-2\end{bmatrix} \\
v_3 &= x_3 - P_{\langle v_1,v_2\rangle}x_3 = x_3 - \frac{4}{2}v_1 - \frac{-3}{9}v_2 = \frac{4}{3}\begin{bmatrix}-1\\4\\1\end{bmatrix}
\end{aligned}$$

Entonces $\{v_1, v_2, v_3\}$ es una base ortogonal de $\mathbb{R}^3$.

Demostración (Existencia de la proyección ortogonal): Sea V un EVPI de dimensión n. Sea $E \subseteq V$ sub espacio con base $\{x_1, \cdots, x_m\}$. Por la ortogonalización Gram-Schmidt se puede construir una base ortogonal $\{v_1, \cdots, v_m\}$ de E y por tanto, por la proposición anterior, existe la proyección ortogonal de $v \in V$ sobre E. En particular, es de la forma,

$$P_E v = \sum_{k=1}^{r} \frac{(v, v_k)}{\|v_k\|^2} v_k$$

■

Por lo tanto, hemos visto que la proyección ortogonal existe, es única y se puede determinar.

3.5. Complementos Ortogonales

Definición: *Sea V un EVPI y sea E un subespacio. El* ***complemento ortogonal*** *de E es el conjunto,*

$$E^{\perp} := \{x \in V : x \perp v, \forall v \in E\} = \{x \in V : x \perp E\}$$

Observación: $E^{\perp}$ es un subespacio vectorial, puesto que si consideramos $x, y \perp E$ y por la linealidad del producto interno en el primer argumento se cumple que,

$$\alpha x + \beta y \perp E$$

Afirmación: *Si $E \subseteq V$ es subespacio, entonces $V = E \oplus E^{\perp}$.*

Demostración: Sea $v \in V$. Luego la proyección ortogonal de v sobre E cumple que,

$$P_E v \in V \ \text{ y } \ v - P_E v \perp E$$

Consideremos $w := v - P_E v$. Probemos que w es la proyección ortogonal de v sobre $E^\perp$. Claramente $w \in E^\perp$, por lo que solo debemos ver que $P_E v \perp E^\perp$. Sea $x \in E^\perp$, entonces para todo $y \in E$ se tiene que,

$$(y, x) = 0$$

Tomando $y = P_E v$ se tiene que $(P_E v, x) = 0$. Por tanto se cumple que $(P_E v, x) = 0$ para todo $x \in E^\perp$, de forma que $P_E v \perp E^\perp$. Como la proyección es única se tiene que,

$$P_{E^\perp} v = v - P_E v \Leftrightarrow v = P_E v + P_{E^\perp} v$$

Es decir, todo $v \in V$ se escribe de forma única por una suma de un elemento de E y de un elemento de $E^\perp$, demostrando la afirmación.

Proposición: *Sea V un EVPI. Sea $E \subseteq V$ subespacio, entonces $(E^\perp)^\perp = E$.*

Demostración: Sea $v \in E$. Luego para todo $w \in E^\perp$ se tiene que,

$$v \perp w$$

Por definición de ortogonalidad se tiene que $v \in (E^\perp)^\perp$. Luego $E \subseteq (E^\perp)^\perp$.

Para probar la igualdad notemos que,

$$V = E \oplus E^\perp = E^\perp \oplus (E^\perp)^\perp$$

Entonces,

$$\dim V = \dim E + \dim E^\perp = \dim E^\perp + \dim (E^\perp)^\perp$$

Por lo tanto $\dim E = \dim (E^\perp)^\perp$. Si V es un espacio vectorial finito se cumple que,

$$E = (E^\perp)^\perp$$

Demostrando la proposición. ■

Aplicación: Sea A matriz real de $m \times n$, sean $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$. Se cumple que,

$$\begin{aligned} y^T A x &= \sum_{i=1}^{m} \sum_{j=1}^{n} y_i A_{ij} x_j \\ &= \sum_{j=1}^{n} \sum_{i=1}^{m} x_j (A^T)_{ji} y_i = x^T A^T y \end{aligned}$$

En particular,

$$(Ax, y)_{\mathbb{R}^m} = (A^T y, x)_{\mathbb{R}^n}$$

donde consideramos $(\cdot,\cdot)_{\mathbb{R}^m}, (\cdot,\cdot)_{\mathbb{R}^n}$ por el producto habitual en $\mathbb{R}^m$ y $\mathbb{R}^n$ respectivamente.

Afirmación: $KerA = (ImA^T)^{\perp}$.

Demostración: Sea $x \in \mathrm{Ker}A$. Sea $z \in \mathrm{Im}A^T$ luego existe $y \in \mathbb{R}^m$ tal que $A^T y = z$, entonces se tiene que,

$$(z, x)_{\mathbb{R}^n} = (A^T y, x)_{\mathbb{R}^n} = (\underbrace{Ax}_{=0}, y)_{\mathbb{R}^m} = (0, y)_{\mathbb{R}^m}$$

Por lo tanto, para todo $z \in \mathrm{Im}A^T$ se tiene que $(z, x)_{\mathbb{R}^n} = 0$, es decir, $x \in (\mathrm{Im}A^T)^{\perp}$ y entonces $\mathrm{Ker}A \subseteq (\mathrm{Im}A^T)^{\perp}$

Probemos la igualdad. Sea $x \in (\mathrm{Im}A^T)^{\perp}$, entonces para todo $z \in \mathrm{Im}A^T$ se cumple que,

$$(z, x)_{\mathbb{R}^n} = 0$$

Sea $y \in \mathbb{R}^m$, luego $A^T y \in \mathrm{Im}A^T$ y por lo tanto,

$$(Ax, y)_{\mathbb{R}^m} = (A^T y, x)_{\mathbb{R}^n} = 0$$

Finalmente se tiene que $(Ax, y) = 0$ para todo $y \in \mathbb{R}^m$, es decir, $Ax = 0$ y entonces $x \in \mathrm{Ker}A$. De esta forma $\mathrm{Ker}A = (\mathrm{Im}A^T)^{\perp}$. ■

Si $A : \mathbb{R}^n \to \mathbb{R}^m$, entonces $A^T : \mathbb{R}^m \to \mathbb{R}^n$. Si $\mathrm{Im}A^T \subseteq \mathbb{R}^n$ es un subespacio de $\mathbb{R}^n$, entonces se cumple que,

$$\mathbb{R}^n = \mathrm{Im}A^T \oplus (\mathrm{Im}A^T)^{\perp}$$

De la afirmación anterior se concluye,

$$\mathbb{R}^n = \mathrm{Ker}A \oplus \mathrm{Im}A^T$$

3.6. Operador Adjunto

Definición: *Sea A una matriz $n \times m$. Definimos la matriz **adjunta** por $A^* = \overline{A^T}$.*

Observación: Si A es una matriz real, entonces $A^* = A^T$.

Proposición: *Sean A, B una matriz $n \times m$ y $m \times k$ respectivamente. Entonces $(AB)^* = B^* A^*$.*

Demostración: Por definición,

$$\begin{aligned}
(B^* A^*)_{ij} &= \sum_{r=1}^{n} (B^*)_{ir} (A^*)_{rj} \\
&= \sum_{r=1}^{n} \overline{(A_{jr} B_{ri})} \\
&= \overline{\sum_{r=1}^{n} A_{jr} B_{ri}} \\
&= \overline{(AB)}_{ji} \\
&= ((AB^) *)_{ij}
\end{aligned}$$

Por lo tanto $(AB)^* = B^*A^*$. ■

Lema: *Sean $A \in \mathbb{C}^{m\times n}, x \in \mathbb{C}^n, y \in \mathbb{C}^m$. Entonces,*

$$(Ax, y)_{\mathbb{C}^m} = (x, A^*y)_{\mathbb{C}^n}$$

donde $(\cdot,\cdot)_{\mathbb{C}^m}, (\cdot,\cdot)_{\mathbb{C}^n}$ son el producto interno habitual de $\mathbb{C}^m$ y $\mathbb{C}^n$ respectivamente.

Demostración: Sean $x \in \mathbb{C}^n$ e $y \in \mathbb{C}^m$, luego se cumple que,

$$\begin{aligned}(Ax, y)_{\mathbb{C}^m} &= y^*Ax \\ &= \sum_{i=1}^{m} \overline{y}_i (Ax)_i \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} \overline{y}_i A_{ij} x_j \\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{m} \overline{(A^*)_{ji} y_i}\right) x_j \\ &= \sum_{j=1}^{n} \overline{(A^*y)}_j x_j = (x, A^*y)_{\mathbb{C}^n}\end{aligned}$$

Demostrando la proposición. ■

Observación: Pensemos en B, A matrices tales que,

$$\mathbb{C}^k \xrightarrow{B} \mathbb{C}^n \xrightarrow{A} \mathbb{C}^m$$

Es decir, $A \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{n\times k}$. Entoncdes se cumple,

$$(ABx, y)_{\mathbb{C}^m} = (Bx, A^*y)_{\mathbb{C}^n} = (x, B^*A^*y)_{\mathbb{C}^k} = (x, (AB)^*y)_{\mathbb{C}^k}$$

Definición: *Sean V, W EVPI y sea $A : V \to W$ una función lineal. Definimos el la función adjunta $A^* : W \to V$ como la única función que satisface que,*

$$(v, A^*w)_V = (Av, w)_W$$

para todo $v \in V, w \in W$.

Afirmación: *El operador A^* existe, es lineal y es único.*

Demostración:

- **Existencia:** Sean $\{v_1, \cdots, v_n\}$ y $\{w_1, \cdots, w_m\}$ base ortonormales de V y W respectivmanete. Definimos la matriz $M \in \mathbb{C}^{m\times n}$ definida por la matriz A en las bases ortonormales. Definimos $B := M^*$.

  Afirmación: *B es el operador adjunto de A.*

Demostración: Claramente $B : W \to V$. Probemos que para todo $v \in V, w \in W$ se tiene que,

$$(Av, w)_W = (v, Bw)_V$$

Notemos que,

$$\begin{aligned}(Av_i, w_j)_W &= \left(\sum_{k=1}^{m} M_{ik} w_k, w_j\right) \\ &= \sum_{k=1}^{m} M_{ik}(w_k, w_j) = M_{ij}\end{aligned}$$

para todo $i = 1, \cdots, n$ y $j = 1, \cdots, m$. Por otro lado,

$$\begin{aligned}(v_i, Bw_j)_V &= \left(v_i, \sum_{l=1}^{n} \overline{M}_{lj} v_l\right) \\ &= \sum_{l=1}^{n} M_{lj}(v_i, v_l) = M_{ij}\end{aligned}$$

para todo $i = 1, \cdots, n$ y $j = 1, \cdots, m$. Por lo tanto, para todo $v \in V, w \in W$ se tiene que,

$$(Av, w)_W = (v, Bw)_V$$

De esta forma existe el operador adjunto de A.

- **Unicidad:** Supongamos que A^*, B^* satisfacen las condiciones de operador adjunto. Sea $w \in W$ fijo, entonces se tiene que para todo $v \in V$ se cumple que,

$$(v, A^*w)_V = (Av, w)_W = (v, B^*w)_V \Leftrightarrow (v, A^*w - B^*w) = 0$$

  Es decir, $A^*w = B^*w$, en particular se cumple para todo $w \in W$. Por lo tanto $A^* = B^*$.

- **Lineal:** Sean $w_1, w_2 \in W$ y β una constante. Luego para todo $v \in V$ se cumple que,

$$\begin{aligned}(v, A^*(\beta w_1 + w_2))_V &= (Av, \beta w_1 + w_2)_W \\ &= \overline{\beta}(Av, w_1)_W + (Av, w_2)_W \\ &= \overline{\beta}(v, A^*w_1)_V + (v, A^*w_2)_V \\ &= \overline{\beta}(v, \beta A^*w_1)_V + (v, \beta A^*w_1 + A^*w_2)_V\end{aligned}$$

  Entonces,

$$A^*(\beta w_1 + w_2) = \beta A^*w_1 + w_2$$

Demostrando la afirmación. ■

Proposición: *Sean V, W EVPI y sean $A, B : V \to W$ matrices (operadores lineales) de $m \times n$. Entonces se cumple,* *(1)* $(A+B)^* = A^* + B^*$.

(2) $(\alpha A)^* = \overline{\alpha} A^*$ *para todo* $\alpha \in \mathbb{C}$.

(3) $(A^*)^* = A$.

(4) $(y, Ax)_W = (A^*y, x)_V$ *para todo* $x \in V, y \in W$ *donde* $(\cdot,\cdot)_W, (\cdot,\cdot)_V$ *son productos internos cualquiera de* W *y* V *respectivamente.*

Demostración:

(1) Por definición para todo $i = 1, \cdots, m$ y $j = 1, \cdots, n$ se tiene que,

$$(A+B)^*_{ij} = (\overline{A+B})_{ji} = \overline{A}_{ji} + \overline{B}_{ji} = (A)^*_{ij} + (B)^*_{ij}$$

Luego $(A+B)^* = A^* + B^*$.

(2) Por definición para todo $i = 1, \cdots, m$ y $j = 1, \cdots, n$ y para todo $\alpha \in \mathbb{C}$ se tiene que,

$$(\alpha A)^*_{ij} = \overline{\alpha A}_{ji} = \overline{\alpha}\overline{A}_{ji} = \overline{\alpha} A^*_{ij}$$

Luego $(\alpha A)^* = \overline{\alpha} A^*$.

(3) Por definición para todo $i = 1, \cdots, m$ y $j = 1, \cdots, n$ se tiene que,

$$(A^*)^*_{ij} = \overline{A^*_{ji}} = \overline{\overline{A}_{ij}} = A_{ij}$$

Luego $(A^*)^* = A$.

(4) (**faltar ver que si** $\widetilde{A}^* = \widetilde{A^*}$) Consideremos bases ortonormales de V y de W. Sin pérdida de generalidad supongamos que la matriz A está en su forma de las bases ortonormales. Sean $x \in V$ e $y \in W$ con coordenadas,

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

en la base ortonormal de V y W respectivamente (estamos considerando abuso de notación). Luego se tiene que,

$$\begin{aligned}
(y, Ax)_W = (Ax)^* \cdot y &= \sum_{i=1}^{m} \overline{(Ax)}_i y_i \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} \overline{A_{ij}x_j} y_i = \sum_{i=1}^{m}\sum_{j=1}^{n} y_i \overline{A_{ij}x_j} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} y_i A^*_{ji}\overline{x_j} = \sum_{j=1}^{n} \overline{x}_j \sum_{i=1}^{m} A^*_{ji} y_i \\
&= \sum_{j=1}^{n} \overline{x}_j \left(A^* y\right)_j = x^* \cdot A^* y = (A^*y, x)_V
\end{aligned}$$

Por lo tanto $(y, Ax)_W = (A^*y, x)_V$.

Demostrando la proposición. ■

Nota: La propiedad más importantes del resultado anterior es que para todo $x \in V$ y $y \in W$ se cumple,

$$(y, Ax)_W = (A^*y, x)_V$$

Para cualquier producto interno definidos en W y en V. Por lo que no es necesario determinar una base ortonormal.

Teorema: *Sean V, W EVPI. Sea $A : V \to W$ operador lineal, entonces se cumple,*

(1) $KerA^* = (ImA)^\perp$

(2) $KerA = (ImA^*)^\perp$

(3) $ImA = (KerA^*)^\perp$

(4) $ImA^* = (KerA)^\perp$

Demostración: Vamos a demostrar el teorema en dos partes. En la primera demostraremos la afirmación (2), en la segunda a partir de (2), demostraremos el resto de afirmaciones.

Primera parte: Sea $x \in \mathrm{Ker}A$. Sea $z \in \mathrm{Im}A^*$, luego existe $y \in W$ tal que $A^*y = z$, entonces,

$$(z, x)_V = (A^*y, x)_V = (y, \underbrace{Ax}_{=0})_W = 0$$

Por lo tanto, para todo $z \in \mathrm{Im}A^*$ se cumple $(z, x)_V = 0$, es decir, $x \in (\mathrm{Im}A^*)^\perp$ y luego $\mathrm{Ker}A \subseteq (\mathrm{Im}A^*)^\perp$.

Probemos la igualdad. Sea $x \in (\mathrm{Im}A^*)^\perp$, entonces para todo $z \in \mathrm{Im}A^*$ se cumple que,

$$(z, x)_V = 0$$

Sea $y \in W$, entonces $A^*y \in \mathrm{Im}A^*$, y luego se cumple,

$$0 = (A^*y, x)_V = (y, Ax)_W$$

Por lo tanto, para todo $y \in W$ se cumple que,

$$(y, Ax)_W = 0$$

Es decir, $Ax = 0$ y por tanto $x \in \mathrm{Ker}A$. Demostrando que $\mathrm{Ker}A = (\mathrm{Im}A^*)^\perp$.

Segunda parte: Observemos lo siguiente,

$$\mathrm{Ker}A = (\mathrm{Im}A^*)^\perp \Leftrightarrow (\mathrm{Ker}A)^\perp((\mathrm{Im}A^*)^\perp)^\perp = \mathrm{Im}A^*$$

Por tanto $\mathrm{Im}A^* = (\mathrm{Ker}A)^\perp$. Para concluir el resto de afirmaciones basta considerar $(A^*)^* = A$.

Demostrando el teorema. ■

Ejemplo: Consideremos la matriz,

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 2 & 4 & 3 \end{bmatrix}$$

Queremos encontrar la proyección ortogonal de los siguientes subespacios: $\mathrm{Ker}A$, $\mathrm{Im}A$, $\mathrm{Ker}A^*$ y $\mathrm{Im}A^*$.

Para ello usaremos el teorema anterior. Si A es una matriz real, entonces $A^* = A^T$, por lo que,

$$A^* = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 3 & 4 \\ 1 & 2 & 3 \end{bmatrix}$$

- **Kernel de A:** Tenemos que determinar la imagen de A^*. Notemos que,

$$A^* \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y+2z \\ x+3y+4z \\ x+2y+3z \end{bmatrix} = x \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{=:v_1} + y \underbrace{\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}}_{=:v_2} + z \underbrace{\begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}}_{=:v_3}$$

  Por lo que,

$$\mathrm{Im}A^* = \langle v_1, v_2, v_2 \rangle$$

  Observemos que $v_1 + v_2 = v_3$, es decir que la colección $\{v_1, v_2, v_3\}$ es linealmente dependiente y debemos eliminar v_3 obteniendo la colección independiente $\{v_1, v_2\}$. Por lo tanto,

$$(\mathrm{Ker}A)^\perp = \mathrm{Im}A^* = \left\langle \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \right\rangle$$

- **Imagen de A:** Determinemos el Kernel de A^*. Por definición,

$$A^* \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$\Leftrightarrow x + z = 0, \ \ y + z = 0$$

  Por lo tanto,

$$\mathrm{Ker}A^* = \left\langle \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\rangle$$

  De esta forma,

$$(\mathrm{Im}A)^\perp = \mathrm{Ker}A^* = \left\langle \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\rangle$$

- **Kernel e imagen de A^*:** Para determinar la proyección ortogonal del kernel e imagen de A^* vamos a realizar el siguiente análisis. Sean $x, y \in \mathbb{R}^n$, entonces la proyección ortogonal de x sobre y es:

$$\begin{aligned} P_y x &= \underbrace{\frac{(x,y)}{\|y\|^2}}_{\text{escalar}} y \\ &= y\frac{(x,y)}{\|y\|^2} \\ &= y\frac{y^T x}{\|y\|^2} \\ &= \frac{1}{\|y\|^2} y y^T x \end{aligned}$$

Observemos que yy^T es una matriz independiente de x. Por lo tanto P_y tiene matriz asociada $\frac{1}{\|y\|^2} yy^T$. Entonces la matriz de la proyección ortogonal $P_{\text{Ker}A^T}$ es,

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

Dimensiones: Consideremos $A : \mathbb{R}^n \to \mathbb{R}^m$. Entonces se satisface que,

$$\begin{aligned} \text{número de columnas de } A = n &= \dim \text{Im}A + \dim \text{Ker}A \\ \text{número de filas de } A = m &= \dim \text{Im}A^T + \dim \text{Ker}A^T \end{aligned}$$

Usando el hecho de que,

$$\begin{aligned} \mathbb{R}^n &= \text{Im}A \oplus \text{Ker}A^T \\ \mathbb{R}^m &= \text{Im}A^T \oplus \text{Ker}A \end{aligned}$$

Entonces por otro lado se cumple,

$$\begin{aligned} n &= \dim \text{Im}A^T + \dim \text{Ker}A \\ m &= \dim \text{Im}A + \dim \text{Ker}A^T \end{aligned}$$

Por lo tanto se cumple que,

$$\begin{aligned} \dim \text{Im}A &= \dim \text{Im}A^T \\ \dim \text{Ker}A &= \dim \text{Ker}A^T \end{aligned}$$

Volviendo al ejemplo ocurre que $m = n = 3$. Si sabemos que $\dim \text{Ker}A^* = 1$, entonces $\dim \text{Ker}A = 1$. Determinemos su generador. Por definición,

$$A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow x - y = 0,\ z + 2y = 0$$

Por lo tanto,

$$(\mathrm{Im}A^*)^{\perp} = \mathrm{Ker}A = \left\langle \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\rangle$$

Ahora podemos determinar la matriz asociada a$P_{\mathrm{Ker}A}$,

$$\frac{1}{6}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}\begin{bmatrix} 1 & 1 & -2 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 1 & 1 & -2 \\ 1 & 1 & -2 \\ -2 & -2 & 4 \end{bmatrix}$$

Ahora para determinar el espacio ortogonal a KerA^*, debemos notar que $\mathbb{R}^3 = \mathrm{Ker}A \oplus \mathrm{Im}A^T = \mathrm{Ker}A^T \oplus \mathrm{Im}A$. Entonces se cumple que,

$$P_{\mathrm{Ker}A} + P_{\mathrm{Im}A^T} = \mathrm{I}$$
$$P_{\mathrm{Ker}A^T} + P_{\mathrm{Im}A} = \mathrm{I}$$

Por lo tanto,

$$P_{\mathrm{Im}A^T} = \mathrm{I} - P_{\mathrm{Ker}A} = \frac{1}{6}\begin{bmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & -4 \end{bmatrix}$$
$$P_{\mathrm{Im}A} = \mathrm{I} - P_{\mathrm{Ker}A^T} = \frac{1}{3}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

Finalmente para obtener el generador de ImA basta con tomar dos vectores x_1, x_2 tales que $P_{\mathrm{Im}A}x_1, P_{\mathrm{Im}A}x_2 \in \mathrm{Im}A$ sean linealmente independiente. En particular podemos tomar e_1, e_2 de la base canónica, y finalmente obtener

$$(\mathrm{Ker}A^*)^{\perp} = \mathrm{Im}A = \left\langle \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\rangle$$

3.7. Isometrías

Definición: *Sea X un conjunto no vacío. Una función* ***distancia*** $d : X \times X \to [0, \infty)$ *con respecto a X, es una función que satisface:*

(1) ***Simetría:*** $d(x, y) = d(y, x)$ *para todo* $x, y \in X$.

(2) ***Desigualdad triangular:*** $d(x, y) \leq d(x, z) + d(z, y)$ *para todo* $x, y, z \in X$.

(3) $d(x, y) = 0$ *si y sólo si* $x = y$.

Al par (X, d) le llamamos ***espacio métrico****.*

Ejemplo: Sea X un EVPI. Definimos la función $d(x, y) := \|x - y\|$

Definición: *Sean $(X, d_X), (Y, d_Y)$ espacio métricos. Sea $f : X \to Y$ una función. Decimos que f es una **isometría** si preserva las distancias, es decir, para todo $x, y \in X$ se cumple que,*

$$d_Y(f(x), f(y)) = d_X(x, y)$$

Observación: Si f es isometría, entonces f es inyectiva puesto que si $f(x) = f(y)$, entonces,

$$0 = d_Y(f(x), f(y)) = d_X(x, y)$$

Es decir, $x = y$.

Lema: *Sea V un EVPI. Definimos $d(x, y) := \|x - y\|$ para todo $x, y \in V$. Entonces (V, d) es un espacio métrico.*

Demostración: Claramente d es una función de la forma $d : V \times V \to [0, \infty)$. Probemos los tres axiomas de distancia.

(1) Sean $x, y \in V$, entonces,

$$d(x, y) = \|x - y\| = \|y - x\| = d(y, x)$$

Probando la simetría.

(2) Sean $x, y, z \in V$, entonces por definición de norma se cumple que,

$$d(x, y) = \|x - y\| = \|(x - z) + (z - y)\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$$

Probando la desigualdad triangular.

(3) Observemos que,

$$d(x, y) = 0 \Leftrightarrow \|x - y\| = 0 \Leftrightarrow x = y$$

Por lo tanto d es una función distancia y (V, d) es un espacio métrico. ■

Nota: Sea V un EVPI. La distancia definida por $d(x, y) := \|x - y\|$ para todo $x, y \in V$ le llamaremos **distancia inducida**.

Lema: *Sean V, W EVPI y sea $A : V \to W$ operador lineal. Entonces las siguientes afirmaciones son equivalentes:*

(1) A es isometría.

*(2) A **preserva la norma:** Para todo $x \in V$ se cumple $\|Ax\| = \|x\|$.*

*(3) A **preserva el producto interno:** Para todo $x, y \in V$ se tiene que $(Ax, Ay)_W = (x, y)_V$.*

Demostración:

- **(1) es equivalente (2):** Supongamos que A es isometría, entonces se cumple que,

  $$\|Ax - Ay\| = \|x - y\|$$

  para todo $x, y \in V$. Tomando $y = 0$ se obtiene que $\|Ax\| = \|x\|$ para todo $x \in V$.

  Probemos la otra dirección. Si para todo $x \in V$ se tiene que $\|Ax\| = \|x\|$, entonces,

  $$\|A(x - y)\| = \|x - y\|$$

  Entonces claramente A es isometría.

- **(2) es equivalente (3):** Primero supongamos que A preserva el producto interno. Entonces, para todo $x \in V$ se tiene que,

  $$\|Ax\|^2 = (Ax, Ax)_W = (x, x)_V = \|x\|^2$$

  Por lo tanto $\|Ax\| = \|x\|$ para todo $x \in V$.

  Para la otra dirección usaremos la polarización del producto interno. Por lo que debemos probar para el caso real y el caso imaginario.

  Recordatorio: Un operador lineal es una función definida sobre dos espacios vectoriales definido en un mismo cuerpo que cumple la propiedad lineal.

  Caso real: Supongamos que se preserva la norma y que V, W son EVPI definidos sobre $\mathbb{R}$. Entonces se tiene que,

  $$\begin{aligned}(Ax, Ay)_W &= \frac{1}{4}(\|Ax + Ay\|^2 - \|Ax - Ay\|^2)\\ &= \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) = (x, y)_V\end{aligned}$$

  Caso complejo: Supongamos que se preserva la norma y que V, W son EVPi definidos sobre $\mathbb{C}$. Entonces se tiene que,

  $$\begin{aligned}(Ax, Ay)_W &= \frac{1}{4}(\|Ax + Ay\|^2 - \|Ax - Ay\|^2 + i\|Ax + iAy\|^2 - i\|Ax - iAy\|^2)\\ &= \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2) = (x, y)_V\end{aligned}$$

  De forma que se preserva el producto interno en general.

Demostrando el lema. ■

Proposición: *Sea V EVPI. Sean E, F de tal forma que están en suma direcra ortogonalmente con respecto a V, es decir, $V = E \oplus F$ y $E \perp F$. Sean $\alpha, \beta \in \mathbb{C}$ tales que $|\alpha| = |\beta| = 1$. Definimos,*

$$A_{\alpha\beta} := \alpha P_E + \beta P_F$$

Entonces $A_{\alpha\beta}$ es isometría.

Demostración: Sea $x \in V$, supongamos que $\{e_1, \cdots, e_m\}$ es base ortonormal de E y $\{f_1, \cdots, f_s\}$ es base ortonormal de F, entonces x se puede escribir de la siguiente forma:

$$x = \alpha_1 e_1 + \cdots + \alpha_m e_m + \beta_1 f_1 + \cdots + \beta_s f_s \in V$$

Luego se tiene que,

$$\begin{aligned}\|A_{\alpha\beta}x\|^2 &= \|\alpha P_E x + \beta P_F x\|^2 \\ &= |\alpha|^2 \|P_E x\| + |\beta|^2 \|P_F x\|^2 \\ &= \|P_E x\|^2 + \|P_F x\|^2\end{aligned}$$

Dado que $E \perp F$. Notemos que,

$$\begin{aligned}P_E x &= \alpha_1 e_1 + \cdots + \alpha_m e_m \\ P_F x &= \beta_1 f_1 + \cdots + \beta_s f_s\end{aligned}$$

Puesto que $e_i \perp f_j$ para todo $i = 1, \cdots, m$ y $j = 1, \cdots, s$. Luego se tiene que,

$$\begin{aligned}\|P_E x\|^2 + \|P_F x\|^2 &= \sum_{i=1}^{m} |\alpha_i|^2 \|e_i\|^2 + \sum_{i=1}^{s} |\beta_i| \|f_i\|^2 \\ &= \|\alpha_1 e_1 + \cdots + \alpha_m e_m + \beta_1 f_1 + \cdots + \beta_s f_s\|^2 = \|x\|^2\end{aligned}$$

Aplicando las base ortonormales y que $e_i \perp f_j$. Por lo tanto,

$$\|A_{\alpha\beta}x\| = \|x\|$$

para todo $x \in V$, es decir, $A_{\alpha\beta}$ es isometría. ■

Definición: *Sea X, Y EVPI. Sea $U : X \to Y$ operador lineal. Decimos que es un operador **unitario** si es isometría y es invertible.*

Afirmación: *Un operador es isométrico si y sólo si $U^*U = I$.*

Demostración: Sean d_X, d_Y las medidas inducidas con respecto a los EVPI X e Y. Supongamos que U es isometría, luego se tiene que,

$$(Ux, Uy)_Y = (x, y)_X \Leftrightarrow (x, U^*Uy)_X = (x, y)_X$$

para todo $x, y \in X$. Aquí concluimos que $U^*U = I$.

Supongamos ahora que $U^*U = I$. Entonces para todo $x, y \in X$ se tiene que,

$$(x, y)_X = (x, U^*Uy)_X = (Ux, Uy)_Y$$

Por lo tanto U es isometría. ■

Definición: *Sea M una matriz. Decimos que es*

- ***Unitario:*** *Si $M^*M = I$ y M es cuadrado.*
- ***Ortogonal:*** *Si es unitario y tiene entradas reales.*

Observación: Si M es ortogonal, entonces M es invertible y en particular, $M^* = M^{-1}$. Si M tiene entradas reales, se concluye que $M^T = M^{-1}$.

Lema: *Sean X, Y EVPI. Sea $U : X \to Y$ operador lineal. Entonces U es unitario si y sólo si es isometría y $\dim X = \dim Y$.*

Demostración: Supongamos que U es isometría tal que $\dim X = \dim Y$. Sea M matriz asociada a U, entonces se cumple que $M^*M = I$ al ser isometría, si M es cuadrado, entonces claramente es invertible, es más,

$$M^* = M^{-1}$$

Esto implica que U es invertible y por tanto es unitario.

Supongamos ahora que U es unitario, entonces claramente es isometría. Si U es invertible, se tiene que X e Y están en isomorfismo, es decir, $\dim X = \dim Y$. ■

Observación: Sea $U : \mathbb{C}^n \to \mathbb{C}^m$ operador. Si U es isometría, entonces se cumple que si M es matriz asociada, entonces $M^*M = I$, es decir,

$$(\text{columna } i \text{ de } M)^*(\text{columna } j \text{ de } M) = \delta_{ij}$$

Por lo tanto, las columnas de la matriz asociada a U son ortonormales, es más, las columnas de U forman una base ortonormal en $\mathbb{C}^n$.

Afirmación: *Si la matriz asociada a U son ortonormales, entonces U es isometría.*

Demostración: Sea M matriz asociada a U. Observemos que para todo $i = 1, \cdots, n$ y $j = 1, \cdots, m$ se cumple que,

$$(Ue_i, Ue_j)_{\mathbb{C}^m} = \delta_{ij} = (e_i, e_j)$$

puesto que Ue_i es la columna i de U. Por lo tanto, en la base canónica se cumple que,

$$(Ux, Uy)_{\mathbb{C}^m} = (x, y)_{\mathbb{C}^n}$$

Es decir, U es isometría. ■

Lema: *Sean $U : X \to Y$, $V : Y \to Z$ operadores lineales. Entonces,*

(1) U es unitario si y sólo si $U^{-1} = U^*$.

(2) Si U es unitario, entonces U^{-1} es unitario.

(3) Si U es isometría y $\{v_1, \cdots, v_n\}$ es una base ortonormal, entonces $\{Uv_1, \cdots, Uv_n\}$ son ortonormales. En particular, si U es unitario, entonces $\{Uv_1, \cdots, Uv_n\}$ es base ortonormal de Y.

(4) Si U, V son isometrías, entonces $V \circ U$ es isometría.

(5) Si U, V son unitarios, entonces $V \circ U$ es unitario.

terminar Demostración:

-

(2) Si U es unitario, entonces $U^* = U^{-1}$, en particular,

$$U^*U = UU* = I$$

Entonces U^* es isometría y además es invertible, por lo que U^{-1} es unitario.

(3) Sea U isometría y $\{v_1, \cdots, v_n\}$ base ortonormal de X. Luego,

$$(Uv_i, Uv_j)_Y = (v_i, v_j)_X = 0$$
$$(Uv_i, Uv_i)_Y = (v_i, v_i)_X = 1$$

Para todo $i \neq j$. Por lo que $\{Uv_1, \cdots, Uv_n\}$ es ortonormal. Si además U es unitario, entonces $\dim X = \dim Y$, luego $\{Uv_1, \cdots, Uv_n\}$ es base ortonormal de Y.

(4) Sean U, V isometría, luego para todo $x, y \in X$ se tiene que,

$$(VUx, VUy)_Z = (Ux, Uy)_Y = (x, y)_X$$

Luego $VU = V \circ U$ es isometría.

- Sean U, V unitarios. Claramente $V \circ U$ es isometría, y para ver la invertibilidad se observa que,

$$VU(VU)^* = VUU^*V = I$$
$$(VU)^*VU = U^*V^*VU = I$$

Es decir, $(VU)^* = (VU)^{-1}$. Por lo tanto $VU = V \circ U$ es unitario.

-

Ejemplo: Consideremos la matriz,

$$\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}$$

Observemos que es unitario puesto que las columnas son ortonormales y además es una matriz cuadrada.

Ejemplo: Sea $U : X \to Y$ operador lineal donde $\dim X = \dim Y = n$. Si $\{x_1, \cdots, x_n\}, \{y_1, \cdots, y_n\}$ son bases ortonormales de X e Y respectivamente tales que $Ux_j = y_j$ para todo $j = 1, \cdots, n$. Se puede verificar que U es unitario. La isometría se cumple puesto que,

$$(Ux_i, Ux_j)_Y = (y_i, y_j) = \delta_{ij} = (x_i, x_j)$$

para todo $i, j = 1, \cdots, n$. y dado que la dimensión de X e Y es igual, se tiene que U es unitario.

Proposición: *Sea $U : X \to Y$ operador lineal unitario. Entonces,* *(1)* $|\det U| = 1$. *Si además es ortogonal, entonces* $\det U = \pm 1$.

(2) Si λ *es valor propio de* U*, entonces* $|\lambda| = 1$.

Demostración:

(1) Si U es unitario, entonces $U^*U = I$, luego,

$$1 = \det(U^*U) = \det U^* \det U$$

Afirmación: $\det U^* = \overline{\det U}$.

Demostración: Usando la fórmula de determinante sobre U^* se observa que,

$$\begin{aligned}\det U^* &= \det \overline{U}^T = \det \overline{U} \\ &= \sum_{P \in S_n} \prod_{j=1}^{n} \overline{U}_{jP(j)} \mathrm{sgn}(P) \\ &= \overline{\sum_{P \in S_n} \prod_{j=1}^{n} U_{jP(j)} \mathrm{sgn}(P)} = \overline{\det U}\end{aligned}$$

Por lo tanto,

$$\det U^* = \overline{\det U}$$

■

De esta forma se tiene que,

$$1 = \det U^* \det U = \overline{\det U} \det U = |\det U|^2$$

Por lo que $|\det U| = 1$.

(2) Sea λ valor propio. Entonces $Ux = \lambda x$ con x vector propio no nulo. Luego se cumple que,

$$\|x\| = \|Ux\| = \|\lambda x\| = |\lambda| \|x\|$$

Por tanto, $|\lambda| = 1$.

Demostrando la proposición. ■

Definición: *Sean* A, B *matrices cuadradas. Decimos que son* ***unitariamente equivalentes*** *si y sólo si existe* U *unitario tal que,*

$$A = U^*BU$$

Proposición: *Sea* A *una matriz de* $n \times n$*. Entonces* A *es unitariamente equivalente a una matriz* D *diagonal si y sólo si hay una base ortogonal de autovectores de* A*.*

Demostración: Supongamos que A es unitaramente equivalente a una matriz diagonal D. Es decir, existe U tal que,

$$A = U^* D U$$

En particular, D está compuesto por valores propios de A y U está conformodado por vectores ortogonales y $U^* = U^{-1}$, si además las columnas de U son vectores propios de A, entonces existe una base ortogonal de vectores propios de A.

Supongamos que existe una base ortogonal de vectores propios de A, entonces existe una matriz X con columnas vectores propios de A y ortogonales tal que,

$$A = X^{-1} D X$$

donde D es diagonal con elementos los valores propios de A. Veamos que X es unitario. Observemos que,

$$(Xe_i, Xe_j) = \delta_{ij} = (e_i, e_j)$$

para todo $i, j = 1, \cdots, n$. Por tanto X es isométrico y luego es unitario. Es decir, $X^* = X^{-1}$ y por tanto se cumple que,

$$A = X^* D X$$

Es decir, A es unitariamente equivalente a una matriz diagonal. ■

Observación: Sean A, Bunitariamente equivalente, entonces existe U unitario tal que,

$$A = U^* B U$$

El determinante de A es igual a B puesto que,

$$\begin{aligned} \det A &= \det U^* \det B \det U \\ &= |\det U| \det B = \det B \end{aligned}$$

Y los valores propios de A son iguales a B, puesto que si λ es valor propio de A, se tiene que,

$$Av = \lambda v \Leftrightarrow U^* B U v = \lambda v \Leftrightarrow B(Uv) = \lambda(Uv)$$

Donde $Uv \neq 0$ puesto que U es invertible y $v \neq 0$.

Ejemplo: Verifiquemos que matrices son unitariamente equivalente.

- Sean las matrices,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix}$$

  Los determinante son:

$$\det A = 1, \quad \det B = -1, \quad \det C = -\frac{1}{4}$$

  Entonces claramente ninguno es unitariamente equivalente con el otro.

- Sean las matrices,
$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
Luego los determinantes son:
$$\det A = 2, \quad \det B = 0$$
Por lo que no son unitaramente equivalente.

- **revisar** Sean las matrices,
$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & i \end{bmatrix}$$
Podemos ver que $\det A = \det B = 1$. Veamos que pasa con los valores propios. Observemos que,
$$p_A(x) = (1 - x)(x^2 + 1) = p_B(x)$$
Es decir, A y B tienen los mismos valores propios. Verifiquemos que son unitaramente equivalentes. Para ello primero encontremos U invertible tal que,
$$A = UBU^{-1} \Leftrightarrow AU^{=}U^{B}$$
Si $U^{=} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$, entonces,
$$\begin{aligned} Av_1 &= v_1 \\ Av_2 &= -iv_2 \\ Av_3 &= iv_3 \end{aligned}$$
Es decir, U está formado por vectores propios. Necesitamos determinar U tal que $U^* = U^{-1}$. En particular U tiene dos formas posibles
$$\begin{bmatrix} 0 & \gamma & \beta \\ 0 & -i\gamma & i\beta \\ \alpha & 0 & 0 \end{bmatrix} \quad \text{ó} \quad \begin{bmatrix} 0 & \gamma & \beta \\ 0 & i\gamma & -i\beta \\ \alpha & 0 & 0 \end{bmatrix}$$
Tomemos la segunda forma. Sean α, γ, β tales que las columnas de sean ortonormales, entonces,
$$\begin{aligned} |\alpha| &= 1 \\ |\gamma|^2 + |i\gamma|^2 &= 1 \\ |\beta|^2 + |i\beta|^2 &= 1 \end{aligned}$$
Y que,
$$\overline{\beta}\gamma - \overline{\gamma} = 0$$
Entonces $\alpha = 1, \beta = \gamma = \frac{1}{\sqrt{2}}$ obtenemos U unitario y por tanto,
$$A = U^*BU$$
Es decir, A y B son unitariamente equivalentes.

3.8. Triangularización y Diagonalización sobre productos internos

Teorema: *Sea $A : X \to X$ operador lineal donde X es un EVPI sobre $\mathbb{C}$. Entonces existe una base ortonormal $u_1, \cdots, u_n$ de X donde A tiene matriz triangular superior.*

Demostración: Demostraremos por inducción sobre la dimensión de X.

Caso base: Si $\dim X = 1$, entonces es evidente que toda matriz es triangular superior.

Caso inductivo: Supongamos que todo operador sobre un espacio de dimensión $n-1$, hay una base ortonormal de X tal que el operador tiene fora triangular superior.

Consideremos $A : X \to X$ con $\dim X = n$. Sabemos que A tiene n valores propios, sea $\lambda_1 \in \mathbb{C}$ valor propio con $u_1 \neq 0$ vector propio, por lo que $Au_1 = \lambda_1 u_1$. Sin pérdida de generalidad supongamos que $\|u_1\| = 1$. Sea $E_1 := \langle u_1 \rangle^{\perp}$, y sean $v_2, \cdots, v_n$ una base ortonormal de E_1.

Observemos que $u_1, v_2, \cdots, v_n$ es base de X, por lo que A tiene asociado una matriz en esta base. Si $Au_1 = \lambda_1 u_1$. En particular, la matriz asociada de A en esta base tiene forma:

$$\begin{bmatrix} \lambda_1 & * \\ 0 & A_1 \end{bmatrix}$$

donde $*$ son elementos cualquiera y A_1 es una matriz. Observemos que A_1 es un operador que manda E_1 a si mismo, si $\dim E_1 = n-1$, entonces por la hipótesis inductiva existe una base ortonormal $u_2, \cdots, u_n$ tal que A_1 tiene forma triangular, finalmente reemplazamos esa forma triangular en la matriz asociada de A para obtener,

$$\begin{bmatrix} \lambda_1 & * \\ 0 & T_1 \end{bmatrix}$$

Donde T_1 es una matriz triangular superior. Por lo tanto la base existe una base ortonormal $u_1, \cdots, u_n$ tal que la matriz asociada a A es triangular superior. ■

Observación: Si A es operador lineal y u_1 es valor propio, entonces $\langle u_1 \rangle^{\perp}$ no es invariante en general, es decir,

$$A(\langle u_1 \rangle^{\perp}) \not\subseteq \langle u_1 \rangle^{\perp}$$

en general.

Teorema: *Sea $A : X \to X$ operador lineal donde X es un EVPI sobre $\mathbb{R}$ de dimensión n. Supongamos que todos los valores propios de A son reales, en este caso existe base ortonormal tal que A es de forma triangular superior en esta base.*

Demostración:

Teorema: *Sea $A : X \to X$ operador lineal con X EVPI. Si A es autoadjunto entonces,*

(1) Todos los valores propios de A son reales.

(2) Existe una base ortonormal de vectores propios de A.

Demostración:

(1) Sea $u_1, \cdots, u_n$ base ortonormal de X tal que A tiene matriz triangular T.

Afirmación: *El operador adjunto A^* en la base $u_1, \cdots, u_n$ tiene matriz asociada T^*.*

Demostración: Sea X la matriz de cambio de coordenada, por lo que,

$$A = X^{-1}TX$$

Entonces,

$$A^* = X^*T^*(X^{-1})^*$$

terminar

■

Entonces se cumple uqe,

$$T_{ji} = (Au_j, u_i)$$

terminar

Finalmente se tiene que $T = T^*$, es decir que la diagonal de T contiene solo valores reales y por tanto los valores propios de A son reales.

(2)

Corolario: *Si M es una matriz autoadjunto, entonces M es unitariamente equivalente a una matriz diagonal con entradas reales.*

Terminar

Observación: Sea A operador autoadjunto y λ un valor propio, luego,

$$(Ax, x) = (\lambda x, x) = \lambda \|x\|^2$$

Por otro lado,

$$(Ax, x) = (x, A^*x) = (x, Ax) = (x, \lambda x) = \overline{\lambda}\|x\|^2$$

Entonces $\overline{\lambda} = \lambda$, es decir, los valores propios de A son reales. Esto se satisface en espacios vectoriales de dimensión infinita.

Proposición: *Sea A operador autoadjunto. Vectores propios de valores propios distintos, son ortogonales.*

Demostración: Sean λ, μ valores propios distintos. Sean v_λ, v_μ vectores propios, luego se tiene que,

$$\begin{aligned}
\lambda(v_\lambda, v_\mu) &= (\lambda v_\lambda, v_\mu) \\
&= (Av_\lambda, v_\mu) \\
&= (v_\lambda, A^*v_\mu) \\
&= (v_\lambda, Av_\mu) \\
&= (v_\lambda, \mu v_\mu) \\
&= \overline{\mu}(v_\lambda, v_\mu)
\end{aligned}$$

Dado que A es autoadjunto, se tiene que todos los valores propios son reales y por tanto,

$$\lambda(v_\lambda, v_\mu) = \mu(\lambda v_\lambda, v_\mu) \Leftrightarrow (\lambda - \mu)(v_\lambda, v_\mu) = 0$$

Si $\lambda \neq \mu$, entonces se tiene que $(v_\lambda, v_\mu) = 0$. Por lo tanto, los vectores propios de λ y de μ son ortonormales. ■

Definición: *Sea $N : X \to X$ un operador lineal con X un EVPI. Decimos que es un **operador normal** si $NN^* = N^*N$.*

Ejemplo:

- Todo operador N diagonal es normal. Puesto que si $N = \text{diag}(\lambda_1, \cdots, \lambda_n)$, entonces

$$NN^* = \text{diag}(|\lambda_1|^2, \cdots, |\lambda_n|^2) = N^*N$$

- Si N es un operador autoadjunto, entonces es normal puesto que,

$$NN^* = N2 = N^*N$$

- Si $N^* = -N$, entonces es normal puesto que,

$$NN^* = -N^2 = N^*N$$

- Si $N^* = AN$ donde A conmuta con N. Entonces N es autoadjunto puesto que,

$$NN^* = NAN = ANN = N^*N$$

Teorema: *Sea $N : X \to X$ un operador normal sobre X un EVPI sobre $\mathbb{C}$. Entonces exsite un base ortonormal de vectores propios de N.*

Demostración: Como X es un EVPI definido en $\mathbb{C}$, entonces existe U unitario tal que,

$$N = U^*TU$$

donde T es matriz triangular. Si demostramos que T es una matriz diagonal, entonces N tiene una base ortonormal de vectores propios de N.

Afirmación: *T es una matriz diagonal.*

Demostración: Si N es un operador normal, entonces,

$$\begin{aligned} NN^* = N^*N &\Leftrightarrow (U^*TU)(U^*TU)^* = (U^*TU)^*(U^*TU) \\ &\Leftrightarrow U^*TT^*U = U^*T^*TU \\ &\Leftrightarrow TT^* = T^*T \end{aligned}$$

Es decir, T es normal. Demostremos que T es diagonal. Supongamos que $T = (a_{ij})_{i,j=1^n}$ donde $a_{ij} = 0$ si $i > j$. Estudiemos $(TT^*)_{11}, (T^*T)_{11}$. Por un lado se tiene,

$$\begin{aligned}(TT^*)_{11} &= \sum_{j=1}^{n} a_{1j}(a_{j1})^* \\ &= \sum_{j=1}^{n} |a_{1j}|^2\end{aligned}$$

Y por otro lado,

$$\begin{aligned}(T^*T)_{11} &= \sum_{j=1}^{n}(a_{1j})^* a_{j1} \\ &= \sum_{j=1}^{n} |a_{j1}|^2 = |a_{11}|^2\end{aligned}$$

Por lo tanto,

$$(TT^*)_{11} = (T^*T)_{11} \Leftrightarrow \sum_{j=1}^{n} |a_{1j}|^2 = |a_{11}|^2$$

Esto implica que $a_{1j} = 0$ para todo $j = 2, \cdots, n$. Estudiemos ahora $(TT^*)_{22}, (T^*T)_{22}$. Por un lado tenemos que,

$$(TT^*)_{22} = |a_{22}|^2$$

Y por otro lado,

$$(TT^*)_{22} = \sum_{j=2}^{n} |a_{2j}|^2$$

Entonces concluimos que $a_{2j} = 0$ para todo $j = 3, \cdots, n$. Iterativamente concluimos que $a_{ij} = 0$ para todo $j < i$ para todo $i = 1, \cdots, n$. Por lo que T es una matriz triangular superior e inferior, es decir, T es diagonal. ■

Finalmente tenemos que N es unitariamente equivalente a una matriz diagonal, por lo tanto existe una base ortonormal de vectores propios de N. ■

Ejemplo: Consideremos la matriz,

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

Vamos a demostrar que N es normal mediante otro método sin verificar de forma directa.

De forma general, sea N un operador normal y sea $\lambda \in \mathbb{C}$ una constante.

Afiramción: *$N + \lambda I$ es un operador normal.*

Demostración: Por definición,

$$\begin{aligned}(N + \lambda I)(N + \lambda I)^* &= (N + \lambda I)(N^* + \lambda^* I) \\ &= NN^* + \lambda^* N + \lambda N^* + \lambda\lambda^* \\ &= N^*N + \lambda^* N + \lambda N^* + \lambda\lambda^* \\ &= (N + \lambda I)^*(N + \lambda I)\end{aligned}$$

Luego $N + \lambda I$ es un operador lineal. ■

Siguiendo con el ejemplo, observemos que,

$$N = \underbrace{\begin{bmatrix} 0 & -3 & -1 \\ 3 & 0 & -3 \\ 1 & 3 & 0 \end{bmatrix}}_{=:B} + 2I$$

Se puede ver que $B = -B^*$, entonces B es normal y por tanto, $N = B + 2I$ es normal.

Proposición: *Sea $N : X \to X$ operador normal sobre X un EVPI. Entonces para todo $x \in X$ se tiene que $\|Nx\|^2 = \|N^*x\|^2$.*

Demostración: Sea $x \in X$, entonces,

$$\begin{aligned}\|Nx\|^2 &= (Nx, Nx) \\ &= (x, N^*Nx) \\ &= (x, NN^*x) \\ &= (N^*x, N^*x) = \|N^*x\|^2\end{aligned}$$

Demostrando la proposición. ■

Definición: *Sea $A : X \to X$ operador lineal autoadjunto. Decimos que A es **definido positivo** si para todo $x \in X$ no nulo se tiene que $(Ax, x) > 0$. Decimos que es **semidefinido posito** si para todo $x \in X$ no nulo se tiene que $(Ax, x) \geq 0$.*

Teorema: *Sea $A = A^*$ operador sobre un EVPI. Entonces,*

(1) $A > 0$ si y sólo si para todo valor propio de A se tiene que $\lambda > 0$.

(2) $A \geq 0$ si y sólo si para todo valor propio de A se tiene que $\lambda \geq 0$.

terminar

Lema: *Si $A = A^* \geq 0$, entonces hay un único $B = B^* \geq 0$ tal que $B^2 = A$. En tal caso denotamos $B = \sqrt{A}$.*

Demostrar: Si A es autoadjunto, entonces existe una base ortonormal de vectores propios de A, digamos que $u_1, \cdots, u_n$ es esta base con valores propios $\lambda_1, \cdots, \lambda_n$, donde $\lambda_j \geq 0$ para todo $j = 1, \cdots, n$ por el teorema anterior.

Definimos el operador lineal B por $Bu_j := \sqrt{\lambda_j}u_j$ para todo $j = 1, \cdots, n$.

Afirmación: *B es única.*

Demostración: trermnar

Definición: *Sean X, Y EVPI. Sea $A : X \to Y$ operador lineal. Definimos el módulo A por,*

$$|A| : X \to X$$
$$|A| := \sqrt{A^*A}$$

Afirmación: *El módulo está bien definido.*

Demostración: Observemos que,

$$(A^*A)^* = A^*A$$

Por lo que A^*A es autoadjunto, probemos que $A^*A \geq 0$. Sea λ valor propio de A^*A, entonces si x es valor propio de λ, se tiene que

$$(A^*Ax, x) = \lambda(x, x) = \lambda\|x\|^2$$

Por otro lado,

$$(A^*Ax, x) = (Ax, Ax) = \|Ax\|^2 \geq 0$$

Por lo que $\lambda \geq 0$, por tanto A^*A y por tanto el módulo de A está bien definido. ■

Afirmación: *Para todo $X \geq 0$ operador autadjunto, se satisface que $B = \sqrt{X}$ también es autoadjunto.*

Demostracióntermiar

Proposición: *Sean X, Y EVPI. Sea $A : X : Y$ un operador lineal, entonces para todo $x \in X$ se cumple que,*

$$\||A|x\| = \|Ax\|$$

Demostración: Por definición,

$$\||A|x\|^2 = (|A|x, |A|x) = (|A|^*|A|x, x)$$

Si $|A| \geq 0$ y A^*A es autoadjunto, entonces $|A| = |A|^*$, por lo que,

$$\||A|x\|^2 = (|A|^2x, x) = (A^*Ax, x) = (Ax, Ax) = \|Ax\|^2$$

Por lo tanto, para todo $x \in X$ se tiene que,

$$\||A|x\| = \|Ax\|$$

■

Proposición:

Teorema (Descomposición polar de matrices): *Sea $A : X \to X$ operador lineal (o matriz cuadrada), entonces existe U unitario tal que $A = U|A|$.*

terminar

Ejemplo:

3.9. Valores Singulares

falta algo

Definición: *Sea $A : X \to Y$ operador lineal entre EVPI. Los* ***valores singulares*** *de A son los valores propios de $|A|$.*

Notación: Sea A operador y $|A|$ su módulo. Denotaremos los valores singulares de A por $\sigma_1, \cdots, \sigma_n$.

Ejemplo: Sea $A : X \to X$ autoadjunto, entonces los valores propios de A son reales. Notemos que,

3.10. Movimiento rígido

4. Forma Bilineales y Cuadráticas

Definición: *Sea V un espacio vectorial real. Una* ***forma bilineal real*** *sobre V es una función $L : V \times V \to \mathbb{R}$ tal que $L = L(x, y)$ es lineal en las dos variables.*

Ejemplo: Sea $V = \mathbb{R}^n$.

- Sea $L_1(x, y) = (x, y)$ donde $(\cdot, \cdot)$ es un producto interno cualquiera. Claramente L_1 es una forma bilineal.
- Sea $A \in \mathbb{R}^{n\times n}$. Consideremos $L_2(x, y) = x^T Ay$, entonces L_2 es una forma bilineal debido a que $L_2(x, y) = (Ay, x)$.
- Consideremos $L_3(x, y) = x_1y_2$. Entonces es una forma bilineal debido a que en la primer argumento se cumple,

$$L_3(cx + z, y) = (cx_1 + z_1)y_2 = cx_1y_2 + z_1y_2 = cL_3(x, y) + L_3(z, y)$$

  Y en el segundo argumento se cumple,

$$L_3(x, cy + z) = x_1(cy_2 + z_2) = cx_1y_2 + z_1z_2 = cL_3(x, y) + L_3(x, z)$$

tetrm inar