Índice general

SEMANA 1:

Ingeniería Matemática
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Matrices

1.1 Definiciones básicas

DEFINICIÓN (MATRIZ) Una **matriz** A, de m filas y n columnas con coeficientes en el cuerpo $\mathbb{K}$ (en este apunte $\mathbb{K}$ será $\mathbb{R}$ ó $\mathbb{C}$) es una tabla de doble entrada:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \qquad a_{ij} \in \mathbb{K}, \quad \forall i = 1, \ldots, m, \ j = 1, \ldots, n.$$

Notamos también la matriz como $A = (a_{ij})$ y denominamos $\mathcal{M}_{mn}(\mathbb{K})$ al conjunto de todas las matrices de m filas y n columnas con coeficientes en el cuerpo $\mathbb{K}$.

DEFINICIÓN (IGUALDAD DE MATRICES) Dadas dos matrices $A \in \mathcal{M}_{mn}(\mathbb{K}), B \in \mathcal{M}_{m'n'}(\mathbb{K})$, diremos que son **iguales** si y sólo si:

- $m = m'$, $n = n'$, y
- Para todo $i = 1, \ldots, m$, $j = 1, \ldots, n$ se cumple $a_{ij} = b_{ij}$

Un caso especial de matrices son aquellas de $n \times 1$. Estas matrices se llamarán posteriormente **vectores** de $\mathbb{K}^n$. Así nuestros vectores serán matrices de una sola columna.

Construimos una estructura algebraica sobre $\mathcal{M}_{mn}(\mathbb{K})$ a partir de las operaciones definidas en el cuerpo $\mathbb{K}$. Se define la **suma** de dos matrices como sigue:

$$\forall A, B \in \mathcal{M}_{mn}(\mathbb{K}) : \quad A + B = (a_{ij} + b_{ij})$$

Por ejemplo, en $(\mathbb{R}, +, \cdot)$:

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

Es fácil verificar que

Proposición 1.1. $(\mathcal{M}_{mn}(\mathbb{K}), +)$ *tiene estructura de grupo Abeliano.*

Demostración. La suma es asociativa y conmutativa, esto se hereda de las operaciones en el cuerpo, dado que la suma es coordenada a coordenada

El **neutro aditivo** es

$$\mathbf{0} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} \in \mathcal{M}_{mn}(\mathbb{K}).$$

El **inverso aditivo** de $A = (a_{ij})$ es $-A = (-a_{ij})$. Por ejemplo, en $\mathcal{M}_{23}(\mathbb{C})$:

$$\begin{pmatrix} i & 0 & 0 \\ 1 & -i & 0 \end{pmatrix} + \begin{pmatrix} -i & 0 & 0 \\ -1 & i & 0 \end{pmatrix} = \begin{pmatrix} i-i & 0+0 & 0+0 \\ 1-1 & -i+i & 0+0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \mathbf{0}.$$

Luego $-\begin{pmatrix} i & 0 & 0 \\ 1 & -i & 0 \end{pmatrix} = \begin{pmatrix} -i & 0 & 0 \\ -1 & i & 0 \end{pmatrix}$. □

Definición (Producto de matrices) Dadas $A = (a_{ij}) \in \mathcal{M}_{mr}(\mathbb{K})$, $B = (b_{ij}) \in \mathcal{M}_{rn}(\mathbb{K})$ se define el **producto** $C = AB$ como aquella matriz $C \in \mathcal{M}_{mn}(\mathbb{K})$ tal que

$$c_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj}, \quad i = 1, \dots, m, \ j = 1, \dots, n.$$

Claramente $C \in \mathcal{M}_{mn}(\mathbb{K})$.

Ejemplo 1.1.

1. En $(\mathbb{R}, +, \cdot)$,

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \in \mathcal{M}_{23}(\mathbb{R}), \quad B = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 2 & -1 & 0 \\ 1 & 0 & -1 & 0 \end{pmatrix} \in \mathcal{M}_{34}(\mathbb{R})$$

$$\implies C = A \cdot B = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 2 & -1 & 0 \\ 1 & 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 6 & -4 & 1 \end{pmatrix} \in \mathcal{M}_{24}(\mathbb{R}).$$

2. En $(\mathbb{C}, +, \cdot)$,

$$A = (-i, \ i, \ 0) \in \mathcal{M}_{13}(\mathbb{C})$$

$$B = \begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix} \in \mathcal{M}_{31}(\mathbb{C})$$

$$\implies \quad C = AB = -i \cdot i + i \cdot 1 + 0 \cdot 0 = 1 + i \in \mathcal{M}_{11}(\mathbb{C}).$$

Observación: La multiplicación de matrices **no es conmutativa**. Por ejemplo en $\mathcal{M}_{22}(\mathbb{R})$:

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

Dos propiedades importantes de la multiplicación son las siguientes:

Proposición 1.2.

1. **Asociatividad:** *Si* $A \in \mathcal{M}_{mn}(\mathbb{K}), B \in \mathcal{M}_{nq}(\mathbb{K}), C \in \mathcal{M}_{qs}(\mathbb{K}),$ *entonces:*

$$A(BC) = (AB)C \in \mathcal{M}_{ms}(\mathbb{K}).$$

2. **Distributividad con respecto a la suma***: Dadas* $A \in \mathcal{M}_{mn}(\mathbb{K}), B, C \in \mathcal{M}_{ns}(\mathbb{K}),$ *entonces*

$$A(B + C) = AB + AC \in \mathcal{M}_{ms}(\mathbb{K}).$$

De igual manera se tiene la distributividad por el otro lado.

Demostración. Demostremos la distributibidad, quedando la asociatividad de ejercicio. Denominando $E = A(B + C)$, se tiene:

$$e_{ij} = \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) \quad \forall i \in \{1, ..., m\}, j \in \{1, ..., s\}.$$

Como la multiplicación distribuye con respecto a la suma en $\mathbb{K}$:

$$\begin{aligned} e_{ij} &= \sum_{k=1}^{n}(a_{ik}b_{kj} + a_{ik}c_{kj}) \\ &= \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}. \end{aligned}$$

De la definición de multiplicación matricial, se obtiene $E = AB + AC$. $\square$

Un caso particular muy importante es el de las matrices **cuadradas**, es decir con igual número de filas y columnas. $(\mathcal{M}_{nn}(\mathbb{K}), \cdot)$ admite un neutro multiplicativo, denominado **matriz identidad:**

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (\delta_{ij}), \quad \text{con } \delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

En efecto, dada $A \in \mathcal{M}_{nn}(\mathbb{K})$:

$$AI = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$= \left(\sum_{k=1}^{n} a_{ik}\delta_{kj} \right) = (a_{ij}\delta_{jj}) = (a_{ij}) = A.$$

Concluímos entonces que

Corolario 1.1. $(\mathcal{M}_{nn}(\mathbb{K}), +, \cdot)$ *es un anillo con unidad (existe neutro para* $\cdot$*).*

Dado que $(\mathcal{M}_{nn}(\mathbb{K}), \cdot)$ tiene un elemento neutro I, ¿Existe el inverso multiplicativo de una matriz en $(\mathcal{M}_{nn}(\mathbb{K}), \cdot)$?.

Definición (Matriz invertible) $A \in \mathcal{M}_{nn}(\mathbb{K})$ es **invertible** si y sólo si existe $B \in \mathcal{M}_{nn}(\mathbb{K})$ tal que:

$$AB = BA = I. \tag{1.1}$$

Proposición 1.3. *De existir una matriz B que satisfaga (1.1), esta es única. Por ende la notaremos $B = A^{-1}$.*

Demostración. Sean $B_1, B_2 \in \mathcal{M}_{nn}(\mathbb{K})$ que satisfacen (1.1). Luego

$$B_1 = B_1 I = B_1(AB_2).$$

Usando la asociatividad de $\cdot$ tenemos,

$$B_1 = (B_1 A)B_2,$$

pero como $B_1 A = I$, se concluye que

$$B_1 = B_2.$$

□

Cabe señalar que para que A sea invertible se requiere que exista una matriz B que sea a la vez inversa por izquierda y por derecha. Probaremos más adelante que **es suficiente que A tenga inversa por un solo lado**.

Ejemplo 1.2.
No todas las matrices cuadradas tienen inverso multiplicativo. En $\mathcal{M}_{22}(\mathbb{R})$, la matriz $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ no tiene inverso. En efecto, si existiese el inverso, digamos $\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$, debería verificarse:

$$\begin{aligned} & \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \iff \quad & \begin{pmatrix} x_1 + 2x_3 & x_2 + 2x_4 \\ 2x_1 + 4x_3 & 2x_2 + 4x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

$$\iff \quad \begin{aligned} x_1 \; + 2x_3 &= 1 \\ x_2 \; + 2x_4 &= 0 \\ 2x_1 + 4x_3 &= 0 \\ 2x_2 + 4x_4 &= 1 \end{aligned}$$

De la primera ecuación multiplicada por 2, obtenemos: $2(x_1 + 2x_3) = 2$. Pero de la tercera ecuación, $2x_1 + 4x_3 = 0$. De esta contradicción concluímos que el sistema **no** tiene solución.

En otros casos sí existe inverso, por ejemplo, para la matriz $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ en $\mathcal{M}_{22}(\mathbb{R})$, su matriz inversa es:

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

En efecto:

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

O bien, para $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ en $\mathcal{M}_{22}(\mathbb{C})$, su matriz inversa es:

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

lo que se verifica rápidamente.

1.2 Matrices Particulares

DEFINICIÓN Diremos que $A \in \mathcal{M}_{nn}(\mathbb{K})$ es:

1. **Diagonal** si y sólo si $a_{ij} = 0$ para todo $i \neq j$:

$$A = \begin{pmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}.$$

   En este caso la matriz es notada $A = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn})$.

2. **Triangular superior** si y sólo si $a_{ij} = 0$ si $i > j$:

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

3. **Triangular inferior** si y sólo si $a_{ij} = 0$ si $i < j$:

$$A = \begin{pmatrix} a_{11} & 0 & & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ & \vdots & & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{pmatrix}.$$

Ejemplos:

$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ son matrices diagonales

$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \ B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & -1 & 3 \end{pmatrix}$ son matrices triangulares, superior e inferior respectivamente.

Ejercicio 1.1: Una propiedad que queda propuesta es que si A es diagonal con $a_{ii} \neq 0, \;\; \text{para } i = 1, \ldots, n$ entonces es invertible, y su inversa también es diagonal con $(A^{-1})_{ii} = \frac{1}{a_{ii}}$.

Notación: Dada una matriz $A \in \mathcal{M}_{mn}(\mathbb{K})$ notaremos su i-ésima fila como:

$$A_{i\bullet} = (a_{i1} a_{i2} \ldots a_{in})$$

y su j-ésima columna:

$$A_{\bullet j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

Podemos escribir entonces la matriz: $A = (A_{\bullet 1}, A_{\bullet 2}, ..., A_{\bullet n})$, denominada notación por columnas. O bien:

$$A = \begin{pmatrix} A_{1\bullet} \\ A_{2\bullet} \\ \vdots \\ A_{m\bullet} \end{pmatrix}, \qquad \text{correspondiente a la notación por filas.}$$

Ejercicio 1.2: Con respecto a esta notación, es un buen ejercicio estudiar cómo el producto de matrices afecta las columnas y filas de las matrices a multiplicar. Más precisamente: Sean $A \in \mathcal{M}_{mn}(\mathbb{K}), B \in \mathcal{M}_{np}(\mathbb{K})$ y describamos B por sus columnas $B = (B_{\bullet 1}, B_{\bullet 2}, ..., B_{\bullet p})$, demostrar que entonces

$$AB = \left(A(B_{\bullet 1}), ..., A(B_{\bullet p})\right),$$

es decir la primera columna de AB es la matriz A por la primera columna de B, etc.

¿Qué hay de las filas de AB?

Definición (Matriz ponderada) Dada una constante $\lambda \in \mathbb{K}$, definimos la matriz A **ponderada** por λ:

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

Ejemplo:

En $\mathcal{M}_{23}(\mathbb{R})$:

$$3 \cdot \begin{pmatrix} 1 & -1 & 0 \\ 2 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -3 & 0 \\ 6 & 6 & 3 \end{pmatrix}$$

Veamos ahora que sucede al multiplicar por la derecha o izquierda una matriz por otra diagonal:

$$DA = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 4 \\ 6 & 0 & 3 \end{pmatrix}$$

constatamos que la primera fila aparece multiplicada por $d_{11} = 2$ y la segunda por $d_{22} = 3$.

$$A\bar{D} = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 6 \\ 4 & 0 & 3 \end{pmatrix}$$

Aquí, la primera columna de A aparece multiplicada por 2, la segunda por 1, la tercera por 3.

En general:

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix} = \begin{pmatrix} \lambda_1 a_{11} & \cdots & \lambda_1 a_{1p} \\ \lambda_2 a_{21} & \cdots & \lambda_2 a_{2p} \\ \vdots & & \vdots \\ \lambda_n a_{n1} & \cdots & \lambda_n a_{np} \end{pmatrix}$$

$$\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{p1} & \cdots & b_{pn} \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 b_{11} & \lambda_2 b_{12} & \cdots & \lambda_n b_{1n} \\ \vdots & \vdots & & \vdots \\ \lambda_1 b_{p1} & \lambda_2 b_{p2} & \cdots & \lambda_n b_{pn} \end{pmatrix}$$

De forma más compacta,

Proposición 1.4. *Si* $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \in \mathcal{M}_{nn}(\mathbb{K})$, $A \in \mathcal{M}_{np}(\mathbb{K})$, $B \in \mathcal{M}_{mn}(\mathbb{K})$, *se tiene que*

$$DA = \begin{pmatrix} \lambda_1 \; A_{1\bullet} \\ \vdots \\ \lambda_n \; A_{n\bullet} \end{pmatrix}, \tag{1.2}$$

$$BD = (\lambda_1 B_{\bullet 1}, ..., \lambda_n B_{\bullet n}). \tag{1.3}$$

Demostración. Probaremos (1.2), mientras que (1.3) queda propuesta como ejercicio

$$\text{Sea} \quad D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = (d_{ij}); \quad \text{con } d_{ij} = \begin{cases} \lambda_i & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

luego,

$$(DA)_{ij} = \sum_{k=1}^{n} d_{ik} a_{kj} = d_{ii} a_{ij} = \lambda_i a_{ij},$$

y por lo tanto

$$AD = \begin{pmatrix} \lambda_1 a_{11} & ... & \lambda_1 a_{1p} \\ & \vdots & \\ \lambda_n a_{n1} & \cdots & \lambda_n a_{np} \end{pmatrix}.$$

Otra propiedad, que utilizaremos más adelante, es la siguiente:

Proposición 1.5. *El producto de matrices triangulares inferiores (superiores) es triangular inferior (superior).*

Demostración. Verifiquemos el caso triangular superior. Sean

$$T_1 = \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}, \; T_2 = \begin{pmatrix} b_{11} & \cdots & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & b_{nn} \end{pmatrix}.$$

Luego $C = T_1 \cdot T_2 = (c_{ij})$, donde: $c_{ij} = \sum\limits_{j=1}^{n} a_{ik} b_{kj}$. Como T_1 y T_2 son triangulares superiores: $(a_{i\ell} = 0) \wedge (b_{i\ell} = 0)\ \forall \ell < i$. Supongamos $j < i$, luego

$$c_{ij} = \sum_{k=1}^{i-1} a_{ik} b_{kj} + \sum_{k=i}^{n} a_{ik} b_{kj}.$$

En el primer término $k < i \Rightarrow a_{ik} = 0$. En el segundo término $j < i \leq k \Rightarrow b_{kj} = 0$, luego $\forall j < i\ c_{ij} = 0$. Es decir, la matriz C es también triangular superior.

Notemos además que $(T_1 T_2)_{ii} = a_{ii} b_{ii}$ es decir los elementos diagonales del producto de dos triangulares superiores es el producto de los elementos diagonales correspondientes. □

1.3 Potencias, traspuestas e inversas

Definimos por recurrencia las potencias de una matriz cuadrada, como sigue:

Definición (Potencias de una matriz) Dada $A \in \mathcal{M}_{nn}(\mathbb{K})$

$$A^0 = I, \quad A^n = AA^{n-1}, \quad n \geq 1.$$

Ejemplo:

Por ejemplo; dada $A \in \mathcal{M}_{33}(\mathbb{R})$:

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

se tendrá

$$A^2 = AA = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 0 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 0 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 4 \\ 0 & 2 & 4 \\ 4 & 3 & 6 \end{pmatrix}$$

$$A^3 = AA^2 = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 0 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 4 & 2 & 4 \\ 0 & 2 & 4 \\ 4 & 3 & 6 \end{pmatrix} = \begin{pmatrix} 8 & 8 & 16 \\ 8 & 4 & 8 \\ 12 & 10 & 20 \end{pmatrix}.$$

Definición (Traspuesta) Dada una matriz $A = (a_{ij}) \in \mathcal{M}_{mn}(\mathbb{K})$, se define la **traspuesta** de A como aquella matriz de $n \times m$ que denotaremos por A^t tal que $(A^t)_{ij} = a_{ji}$.

Esto corresponde a intercambiar el rol de las filas y columnas. Más claramente, la primera fila de A^t es la primera columna de A y así sucesivamente.

Ejemplo:

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}, \quad A^t = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

$$A = (1, \ -1, \ 0, \ 0, \ 1), \quad A^t = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 4 \end{pmatrix}, \quad A^t = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 4 \end{pmatrix}.$$

En el último caso vemos que $A = A^t$.

Definición (Matriz simétrica) Diremos que una matriz $A \in \mathcal{M}_{nn}(\mathbb{K})$ es **simétrica** si y sólo si

$$A = A^t$$

Es fácil verificar que A es simétrica si y sólo si:

$$a_{ij} = a_{ji} \qquad \forall i, j = 1, .., n$$

Algunas propiedades de la trasposición de matrices son las siguientes:

Proposición 1.6.

1. $(A^t)^t = A, \quad \forall A \in \mathcal{M}_{mn}(\mathbb{K})$.
2. $(AB)^t = B^t A^t$.
3. *Si* $D \in \mathcal{M}_{nn}(\mathbb{K})$ *es diagonal, entonces* $D^t = D$.

Demostración. Probaremos 2: Sea $C = AB$, $C^t = (AB)^t$. En la posición (i, j) de la matriz C^t aparece el término c_{ji} de la matriz C: $c_{ji} = \sum\limits_{k=1}^{n} a_{jk} b_{ki}$ Consideremos $Z = B^t\ A^t$. El término (i, j) de la matriz Z esta dado por:

$$z_{ij} = \sum_{k=1}^{n} (B^t)_{ik} (A^t)_{kj} = \sum_{k=1}^{n} b_{ki} a_{jk} = \sum_{k=1}^{n} a_{jk} b_{ki} = (AB)_{ji} = (C^t)_{ij},$$

luego $Z = C^t$, y por lo tanto $(AB)^t = B^t A^t$. $\square$

Algunas propiedades elementales de matrices invertibles son las siguientes

Proposición 1.7.
Sean $A, B \in \mathcal{M}_{nn}(\mathbb{K})$ *invertibles entonces:*

1. *La inversa de* A *es invertible y* $(A^{-1})^{-1} = A$.
2. *El producto* AB *es invertible y* $(AB)^{-1} = B^{-1} A^{-1}$.
3. $\forall n \geq 0,\ (A^n)^{-1} = (A^{-1})^n$.
4. A^t *es invertible y* $(A^t)^{-1} = (A^{-1})^t$.

Demostración. Veamos:

La propiedad 1 se prueba directamente de la definición y de la unicidad de la inversa (Proposición 1.3).

Para 2 se tiene que,

$$AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$$

De manera análoga se prueba $(B^{-1}A^{-1})AB = I$ Como la inversa es única $(AB)^{-1} = B^{-1}A^{-1}$.

Para 3, procederemos por inducción. Se verifica para $n \in \{0,1\}$. Supongamos cierto el resultado para $n \geq 1$ y veamos:

$$(A^{n+1})^{-1} = (A^n \cdot A)^{-1} = A^{-1}(A^n)^{-1},$$

por hipótesis de inducción

$$(A^{n+1})^{-1} = A^{-1}(A^{-1})^n = (A^{-1})^{n+1}.$$

Para probar 4, notemos que $AA^{-1} = I$, así trasponiendo nos dá $(AA^{-1})^t = I^t$, lo que implica a su vez $(A^{-1})^t A^t = I$.

Igualmente se obtiene que $A^t(A^{-1})^t = I$. Finalmente, por unicidad de la inversa: $(A^t)^{-1} = (A^{-1})^t$. □

Ejercicio 1.3: En general el problema de saber si una matriz es invertible o no es un problema difícil. Tendremos que esperar a saber resolver ecuaciones lineales para obtener un algoritmo eficiente que nos permitirá decidir si una matriz en particular es invertible y obtener su inversa.

Por el momento nos contentaremos con el siguiente resultado, propuesto como ejercicio.

Sea $A = (a_{ij})$ una matriz de **permutación**, es decir una matriz de $n \times n$ con solo 0 y 1 tal que tiene un y sólo un 1, por cada fila y columna.

Se tiene que A es invertible y su inversa es $A^{-1} = A^t$.

Guía de Ejercicios

1. Demuestre la asociatividad de la multiplicación de matrices. Es decir, si $A \in \mathcal{M}_{mn}(\mathbb{K})$, $B \in \mathcal{M}_{nq}(\mathbb{K})$ y $C \in \mathcal{M}_{qs}(\mathbb{K})$, entonces:

$$A(BC) = (AB)C \in \mathcal{M}_{ms}(\mathbb{K}).$$

2. Pruebe que si A es diagonal ($A = \mathrm{diag}(a_{11}, \ldots, a_{nn}) \in \mathcal{M}_{nn}(\mathbb{K})$), entonces

$$A \text{ es invertible} \iff a_{ii} \neq 0, \text{ para } i \in \{1, \ldots, n\}.$$

Demuestre además que, en caso de ser invertible, su inversa es diagonal con $(A^{-1})_{ii} = \frac{1}{a_{ii}}$.

3. Sean $A \in \mathcal{M}_{mn}(\mathbb{K}), B \in \mathcal{M}_{np}(\mathbb{K})$ y describamos B por sus columnas $B = (B_{\bullet 1}, B_{\bullet 2}, ..., B_{\bullet n})$, demuestre que entonces

$$AB = \left(A(B_{\bullet 1}), ..., A(B_{\bullet n})\right).$$

4. Sea $A = (a_{ij})$ una matriz de **permutación**, es decir una matriz de $\mathcal{M}_{nn}(\mathbb{K})$ con sólo 0's y 1's tal que tiene un y sólo un 1, por cada fila y columna. Pruebe que A es invertible y su inversa es $A^{-1} = A^t$.

5. Se dice que $P \in \mathcal{M}_{nn}(\mathbb{K})$ es una matriz de proyección si $P = P^2$.

(a) Pruebe que si P es matriz de proyección, entonces $I_n - P$ es matriz de proyección, donde I_n es la matriz identidad en $\mathcal{M}_{nn}(\mathbb{K})$.

(b) Pruebe que P es matriz de proyección si y sólo si $P^2(I_n - P) = \mathbf{0}$ y $P(I_n - P)^2 = \mathbf{0}$.

(c) Encuentre $P \in \mathcal{M}_{22}(\mathbb{K})$ tal que $P \neq P^2$ y $P^2(I_2 - P) = \mathbf{0}$.
Indicación: Considere matrices con coeficientes en $\{0, 1\}$.

Guía de Problemas

P1. Sea $D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \in \mathcal{M}_{nn}(\mathbb{R})$ diagonal, con $d_1, \ldots, d_n$ distintos y $A, B, M, S \in \mathcal{M}_{nn}(\mathbb{R})$.

(a) Pruebe que si $MD = DM$, entonces M es diagonal.

(b) Sea S invertible, tal que $S^{-1}AS$ y $S^{-1}BS$ son diagonales. Pruebe que $AB = BA$. *Indicación:* Recuerde que el producto de matrices diagonales conmuta.

(c) Sea S invertible tal que $S^{-1}AS = D$. Suponiendo que $AB = BA$, verifique que $S^{-1}AS$ y $S^{-1}BS$ conmutan y concluya que $S^{-1}BS$ es diagonal.

P2. **(a)** Demuestre que si A, B y $(A+B^{-1})$ son matrices invertibles, entonces $(A^{-1}+B)$ también es invertible y su inversa es $A(A+B^{-1})^{-1}B^{-1}$.

(b) Sea la matriz de n y n columnas triangular superior a coeficientes reales siguiente:

$$C = \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & 1 & 0 & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & 1 & 1 & 0 \\ \vdots & \dots & \dots & \dots & 0 & 1 & 1 \\ \vdots & \dots & \dots & \dots & 0 & 0 & 1 \end{pmatrix}.$$

Sea $N = C - I$, donde I es la matriz identidad de $n \times n$. Demuestre que $N^n = 0$.

(c) Demuestre que para las matrices C y N definidas en (b), se tiene que C es invertible y su inversa es:

$$C^{-1} = I - N + N^2 - N^3 + \cdots + (-1)^{n-1}N^{n-1}.$$

P3. **(a)** Sea $A = (a_{ij}) \in \mathcal{M}_{nn}(\mathbb{R})$ una matriz cualquiera. Se define la *traza* de A, denotada por $\operatorname{tr}(A)$, como la suma de los elementos de la diagonal principal, es decir,

$$\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Por otra parte, se define la función $f : \mathcal{M}_{nn}(\mathbb{R}) \to \mathbb{R}$, donde

$$f(A) = \operatorname{tr}(AA^t).$$

Pruebe que:

(1) Dadas $A, B \in \mathcal{M}_{nn}(\mathbb{R})$,

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

(2) $f(A) \geq 0$, $\forall A \in \mathcal{M}_{nn}(\mathbb{R})$, además muestre que

$$f(A) = 0 \iff A = \mathbf{0},$$

donde $\mathbf{0}$ es la matriz nula de $\mathcal{M}_{nn}(\mathbb{R})$.

(3) $f(A) = \operatorname{tr}(A^tA)$.

(b) Sea $M \in \mathcal{M}_{mn}(\mathbb{R})$ una matriz tal que la matriz $(M^tM) \in \mathcal{M}_{nn}(\mathbb{R})$ es invertible. Definamos la matriz $P \in \mathcal{M}_{mm}(\mathbb{R})$ como

$$P = I_m - M(M^tM)^{-1}M^t,$$

donde I_M es la matriz identidad de orden m.

Pruebe que:

(1) $P^2 = P$. Muestre además que $PM = \mathbf{0}$, donde $\mathbf{0} \in \mathcal{M}_{mn}(\mathbb{R})$ es la matriz nula.

(2) La matriz (M^tM) es simétrica y muestre que la matriz P es también simétrica.

Ingeniería Matemática
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Matrices

1.4 Matrices elementales

Como veremos la resolución de sistemas de ecuaciones via eliminación de variables corresponde a premultiplicar (multiplicar por la izquierda) una cierta matriz por matrices elementales que estudiaremos a continuación. Hay dos tipos de matrices elementales: **elemental de permutación** y de **suma**.

DEFINICIÓN (MATRIZ ELEMENTAL DE PERMUTACIÓN) Una matriz elemental de permutación tiene la siguiente estructura:

$$I_{pq} = \begin{pmatrix} 1 & & & & & & & & & 0 \\ & \ddots & & & & & & 0 & & \\ & & 1 & & & & & & & \\ & & & 0 & \cdots & \cdots & \cdots & 1 & & \\ & & & & 1 & & & & & \\ & & & \vdots & & 1 & & \vdots & & \\ & & & & & & 1 & & & \\ & & & 1 & \cdots & \cdots & \cdots & 0 & & \\ & & & & & & & & 1 & \\ & 0 & & & & & & & & \ddots \\ 0 & & & & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ \text{fila } p \\ \\ \\ \\ \text{fila } q \\ \\ \\ \\ \end{array}$$

La matriz I_{pq} se construye a partir de la identidad, permutando el orden de las filas p y q.

Ejemplo 1.3.
En $M_{44}(\mathbb{R})$:

$$I_{24} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad I_{13} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Notemos que toda matriz elemental de permutación es una matriz de permutación, pero no al revés. ¿Puedes dar un ejemplo?.

Veamos ahora lo que sucede al multiplicar una matriz A, por la derecha o izquierda, por una matriz elemental de permutación I_{pq}: En el ejemplo anterior, sea $A = (a_{ij}) \in M_{43}(\mathbb{R})$

$$I_{24}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{41} & a_{42} & a_{43} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}.$$

El resultado consiste en permutar las filas 2 y 4 de A. Análogamente, sea $B \in M_{34}(\mathbb{R})$:

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{14} & b_{13} & b_{12} \\ b_{21} & b_{24} & b_{23} & b_{22} \\ b_{31} & b_{34} & b_{33} & b_{32} \end{pmatrix}.$$

la matriz resultante es B con las columnas 2 y 4 permutadas. En general,

Propiedad 1.1.
Dadas $I_{pq} \in \mathcal{M}_{nn}(\mathbb{K}), A \in \mathcal{M}_{ns}(\mathbb{K})$ y $B \in M_{qn}(\mathbb{K})$:

1. *$I_{pq}A$ corresponde a la matriz A con las filas p y q permutadas.*
2. *BI_{pq} corresponde a las matriz B con las columnas p y q permutadas.*

Demostración. Recordemos que, por ser una matriz de permutación (Ejercicio 1.3), I_{pq} es invertible y además $I_{pq}^{-1} = I_{pq}$. En efecto, el multiplicar por la izquierda I_{pq} por si misma, corresponde a intercambiar sus filas p y q, con lo cual se obtiene la identidad.$\square$

Tomemos ahora la matriz $A \in M_{44}(\mathbb{R})$ y sea:

$$E_{2,4}(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \lambda & 0 & 1 \end{pmatrix}$$

Esta matriz se construye a partir de la identidad colocando el valor λ en la posición (4,2) (notar que sólo la fila 4 es diferente de la identidad).

Al multiplicar por la izquierda A por $E_{2,4}(\lambda)$:

$$E_{2,4}(\lambda)\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \lambda & 0 & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ \lambda a_{21} + a_{41} & \lambda a_{22} + a_{42} & \lambda a_{23} + a_{43} \end{pmatrix}$$

La matriz, $E_{2,4}(\lambda)A$ es exactamente la matriz A, excepto por la fila 4, la cual se obtiene de sumar la fila 4 de A más la fila 2 de A ponderada por λ.

En general,

Definición (Matriz elemental) Definimos la matriz elemental $E_{p,q}(\lambda) \in M_{nn}(\mathbb{K})$ como:

$$E_{p,q}(\lambda) = \begin{array}{c} \begin{array}{ccccccccc} & & \text{col. } p & & & \text{col. } q & & & \\ & & \downarrow & & & \downarrow & & & \end{array} \\ \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & 0 & & \\ & & 1 & & & & & & \\ & & & \ddots & & & & & \\ & & \vdots & & 1 & & & & \\ 0 & \cdots & \lambda & \cdots & & 1 & & & \\ \vdots & & \vdots & & & & 1 & & \\ \vdots & & \vdots & & & & & \ddots & \\ 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} \end{array} \qquad \begin{array}{c} \lambda \in \mathbb{K} \\ p < q \end{array}.$$

Propiedad 1.2. *Dada una matriz $A \in M_{ns}(\mathbb{K})$; se tiene:*

$$C = E_{p,q}(\lambda) \cdot A = \begin{pmatrix} 1 & & & & & & & 0 \\ & \ddots & & & & & 0 & \\ & & 1 & & & & & \\ & & & \cdots & & & & \\ & & \vdots & & 1 & & & \\ & 0 & \lambda & \cdots & & 1 & & \\ & & & & & & 1 & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1s} \\ \vdots & & & \vdots \\ a_{p1} & \cdots & \cdots & a_{ps} \\ \vdots & & & \vdots \\ a_{q1} & \cdots & \cdots & a_{qs} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{ns} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & \cdots & a_{1s} \\ \vdots & & \vdots \\ a_{p-11} & \cdots & a_{p-1s} \\ \vdots & & \vdots \\ \lambda a_{p1} + a_{q1} & \cdots & \lambda a_{ps} + a_{qs} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{ns} \end{pmatrix} \begin{array}{c} \\ \\ \\ \\ \leftarrow q \\ \\ \\ \end{array}.$$

o, en notación por filas:

$$
\begin{aligned}
C_{i\bullet} &= A_{i\bullet} \qquad \forall i \neq q \\
C_{q\bullet} &= \lambda A_{p\bullet} + A_{q\bullet}
\end{aligned}
$$

Se tiene entonces que el efecto, sobre A, de la premultiplicación por $E_{p,q}(\lambda)$ es una matriz que sólo difiere de A en la fila q: Esta es la suma de la fila p ponderada por λ y la fila q.

Es importante observar que la matriz $E_{p,q}(\lambda)$ es **triangular inferior**, que tiene unos en la diagonal y además:

Proposición 1.8. $E_{p,q}(\lambda)$ *es* **invertible**. *Su inversa es la siguiente:*

$$
(E_{p,q}(\lambda))^{-1} = \begin{array}{c}
\begin{pmatrix}
1 & & & & & & & & \\
 & \ddots & & & & & 0 & & \\
 & & 1 & & & & & & \\
 & & & \ddots & & & & & \\
 & & \vdots & & 1 & & & & \\
0 & \cdots & -\lambda & \cdots & & 1 & & & \\
\vdots & & \vdots & & & & 1 & & \\
\vdots & & \vdots & & & & & \ddots & \\
0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1
\end{pmatrix} \\
\begin{array}{ccccccccc}
 & & \uparrow & & & \uparrow & & & \\
 & & p & & & q & & &
\end{array}
\end{array}
$$

Demostración. En efecto, sea $C = E_{p,q}(-\lambda)E_{p,q}(\lambda)$. Se tendrá que C es la matriz cuya fila q es la suma de la fila p de $E_{p,q}(\lambda)$, ponderada por $-\lambda$ y la fila q de $E_{p,q}(\lambda)$. Es decir:

$$
\begin{aligned}
C_{q\bullet} &= -\lambda(0\ldots 0 \underset{\underset{p}{\uparrow}}{1} 0\ldots 0) + (0\ldots \underset{\underset{p}{\uparrow}}{\lambda} 0\ldots 0 \underset{\underset{q}{\uparrow}}{\beta} 0\ldots 0) \\
&= (0\ldots 0 \underset{\underset{p}{\uparrow}}{-\lambda} 0\ldots 0) + (0\ldots \underset{\underset{p}{\uparrow}}{\lambda} 0\ldots 0 \underset{\underset{q}{\uparrow}}{1} 0\ldots 0).
\end{aligned}
$$

Luego: $C_{q\bullet} = (0..,0..,1..,0..,0)$ además $\forall i \neq q$ $C_{i\bullet}$ es la i-ésima fila de la identidad. Es decir $C = I$, la matriz identidad $\square$

1.5 Sistemas lineales y escalonamiento de matrices

Las matrices elementales son de gran utilidad en la resolución de sistemas de ecuaciones lineales.

Ejemplo 1.4.
Consideremos, a modo de ejemplo, el sistema de ecuaciones

$$\begin{array}{rrrrcl} x_1 & +2x_2 & +x_3 & +x_4 & = & 2 \\ -2x_1 & +3x_2 & -x_3 & & = & -1 \\ x_1 & & +x_3 & +x_4 & = & 1 \end{array} \qquad \begin{array}{c} (1) \\ (2) \\ (3) \end{array}$$

Para resolverlo, utilizamos la eliminación de variables, pero en forma ordenada, desde la primera variable de la izquierda y desde arriba hacia abajo:

1. Eliminamos la variable x_1 en las ecuaciones (2) y (3): para ello multiplicamos la primera por dos y la sumamos a la segunda obteniendo:

   $$\begin{array}{rrrrcl} x_1 & +2x_2 & +x_3 & +x_4 & = & 2 \\ & 7x_2 & +x_3 & +2x_4 & = & 3 \\ x_1 & & +x_3 & +x_4 & = & 1 \end{array}$$

   Luego, multiplicamos la primera por -1 y la sumamos a la tercera:

   $$\begin{array}{rrrrcr} x_1 & +2x_2 & +x_3 & +x_4 & = & 2 \\ & 7x_2 & +x_3 & +2x_4 & = & 3 \\ & -2x_2 & & & = & -1 \end{array}$$

2. Continuamos ahora con x_2, pero a partir de la segunda ecuación. Multiplicando la segunda por $\frac{2}{7}$ y sumándola a la tercera se obtiene:

   $$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 2 \\ 7x_2 + x_3 + 2x_4 &= 3 \\ \frac{2}{7}x_3 + \frac{4}{7}x_4 &= -\frac{1}{7} \end{aligned}$$

   Ya no es posible eliminar más variables. Ahora, desde la última hasta la primera ecuación despejamos en función de x_4:

   $$\begin{array}{ll} x_3 & = -\frac{4}{2}x_4 - \frac{1}{2} = -2x_4 - \frac{1}{2} \\ x_2 & = \frac{1}{7}(-x_3 - 2x_4 + 3) = \frac{1}{7}(2x_4 + \frac{1}{2} - 2x_4 + 3) = \frac{1}{2} \\ x_1 & = -2x_2 - x_3 - x_4 + 2 = -2\frac{1}{2} + 2x_4 + \frac{1}{2} - x_4 + 2 = x_4 + \frac{3}{2} \end{array}$$

   Obteniéndose la solución:

   $$\begin{array}{rcl} x_1 & = & \frac{3}{2} + x_4 \\ x_2 & = & \frac{1}{2} \\ x_3 & = & -\frac{1}{2} - 2x_4 \\ x_4 & = & x_4 \end{array}$$

Así, para cualquier valor real de x_4, obtenemos una solución del sistema. Existen entonces, infinitas soluciones, dependiendo de la variable x_4. La variable x_4 se denomina **independiente o libres**.

Veamos ahora, en general, qué es un sistema de ecuaciones,

DEFINICIÓN (SISTEMA DE ECUACIONES) Un sistema de m ecuaciones y n incógnitas consiste en el siguiente conjunto de ecuaciones en las variables $x_1, ..., x_n \in \mathbb{K}$:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad\qquad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

en donde los coeficientes, a_{ij}, y el lado derecho, b_j, son elementos del cuerpo $\mathbb{K}$.

Definiendo la matriz:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M_{mn}(\mathbb{K})$$

la m-tupla (lado derecho) y la n tupla de incógnitas

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

Podemos escribir el sistema matricialmente:

$$Ax = b,$$

Realizar el procedimiento de eliminación de variables descrito en el ejemplo precedente, con el fin de resolver el sistema, es equivalente a producir ceros en la matriz **aumentada** con la columna lado derecho, $(A|b) \in M_{m(n+1)}(\mathbb{K})$. En el ejemplo anterior:

$$(A|b) = \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ -2 & 3 & -1 & 0 & -1 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Eliminar x_1 de la segunda ecuación es equivalente a producir un cero en la posición (2,1) de $(A|b)$. Para ello se multiplica la primera fila por 2 y se suma a la segunda fila. Para eliminar x_1 de la tercera ecuación se multiplica la primera fila por -1 y se suma a la tercera

$$(A|b) \to \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 7 & 1 & 2 & 3 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 7 & 1 & 2 & 3 \\ 0 & -2 & 0 & 0 & -1 \end{pmatrix}$$

Eliminar x_2 en la tercera ecuación a partir de la segunda es equivalente a multiplicar la segunda fila por $\frac{2}{7}$ y sumarla a la tercera:

$$\to \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 7 & 1 & 2 & 3 \\ 0 & 0 & \frac{2}{7} & \frac{4}{7} & -\frac{1}{7} \end{pmatrix} = (\tilde{A}|\tilde{b})$$

Así, el sistema inicial es equivalente al sistema $\tilde{A}x = \tilde{b}$.

Conviene señalar que en el procedimiento anterior la operación de base ha sido:

Sumar a una fila q, la fila p ponderada por un número $\lambda \in \mathbb{K}$

De la definición de matrices elementales sabemos que esto es equivalente a premultiplicar por la izquierda por la matriz $E_{p,q}(\lambda)$. Veamos esto en el mismo ejemplo.

Ejemplo 1.5.

1. Producir un cero en la posición (2,1) de $(A|b)$:

$$\begin{aligned} E_{1,2}(2)(A|b) &= \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ -2 & 3 & -1 & 0 & -1 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 7 & 1 & 2 & 3 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix}. \end{aligned}$$

2. Producir un cero en la posición (3,1):

$$\begin{aligned} E_{1,3}(-1)E_{1,2}(2)(A|b) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 7 & 1 & 2 & 3 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 1 & 1 & 1 \\ 0 & 7 & 1 & 2 & 3 \\ 0 & -2 & 0 & 0 & -1 \end{pmatrix} \end{aligned}$$

3. Producir un cero en la posición (3,2) desde la posición (2,2):

$$\begin{aligned} E_{2,3}(\tfrac{2}{7})E_{1,3}(-1)E_{1,2}(2)(A|b) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{2}{7} & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 7 & 1 & 2 & 3 \\ 0 & -2 & 0 & 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 7 & 1 & 2 & 3 \\ 0 & 0 & \frac{2}{7} & \frac{4}{7} & -\frac{1}{7} \end{pmatrix}. \end{aligned}$$

Se concluye entonces que la operación de eliminación de variable puede realizarse mediante la pre-multiplicación de $(A|b)$ por matrices elementales. Hay casos en que también es necesario utilizar matrices de permutación de filas. Por ejemplo, si se tiene:

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

Vemos que, como $a_{22} = 0$, no es posible "producir" ceros en la segunda columna, a partir de a_{22}. Luego intercambiamos el orden de las filas (claramente esto no cambia el sistema de ecuaciones asociado). Por ejemplo, colocamos la cuarta fila en la segunda posición y la segunda en la cuarta. Esto es equivalente a premultiplicar por I_{24}:

$$I_{24}(A|b) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

lo cual nos permite seguir produciendo ceros. Consideremos ahora $A \in M_{mn}(\mathbb{K})$, definimos la matriz **escalonada** asociada a la matriz A, como $\tilde{A} \in M_{mn}(\mathbb{K})$, tal que:

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \cdots & & \tilde{a}_{1i_2} & \cdots & \cdots & \cdots & \tilde{a}_{1n} \\ 0 & \cdots & 0 & \tilde{a}_{2i_2} & & & & \vdots \\ \vdots & & \vdots & \vdots & \ddots & & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \tilde{a}_{si_s} & \cdots & \tilde{a}_{sn} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & & \vdots & & & & \vdots \\ 0 & \cdots & & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

donde los elementos $\tilde{a}_{11} \neq 0$, $\tilde{a}_{2i_2} \neq 0, ..., \tilde{a}_{si_s} \neq 0$, se denominan **pivotes**. Notar que hemos supuesto que la primera columna no tiene ceros. De no ser así, quiere decir que la primera variable no juega ningun rol, y por lo tanto si el sistema tiene solución esta variable queda de inmediato libre. Supondremos en lo que sigue que la primera columna no es cero.

Observación: No hay una única manera de escalonar una matriz. En efecto en el último ejemplo podríamos haber permutado las filas 2 y 3 en vez de las 2 y 4.

Hay dos cosas que son importantes de resaltar a este respecto:

1. Desde el punto de vista de sistemas de ecuaciones no hay diferencia entre un escalonamiento u otro: todos dan el mismo conjunto solución (probablemente descrito de distinta manera).
2. Es preferible por otras razones (teóricas, como son el cálculo del determinante y la determinación de matrices definidas positivas) tratar de no utilizar permutaciones .

La matriz $\tilde{A}$ se obtiene mediante la premultiplicación de A por matrices elementales:

$$\tilde{A} = \left(\prod_j E_j\right) A,$$

donde E_j es una matriz elemental de suma o de permutación de filas.

Además, recordemos que las matrices elementales son invertibles. Esta propiedad es crucial para probar que el sistema original, $Ax = b$, y el obtenido después del escalonamiento, $\tilde{A}x = \tilde{b}$, son equivalentes (tienen idéntico conjunto de soluciones). En efecto:

Proposición 1.9. *Dada una matriz C, invertible entonces:*

$$a \in \mathbb{K}^n \textit{ es solución de } Ax = b \iff a \textit{ es solución de } (CA)x = Cb.$$

Demostración. La demostración es simple:

$\Longrightarrow$) Si $Aa = b$, entonces $C(Aa) = Cb$ y por ende $(CA)a = Cb$.

$\Longleftarrow$) Supongamos $(CA)a = Cb$. Como C es invertible se tiene $C^{-1}(CA)a = C^{-1}(Cb)$, lo cual implica que $Aa = b$. $\square$

Como las matrices elementales son invertibles y el producto de matrices invertibles también lo es, y por lo tanto podemos usar la Proposición 1.9 con $C = \prod_j E_j$, para concluir que los sistemas $Ax = b$ y $\tilde{A}x = \tilde{b}$ son equivalentes.

Guía de Ejercicios

1. Encuentre un ejemplo de una matriz que sea de permutación, pero no sea matriz elemental de permutación.

2. Dadas $I_{pq} \in \mathcal{M}_{nn}(\mathbb{K})$ y $B \in \mathcal{M}_{qn}(\mathbb{K})$, pruebe que BI_{pq} corresponde a las matriz B con las columnas p y q permutadas.
Indicación: Puede usar la parte 1 de la Proposición 1.1.

3. Escriba los siguientes sistemas de ecuaciones en las variables x_i, de forma matricial $Ax = b$, especificando claramente A, b y x con sus dimensiones:

(a) $$\begin{cases} 2x_1 + 4x_2 - x_3 + x_5 &= 0 \\ x_1 - 16x_2 + x_3 + 3x_4 + 6x_5 &= 2 \\ \frac{1}{3}x_2 - 10x_3 + x_4 - 6x_5 &= 2x_3 \\ x_1 - x_3 - x_4 + \frac{3}{4}x_5 &= 1 \end{cases}$$

(b) $$\begin{cases} x_4 + x_1 &= -3 \\ 4x_1 - 7x_2 + \pi x_3 - 12x_4 &= 1 \\ -11 &= x_3 \\ \frac{1}{3}(x_1 - x_3) + x_4 - x_5 &= 20 \end{cases}$$

(c) $$\begin{cases} x_1 - (\beta - 1)x_2 + x_3 - \alpha x_4 &= 2 \\ -7\alpha x_2 + \pi x_3 &= \alpha + \beta \\ \frac{\beta}{3}x_1 - x_2 + x_4 &= \frac{\beta}{3} \end{cases}, \text{ con } \alpha, \beta \in \mathbb{R}.$$

(d) $$\begin{cases} 2x_1 &= 1 \\ -4x_2 &= 2 \\ 9x_3 &= 0 \\ \frac{3}{2}x_4 &= -1 \end{cases}$$

4. Escriba explícitamente las siguientes multiplicaciones de matrices:

(a) $I_{12}E_{13}(1, -1)$, en $\mathcal{M}_{44}(\mathbb{R})$.

(b) $I_{34}E_{23}(i, 2)I_{34}$, en $\mathcal{M}_{44}(\mathbb{C})$.

(c) $E_{12}(2, 1)E_{13}(1, -1)$, en $\mathcal{M}_{33}(\mathbb{R})$.

(d) $E_{23}(0, 3i)^{-1}E_{12}(-1, i)E_{34}(\frac{1}{2}, -3)$, en $\mathcal{M}_{44}(\mathbb{C})$.

Guía de Problemas

P1. Sea $J \in \mathcal{M}_{nn}(\mathbb{R})$ tal que $J = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$ y $e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

(a) Pruebe que $Je = ne$ y $J^2 = nJ$.

(b) Sea $\alpha \neq \frac{1}{n}$. Calcule β tal que $(I - \alpha J)^{-1} = I + \beta J$.

(c) Sea $\alpha = \frac{1}{n}$, verifique que $(I - \alpha J)e = \mathbf{0}$.

P2. **(a)** **(1)** Sean A, B matrices de $n \times m$ con coeficientes reales. Pruebe que

$$A = B \iff \forall x \in \mathcal{M}_{m1}(\mathbb{R}), \forall y \in \mathcal{M}_{n1}(\mathbb{R}) \quad x^t A y = x^t B y.$$

Indicación: Calcule $e_j^t A e_i$ donde $e_j = (I_m)_{j\bullet}$ y $e_i = (I_n)_{i\bullet}$. I_m y I_n son las identidades de tama nos m y n respectivamente.

(2) Sean A, B matrices de $n \times n$ simétricas con coeficientes reales. Pruebe que

$$A = B \iff \forall x \in \mathcal{M}_{n1}(\mathbb{R}) \quad x^t A x = x^t B x.$$

(b) Demuestre que si una matriz cuadrada A verifica $A^k = I$, para algún natural $k \geq 1$, entonces A es invertible y determine su inversa.

(c) Encuentre todas las matrices de la forma $A = \begin{pmatrix} x & y \\ 0 & y \end{pmatrix}$ que cumplen $A^2 = I$ $(x, y, z \in \mathbb{R})$.

P3. Considere el siguiente sistema lineal:

$$\begin{array}{rcrcrcrcr}
-x_1 & & & + & x_3 & + & 2x_4 & = & 1\,, \\
x_1 & - & x_2 & + & x_3 & - & 2x_4 & = & -1\,, \\
-x_1 & - & 2x_2 & + & x_3 & + & 3x_4 & = & 2\,, \\
-x_1 & - & 4x_2 & + & 2x_3 & + & 4x_4 & = & 5\,.
\end{array}$$

(a) Escriba el sistema de forma matricial $Ax = b$. Especifique claramente A, b y x, junto con sus dimensiones.

(b) Resuelva es sistema lineal *escalonando* la matriz aumentada $(A|b)$.

P4. Considere el siguiente sistema lineal:

$$\begin{array}{rcrcrcrcr}
x_1 & + & 2x_2 & + & 3x_3 & - & x_4 & = & 1\,, \\
-x_1 & + & x_2 & & & + & x_4 & = & 1\,, \\
 & & 3x_2 & + & 3x_3 & & & = & 0\,, \\
2x_1 & + & x_2 & + & 3x_3 & - & 2x_4 & = & -2\,.
\end{array}$$

(a) Escriba el sistema de forma matricial $Ax = b$. Especifique claramente A, b y x, junto con sus dimensiones.

(b) Resuelva es sistema lineal *escalonando* la matriz aumentada $(A|b)$.

1.6 Solución general de sistemas lineales

Dado el sistema $Ax = b, A \in M_{mn}(\mathbb{K}), b \in \mathbb{K}^m, x \in \mathbb{K}^n$, al escalonarlo (escalonando $(A|b)$) obtenemos:

$$(\tilde{A}|\tilde{b}) = \begin{pmatrix} \tilde{a}_{11} & \cdots & & \tilde{a}_{1i_2} & \cdots & \cdots & \cdots & \tilde{a}_{1n} & \tilde{b}_1 \\ 0 & \cdots & 0 & \tilde{a}_{2i_2} & & & & \vdots & \vdots \\ \vdots & & \vdots & \vdots & \ddots & & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \tilde{a}_{si_s} & \cdots & \tilde{a}_{sn} & \tilde{b}_s \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \tilde{b}_{s+1} \\ \vdots & & & \vdots & & & & \vdots & \vdots \\ 0 & \cdots & & 0 & \cdots & 0 & \cdots & 0 & \tilde{b}_m \end{pmatrix}.$$

donde los elementos $\tilde{a}_{11}, \tilde{a}_{2i_2}, \ldots, \tilde{a}_{si_s}$ son no nulos. Claramente, si existe un índice $j \geq s+1$ tal que $\tilde{b}_j \neq 0$, el sistema no tiene solución, ya que se tendría una ecuación incompatible: $0 = \tilde{b}_j \neq 0$.

Luego, el sistema tiene solución si y solo si $\tilde{b}_k = 0 \ \forall k \geq s+1$. En efecto, si $\tilde{b}_k = 0$, $\forall k \geq s+1$, se obtiene la(s) solución(es) despejando desde la s-ésima ecuación hasta la primera en función de las variables independientes. En este caso diremos que el sistema es **compatible**. Cada una de las diferencias en el número de columnas que son cero bajo las filas sucesivas, las llamaremos **peldaños** o **escalones**. Así el peldaño que se produce entre la fila k y $k+1$ es $i_{k+1} - i_k$. Notar que el último peldaño es: $n + 1 - i_s$, es decir los peldaños se miden en la matriz aumentada. La importancia de estos peldaños es la siguiente:

Proposición 1.10. *Si existe solución para el sistema $Ax = b$ y en la matriz $\tilde{A}$ hay algún pelda no de largo mayor o igual a dos, entonces existe* **más de una** *solución.*

Demostración. En efecto. Como por cada fila no nula de la matriz escalonada podemos despejar a lo más una variable, tendremos entonces que dejar libres tantas variables como número de peldaños, menos uno. □

Un corolario directo del resultado anterior es

Corolario 1.2. *Si el sistema $Ax = b$ es tal que $n > m$ (número de incógnitas mayor que el número de ecuaciones), entonces tiene infinitas soluciones, si es compatible.*

Demostración. En efecto, como $n > m$ siempre existe en la matriz escalonada, $\tilde{A}$, un peldaño de largo superior o igual a dos. De no ser así se tendría

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ 0 & \tilde{a}_{22} & \cdots & \tilde{a}_{2n} \\ \vdots & 0 & & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \tilde{a}_{mn} \end{pmatrix}$$

de donde $m = n$.

Un caso particular importante es cuando el lado derecho del sistema es nulo, $b = 0$. Hablamos entonces de un **sistema homogéneo**, $Ax = 0$. Vemos que, en este caso, siempre existe al menos una solución, la **trivial** $x = 0 \in \mathbb{K}^n$. En otras palabras sabemos de antemano que todo sistema homogéneo es compatible.

Podemos resumir nuestro estudio de sistemas en el cuadro siguiente:

	Sistema Homogéneo ($b = 0$)		Sistema no-Homogéneo ($b \neq 0$)	
Dimensiones	$n \leq m$	$n > m$	$n \geq m$	$n > m$
Número Soluciones	$1, \infty$	∞	$0, 1, \infty$	$0, \infty$

Ejemplo 1.6.
Resolvamos el sistema

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & 2 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & -1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Escalonando:

$$(A|b) \to \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 3 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -2 & 0 & 0 & -1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 3 & 1 & 0 & 1 & 1 \\ 0 & 0 & \frac{2}{3} & 0 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & \frac{2}{3} & 0 & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \to$$

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

Obtenemos que los peldaños medidos desda la fila 1 son sucesivamente: 1,1,2,1. De la última ecuación despejamos $x_5 = -1$.

Con esta información en la tercera ecuación dejamos libre x_4 y despejamos x_3. En la segunda ecuación despejamos x_2 y finalmente de la primera despejamos x_1, para obtener.

$$\begin{aligned}
&x_5 = -1\\
&x_4 : \textbf{libre}\\
&x_3 = \frac{3}{2}(-\frac{1}{3} - \frac{2}{3}x_5) = -\frac{1}{2} - x_5 = -\frac{1}{2} + 1 = \frac{1}{2}\\
&x_2 = \frac{1}{3}(1 - x_3 - x_5) = \frac{1}{3}(1 - \frac{1}{2} + 1) = \frac{1}{2}\\
&x_1 = 1 - x_2 - x_3 - x_4 - x_5 = 1 - \frac{1}{2} - \frac{1}{2} - x_4 + 1 = 1 - x_4
\end{aligned}$$

La solución general es: $x_1 = 1 - x_4, x_2 = \frac{1}{2}, x_3 = \frac{1}{2},\ x_4 = x_4, x_5 = -1$ Existen entonces infinitas soluciones, una por cada valor de la variable independiente $x_4 \in \mathbb{R}$.

1.7 Sistemas cuadrados y Algoritmo de Gauss

Supongamos que el número de ecuaciones es igual al número de incógnitas $(n = m)$, en este caso el sistema se escribe:

$$Ax = b, \quad A \in M_{nn}(\mathbb{K}),\ x, b \in \mathbb{K}^n$$

Escalonemos el sistema y sea $\tilde{A}$ la matriz escalonada, sin considerar el nuevo lado derecho $\tilde{b}$:

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \cdots & \tilde{a}_{1n} \\ & \ddots & \vdots \\ 0 & & \tilde{a}_{nn} \end{pmatrix}$$

donde entre los elementos de la diagonal, $\tilde{a}_{ii}$, podría haber algunos nulos. Señalemos que en el caso de matrices cuadradas el proceso de escalonamiento se denomina **Algoritmo de Gauss**. Un resultado importante, que relaciona matrices invertibles, solución de sistemas y el algoritmo de Gauss, es el siguiente:

Teorema 1.1. *Sea $A \in \mathcal{M}_{nn}(\mathbb{K})$, entonces las proposiciones siguientes son equivalentes:*

1. *A es invertible.*
2. *$\forall b \in \mathbb{K}^n$, $Ax = b$ tiene solución única.*
3. *$\prod_{i=1}^{n} \tilde{a}_{ii} \neq 0$.*

Demostración. Utilizamos un argumento de transitividad para probar solamente que:

($1 \Rightarrow 2$). Si A es invertible consideremos entonces A^{-1}. Así $(A^{-1}A)x = A^{-1}b$ y luego $x = A^{-1}b$, por ende el sistema tiene solución y claramente esta es la única.

En efecto, si $x^1, x^2 \in \mathbb{K}^n$ son soluciones, entonces $Ax^1 = Ax^2 = b$ y se tendrá multiplicando por A^{-1} que $(A^{-1}A)x^1 = (A^{-1}A)x^2$, pero entonces $x^1 = x^2$.

($2 \Rightarrow 3$). Sea $\tilde{A}$ la matriz escalonada asociada a A y supongamos que para algún i: $\tilde{a}_{ii} = 0$. Esto quiere decir que alguno de los escalones tiene largo mayor o igual que 2 y por lo tanto el sistema homogéneo $Ax = 0$ tiene infinitas soluciones (por Proposición 1.10) lo que es una contradicción. Por lo tanto para todo $i = 1, \ldots, n$ $\tilde{a}_{ii} \neq 0$ o equivalentemente $\prod_{i=1}^{n} \tilde{a}_{ii} \neq 0$.

($3 \Rightarrow 1$). Supongamos $\prod\limits_{i=1}^{n} \tilde{a}_{ii} \neq 0$, esto quiere decir que los elementos diagonales de $\tilde{A}$ son todos no nulos.

Notemos que $\tilde{A} = (\prod_j E_j)A$, donde las matrices E_j son elementales y por lo tanto invertibles. Podríamos ahora escalonar hacia arriba la matriz $\tilde{A}$ para obtener de manera similar una matriz diagonal D. Esto lo podemos hacer de manera que D tenga por diagonal la misma que $\tilde{A}$. Es decir

$$D = \begin{pmatrix} \tilde{a}_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \tilde{a}_{nn} \end{pmatrix} = \left(\prod_j E_j\right) A \left(\prod_k F_k\right),$$

donde las matrices F_k son del mismo estilo que las matrices elementales que hemos usado. Sabemos entonces que las matrices D, $E = \prod_j E_j$ y $F = \prod_k F_k$ son invertibles y como $A = E^{-1}DF^{-1}$ concluimos que A también es invertible.

Como corolario importante de lo demostrado en esta parte es que:

Corolario 1.3. *Una matriz triangular superior es invertible si y sólo si todos sus elementos diagonales son distintos de cero.*

Observación: Notemos que pasando por la traspuesta se deduce que si A es triangular inferior entonces es invertible si y solo si todos sus elementos diagonales son distintos de cero. Más aún la inversa de una triangular inferior es triangular inferior. En efecto, si aplicamos el algoritmo de escalonamiento sin permutar obtendremos:

$$\tilde{A} = \begin{pmatrix} a_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{pmatrix} = \left(\prod_{j=1}^{r} E_j\right) A,$$

donde r es el número de pasos tomados por el algoritmo y las matrices elementales E_j son del tipo $E_{pq}(\lambda, 1)$. Notar que los pivotes son los elementos de la diagonal de

A. Como la inversa de cada E_j es del mismo tipo se deduce que

$$A = \left(\prod_{j=r}^{1} (E_j)^{-1} \right) \tilde{A}.$$

$$\Longrightarrow \quad A^{-1} = (\tilde{A})^{-1} \left(\prod_{j=1}^{r} E_j \right),$$

y por lo tanto la inversa de A es triangular inferior y además sus elementos diagonales son $1/a_{ii}$ $i = 1, .., n$. Notar que el producto de las inversas en la expresión anterior está tomado en sentido opuesto:

$$\left(\prod_{j=1}^{r} E_j \right)^{-1} = \prod_{j=r}^{1} (E_j)^{-1}.$$

Pasando por la traspuesta podemos deducir que la inversa de una triangular superior también es triangular superior.

1.8 Cálculo de la inversa

Supongamos que dada una matriz cuadrada A de tamaño n, buscamos una matriz Z de $n \times n$ tal que $AZ = I$. Si describimos Z por sus columnas, esto es $Z = (Z_{\bullet 1} \cdots Z_{\bullet n})$ entonces la ecuación $AZ = I$ es equivalente a los n sistemas

$$AZ_{\bullet i} = e_i, \quad \text{con } i \in \{1, \ldots, n\}, \text{ y } e_i \text{ la i-ésima columna de } I.$$

Probemos ahora que,

Proposición 1.11. *Si los n sistemas anteriores tienen solución, entonces A es invertible y además $A^{-1} = Z$.*

Demostración. Para ello probemos primero que $\forall b \in \mathbb{K}^n$ el sistema $Ax = b$ tiene solución.

En efecto sea $b = (b_1 \cdots b_n)^t$ un elemento de $\mathbb{K}^n$, y consideremos $a = b_1 Z_{\bullet 1} + \cdots b_n Z_{\bullet n} \in \mathbb{K}^n$. Utilizando las propiedades de la suma, multiplicación y ponderación de matrices se obtiene que

$$\begin{aligned} Aa &= A(b_1 Z_{\bullet 1} + b_2 AZ_{\bullet 2} + \cdots b_n Z_{\bullet n}) = b_1 AZ_{\bullet 1} + b_2 AZ_{\bullet 2} + \cdots b_n AZ_{\bullet n} \\ &= b_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + b_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + \cdots b_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix} = b, \end{aligned}$$

y por lo tanto a es una solución del sistema $Ax = b$.

Si A no fuese invertible esto quiere decir que al escalonar A, se tendrá que algún $\tilde{a}_{ii} = 0$. Consideremos las filas i e $i+1$ de $\tilde{A}$:

$$\tilde{A} = \begin{pmatrix} \vdots \\ \tilde{A}_{\bullet i} \\ \tilde{A}_{\bullet i+1} \\ \vdots \end{pmatrix} = \begin{pmatrix} & & & \vdots & & & \\ 0 & \cdots & 0 & \tilde{a}_{ii} = 0 & \tilde{a}_{ii+1} & \cdots & \tilde{a}_{in} \\ 0 & \cdots & 0 & 0 & \tilde{a}_{i+1i+1} & \cdots & \tilde{a}_{i+1n} \\ & & & \vdots & & & \end{pmatrix}.$$

Si $\tilde{a}_{ii+1} = 0$, $\tilde{a}_{i+1i+1} \neq 0$ habríamos permutados las filas y por lo tanto la matriz escalonada tendría un 0 en la posición $(i+1, i+1)$, lo que es una contradicción. Si $\tilde{a}_{ii+1} \neq 0$ entonces podemos utilizar este elemento para escalonar hacia abajo y por lo tanto $\tilde{a}_{i+1i+1} = 0$. En cualquier caso se deduce que $\tilde{a}_{i+1i+1} = 0$, y por un argumento de inducción que $\tilde{a}_{nn} = 0$, esto es toda la última fila de $\tilde{A}$ debe ser 0. Consideremos ahora la matriz invertible $C = \prod\limits_j E_j$, donde $\tilde{A} = CA$, es decir la matriz que permite pasar de A a $\tilde{A}$. Definamos

$$b = C^{-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Si consideramos el sistema $Ax = b$ tendremos que al escalonar la matriz aumentada $(A|b)$ se obtendrá una matriz escalonada (usando las mismas operciones elemementales)

$$(\tilde{A}|\tilde{b}) = C(A|b) = \begin{pmatrix} & \vdots & & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix},$$

lo que representa un sistema incompatible y por lo tanto $Ax = b$ no puede tener solución. Hemos obtenido una contradicción y por lo tanto la única posibilidad es que $\forall i = 1, .., n$ $\tilde{a}_{ii} \neq 0$, y por lo tanto A es invertible. Ahora de $AZ = I$ se obtiene premultiplicando por A^{-1} que $Z = A^{-1}$.

Notemos que de paso se ha probado el siguiente resultado interesante

Corolario 1.4. *Una matriz $A \in \mathcal{M}_{nn}(\mathbb{K})$ es invertible si y sólo si todo sistema $Ax = b$ tiene solución.*

Lo que a su vez es equivalente a **todo sistema $Ax = b$ tiene solución única**, por lo probado anteriormente.

Ejercicio 1.4: ¿Qué puede decir de una matriz cuadrada A para la cual todo sistema $Ax = b$ tiene a lo más una solución?. ¿Debe ser A invertible?. Pruébelo.

Para saber si una matriz es invertible hemos entonces probado que basta estudiar n sistemas lineales. Para el cálculo de la inversa podemos ir un poco más lejos. Consideremos el siguiente ejemplo:

Ejemplo 1.7.

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

luego:

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

Escalonando:

$$(A|I) \to \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & -1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & -2 & 2 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}$$

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

En este paso ya estamos en condiciones de resolver los tres sistemas. Pero podemos aún pivotear, ahora "sobre" la diagonal; sin alterar la solución (ya que multiplicamos por matrices invertibles):

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

$$\to \begin{pmatrix} 1 & 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & -2 & 0 & -1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 2 & 1 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{4} & \frac{3}{4} \\ 0 & -2 & 0 & -1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 2 & 1 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

Finalmente, premultiplicando por la matriz invertible:

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

Se tiene:

$$(I|\tilde{I}) = \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{4} & \frac{3}{4} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{4} & -\frac{1}{4} \end{pmatrix}$$

Obteniendo los sistemas de solución trivial:

$$I \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, I \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ \frac{1}{4} \end{pmatrix}, I \begin{pmatrix} x_{13} \\ x_{23} \\ x_{33} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \end{pmatrix},$$

de donde:

$$X = A^{-1} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{2} & -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & -\frac{1}{4} \end{pmatrix}$$

En efecto:

$$AA^{-1} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{2} & -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & -\frac{1}{4} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Luego para calcular la inversa de $A \in M_{nn}(\mathbb{K})$ basta aplicar el Algoritmo de Gauss sobre la matriz $(A|I)$. Una vez que se ha obtenido la matriz triangular superior $\tilde{A}$, realizar el mismo procedimiento para los elementos sobre la diagonal. Finalmente, se premultiplica por una matriz diagonal para obtener la identidad en el lugar de $\tilde{A}$.

¿Qué sucede si una matriz A no es invertible?

Sabemos que necesariamente aparecerá, al escalonar, un pivote nulo irreparable.

Ejemplo 1.8.

$$(A|I) = \begin{pmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 & 1 \end{pmatrix} \text{ y } \begin{pmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

y los sistemas $\tilde{A}x = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$, $\tilde{A}x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ son incompatibles.

Pero, en ocasiones, hay pivotes nulos “reparables”.

Ejemplo 1.9.

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

Acá el elemento $\tilde{a}_{22} = 0$, pero podemos, premultiplicando por la matriz de permutación I_{23}, intercambiar las filas 2 y 3 obteniendo:

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

$$\to \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{pmatrix} \implies A^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Ejemplo y ejercicio importante: Descomposición LDU.

Supongamos que al escalonar $A \in M_{nn}(\mathbb{K})$ no encontramos pivotes nulos (¡no permutamos filas!). Se obtiene entonces

$$\begin{pmatrix} \tilde{a}_{11} & \cdots & \tilde{a}_{1p} \\ & \ddots & \\ 0 & & \tilde{a}_{nn} \end{pmatrix} = \left(\prod_j E_j \right) A$$

donde cada una de las matrices E_j es de la forma:

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & 0 & \\ & \lambda & 1 & & \\ \cdots & \cdots & & \ddots & \\ 0 & & & & 1 \end{pmatrix}, \quad \lambda \in \mathbb{K}$$

1. Pruebe que la matriz $\prod\limits_j E_j$ es invertible y su inversa es triangular inferior con 1's en la diagonal. Digamos:

$$\left(\prod_j E_j \right)^{-1} = L = \begin{pmatrix} 1 & & & 0 \\ \ell_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \ell_{n1} & \cdots & \ell_{nn-1} & 1 \end{pmatrix}$$

Despejamos entonces A:

$$A = \left(\prod_j E_j \right)^{-1} \begin{pmatrix} \tilde{a}_{n} & \cdots & \tilde{a}_{1n} \\ & \vdots & \\ 0 & & \tilde{a}_{nn} \end{pmatrix}$$

1. Concluya que si al escalonar A no se realizan permutaciones, A puede factorizarse como el producto de una matriz L, triangular inferior con unos en la diagonal, y una matriz U, triangular superior, donde los pivotes figuran en la diagonal. Esta factorización se llama *descomposición LU*.

2. Recordando que $\prod_{i=1}^{n} \tilde{a}_{ii} \neq 0$, demuestre que:

$$A = \begin{pmatrix} 1 & & & \\ \ell_{21} & \ddots & 0 & \\ \vdots & & \ddots & \\ \ell_{n1} & \cdots & \ell_{nn-1} & 1 \end{pmatrix} \begin{pmatrix} \tilde{a}_{11} & & 0 \\ & \ddots & \\ 0 & & \tilde{a}_{nn} \end{pmatrix} \begin{pmatrix} 1 & \frac{\tilde{a}_{12}}{\tilde{a}_{11}} & \cdots & \frac{\tilde{a}_{1n}}{\tilde{a}_{11}} \\ \vdots & \ddots & & \vdots \\ & & 1 & \frac{\tilde{a}_{n-1n}}{\tilde{a}_{n-1n}} \\ 0 & \cdots & & 1 \end{pmatrix}. \quad (1.4)$$

Es decir, A admite una factorización $A = LDU$, en donde L es triangular inferior con unos en la diagonal, D es diagonal (formada por los pivotes del escalonamiento) y U es triangular superior con diagonal de unos. Esta factorización se llama *descomposición LDU*.

3. Demuestre que si A admite la descomposición LDU de (1.5), entonces la descomposición es única.

4. Demuestre además que si A es simétrica, entonces $L = U^t$.

5. Encuentre la descomposición LDU de

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

Guía de Ejercicios

1. ¿Qué puede decir de una matriz cuadrada A para la cual todo sistema $Ax = b$ tiene a lo más una solución?. ¿Debe ser A invertible?. Pruébelo.

2. Supongamos que al escalonar $A \in M_{nn}(\mathbb{K})$ no encontramos pivotes nulos (¡no permutamos filas!). Se obtiene entonces

$$\begin{pmatrix} \tilde{a}_{11} & \cdots & \tilde{a}_{1p} \\ & \ddots & \\ 0 & & \tilde{a}_{nn} \end{pmatrix} = (\prod_j E_j)A$$

donde cada una de las matrices E_j es de la forma $E_{pq}(\lambda, 1)$, con $\lambda \in \mathbb{K}$.

a) Pruebe que la matriz $\prod_j E_j$ es invertible y su inversa es triangular inferior con 1's en la diagonal. Digamos:

$$(\prod_j E_j)^{-1} = L = \begin{pmatrix} 1 & & & 0 \\ \ell_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \ell_{n1} & \cdots & \ell_{nn-1} & 1 \end{pmatrix}$$

Despejamos entonces A:

$$A = (\prod_j E_j)^{-1} \begin{pmatrix} \tilde{a}_n & \cdots & \tilde{a}_{1n} \\ & \vdots & \\ 0 & & \tilde{a}_{nn} \end{pmatrix}$$

a) Concluya que si al escalonar A no se realizan permutaciones, A puede factorizarse como el producto de una matriz L, triangular inferior con unos en la diagonal, y una matriz U, triangular superior, donde los pivotes figuran en la diagonal. Esta factorización se llama *descomposición LU*.

b) Recordando que $\prod_{i=1}^{n} \tilde{a}_{ii} \neq 0$, demuestre que:

$$A = \begin{pmatrix} 1 & & & \\ \ell_{21} & \ddots & 0 & \\ \vdots & & \ddots & \\ \ell_{n1} & \cdots & \ell_{nn-1} & 1 \end{pmatrix} \begin{pmatrix} \tilde{a}_{11} & & 0 \\ & \ddots & \\ 0 & & \tilde{a}_{nn} \end{pmatrix} \begin{pmatrix} 1 & \frac{\tilde{a}_{12}}{\tilde{a}_{11}} & \cdots & \frac{\tilde{a}_{1n}}{\tilde{a}_{11}} \\ \vdots & \ddots & & \vdots \\ & & 1 & \frac{\tilde{a}_{n-1n}}{\tilde{a}_{n-1n}} \\ 0 & \cdots & & 1 \end{pmatrix}. \tag{1.5}$$

Es decir, A admite una factorización $A = LDU$, en donde L es triangular inferior con unos en la diagonal, D es diagonal (formada por los pivotes del escalonamiento) y U es triangular superior con diagonal de unos. Esta factorización se llama *descomposición LDU*.

c) Demuestre que si A admite la descomposición LDU de (1.5), entonces la descomposición es única.

d) Demuestre además que si A es simétrica, entonces $L = U^t$.

e) Encuentre la descomposición LDU de

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

Guía de Problemas

P1. **(a)** Sea la matriz de coeficientes reales $A = \begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix}$. Demueste que si la ecuación $Ax = 0$ tiene una única solución entonces $(a \neq b) \wedge (a \neq c) \wedge (b \neq c)$.

(b) Considere el sistema lineal a coeficientes reales siguiente

$$\begin{array}{rcrcrcrcl} -x_1 & & & + & \alpha x_3 & + & \beta x_4 & = & 0\,, \\ & & x_2 & + & \alpha x_3 & + & \beta x_4 & = & 0\,, \\ \alpha x_1 & + & \beta x_2 & & & & & = & 0\,, \\ \beta x_1 & + & \beta x_2 & + & \alpha x_3 & & & = & 0\,. \end{array}$$

Determine las condiciones sobre los parámetros reales α y β que garanticen que el sistema tenga una única solución.

P2. **(a)** Considere las matrices cuadradas $A, B \in \mathcal{M}_{nn}(\mathbb{R})$. Demuestre que

1) A es invertible si y sólo si AA^t es invertible.
2) Si $A^2 = A$ y $B = I - A$ entonces $B^3 = B$.

Si A es invertible, utilice las condiciones dadas para calcular las matrices A y B.

(b) Considere los vectores $u, v \in \mathcal{M}_{n1}(\mathbb{R}) \setminus \{\mathbf{0}\}$. Se define la matriz $A \in \mathcal{M}_{nn}(\mathbb{R})$ por $A = uv^t$.

1) Pruebe que $\forall x \in \mathcal{M}_{n1}(\mathbb{R}),\ Ax = \mathbf{0} \iff v^t x = 0$.
Indicación: Observe que $v^t x \in \mathbb{R}$.
2) Encuentre el número de variables libres en la resolución del sistema $Ax = 0$.

¿Es A invertible?

P3. Considere

$$A = \begin{pmatrix} 1 & 2 & 3 & -1 \\ -1 & 1 & 0 & 1 \\ 0 & 3 & 3 & \alpha - 3 \\ 2 & 1 & \alpha & -2 \end{pmatrix} \quad b = \begin{pmatrix} \beta \\ \beta^2 \\ 0 \\ -2 \end{pmatrix}.$$

Estudiaremos las soluciones del sistema $Ax = b$, con $x \in \mathcal{M}_{41}(\mathbb{R})$. Encuentre los valores de α y β de manera que:

- El sistema no tenga solución.
- El sistema tenga una única solución. Encuentre la matriz inversa de A y calcule la única solución del sistema.
- El sistema tenga infinitas soluciones y encuentre sus soluciones. Determine además el número de variables independientes.

SEMANA 4:

Geometría lineal en $\mathbb{K}^n$

2.1 Definiciones básicas

Sea $\mathbb{K}$ un cuerpo. Anotaremos los vectores de $\mathbb{K}^n$ como columnas[1], es decir,

$$\mathbb{K}^n = M_{n,1}(\mathbb{K}) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in \mathbb{K},\ i = 1, \dots, n \right\}.$$

Recordar que en $\mathbb{K}^n = M_{n,1}(\mathbb{K})$ tenemos una *suma* (de matrices columna) definida por:

si $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{K}$, entonces $\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$,

y con ella resulta que $(\mathbb{K}^n, +)$ es un grupo Abeliano, con neutro $\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ y opuestos

$$-\mathbf{x} = -\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix}.$$

Agregamos ahora la operación *Multiplicación por Escalar* definida como sigue:

Definición Si $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ y $\lambda \in \mathbb{K}$, entonces se define una función $\begin{array}{ccc} \mathbb{K} \times \mathbb{K}^n & \to & \mathbb{K}^n \\ (\lambda, \mathbf{x}) & \mapsto & \lambda \mathbf{x} \end{array}$, en donde $\lambda \mathbf{x} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$. Se le llama *ponderación* o *multiplicación por escalar.*

[1] Para fijar ideas puede pensar en lo que sigue en $\mathbb{K} = \mathbb{R}$ y $n = 2$ ó 3, que con seguridad le serán familiares de sus cursos de Cálculo. Sin embargo, salvo cuando se diga explícitamente lo contrario, $\mathbb{K}$ puede ser cualquier cuerpo, y n un natural mayor o igual a 1.

Observación: Llamaremos *vectores columna*, o simplemente *vectores*, a los elementos de $\mathbb{K}^n$, y *escalares* a los elementos de $\mathbb{K}$. De aquí el nombre de esta operación.

Proposición 2.1. *Se tienen las siguientes propiedades:*

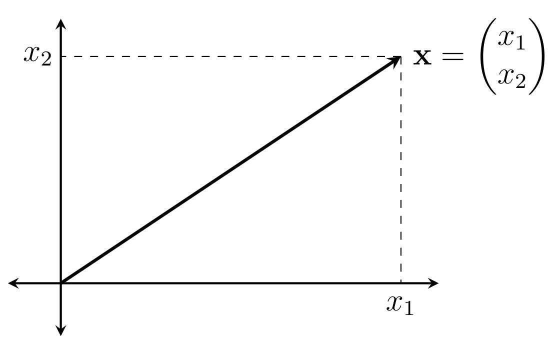
1. $(\forall \mathbf{x} \in \mathbb{K}^n)\ 1 \cdot \mathbf{x} = \mathbf{x}$
2. $(\forall \lambda_1, \lambda_2 \in \mathbb{K})\ (\forall \mathbf{x} \in \mathbb{K}^n)\ \lambda_1(\lambda_2 \mathbf{x}) = (\lambda_1 \lambda_2)\mathbf{x}$
3. $(\forall \lambda_1, \lambda_2 \in \mathbb{K})\ (\forall \mathbf{x} \in \mathbb{K}^n)\ (\lambda_1 + \lambda_2)\mathbf{x} = \lambda_1 \mathbf{x} + \lambda_2 \mathbf{x}$
4. $(\forall \lambda \in \mathbb{K})\ (\forall \mathbf{x}, \mathbf{y} \in \mathbb{K}^n)\ \lambda(\mathbf{x} + \mathbf{y}) = \lambda \mathbf{x} + \lambda \mathbf{y}$.
5. $(\forall \lambda \in \mathbb{K})\ (\forall \mathbf{x} \in \mathbb{K}^n)\ \lambda \mathbf{x} = \mathbf{0} \iff \lambda = 0 \vee \mathbf{x} = \mathbf{0}$.

Demostración. Las demostraciones de estos hechos son cálculos elementales y quedan como un ejercicio muy simple para el alumnos.

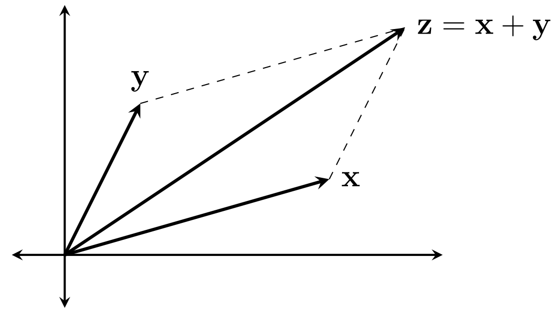
Queda además el lector encargado de verificar que esta última propiedad se puede también obtener como consecuencia de las propiedades 1. a 4. , más los hechos de que $(\mathbb{K}^n, +)$ es un grupo Abeliano y $\mathbb{K}$ un cuerpo. $\square$

Ejemplos: Caso $\mathbb{K} = \mathbb{R}$.

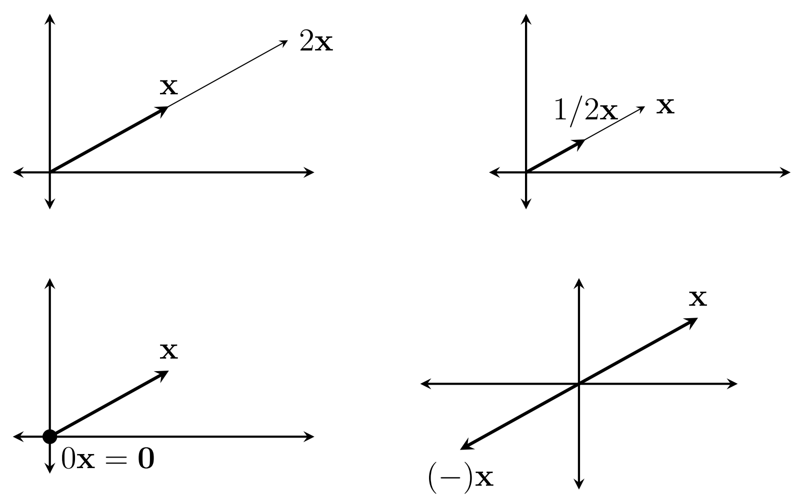
- $n = 1$: El conjunto de vectores es $\mathbb{K}^1 = \mathbb{R}$, la recta real, el $+$ es la suma usual en $\mathbb{R}$ y la multiplicación por escalar es el producto en $\mathbb{R}$.
- $n = 2$: El conjunto de vectores es $\mathbb{K}^2 = \mathbb{R}^2$, el "plano $x_1 - x_2$".



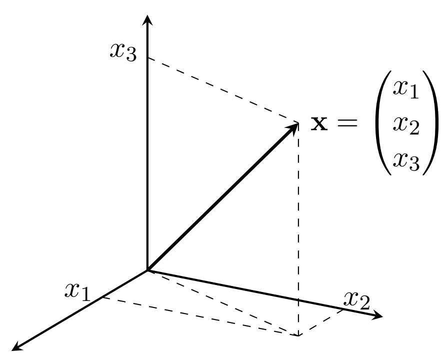
Un vector $\mathbf{x}$ representará ambos, un punto del plano y una flecha partiendo del origen que termina en dicho punto. La suma de elementos de $\mathbb{R}^2$ corresponde entonces a la suma usual de flechas en el plano:



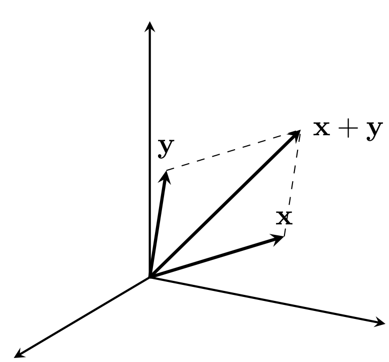
y del mismo modo se representa la ponderación:



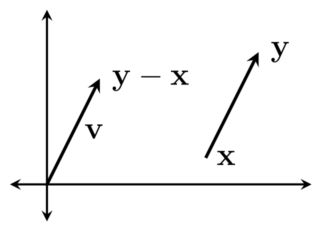
- $n = 3$: La representación de vectores en $\mathbb{R}^3$ se hace en el espacio 3-dimensional:



La ponderación se representa de modo similar al caso plano, y la suma de vectores también corresponde a la diagonal del paralelógramo definido por los sumandos:

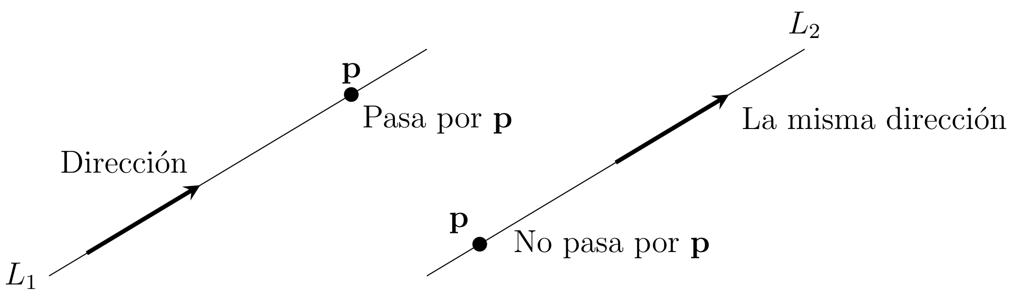


Observación: A veces conviene (por ejemplo, para mayor claridad de una figura) dibujar un vector como una flecha que parte de algún otro punto distinto del origen. A modo de ejemplo consideremos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ y sea $\mathbf{v} = \mathbf{y} - \mathbf{x}$. Se ve fácilmente que $\mathbf{v}$ será una flecha partiendo en el origen que tiene el mismo tamaño y apunta hacia el mismo lado que una flecha que parta en $\mathbf{x}$ y termine en $\mathbf{y}$. Dado ésto, muchas veces no nos molestamos en dibujar $\mathbf{v}$ partiendo del origen, sino que sólo dibujamos la flecha de $\mathbf{x}$ a $\mathbf{y}$.



2.2 Rectas en $\mathbb{K}^n$

A partir de nuestra intuición geométrica rescatamos del concepto de recta (en el plano o en el *espacio*) dos elementos importantes: que una recta define una *dirección de movimiento* y que *pasa por algún lugar*.

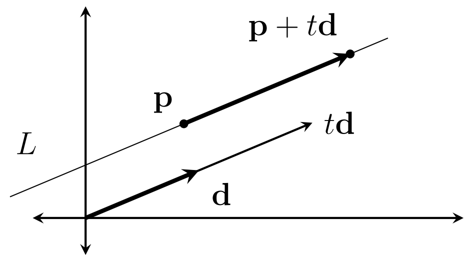


Las rectas L_1 y L_2 del dibujo *llevan la misma dirección*, y lo que las diferencia es que *pasan por distintos puntos*. De hecho L_2 es L_1 *desplazada*. Intentamos entonces *definir* una recta en $\mathbb{K}^n$ como sigue:

Definición (Recta) Sean $\mathbf{d} \in \mathbb{K}^n \setminus \{\mathbf{0}\}$ un vector y $\mathbf{p} \in \mathbb{K}^n$ un punto dado. La recta L que pasa por $\mathbf{p}$ y va en la dirección de $\mathbf{d}$ es el siguiente subconjunto de $\mathbb{K}^n$:

$$L = L_{\mathbf{p},\mathbf{d}} = \{\mathbf{v} \in \mathbb{K}^n \mid \mathbf{v} = \mathbf{p} + t\mathbf{d},\ t \in \mathbb{K}\}.$$

Se dice que $\mathbf{p}$ es una *posición* de L y que $\mathbf{d}$ es un *vector director* de L.



Ejercicio 2.1: Un par de observaciones importantes:

1. Muchas posiciones definen la misma recta. Probar que, dado $\mathbf{d} \in \mathbb{K}^n \setminus \{\mathbf{0}\}$ fijo, se tendrá $\mathbf{q} \in L_{\mathbf{p},\mathbf{d}} \iff L_{\mathbf{q},\mathbf{d}} = L_{\mathbf{p},\mathbf{d}}$. (Las posiciones de una recta L son sus puntos.)

2. Muchos vectores directores definen una misma recta. Probar que, dado $\mathbf{p} \in \mathbb{K}^n$ fijo, se tendrá $L_{\mathbf{p},\mathbf{d}} = L_{\mathbf{p},\overline{\mathbf{d}}} \iff (\exists \lambda \in \mathbb{K} \setminus \{0\})\,\overline{\mathbf{d}} = \lambda \mathbf{d}$. (Los vectores directores de una misma recta son los múltiplos no nulos de un vector director dado.)

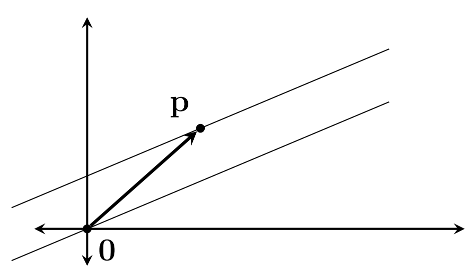
El Ejercicio 2.1.2 motiva la siguiente definición:

Definición (Vectores paralelos) Sean $\mathbf{v}, \mathbf{w} \in \mathbb{K}^n$. El vector $\mathbf{v}$ se dice *paralelo* a $\mathbf{w}$ (lo que se anota $\mathbf{v}\|\mathbf{w}$) ssi $(\exists \lambda \in \mathbb{K} \setminus \{0\})\, \mathbf{w} = \lambda \mathbf{v}$.

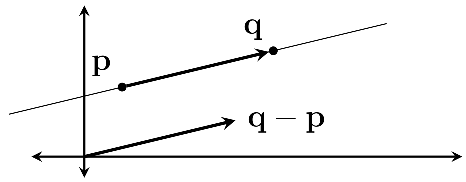
Usando esta definición, el Ejercicio 2.1.2 puede ser replanteado como

$$\overline{\mathbf{d}} \text{ es un vector director de } L_{\mathbf{p},\mathbf{d}} \iff \overline{\mathbf{d}}\|\mathbf{d}.$$

Ejercicio 2.2: 1. Si $A \subseteq \mathbb{K}^n$ y $\mathbf{p} \in \mathbb{K}^n$, definimos la *traslación* de A en el vector $\mathbf{p}$ por $\mathbf{p} + A = \{\mathbf{p} + \mathbf{x} \mid \mathbf{x} \in A\}$. Mostrar que las rectas que pasan por $\mathbf{p} \in \mathbb{K}^n$ son exactamente las traslaciones en $\mathbf{p}$ de las rectas que pasan por el origen $\mathbf{0}$.



2. Sean $\mathbf{p}, \mathbf{q} \in \mathbb{K}^n$, con $\mathbf{p} \neq \mathbf{q}$. Mostrar que la recta de posición $\mathbf{p}$ y vector director $\mathbf{d} = \mathbf{q} - \mathbf{p}$ es la única recta que pasa por ambos puntos, $\mathbf{p}$ y $\mathbf{q}$.



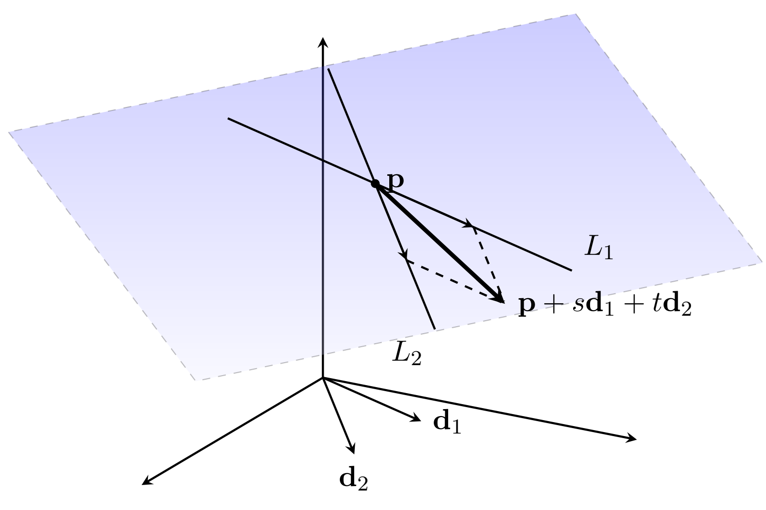
Observación: La ecuación

$$\mathbf{v} = \mathbf{p} + \lambda\mathbf{d}, \quad \lambda \in \mathbb{K}$$

que describe los puntos (vectores) de la recta $L_{\mathbf{p},\mathbf{d}}$ a medida que se varía $\lambda \in \mathbb{K}$ se llama *ecuación vectorial* de la recta $L_{\mathbf{p},\mathbf{d}}$.

2.3 Planos en $\mathbb{K}^n$

Sean $\mathbf{p} \in \mathbb{K}^n$ y $\mathbf{d}_1, \mathbf{d}_2 \in \mathbb{K}^n \setminus \{\mathbf{0}\}$ vectores no paralelos. Apelando a nuestra intuición en el espacio tridimensional, vemos que por $\mathbf{p}$ pasan las dos rectas distintas $L_1 = L_{\mathbf{p},\mathbf{d}_1}$ y $L_2 = L_{\mathbf{p},\mathbf{d}_2}$, y ambas están contenidas en un mismo plano Π.



Observando la forma como los puntos de este plano se obtienen a partir de $\mathbf{p}$, $\mathbf{d}_1$ y $\mathbf{d}_2$ podemos dar la siguiente definición en $\mathbb{K}^n$:

Definición (Plano) El plano que pasa por $\mathbf{p}$ y tiene vectores directores $\mathbf{d}_1$ y $\mathbf{d}_2$ es el subconjunto de $\mathbb{K}^n$:

$$\Pi_{\mathbf{p},\mathbf{d}_1,\mathbf{d}_2} = \{\mathbf{v} \in \mathbb{K}^n \mid \mathbf{v} = \mathbf{p} + s\mathbf{d}_1 + t\mathbf{d}_2\,, s, t \in \mathbb{K}\}$$

Observación:

1. $\mathbf{v} = \mathbf{p} + s\mathbf{d}_1 + t\mathbf{d}_2$, $s, t \in \mathbb{K}$ se llama *ecuación vectorial* del plano $\Pi_{\mathbf{p},\mathbf{d}_1,\mathbf{d}_2}$.

2. Sea $E = \left[\ \mathbf{d}_1 \mid \mathbf{d}_2\ \right]$ la matriz que tiene por columnas a los vectores directores $\mathbf{d}_1$ y $\mathbf{d}_2$. Entonces juntando los parámetros s y t en el vector $\mathbf{r} = \begin{pmatrix} s \\ t \end{pmatrix} \in \mathbb{K}^2$, la ecuación vectorial queda

$$\begin{aligned} \mathbf{v} &= \mathbf{p} + E\mathbf{r} && ,\mathbf{r} \in \mathbb{K}^2 \\ &= \mathbf{p} + \left[\ \mathbf{d}_1 \mid \mathbf{d}_2\ \right] \begin{pmatrix} s \\ t \end{pmatrix} && ,s, t \in \mathbb{K} \end{aligned}$$

3. Notar la similitud con la ecuación de una recta $\mathbf{v} = \mathbf{p} + t\mathbf{d}$, $t \in \mathbb{K}$. En la recta hay un parámetro libre: $t \in \mathbb{K}$ (*dimensión 1*, esto se formalizará mas adelante) y en el plano dos: $s, t \in \mathbb{K}$ (*dimensión 2*).

Ejercicio 2.3: Se deja como ejercicio para el lector probar las siguientes propiedades:

1. $\mathbf{q} \in \Pi_{\mathbf{p},\mathbf{d}_1,\mathbf{d}_2} \Leftrightarrow \Pi_{\mathbf{q},\mathbf{d}_1,\mathbf{d}_2} = \Pi_{\mathbf{p},\mathbf{d}_1,\mathbf{d}_2}$. (Cualquier punto de un plano sirve como su posición.)

2. Si $\mathbf{p} \in \mathbb{K}^n$, y $\mathbf{d}_1, \mathbf{d}_2, \overline{\mathbf{d}_1}, \overline{\mathbf{d}_2} \in \mathbb{K}^n \setminus \{\mathbf{0}\}$, con $\mathbf{d}_1$no paralelo a $\mathbf{d}_2$ y $\overline{\mathbf{d}_1}$ no paralelo a $\overline{\mathbf{d}_2}$, entonces

$$\Pi_{\mathbf{p},\mathbf{d}_1,\mathbf{d}_2} = \Pi_{\mathbf{p},\overline{\mathbf{d}_1},\overline{\mathbf{d}_2}} \Leftrightarrow (\exists A \in M_{2,2}(\mathbb{K}) \text{ invertible}\,)\ \left[\ \overline{\mathbf{d}_1} \mid \overline{\mathbf{d}_2}\ \right] = \left[\ \mathbf{d}_1 \mid \mathbf{d}_2\ \right] A$$

3. Dado $\mathbf{p} \in \mathbb{K}^n$, todo plano que pasa por $\mathbf{p}$ es la traslación en $\mathbf{p}$ de un plano que pasa por el origen, y si trasladamos en $\mathbf{p}$ un plano cualquiera que pasa por el origen, obtenemos un plano que pasa por $\mathbf{p}$.

4. Sean $\mathbf{p}, \mathbf{q}, \mathbf{r}$ tres puntos *no colineales* en $\mathbb{K}^n$. Probar que si $\mathbf{d}_1 = \mathbf{q} - \mathbf{p}$ y $\mathbf{d}_2 = \mathbf{r} - \mathbf{p}$, entonces $\Pi_{\mathbf{p},\mathbf{d}_1,\mathbf{d}_2}$ es el único plano que pasa por $\mathbf{p}$, $\mathbf{q}$ y $\mathbf{r}$.

2.4 Ecuaciones paramétricas y cartesianas de rectas y planos

Rectas

Sean $\mathbf{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} \in \mathbb{K}^n$ y $\mathbf{d} = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} \in \mathbb{K}^n \setminus \{\mathbf{0}\}$. Sea $L = L_{\mathbf{p},\mathbf{d}}$. Recordemos que la ecuación vectorial de L es $\mathbf{v} = \mathbf{p} + t\mathbf{d}$, $t \in \mathbb{K}$. Si anotamos $\mathbf{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ el punto que recorre L cuando t recorre $\mathbb{K}$, resulta:

$$\begin{cases} x_1 &= p_1 + td_1 \\ x_2 &= p_2 + td_2 \\ &\vdots \\ x_n &= p_n + td_n \end{cases} , \qquad t \in \mathbb{K}.$$

Estas son las llamadas *ecuaciones paramétricas* de la recta L. (Notar que $p_1, \ldots, p_n, d_1, \ldots, d_n$ son constantes en $\mathbb{K}$, las coordenadas de la posición y de la dirección de L.)

Supongamos, para simplificar la notación, que las primeras k componentes $d_1, \ldots, d_k$ de d son no nulas, y el resto todas nulas: $d_{k+1} = d_{k+2} = \ldots = d_n = 0$, es decir $\mathbf{d} = \begin{pmatrix} d_1 \\ \vdots \\ d_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

Entonces las ecuaciones paramétricas de la recta L son

$$\begin{aligned} x_1 &= p_1 + td_1 \\ &\vdots \\ x_k &= p_k + td_k \\ \left.\begin{aligned} x_{k+1} &= p_{k+1} \\ &\vdots \\ x_n &= p_n \end{aligned}\right\} & \text{ Constantes.} \end{aligned}$$

Despejando t en cada una de las primeras k ecuaciones, e igualando:

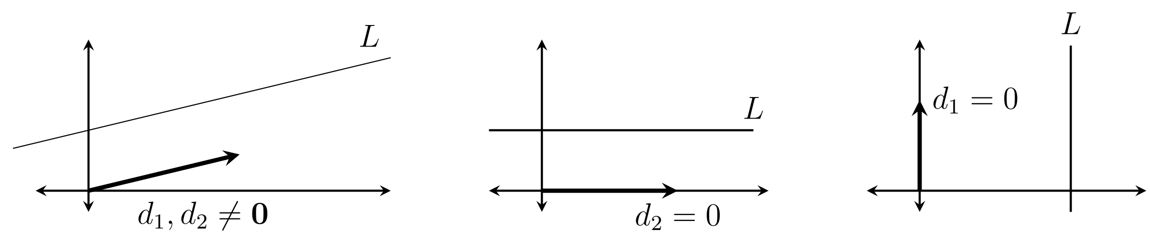
$$\begin{cases} \frac{x_1 - p_1}{d_1} = \frac{x_2 - p_2}{d_2} = \cdots = \frac{x_k - p_k}{d_k} \quad (= t) \\ x_{k+1} = p_{k+1} \\ \vdots \\ x_n = p_n \end{cases}$$

Este es un sistema de $n-1$ ecuaciones lineales para las n incógnitas $x_1, \ldots, x_n$, que son las componentes de un punto cualquiera $\mathbf{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ de L. Las ecuaciones se llaman *ecuaciones Cartesianas* de la recta L. ¿Cómo quedan las ecuaciones cuando $k = 1$?

Ejemplos:

- $n = 2$: *La ecuación Cartesiana* (notar que en este caso queda una sola) de $L = L_{\left(\begin{smallmatrix} p_1 \\ p_2 \end{smallmatrix}\right), \left(\begin{smallmatrix} d_1 \\ d_2 \end{smallmatrix}\right)}$ queda:

$$\begin{aligned} \frac{x_1 - p_1}{d_1} = \frac{x_2 - p_2}{d_2}, &\quad \text{si } d_1, d_2 \neq 0, \\ x_2 = p_2, &\quad \text{si } d_2 = 0, \text{ y} \\ x_1 = p_1, &\quad \text{si } d_1 = 0. \end{aligned}$$



- $n = 3$: Una recta queda descrita por *dos* ecuaciones Cartesianas:

$$\frac{x_1 - p_1}{d_1} = \frac{x_2 - p_2}{d_2} = \frac{x_3 - p_3}{d_3}$$

(en caso que $d_1, d_2, d_3 \neq 0$. Escribir las otras posibilidades.)

Planos (caso $n = 3$)

Para simplificar el análisis veremos sólo el caso $n = 3$ para las ecuaciones Cartesianas de planos. Sea $\Pi = \Pi_{\mathbf{p}, \mathbf{d}_1, \mathbf{d}_2}$ un plano en $\mathbb{K}^3$, donde $\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} \in \mathbb{K}^3$ es una posición de Π, y los vectores directores son los vectores no paralelos en $\mathbb{K}^3 \setminus \{\mathbf{0}\}$, $\mathbf{d}_1 = \begin{pmatrix} d_{11} \\ d_{21} \\ d_{31} \end{pmatrix}$ y $\mathbf{d}_2 = \begin{pmatrix} d_{12} \\ d_{22} \\ d_{32} \end{pmatrix}$. La ecuación vectorial de Π está dada por $\mathbf{v} = \mathbf{p} + s\mathbf{d}_1 + t\mathbf{d}_2$, $s, t \in \mathbb{K}$, y descomponiéndola en sus 3 componentes resultan las *ecuaciones paramétricas* de Π:

$$\begin{cases} x_1 &= p_1 + sd_{11} + td_{12} \\ x_2 &= p_2 + sd_{21} + td_{22} \\ x_3 &= p_3 + sd_{31} + td_{32} \end{cases} \qquad ,s,t \in \mathbb{K}.$$

Pivoteando, es posible eliminar el parámetro s de dos de estas ecuaciones, y luego también t de una de ellas, de modo que nos quedamos con una ecuación final con sólo x_1, x_2, x_3 (sin s y t) de la forma $Ax_1 + Bx_2 + Cx_3 = D$, donde $A, B, C, D \in \mathbb{K}$, con A, B o $C \neq 0$. Esta es la llamada *ecuación Cartesiana* del plano Π.

Observación: Si el plano hubiese estado en $\mathbb{K}^n$, con $n \geq 3$, habríamos obtenido un sistema de $n-2$ ecuaciones cartesianas para el plano. ¿Qué pasa en el caso $n = 2$? ¿Y qué puede decir de $n = 1$?

Sabemos del capítulo de Sistemas Lineales que, recíprocamente, un sistema de una ecuación y tres incógnitas, como la ecuación Cartesiana de Π, tiene dos variables libres (digamos x_2 y x_3) y sus soluciones son de la forma:

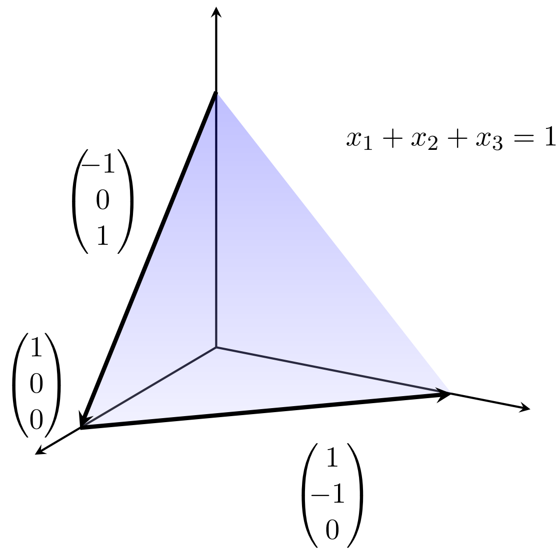
$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} \beta_1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} \gamma_1 \\ 0 \\ 1 \end{pmatrix} \qquad ,x_2, x_3 \in \mathbb{K},$$

con $\begin{pmatrix} \beta_1 \\ 1 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} \gamma_1 \\ 0 \\ 1 \end{pmatrix}$ claramente no paralelos (¿por qué?), así el conjunto de soluciones representa un plano en $\mathbb{K}^3$. Tenemos entonces que los planos en $\mathbb{K}^3$ son exactamente los conjuntos solución de ecuaciones cartesianas como la que escribimos arriba para Π.

Ejemplo 2.1.
El plano $\Pi \subseteq \mathbb{R}^3$ que pasa por el punto $\mathbf{p} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ y con vectores directores $\mathbf{d}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ y $\mathbf{d}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ tiene por ecuaciones paramétricas $\begin{cases} x_1 &= 1 + s - t \\ x_2 &= -s \\ x_3 &= t \end{cases}$ $,s,t \in \mathbb{R}$ de las que se puede fácilmente eliminar s y t para obtener la ecuación Cartesiana $x_1 + x_2 + x_3 = 1$. Por otra parte, si partimos de esta ecuación, despejando x_1 en términos de x_2 y x_3 resulta $x_1 = 1 - x_2 - x_3$, por lo que un punto cualquiera $\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$ que satisfaga esta ecuación tendrá la forma

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 - x_2 - x_3 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \text{ con } x_2, x_3 \in \mathbb{R}$$

tomando valores arbitrarios. El lector reconocerá esto como una ecuación vectorial del mismo plano Π.



2.5 Geometría métrica en $\mathbb{R}^n$

Definiciones y propiedades básicas

Incorporaremos ahora a nuestro estudio conceptos como distancia entre puntos, ángulos, perpendicularidad, etc. Para ello debemos perder un poco de generalidad y restringirnos al cuerpo $\mathbb{K} = \mathbb{R}$ de los números reales. (También se puede hacer todo esto con $\mathbb{K} = \mathbb{C}$, los números complejos, pero esto lo postergaremos para mas adelante.)

Definición (Producto punto en $\mathbb{R}^n$) Sean $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, con $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. Se define el *producto punto* $\langle \mathbf{x}, \mathbf{y} \rangle$ (también se anota $\mathbf{x} \bullet \mathbf{y}$) como el real

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i = \mathbf{x}^t \mathbf{y} = \mathbf{y}^t \mathbf{x} \in \mathbb{R}.$$

Tenemos así una función

$$\begin{array}{rccc} <,> \;:\; & \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & (\mathbf{x}, \mathbf{y}) & \longmapsto & \langle \mathbf{x}, \mathbf{y} \rangle \end{array}$$

llamada *producto punto.*

Proposición 2.2. *Se tienen las siguientes propiedades:*

1. $(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n)\ \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ *(Simetría).*

2. $(\forall \mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}} \in \mathbb{R}^n)\ \langle \mathbf{x} + \overline{\mathbf{x}}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \overline{\mathbf{x}}, \mathbf{y} \rangle$ *y* $\langle \mathbf{x}, \mathbf{y} + \overline{\mathbf{y}} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \overline{\mathbf{y}} \rangle$ *(Bi–aditividad).*

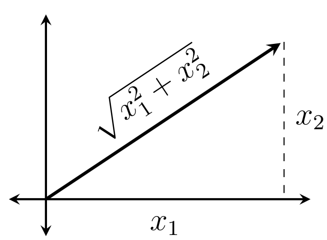
3. $(\forall \lambda \in R)\,(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n)\ \langle \lambda \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \lambda \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle$ *(Bi–homogeneidad).*

4. $(\forall \mathbf{x} \in \mathbb{R}^n)\ \langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ *y* $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}$ *(Positividad).*

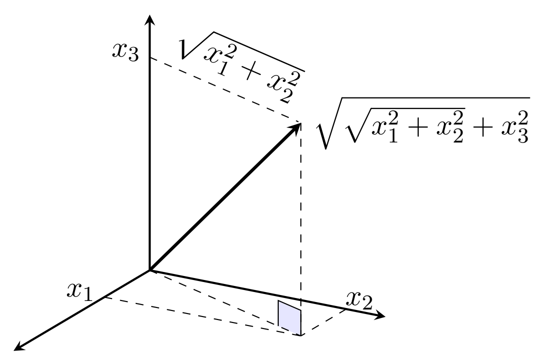
Demostración. La demostración de estas propiedades quedan como ejercicio.

El número $\langle \mathbf{x}, \mathbf{x} \rangle$ que aparece en la propiedad 4. es $\langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^{n} x_i^2$, si $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Jugando con el Teorema de Pitágoras se puede verificar fácilmente que si $n = 2$ ó $n = 3$, $\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ corresponde a la longitud de la flecha representada por el vector $\mathbf{x}$.



$n = 2$



$n = 3$

Aprovechando estos hechos, y con la esperanza de re–obtener el Teorema de Pitágoras en $\mathbb{R}^n$, para todo n, definimos

Definición (Norma Euclidiana) La *norma Euclidiana* (o, simplemente, *norma*) de un vector $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ como $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ (raíz cuadrada ≥ 0).

Tenemos así la función norma

$$\begin{array}{rccc} \|\ \| & : & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & & \mathbf{x} & \longmapsto & \|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \end{array}.$$

Proposición 2.3. *Se tienen las siguientes propiedades:*

1. $(\forall \mathbf{x} \in \mathbb{R}^n)$ $\|\mathbf{x}\| \geq 0$ *y* $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$

2. $(\forall \lambda \in \mathbb{R})\,(\forall \mathbf{x} \in \mathbb{R}^n)$ $\|\lambda \mathbf{x}\| = |\lambda|\, \|\mathbf{x}\|$

3. $(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n)$ $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ *(Desigualdad Triangular.)*

Demostración. Las demostraciones de 1 y 2 son elementales y se dejan como un ejercicio fácil para el lector. La prueba de la Desigualdad Triangular dependerá del siguiente importante Lema:

Lema 1 (Desigualdad de Cauchy–Schwartz).

$$(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n)\ \ |\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|\, \|\mathbf{y}\|\,.$$

Demostración. (Desigualdad de C–S) Notar que si $\mathbf{x} = \mathbf{0} \vee \mathbf{y} = \mathbf{0}$, la desigualdad se verifica trivialmente $(0 \leq 0)$. Supondremos entonces que ambos vectores $\mathbf{x}$ e $\mathbf{y}$ son no nulos. Para $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ fijos, consideremos ahora la función $\varphi : \mathbb{R} \to \mathbb{R}$ definida por $\varphi(t) = \|t\mathbf{x} - \mathbf{y}\|^2$. Claramente $\varphi(t) \geq 0$ para todo $t \in \mathbb{R}$ (y además, si $\mathbf{y}$ no es un múltiplo de $\mathbf{x}$, $\varphi(t) > 0$ para todo $t \in \mathbb{R}$). Por otra parte, $\varphi(t) = \langle t\mathbf{x} - \mathbf{y}, t\mathbf{x} - \mathbf{y} \rangle = \|\mathbf{x}\|^2 t^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle t + \|\mathbf{y}\|^2$. Así, φ es una función polinomial de segundo grado en t, con coeficientes $a = \|\mathbf{x}\|^2$, $b = -2\langle \mathbf{x}, \mathbf{y} \rangle$, $c = \|\mathbf{y}\|^2$ (notar que como $\mathbf{x} \neq \mathbf{0}$, la función es efectivamente de segundo grado, y no de un grado menor). Como $\varphi(t) \geq 0$ para todo $t \in \mathbb{R}$, sabemos que el discriminante $b^2 - 4ac$ será menor o igual a 0 (a lo más una raíz real de multiplicidad 2), es decir $4\langle \mathbf{x}, \mathbf{y} \rangle^2 - 4\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \leq 0$, de donde $\langle \mathbf{x}, \mathbf{y} \rangle^2 \leq \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$, lo que implica la desigualdad deseada.

> **Observación:** Notar que si $\mathbf{y}$ no es múltiplo de $\mathbf{x}$, $\varphi(t) > 0$ para todo $t \in \mathbb{R}$, luego $b^2 - 4ac < 0$, de donde se desprende que $|\langle \mathbf{x}, \mathbf{y} \rangle| < \|\mathbf{x}\|\, \|\mathbf{y}\|$. De aquí que (aún en el caso $\mathbf{x} = \mathbf{0} \vee \mathbf{y} = \mathbf{0}$) se verifica la *igualdad* en Cauchy–Schwartz ssi alguno de los dos vectores es múltiplo del otro. De modo similar se tiene que $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\, \|\mathbf{y}\|$ ssi alguno de los dos vectores es un *múltiplo no negativo* del otro.

Podemos ahora probar la desigualdad triangular. Sean $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Tenemos:

$$\begin{aligned}
\|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 \\
&\leq \|\mathbf{x}\|^2 + 2|\langle \mathbf{x}, \mathbf{y} \rangle| + \|\mathbf{y}\|^2 \\
&\overset{\text{C.–S.}}{\leq} \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\, \|\mathbf{y}\| + \|\mathbf{y}\|^2 \\
&= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2
\end{aligned}$$

$\square$

Observación: De la observación bajo la demostración de la desigualdad de Cauchy–Schwartz podemos concluir que en la desigualdad triangular vale la *igualdad* ssi alguno de los dos vectores es múltiplo no negativo del otro.

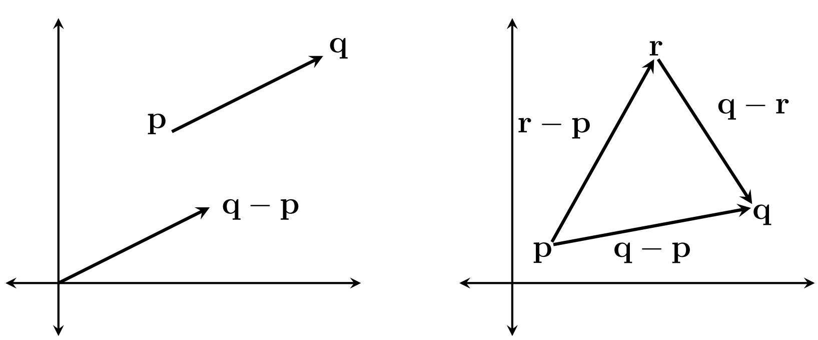
Usaremos ahora la norma de $\mathbb{R}^n$ para definir la noción de distancia entre puntos de este conjunto.

DEFINICIÓN Sean $\mathbf{p}, \mathbf{q}$ dos puntos en $\mathbb{R}^n$. La *distancia entre* $\mathbf{p}$ *y* $\mathbf{q}$ es el número real no negativo $\mathrm{d}(\mathbf{p}, \mathbf{q}) = \|\mathbf{q} - \mathbf{p}\|$.

Observación:

1. Recordando que $\mathbf{q} - \mathbf{p}$ se puede visualizar como el vector (flecha) que va de $\mathbf{p}$ hasta $\mathbf{q}$, resulta que $\mathrm{d}(\mathbf{p}, \mathbf{q})$ es la longitud de esta flecha, es decir, como esperaríamos, la longitud del trazo recto entre $\mathbf{p}$ y $\mathbf{q}$.

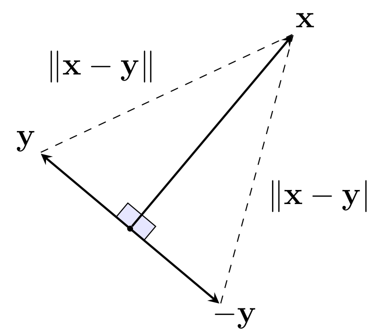
2. Podemos usar la distancia entre puntos para interpretar la desigualdad triangular. Sean $\mathbf{p}, \mathbf{q}, \mathbf{r} \in \mathbb{R}^n$ tres puntos que forman un triángulo en $\mathbb{R}^n$. Entonces: $\|\mathbf{q} - \mathbf{p}\| = \|(\mathbf{q} - \mathbf{r}) + (\mathbf{r} - \mathbf{p})\| \leq \|\mathbf{q} - \mathbf{r}\| + \|\mathbf{r} - \mathbf{p}\|$, es decir $\mathrm{d}(\mathbf{p}, \mathbf{q}) \leq \mathrm{d}(\mathbf{r}, \mathbf{q}) + \mathrm{d}(\mathbf{p}, \mathbf{r})$. En el triángulo $\mathbf{pqr}$ esto significa que la longitud del lado $\overline{\mathbf{pq}}$ es inferior o igual a la suma de los otros dos lados ($\overline{\mathbf{rq}} + \overline{\mathbf{pr}}$), lo que es un viejo teorema de geometría plana.



El producto punto está íntimamente ligado a la noción de *ángulo* entre vectores. Recurramos nuevamente a nuestra intuición geométrica para ver algunos hechos que indican que esta afirmación no es gratuita.

Situación 1:

Consideremos el siguiente diagrama:



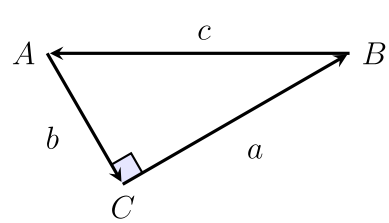
Los vectores $\mathbf{x}$ e $\mathbf{y}$ son no nulos en $\mathbb{R}^n$, y los hemos dibujado en el plano que pasa por el origen y que los contiene. Supongamos que $\mathbf{x}$ es *perpendicular* a $\mathbf{y}$ (vamos a suponer que esto tiene sentido en $\mathbb{R}^n$, y que en el plano de $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{0}$ se cumplen las propiedades usuales que conocemos de la geometría plana del colegio). Esperamos entonces que el triángulo que tiene por vértices $\mathbf{x}$, $\mathbf{y}$ y $-\mathbf{y}$ sea isósceles en $\mathbf{x}$ (la altura es la flecha $\mathbf{x}$), y así

$$
\begin{aligned}
\|\mathbf{x}-\mathbf{y}\| = \|\mathbf{x}+\mathbf{y}\| \quad &\Longleftrightarrow \quad \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle = \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle \\
&\Longleftrightarrow \quad \|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2 \\
&\Longleftrightarrow \quad 4\langle \mathbf{x}, \mathbf{y}\rangle = 0 \\
&\Longleftrightarrow \quad \langle \mathbf{x}, \mathbf{y}\rangle = 0.
\end{aligned}
$$

Aprovechemos esto para definir perpendicularidad en $\mathbb{R}^n$:

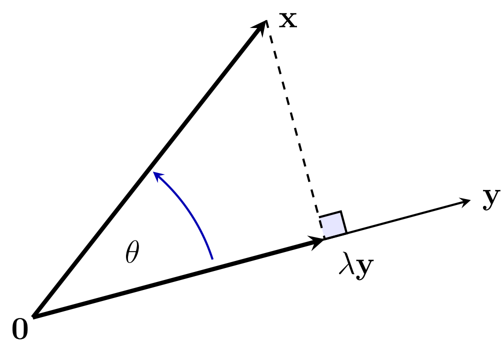
> Definición (Perpendicularidad) Dos vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ se dicen *perpendiculares* u *ortogonales* ssi $\langle \mathbf{x}, \mathbf{y}\rangle = 0$. Lo anotaremos como $\mathbf{x} \perp \mathbf{y}$.

Ejercicio 2.4: Probar el *Teorema de Pitágoras* en $\mathbb{R}^n$: Sean $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^n$, con $(\mathbf{B}-\mathbf{C}) \perp (\mathbf{C}-\mathbf{A})$. Si $a = \mathrm{d}(\mathbf{C}, \mathbf{B}) = \|\mathbf{B}-\mathbf{C}\|$, $b = \mathrm{d}(\mathbf{A}, \mathbf{C}) = \|\mathbf{C}-\mathbf{A}\|$ y $c = \mathrm{d}(\mathbf{B}, \mathbf{A}) = \|\mathbf{A}-\mathbf{B}\|$ son los tres lados del triángulo $\mathbf{ABC}$ (triángulo rectángulo en $\mathbf{C}$, con a, b los catetos y c la hipotenusa), entonces $a^2 + b^2 = c^2$.



Situación 2:

Queremos usar nuestra intuición geométrica para descubrir una relación que nos permita definir el ángulo entre dos vectores de $\mathbb{R}^n$. Consideremos el siguiente diagrama:



Todo transcurre en el plano que pasa por $\mathbf{0}$ y tiene a $\mathbf{x}$ e $\mathbf{y}$ como vectores directores. El punto $\lambda\mathbf{y}$ es la *proyección ortogonal* del punto $\mathbf{x}$ sobre la recta que pasa por $\mathbf{0}$ y tiene a $\mathbf{y}$ como vector director, θ será el ángulo de la flecha $\mathbf{y}$ a la $\mathbf{x}$ (si todo esto tiene sentido en $\mathbb{R}^n$). La cantidad $\|\lambda\mathbf{y}-\mathbf{x}\|$ es la distancia de $\mathbf{x}$ a su proyección. Las relaciones trigonométricas que esperaríamos que se cumplan dicen: $\cos\theta = \frac{\|\lambda\mathbf{y}\|}{\|\mathbf{x}\|} = \frac{|\lambda|\|\mathbf{y}\|}{\|\mathbf{x}\|}$.

Nos gustaría eliminar λ de esta fórmula. Para ello notemos que como $(\lambda\mathbf{y}-\mathbf{x}) \perp \mathbf{y}$, entonces $\langle\lambda\mathbf{y}-\mathbf{x},\mathbf{y}\rangle = 0$, de donde $\lambda\|\mathbf{y}\|^2 - \langle\mathbf{x},\mathbf{y}\rangle = 0$, y así $|\lambda| = \lambda = \frac{\langle\mathbf{x},\mathbf{y}\rangle}{\|\mathbf{y}\|^2}$. (El dibujo sugiere que $\lambda\mathbf{y}$ apunta hacia el mismo lado que $\mathbf{y}$. Analizar lo que pasa si no. Notar que $\theta > \frac{\pi}{2}$ en dicho caso.) Luego $\cos\theta = \frac{\langle\mathbf{x},\mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}$. Usamos esto para definir el ángulo entre vectores de $\mathbb{R}^n$:

Definición (Ángulo entre vectores) Sean $\mathbf{x},\mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Definimos el *ángulo* entre $\mathbf{x}$ e $\mathbf{y}$ como aquel número $\theta \in [0,\pi]$ tal que

$$\cos\theta = \frac{\langle\mathbf{x},\mathbf{y}\rangle}{\|\mathbf{x}\|\,\|\mathbf{y}\|}.$$

Observación:

1. Notar que la desigualdad de Cauchy–Schwartz asegura que $-1 \leq \frac{\langle\mathbf{x},\mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1$, así la definición anterior tiene sentido.

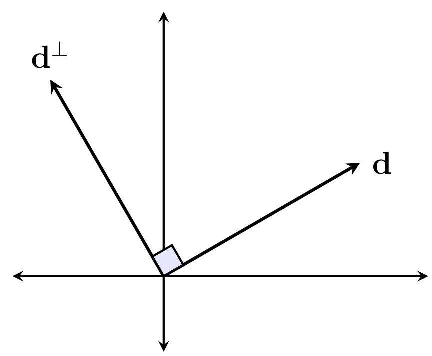
2. Esta definición entrega de manera trivial la famosa fórmula $\langle\mathbf{x},\mathbf{y}\rangle = \|\mathbf{x}\|\,\|y\|\cos\theta$.

2.6 Aplicación a Rectas en $\mathbb{R}^2$ y Planos en $\mathbb{R}^3$.

Rectas en $\mathbb{R}^2$.

Sea $\mathbf{d} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$. Definamos el nuevo vector $\mathbf{d}^\perp = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \in \mathbb{R}^2$. Es un ejercicio fácil verificar que las siguientes propiedades se satisfacen:

1. $\left\|\mathbf{d}^\perp\right\| = \|\mathbf{d}\|$.
2. $\mathbf{d}^\perp$ es ortogonal a $\mathbf{d}$.
3. $\left(\mathbf{d}^\perp\right)^\perp = -\mathbf{d}$.
4. Si $\mathbf{d} \neq \mathbf{0}$, entonces $\quad (\forall \mathbf{v} \in \mathbb{R}^2)\, [\mathbf{v} \perp \mathbf{d} \iff (\exists \lambda \in \mathbb{R})\, \mathbf{v} = \lambda \mathbf{d}^\perp]$.
5. Si $\mathbf{d} \neq \mathbf{0}$, entonces $\quad (\forall \mathbf{v} \in \mathbb{R}^2)\, [\mathbf{v} \perp \mathbf{d}^\perp \iff (\exists t \in \mathbb{R})\, \mathbf{v} = t\mathbf{d}]$.



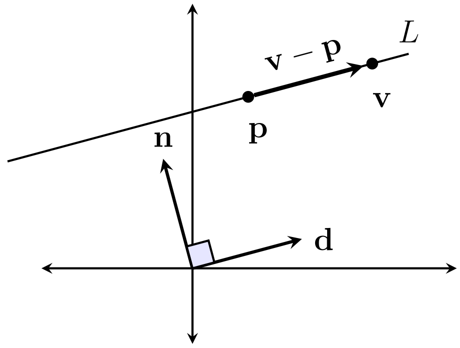
Sea ahora $\mathbf{p} \in \mathbb{R}^2$ y $L = L_{\mathbf{p},\mathbf{d}} = \{\mathbf{v} \in \mathbb{R}^2 \mid \mathbf{v} = \mathbf{p} + t\mathbf{d},\ t \in \mathbb{R}\}$. Tenemos que

$$\begin{aligned} \mathbf{v} \in L &\iff \mathbf{v} = \mathbf{p} + t\mathbf{d},\ t \in \mathbb{R} \\ &\iff (\exists t \in \mathbb{R})\, \mathbf{v} - \mathbf{p} = t\mathbf{d} \\ &\iff \mathbf{v} - \mathbf{p} \perp \mathbf{d}^\perp \\ &\iff \left\langle \mathbf{v} - \mathbf{p}, \mathbf{d}^\perp \right\rangle = 0. \end{aligned}$$

Si llamamos $\mathbf{n} = \mathbf{d}^\perp$ (o cualquier otro vector paralelo a éste), hemos concluido que $\mathbf{v} \in L \Leftrightarrow \langle \mathbf{v} - \mathbf{p}, \mathbf{n} \rangle = 0$, es decir, $L = \{\mathbf{v} \in \mathbb{R}^2 \mid \langle \mathbf{v} - \mathbf{p}, \mathbf{n} \rangle = 0\}$. La ecuación

$$\langle \mathbf{v} - \mathbf{p}, \mathbf{n} \rangle = 0$$

se llama *ecuación normal* de L. El vector $\mathbf{n}$ (es decir, $\mathbf{d}^\perp$ o cualquier otro paralelo a él) se dice *vector normal* a L.



De la ecuación normal de L sale fácilmente una ecuación Cartesiana. Sean $\mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}$, $\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Desarrollando:

$$\langle \mathbf{v} - \mathbf{p}, \mathbf{n} \rangle = 0 \iff \langle \mathbf{v}, \mathbf{n} \rangle - \langle \mathbf{p}, \mathbf{n} \rangle = 0 \iff n_1x_1 + n_2x_2 + (-n_1p_1 - n_2p_2) = 0,$$

es decir

$$ax_1 + bx_2 + c = 0, \qquad \text{con } a = n_1,\ b = n_2 \text{ y } c = -\langle \mathbf{p}, \mathbf{n} \rangle\,. \tag{2.1}$$

> **Ejercicio 2.5:** Concluya que si (2.1) es una ecuación Cartesiana de una recta de $\mathbb{R}^2$, entonces un vector director de esta recta es $\mathbf{d} = \left(\begin{smallmatrix} b \\ -a \end{smallmatrix}\right)$.

Planos en $\mathbb{R}^3$.

Sean $\mathbf{d}_1, \mathbf{d}_2$ vectores no nulos y no paralelos en $\mathbb{R}^3$. Más adelante se prueba que existe un vector $\mathbf{n} = \mathbf{d}_1 \times \mathbf{d}_2 \in \mathbb{R}^3$ con las propiedades:

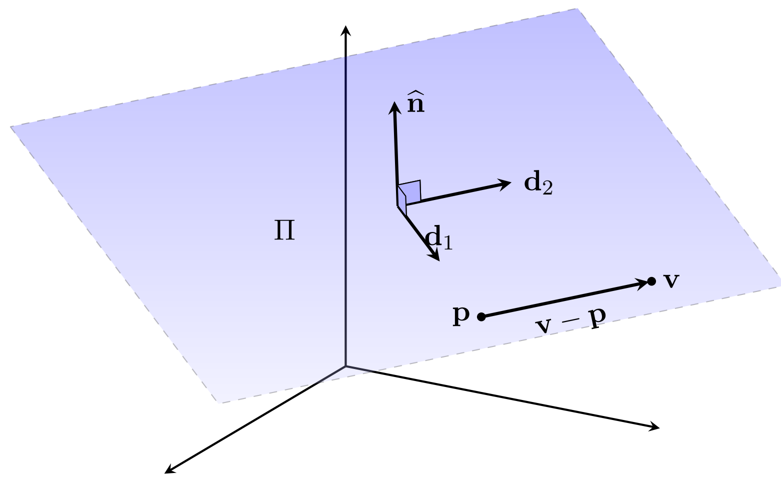
1. $\|\mathbf{n}\| = \|\mathbf{d}_1\|\,\|\mathbf{d}_2\|\,|\operatorname{sen}\theta| \neq 0$, con θ el ángulo entre $\mathbf{d}_1$ y $\mathbf{d}_2$.
2. $\mathbf{n} \perp \mathbf{d}_1 \wedge \mathbf{n} \perp \mathbf{d}_2$.
3. $(\forall \mathbf{v} \in \mathbb{R}^3)\ [\mathbf{v} \perp \mathbf{d}_1 \wedge \mathbf{v} \perp \mathbf{d}_2 \Leftrightarrow (\exists \lambda \in \mathbb{R})\, \mathbf{v} = \lambda \mathbf{n}]$.
4. $(\forall \mathbf{v} \in \mathbb{R}^3)\ [\mathbf{v} \perp \mathbf{n} \Leftrightarrow (\exists s, t \in \mathbb{R})\, \mathbf{v} = s\mathbf{d}_1 + t\mathbf{d}_2]$.

Con estas propiedades, sea $\Pi = \Pi_{\mathbf{p},\mathbf{d}_1,\mathbf{d}_2}$ el plano en $\mathbb{R}^3$ que pasa por un punto $\mathbf{p} \in \mathbb{R}^3$ y de vectores directores $\mathbf{d}_1$ y $\mathbf{d}_2$. Entonces el lector puede demostrar fácilmente que para cualquier $\mathbf{v} \in \mathbb{R}^3$, $\mathbf{v} \in \Pi \Leftrightarrow \langle \mathbf{v} - \mathbf{p}, \mathbf{n} \rangle = 0$. Nuevamente, $\mathbf{n}$ (o cualquier vector paralelo a $\mathbf{n}$) se dirá *vector normal* a Π, y la ecuación

$$\langle \mathbf{v} - \mathbf{p}, \mathbf{n} \rangle = 0$$

se llamará *ecuación normal* de Π.

De modo similar al caso de rectas en $\mathbb{R}^2$, la ecuación normal conduce directamente a la ecuación Cartesiana $Ax_1 + Bx_2 + Cx_3 + D = 0$ para Π. (Donde $\mathbf{n} = \left(\begin{smallmatrix} A \\ B \\ C \end{smallmatrix}\right)$ y $D = -\langle \mathbf{p}, \mathbf{n} \rangle$.)



$\widehat{\mathbf{n}}$
$\mathbf{d}_2$
Π
$\mathbf{d}_1$
$\mathbf{v}$
$\mathbf{p}$
$\mathbf{v}-\mathbf{p}$

Guía de Ejercicios

1. Pruebe las siguientes propiedades de la ponderación por escalar:

(a) $(\forall \lambda_1, \lambda_2 \in \mathbb{K})\ (\forall \mathbf{x} \in \mathbb{K}^n)\ \lambda_1(\lambda_2 \mathbf{x}) = (\lambda_1 \lambda_2)\mathbf{x}$.

(b) $(\forall \lambda \in \mathbb{K})\ (\forall \mathbf{x}, \mathbf{y} \in \mathbb{K}^n)\ \lambda(\mathbf{x} + \mathbf{y}) = \lambda \mathbf{x} + \lambda \mathbf{y}$.

(c) $(\forall \lambda \in \mathbb{K})\ (\forall \mathbf{x} \in \mathbb{K}^n)\ \lambda \mathbf{x} = \mathbf{0} \Leftrightarrow \lambda = 0 \vee \mathbf{x} = \mathbf{0}$.

2. Pruebe que:

(a) Dado $\mathbf{d} \in \mathbb{K}^n \setminus \{\mathbf{0}\}$ fijo, se tendrá $\mathbf{q} \in L_{\mathbf{p},\mathbf{d}} \Leftrightarrow L_{\mathbf{q},\mathbf{d}} = L_{\mathbf{p},\mathbf{d}}$. (Las posiciones de una recta L son sus puntos.)

(b) Dado $\mathbf{p} \in \mathbb{K}^n$ fijo, se tendrá $L_{\mathbf{p},\mathbf{d}} = L_{\mathbf{p},\overline{\mathbf{d}}} \Leftrightarrow (\exists \lambda \in \mathbb{K} \setminus \{0\})\, \overline{\mathbf{d}} = \lambda \mathbf{d}$. (Los vectores directores de una misma recta son los múltiplos no nulos de un vector director dado.)

(c) Si $A \subseteq \mathbb{K}^n$ y $\mathbf{p} \in \mathbb{K}^n$, definimos la *traslación* de A en el vector $\mathbf{p}$ por $\mathbf{p} + A = \{\mathbf{p} + \mathbf{x} \mid \mathbf{x} \in A\}$. Las rectas que pasan por $\mathbf{p} \in \mathbb{K}^n$ son exactamente las traslaciones en $\mathbf{p}$ de las rectas que pasan por el origen $\mathbf{0}$.

(d) Sean $\mathbf{p}, \mathbf{q} \in \mathbb{K}^n$, con $\mathbf{p} \neq \mathbf{q}$. La recta $L_{\mathbf{p},\mathbf{q}-\mathbf{p}}$ es la única recta que pasa por ambos puntos, $\mathbf{p}$ y $\mathbf{q}$.

3. Pruebe que:

(a) $\mathbf{q} \in \Pi_{\mathbf{p},\mathbf{d}_1,\mathbf{d}_2} \Leftrightarrow \Pi_{\mathbf{q},\mathbf{d}_1,\mathbf{d}_2} = \Pi_{\mathbf{p},\mathbf{d}_1,\mathbf{d}_2}$. (Cualquier punto de un plano sirve como su posición.)

(b) Dado $\mathbf{p} \in \mathbb{K}^n$, todo plano que pasa por $\mathbf{p}$ es la traslación en $\mathbf{p}$ de un plano que pasa por el origen, y si trasladamos en $\mathbf{p}$ un plano cualquiera que pasa por el origen, obtenemos un plano que pasa por $\mathbf{p}$.

(c) Sean $\mathbf{p}, \mathbf{q}, \mathbf{r}$ tres puntos *no colineales* en $\mathbb{K}^n$. Probar que si $\mathbf{d}_1 = \mathbf{q} - \mathbf{p}$ y $\mathbf{d}_2 = \mathbf{r} - \mathbf{p}$, entonces $\Pi_{\mathbf{p},\mathbf{d}_1,\mathbf{d}_2}$ es el único plano que pasa por $\mathbf{p}$, $\mathbf{q}$ y $\mathbf{r}$.

4. Pruebe las siguientes propiedades del producto punto:

(a) $(\forall \mathbf{x}, \mathbf{y}, \overline{\mathbf{x}}, \overline{\mathbf{y}} \in \mathbb{R}^n)\ \langle \mathbf{x} + \overline{\mathbf{x}}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \overline{\mathbf{x}}, \mathbf{y} \rangle \wedge \langle \mathbf{x}, \mathbf{y} + \overline{\mathbf{y}} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \overline{\mathbf{y}} \rangle$ (Bi–aditividad).

(b) $(\forall \lambda \in R)\, (\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n)\ \langle \lambda \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \lambda \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle$ (Bi–homogeneidad).

(c) $(\forall \mathbf{x} \in \mathbb{R}^n)\ \langle \mathbf{x}, \mathbf{x} \rangle \geq 0 \quad \wedge \quad \langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}$ (Positividad).

5. Pruebe las siguientes propiedades de la norma Euclidiana:

(a) $(\forall \mathbf{x} \in \mathbb{R}^n)\ \|\mathbf{x}\| \geq 0 \ \wedge\ \|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$.

(b) $(\forall \lambda \in \mathbb{R})\, (\forall \mathbf{x} \in \mathbb{R}^n)\ \|\lambda \mathbf{x}\| = |\lambda|\, \|\mathbf{x}\|$.

Guía de Problemas

P1. **(a)** Considere los puntos $P = \left(\begin{smallmatrix} 2 \\ 1 \\ 2 \end{smallmatrix}\right), Q = \left(\begin{smallmatrix} 0 \\ -1 \\ 1 \end{smallmatrix}\right)$ y $R = \left(\begin{smallmatrix} -1 \\ 1 \\ 5 \end{smallmatrix}\right)$.
Verifique que son puntos **no** colineales y encuentre la ecuación vectorial (o paramétrica) y Cartesiana del plano Π_1 que los contiene.

(b) Dadas las rectas L_1 y L_2 definidas por

$$L_1\colon \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ -3 \\ 0 \end{pmatrix},\ t \in \mathbb{R} \quad \text{y} \quad L_2\colon \begin{cases} 3x + y + 4 = 0 \\ z + 5 = 0 \end{cases}.$$

Verifique que L_1 y L_2 son paralelas y distintas y ecuentre la ecuación vectorial (o paramétrica) y Cartesiana del plano Π_2 que las contiene.

(c) Encuentre la ecuación vectorial de la recta L que se obtiene como la intersección de los plano Π_1 y Π_2 ($L = \Pi_1 \cap \Pi_2$).

(d) Encuentre el punto S de intersección de las rectas L_1 y L, y verifique que S satisface la ecuación Cartesiana del plano Π_1.

P2. **(a)** Verifique que las rectas

$$L_1\colon \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} + t \begin{pmatrix} -1 \\ 2 \\ -5 \end{pmatrix},\ t \in \mathbb{R}, \quad L_2\colon \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} + s \begin{pmatrix} 2 \\ 1 \\ -5 \end{pmatrix},\ s \in \mathbb{R}$$

se intersectan en un único punto. Encuéntrelo.

(b) Sea Π el plano con vectores directores $d_1 = \left(\begin{smallmatrix} 1 \\ 2 \\ 3 \end{smallmatrix}\right)$ y $d_2 = \left(\begin{smallmatrix} 4 \\ 0 \\ 4 \end{smallmatrix}\right)$, que pasa por el punto $P = \left(\begin{smallmatrix} 0 \\ -2 \\ 1 \end{smallmatrix}\right)$, y sea L la recta con vector director $d = \left(\begin{smallmatrix} 1 \\ b \\ 0 \end{smallmatrix}\right)$ que pasa por el punto $Q = \left(\begin{smallmatrix} a \\ 0 \\ 0 \end{smallmatrix}\right)$, donde $a, b \in \mathbb{R}$. Encuentre los valores de los parámetros a, b tales que

- L esté contenida en Π ($L \subset \Pi$),
- L y Π no tengan puntos en común ($L \cap \Pi = \emptyset$). y
- $L \cap \Pi$ contenga exactamente un solo punto.

P3. Sean Π_1 el plano de ecuación $x + y + 2z = 1$, Π_2 el plano de ecuación $-x + y = 2$ y L_1 la recta que pasa por el punto $P_1 = \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right)$ y cuya dirección es $D_1 = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$.

(a) Encuentre la ecuación de la recta L_2, que se obtiene como la intersección de los planos Π_1 y Π_2. Entregue un vector director de dicha recta.

(b) Encuentre el punto P_2 de intersección de la recta L_1 y Π_1.

(c) Calcular el punto P_3 de intersección de L_2 con el plano perpendicular a L_2 que pasa por el punto P_2.

(d) Encuentre la ecuación paramétrica o vectorial de la recta contenida en Π_2 que pasa por el punto P_3 y es perpendicular a L_2.

Geometría

2.7 Producto cruz (o producto vectorial) en $\mathbb{R}^3$

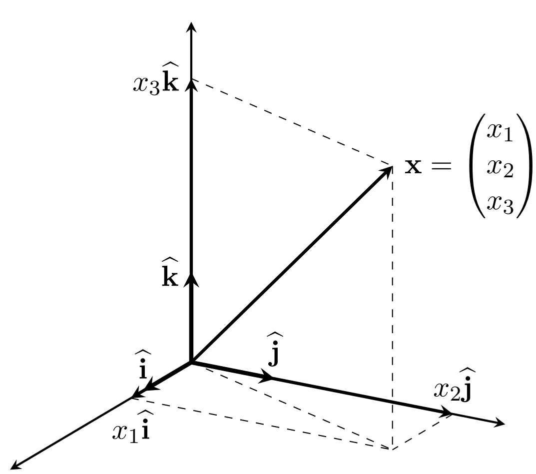
Utilizaremos la siguiente notación, ya tradicional en las aplicaciones de la geometría vectorial. Sean $\widehat{\mathbf{i}}, \widehat{\mathbf{j}}, \widehat{\mathbf{k}} \in \mathbb{R}^3$ los siguientes vectores:

$$\widehat{\mathbf{i}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \widehat{\mathbf{j}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \ \widehat{\mathbf{k}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Las siguientes afirmaciones son evidentes:

1. $\left\|\widehat{\mathbf{i}}\right\| = \left\|\widehat{\mathbf{j}}\right\| = \left\|\widehat{\mathbf{k}}\right\| = 1$ (Estos vectores se dicen *unitarios*.)
2. $\widehat{\mathbf{i}} \perp \widehat{\mathbf{j}} \wedge \widehat{\mathbf{j}} \perp \widehat{\mathbf{k}} \wedge \widehat{\mathbf{k}} \perp \widehat{\mathbf{i}}$ (Son ortogonales entre sí.)
3. $\left(\forall \mathbf{x} = \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right) \in \mathbb{R}^3\right)\ \mathbf{x} = x_1\widehat{\mathbf{i}} + x_2\widehat{\mathbf{j}} + x_3\widehat{\mathbf{k}}$, y esta escritura es única, es decir, si $x_1\widehat{\mathbf{i}} + x_2\widehat{\mathbf{j}} + x_3\widehat{\mathbf{k}} = y_1\widehat{\mathbf{i}} + y_2\widehat{\mathbf{j}} + y_3\widehat{\mathbf{k}}$, entonces $x_1 = y_1 \wedge x_2 = y_2 \wedge x_3 = y_3$.

Observación: Las propiedades 1. y 2. se abrevian diciendo que $\left\{\widehat{\mathbf{i}}, \widehat{\mathbf{j}}, \widehat{\mathbf{k}}\right\}$ es un *conjunto ortonormal* en $\mathbb{R}^3$.



Mas adelante en el curso definiremos la noción de *determinante*, pero ahora consideraremos la siguiente

Notación: Si $a, b, c, d \in \mathbb{R}$, anotamos $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \in \mathbb{R}$.

Con esto podemos dar la definición:

DEFINICIÓN (PRODUCTO CRUZ) Si $\mathbf{x} = \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right)$, $\mathbf{y} = \left(\begin{smallmatrix} y_1 \\ y_2 \\ y_3 \end{smallmatrix}\right) \in \mathbb{R}^3$, definimos el *producto cruz* o *producto vectorial* de $\mathbf{x}$ e $\mathbf{y}$ como el siguiente vector de $\mathbb{R}^3$:

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} \widehat{\mathbf{i}} & \widehat{\mathbf{j}} & \widehat{\mathbf{k}} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} \widehat{\mathbf{i}} - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} \widehat{\mathbf{j}} + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \widehat{\mathbf{k}}.$$

Es decir, resulta el vector de $\mathbb{R}^3$:

$$\mathbf{x} \times \mathbf{y} = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix} \in \mathbb{R}^3.$$

(Obviamente, esta última fórmula no es muy fácil de recordar, y es preferible usar la notación de la definición dada primero, cuya operatoria está ahí mismo definida, para memorizar y calcular $\mathbf{x} \times \mathbf{y}$.)

Hemos así definido una ley de composición interna $\times$ en $\mathbb{R}^3$:

$$\begin{array}{cccc} \times & : & \mathbb{R}^3 \times \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & & (\mathbf{x}, \mathbf{y}) & \longmapsto & \mathbf{x} \times \mathbf{y} \end{array}.$$

Proposición 2.4. *Se tienen las siguientes propiedades*

1. $(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^3)\ \mathbf{x} \times \mathbf{y} \perp \mathbf{x} \wedge \mathbf{x} \times \mathbf{y} \perp \mathbf{y}$.
2. $(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^3)\ \mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$ *($\times$ es anti–conmutativo).*
3. $(\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3)\ \mathbf{x} \times (\mathbf{y} + \mathbf{z}) = \mathbf{x} \times \mathbf{y} + \mathbf{x} \times \mathbf{z} \wedge (\mathbf{x} + \mathbf{y}) \times \mathbf{z} = \mathbf{x} \times \mathbf{z} + \mathbf{y} \times \mathbf{z}$ *($\times$ distribuye sobre $+$).*
4. $(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^3)\ (\forall \lambda \in \mathbb{R})\ (\lambda \mathbf{x}) \times \mathbf{y} = \mathbf{x} \times (\lambda \mathbf{y}) = \lambda(\mathbf{x} \times \mathbf{y})$.
5. $(\forall \mathbf{x} \in \mathbb{R}^3)\ \mathbf{x} \times \mathbf{x} = \mathbf{0}$.

DEMOSTRACIÓN. La demostración de estas propiedades básicas quedan propuestas como ejercicio.

Notemos que con 2, 3, 4 y 5, más el simple cálculo de

$$\widehat{\mathbf{i}} \times \widehat{\mathbf{j}} = \widehat{\mathbf{k}},\ \widehat{\mathbf{j}} \times \widehat{\mathbf{k}} = \widehat{\mathbf{i}},\ \widehat{\mathbf{k}} \times \widehat{\mathbf{i}} = \widehat{\mathbf{j}},$$

se puede recuperar la fórmula con que definimos originalmente el producto cruz:

$$\begin{aligned}
\mathbf{x} \times \mathbf{y} &= (x_1\widehat{\mathbf{i}} + x_2\widehat{\mathbf{j}} + x_3\widehat{\mathbf{k}}) \times (y_1\widehat{\mathbf{i}} + y_2\widehat{\mathbf{i}} + y_3\widehat{\mathbf{k}}) \\
&= x_1y_1\widehat{\mathbf{i}} \times \widehat{\mathbf{i}} + x_1y_2\widehat{\mathbf{i}} \times \widehat{\mathbf{j}} + x_1y_3\widehat{\mathbf{i}} \times \widehat{\mathbf{k}} + x_2y_1\widehat{\mathbf{j}} \times \widehat{\mathbf{i}} + x_2y_2\widehat{\mathbf{j}} \times \widehat{\mathbf{j}} + x_2y_3\widehat{\mathbf{j}} \times \widehat{\mathbf{k}} \\
&\quad + x_3y_1\widehat{\mathbf{k}} \times \widehat{\mathbf{i}} + x_3y_2\widehat{\mathbf{k}} \times \widehat{\mathbf{j}} + x_3y_3\widehat{\mathbf{k}} \times \widehat{\mathbf{k}} \\
&= \mathbf{0} + x_1y_2\widehat{\mathbf{k}} + x_1y_3(-\widehat{\mathbf{j}}) + x_2y_1(-\widehat{\mathbf{k}}) + \mathbf{0} + x_2y_3\widehat{\mathbf{i}} + x_3y_1\widehat{\mathbf{j}} + x_3y_2(-\widehat{\mathbf{i}}) + \mathbf{0} \\
&= (x_2y_3 - x_3y_2)\widehat{\mathbf{i}} + (x_3y_1 - x_1y_3)\widehat{\mathbf{j}} + (x_1y_2 - x_2y_1)\widehat{\mathbf{k}}.
\end{aligned}$$

La siguiente es una fórmula para el producto cruz de 3 vectores. Su demostración es un simple cálculo que omitimos.

$$(\mathbf{x} \times \mathbf{y}) \times \mathbf{z} = \langle \mathbf{x}, \mathbf{z} \rangle\, \mathbf{y} - \langle \mathbf{y}, \mathbf{z} \rangle\, \mathbf{x} \tag{2.2}$$

Notar que el producto cruz NO es conmutativo (pero tiene la propiedad parecida 2) y que desgraciadamente tampoco es asociativo. Para ilustrar cuán lejos está de serlo, está la llamada *Identidad de Jacobi*, que se puede probar fácilmente a partir de la propiedad anterior:

$$\left(\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3\right)\ \mathbf{x} \times (\mathbf{y} \times \mathbf{z}) + \mathbf{y} \times (\mathbf{z} \times \mathbf{x}) + \mathbf{z} \times (\mathbf{x} \times \mathbf{y}) = \mathbf{0} \tag{2.3}$$

Notemos que usando 2.3 y las propiedades anteriores resulta:

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) - (\mathbf{x} \times \mathbf{y}) \times \mathbf{z} = \mathbf{y} \times (\mathbf{x} \times \mathbf{z}),$$

lo que indica que $\mathbf{x} \times (\mathbf{y} \times \mathbf{z})$ no tiene por qué ser igual a $(\mathbf{x} \times \mathbf{y}) \times \mathbf{z}$.

Veamos mas propiedades de "$\times$":

Proposición 2.5.

$$\left(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}\right)\ \|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\|\, \|\mathbf{y}\|\, |\mathrm{sen}\, \theta|\,,$$

donde θ es el ángulo entre $\mathbf{x}$ e $\mathbf{y}$.

Demostración. Esto saldrá de la siguiente identidad, la que pedimos al lector que tenga la paciencia de verificar (es un cálculo simple pero tedioso):

$$\left(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^3\right)\ \langle \mathbf{x}, \mathbf{y} \rangle^2 + \|\mathbf{x} \times \mathbf{y}\|^2 = \|\mathbf{x}\|^2\, \|\mathbf{y}\|^2\,.$$

Una vez establecida esta identidad, y recordando que $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\, \|\mathbf{y}\| \cos(\theta)$, resulta:

$$\begin{aligned}
\|\mathbf{x} \times \mathbf{y}\|^2 &= \|\mathbf{x}\|^2\, \|\mathbf{y}\|^2 - \langle \mathbf{x}, \mathbf{y} \rangle^2 \\
&= \|\mathbf{x}\|^2\, \|\mathbf{y}\|^2 - (\|\mathbf{x}\|\, \|\mathbf{y}\| \cos(\theta))^2 \\
&= \|\mathbf{x}\|^2\, \|\mathbf{y}\|^2\, (1 - \cos^2(\theta)) \\
&= \|\mathbf{x}\|^2\, \|\mathbf{y}\|^2\, \mathrm{sen}^2(\theta).
\end{aligned}$$

$\square$

La siguiente propiedad muestra que el producto cruz representa a la dirección ortogonal a los vectores participantes.

Proposición 2.6. *Sean* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ *vectores no paralelos, y sea* $\mathbf{z} \in \mathbb{R}^3$. *Entonces*

1. $\mathbf{z} \perp \mathbf{x} \ \wedge \ \mathbf{z} \perp \mathbf{y} \implies (\exists \lambda \in \mathbb{R})\ \mathbf{z} = \lambda \mathbf{x} \times \mathbf{y}$.
2. $\mathbf{z} \perp \mathbf{x} \times \mathbf{y} \implies (\exists s, t \in \mathbb{R})\ \mathbf{z} = s\mathbf{x} + t\mathbf{y}$.

DEMOSTRACIÓN. La demostración es un ejercicio, explicado en más detalle a continuación. $\square$

Ejercicio 2.6: El objetivo de este ejercicio es probar la Proposición 2.

1. Sean $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ tres vectores no coplanares con el origen, es decir, que no existe un plano que pasa por el origen y que contiene a los tres vectores $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{z}$. Pruebe que esto equivale a:

$$\begin{array}{rcl} (\forall s, t \in \mathbb{R}) \quad \mathbf{x} & \neq & \mathbf{0} + s\mathbf{y} + t\mathbf{z} \\ \wedge \quad \mathbf{y} & \neq & \mathbf{0} + s\mathbf{x} + t\mathbf{z} \\ \wedge \quad \mathbf{z} & \neq & \mathbf{0} + s\mathbf{x} + t\mathbf{y} \end{array}$$

2. Pruebe que la condición anterior es equivalente a

$$(\forall \alpha, \beta, \gamma \in \mathbb{R})\ \alpha\mathbf{x} + \beta\mathbf{y} + \gamma\mathbf{z} = \mathbf{0} \implies \alpha = \beta = \gamma = 0. \tag{2.4}$$

3. Sea $A \in M_{33}(\mathbb{R})$ la matriz descrita por columnas como $A = \left[\begin{array}{c|c|c} \mathbf{x} & \mathbf{y} & \mathbf{z} \end{array}\right]$. Pruebe que la propiedad 2.9 equivale a que el sistema homogéneo $A\left[\begin{smallmatrix}\alpha\\ \beta\\ \gamma\end{smallmatrix}\right] = \mathbf{0}$ tiene como solución única $\left[\begin{smallmatrix}\alpha\\ \beta\\ \gamma\end{smallmatrix}\right] = \mathbf{0}$, es decir, A es invertible.

4. Concluya que tres vectores $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ no son coplanares con el origen ssi la matriz $A = \left[\begin{array}{c|c|c} \mathbf{x} & \mathbf{y} & \mathbf{z} \end{array}\right]$ es invertible.

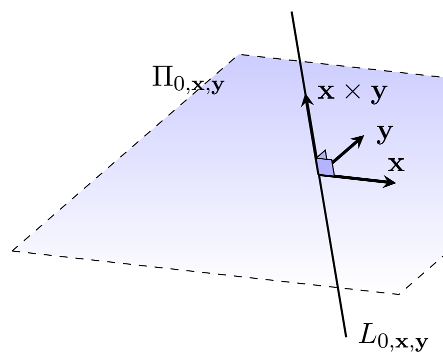
5. Considere ahora $\mathbf{x}$ e $\mathbf{y}$ no nulos y no paralelos. Pruebe que los vectores $\mathbf{x}, \mathbf{y}$ y $\mathbf{x} \times \mathbf{y} \neq 0$ son no coplanares con el origen.
Indicación: Verifique la condición 2.9.

6. Use la conclusión de la parte 2c para probar que

$$\left(\forall \mathbf{z} \in \mathbb{R}^3\right)\ (\exists s, t, \lambda \in \mathbb{R})\ \mathbf{z} = s\mathbf{x} + t\mathbf{y} + \lambda\mathbf{x} \times \mathbf{y}. \tag{2.5}$$

7. Pruebe, usando la propiedad 2.10, la Proposición 2.

Observación: La interpretación geométrica del resultado anterior es: Si $\mathbf{x}, \mathbf{y}$ son vectores no nulos y no paralelos en $\mathbb{R}^3$, $\mathbf{x} \times \mathbf{y}$ define la *línea recta* (que pasa por $\mathbf{0}$) de vectores perpendiculares a $\mathbf{x}$ y a $\mathbf{y}$. Por otra parte los vectores perpendiculares a $\mathbf{x} \times \mathbf{y}$ son exactamente aquellos del *plano* que pasa por el origen y tiene a $\mathbf{x}$ e $\mathbf{y}$ como vectores directores.

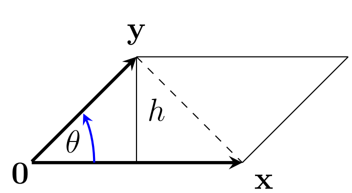


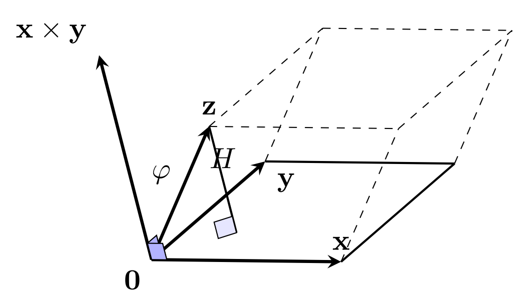
Un par de propiedades geométricas más de "$\times$":

Proposición 2.7. *1. Sean $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ vectores no paralelos. Entonces el área del paralelógramo que definen es $\|\mathbf{x} \times \mathbf{y}\|$.*

2. Sean $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ vectores no coplanares con $\mathbf{0}$. Entonces el volumen del paralelepípedo que definen es $|\langle \mathbf{x} \times \mathbf{y}, \mathbf{z} \rangle|$.

Demostración. (Deberemos confiar en los conceptos intuitivos de *área* y *volumen* que traemos de la geometría escolar).





1. El área del paralelógramo es dos veces el área del triángulo cuyos vértices son $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{0}$. Así, $\mathbf{Area} = 2 \cdot \frac{1}{2} \cdot \text{base} \cdot \text{altura} = 2 \cdot \frac{1}{2} \cdot \|\mathbf{x}\|\, h$, con $h = \|\mathbf{y}\| \operatorname{sen} \theta$. De aquí que $\mathbf{Area} = 2 \cdot \frac{1}{2} \cdot \|\mathbf{x}\|\, \|\mathbf{y}\| \operatorname{sen} \theta = \|\mathbf{x}\|\, \|\mathbf{y}\| \operatorname{sen} \theta = \|\mathbf{x} \times \mathbf{y}\|$.

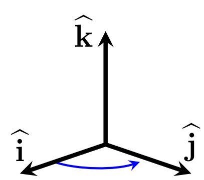
2. De la geometría de colegio, el volumen del paralelepípedo se calcula como el producto del área de su base y su altura. De la parte (a), el área de la base es $\|\mathbf{x} \times \mathbf{y}\|$. Por otra parte la altura H es la proyección de $\mathbf{z}$ sobre la dirección

perpendicular al plano de $\mathbf{x}$ e $\mathbf{y}$, con lo que $H = \|\mathbf{z}\| \cos(\varphi)$, donde φ es el ángulo entre $\mathbf{x} \times \mathbf{y}$ y $\mathbf{z}$. Se obtiene finalmente entonces que $\mathbf{Vol} = \|\mathbf{x} \times \mathbf{y}\| \, \|\mathbf{z}\| \cos(\varphi) = |\langle \mathbf{x} \times \mathbf{y}, \mathbf{z} \rangle|$.

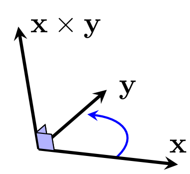
$\square$

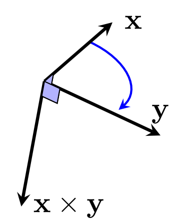
Una última propiedad del producto cruz que es necesario indicar es la llamada *regla de la mano derecha.*

De $\mathbf{x} \times \mathbf{y}$ conocemos que apunta en la línea perpendicular al plano de $\mathbf{x}$ e $\mathbf{y}$, y que tiene por norma $\|\mathbf{x}\| \, \|\mathbf{y}\| \operatorname{sen} \theta$, con θ el ángulo entre $\mathbf{x}$ e $\mathbf{y}$. Estos datos nos dejan aún dos opciones posibles para $\mathbf{x} \times \mathbf{y}$: hacia un lado del plano de $\mathbf{x}$ e $\mathbf{y}$ o hacia el otro. La verdad es que hacia cuál de estos dos lados se dibuja depende de cómo dibujamos los vectores ortonormales $\widehat{\mathbf{i}}, \widehat{\mathbf{j}}, \widehat{\mathbf{k}}$ (es decir, de cómo dibujamos los tres ejes de coordenadas de $\mathbb{R}^3$). Generalmente *se conviene* en dibujarlos de modo que si en el plano de $\widehat{\mathbf{i}}, \widehat{\mathbf{j}}$ (el plano $x_1 - x_2$) hacemos los dedos de la mano derecha girar rotando de $\widehat{\mathbf{i}}$ hacia $\widehat{\mathbf{j}}$, $\widehat{\mathbf{k}}$ apunta en la misma dirección del pulgar.

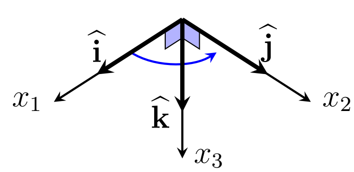


Con esta convención resulta también que, si $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ son no paralelos y hacemos girar los dedos de la mano derecha en el plano definido por $\mathbf{x}$ e $\mathbf{y}$, desde $\mathbf{x}$ hacia $\mathbf{y}$, el pulgar apunta en la dirección de $\mathbf{x} \times \mathbf{y}$.





Hay que indicar, sin embargo, que si alguien prefiriese dibujar los ejes como:

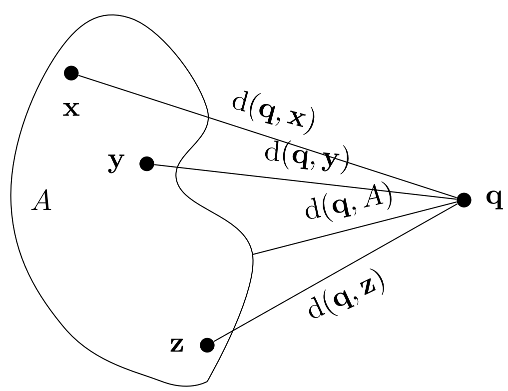


entonces los productos cruz cumplirían una regla de la mano izquierda. Nosotros seguiremos la convención universalmente aceptada, dibujando los ejes del primer modo que mencionamos, y nuestros productos cruz se dibujarán siempre según la *regla de la mano derecha.*

2.8 Distancia entre un punto y un conjunto. Proyecciones.

Definición (Distancia punto–conjunto) Sea $A \subseteq \mathbb{R}^n$ un subconjunto *no vacío* de $\mathbb{R}^n$, y sea $\mathbf{q} \in \mathbb{R}^n$ un punto. Definimos la distancia de $\mathbf{q}$ al conjunto A como:

$$\mathrm{d}(\mathbf{q}, A) = \text{ínf}\,\{\mathrm{d}(\mathbf{q}, \mathbf{x}) \mid \mathbf{x} \in A\}\,.$$



Ejercicio 2.7: Verifique que $\mathrm{d}(\mathbf{q}, A)$ queda bien definida.

Cabe notar que no tiene por qué existir un punto $\mathbf{r} \in A$ tal que $\mathrm{d}(\mathbf{q}, A) = \mathrm{d}(\mathbf{q}, \mathbf{r})$, y si tal $\mathbf{r}$ llega a existir, no tiene por qué ser único (¡Construir ejemplos de estas afirmaciones!).

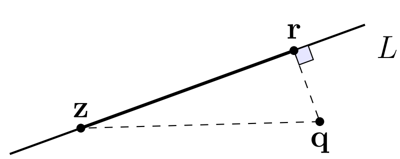
Estudiemos a modo de ejemplo el caso en que $A = L$ es una recta $L_{\mathbf{p},\mathbf{x}}$ en $\mathbb{R}^n$ (recta que pasa por $\mathbf{p}$ y tiene vector director $\mathbf{x}$).

Supongamos que existe un punto $\mathbf{r} \in L$ tal que $\mathbf{q} - \mathbf{r}$ es perpendicular a L. Afirmamos (como nuestra intuición geométrica lo indica) que $\mathrm{d}(\mathbf{q}, L) = \mathrm{d}(\mathbf{q}, \mathbf{r})$. Probemos algo mejor:

$$(\forall \mathbf{z} \in L)\;\; \mathbf{z} \neq \mathbf{r} \Rightarrow \|\mathbf{q} - \mathbf{z}\| > \|\mathbf{q} - \mathbf{r}\|\,,$$

de donde la distancia de $\mathbf{q}$ a L *se alcanza* en *un único* punto $\mathbf{r}$.

En efecto, si $\mathbf{z} \in L$, como $\mathbf{q} - \mathbf{r}$ es normal a L, el triángulo cuyos vértices son $\mathbf{z}$, $\mathbf{r}$ y $\mathbf{q}$ es rectángulo en $\mathbf{r}$, y por Pitágoras (recuerde que usted demostró en el Ejercicio 2.4 que es válido en $\mathbb{R}^n$) $\|\mathbf{q} - \mathbf{z}\|^2 = \|\mathbf{q} - \mathbf{r}\|^2 + \|\mathbf{r} - \mathbf{z}\|^2$. Como $\mathbf{z} \neq \mathbf{r}$, entonces $\|\mathbf{r} - \mathbf{z}\|^2 > 0$, y eso prueba la desigualdad.



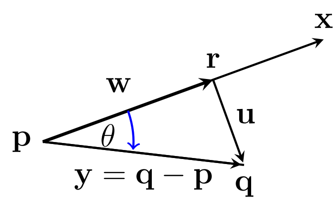
Veremos ahora que tal $\mathbf{r} \in L$ efectivamente existe.

Sea $\widehat{\mathbf{x}} = \frac{1}{\|\mathbf{x}\|}\mathbf{x}$ (*vector unitario* en la dirección de $\mathbf{x}$: claramente $\|\widehat{\mathbf{x}}\| = 1$). Sea $\mathbf{y} = \mathbf{q} - \mathbf{p}$ el vector que conecta $\mathbf{p}$ con $\mathbf{q}$, y sea $\mathbf{w} = \langle \mathbf{y}, \widehat{\mathbf{x}} \rangle \widehat{\mathbf{x}}$. Notar que $\mathbf{w}$ apunta en la misma línea que $\mathbf{x}$, y que

$$\|\mathbf{w}\| = |\langle \mathbf{y}, \widehat{\mathbf{x}} \rangle| \, \|\widehat{\mathbf{x}}\| = \|\mathbf{y}\| \, \|\widehat{\mathbf{x}}\|^2 \, |\cos(\theta)| = \|\mathbf{y}\| \cos(\theta),$$

donde θ es el ángulo entre $\mathbf{x}$ e $\mathbf{y}$ (que lo supondremos en $\left[0, \frac{\pi}{2}\right]$ para no preocuparnos del signo de $\cos(\theta)$. Si este no es el caso, podemos usar $-\mathbf{x}$ en lugar de $\mathbf{x}$ como vector director de L.)

Este vector $\mathbf{w}$ es lo que llamaremos *proyección* del vector $\mathbf{y}$ sobre la dirección definida por el vector $\mathbf{x}$. Notar que si en vez de $\mathbf{x}$ usamos otro vector director $\lambda\mathbf{x}$ de L, resulta exactamente el mismo vector $\mathbf{w}$.



Sea $\mathbf{u} = \mathbf{y} - \mathbf{w} = \mathbf{y} - \langle \mathbf{y}, \widehat{\mathbf{x}} \rangle \widehat{\mathbf{x}}$. Nuestra intuición nos haría esperar que $\mathbf{u} \perp \mathbf{x}$. Verifiquemos que eso es correcto:

$$\langle \mathbf{u}, \mathbf{x} \rangle = \langle \mathbf{y} - \langle \mathbf{y}, \widehat{\mathbf{x}} \rangle \widehat{\mathbf{x}}, \mathbf{x} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle - \langle \mathbf{y}, \widehat{\mathbf{x}} \rangle \langle \widehat{\mathbf{x}}, \mathbf{x} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle - \left\langle \mathbf{y}, \frac{1}{\|\mathbf{x}\|}\mathbf{x} \right\rangle \|\mathbf{x}\| = 0,$$

de donde $\mathbf{u} \perp \mathbf{x}$.

Basta tomar ahora $\mathbf{r} = \mathbf{p} + \mathbf{w} = \mathbf{p} + \langle \mathbf{y}, \widehat{\mathbf{x}} \rangle \widehat{\mathbf{x}}$, es decir

$$\mathbf{r} = \mathbf{p} + \langle \mathbf{q} - \mathbf{p}, \widehat{\mathbf{x}} \rangle \widehat{\mathbf{x}},$$

que tiene la propiedad deseada pues $\mathbf{q} - \mathbf{r} = \mathbf{q} - \mathbf{p} - \mathbf{w} = \mathbf{y} - \mathbf{w} = \mathbf{u} \perp \mathbf{x}$, luego $\mathbf{q} - \mathbf{r} \perp L$.

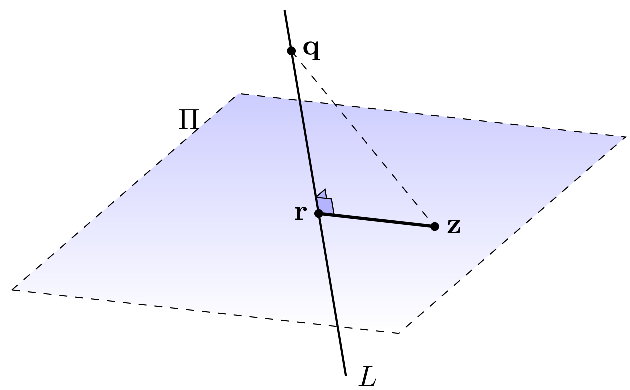
Definición (Proyección ortogonal sobre una recta) Llamamos *proyección ortogonal* del punto $\mathbf{q} \in \mathbb{R}^n$ sobre la recta $L \subseteq \mathbb{R}^n$, al punto $\mathbf{r} \in L$ tal que $\mathbf{q} - \mathbf{r} \perp L$.

Acabamos de probar que dicha proyección $\mathbf{r}$ existe y es el único punto de L que minimiza la distancia a $\mathbf{q}$, y dimos además una fórmula para calcular esta proyección en términos de $\mathbf{q}$, la posición $\mathbf{p}$ de L y su vector director $\mathbf{x}$.

Como segundo ejemplo, vamos a estudiar ahora en detalle el caso en que $n = 3$ y A es un plano $\Pi \subseteq \mathbb{R}^3$.

Ejercicio 2.8: Desarrollar de manera similar, en paralelo, el caso en que $n = 2$ y A es una recta en $\mathbb{R}^2$. Notar que ese problema es un caso particular del que se resolvió antes, pero la solución va por otro camino, usando ecuaciones normales en vez de vectores directores.

Sea L la recta que pasa por $\mathbf{q}$ y que es perpendicular a Π (es decir, que tiene por vector director un vector normal a Π). Veremos en un momento que la intersección entre Π y L consiste en un solo punto $\mathbf{r}$. Nuestra intuición geométrica indica que $\mathrm{d}(\mathbf{q}, \Pi) = \mathrm{d}(\mathbf{q}, \mathbf{r})$, y la demostración de este hecho sigue exactamente las mismas lineas de lo que hicimos al estudiar la proyección de un punto sobre una recta en $\mathbb{R}^n$. (Se prueba la desigualdad $(\forall \mathbf{z} \in \Pi)\ \mathbf{z} \neq \mathbf{r} \Rightarrow \|\mathbf{q} - \mathbf{z}\| > \|\mathbf{q} - \mathbf{r}\|$, etc.)



Definición (Proyección ortogonal sobre un plano) Llamamos *proyección ortogonal* del punto $\mathbf{q} \in \mathbb{R}^3$ sobre el plano $\Pi \subseteq \mathbb{R}^3$, al punto $\mathbf{r} \in \Pi$ tal que $\mathbf{q} - \mathbf{r} \perp \Pi$.

Sabemos que dicho $\mathbf{r}$ es el único punto de Π que minimiza la distancia a $\mathbf{q}$. Lo que no hemos probado aún es que tal punto $\mathbf{r}$ realmente existe, y a eso nos dedicaremos ahora.

Sea $\mathbf{n}$ un vector normal a Π (por ejemplo, si $\Pi = \Pi_{\mathbf{p},\mathbf{d}_1,\mathbf{d}_2}$, $\mathbf{n}$ se puede tomar como $\mathbf{d}_1 \times \mathbf{d}_2$, o si tenemos una ecuación Cartesiana $Ax_1 + Bx_2 + Cx_3 + D = 0$ para Π, $\mathbf{n}$ se puede tomar como $\left(\begin{smallmatrix} A \\ B \\ C \end{smallmatrix}\right)$). La recta L que pasa por $\mathbf{q}$ y es perpendicular a Π tiene entonces por ecuación vectorial:

$$\mathbf{v} = \mathbf{q} + t\mathbf{n}, \qquad t \in \mathbb{R}. \tag{2.6}$$

Sea $\mathbf{p}$ un punto (fijo) cualquiera de Π. Podemos entonces escribir la ecuación normal de Π:

$$\langle \mathbf{v} - \mathbf{p}, \mathbf{n} \rangle = 0. \tag{2.7}$$

Buscamos $L \cap \Pi$, es decir, los $\mathbf{v} \in \mathbb{R}^3$ que satisfacen simultaneamente (2.6) y (2.7). Pero esto es fácil de resolver. En efecto, (2.7) equivale a $\langle \mathbf{v}, \mathbf{n} \rangle = \langle \mathbf{p}, \mathbf{n} \rangle$, y usando (2.6), obtenemos

$$\langle \mathbf{q} + t\mathbf{n}, \mathbf{n} \rangle = \langle \mathbf{p}, \mathbf{n} \rangle \Leftrightarrow \langle \mathbf{q}, \mathbf{n} \rangle + t \|\mathbf{n}\|^2 = \langle \mathbf{p}, \mathbf{n} \rangle \Leftrightarrow t = \frac{\langle \mathbf{p} - \mathbf{q}, \mathbf{n} \rangle}{\|\mathbf{n}\|^2}.$$

Una vez que tenemos la incógnita real t despejada, en (2.6) resulta

$$\mathbf{r} = \mathbf{v} = \mathbf{q} + \frac{\langle \mathbf{p} - \mathbf{q}, \mathbf{n} \rangle}{\|\mathbf{n}\|^2}\mathbf{n} \tag{2.8}$$

como única solución de (2.6) y (2.7).

Hemos de paso encontrado una fórmula (2.8) para la calcular la proyección de $\mathbf{q}$ sobre Π en términos de un punto $\mathbf{p}$ de Π y un vector normal $\mathbf{n}$ a Π. Podemos también calcular la distancia de $\mathbf{q}$ a Π: $\mathrm{d}(\mathbf{q}, \Pi) = \mathrm{d}(\mathbf{q}, \mathbf{r}) = \|\mathbf{r} - \mathbf{q}\| = \frac{|\langle \mathbf{p}-\mathbf{q}, \mathbf{n}\rangle|}{\|\mathbf{n}\|}$. De aquí sale una famosa fórmula para la distancia de un punto a un plano en $\mathbb{R}^3$. Supongamos que Π está descrito por la ecuación Cartesiana $Ax_1 + Bx_2 + Cx_3 + D = 0$, y $\mathbf{q} = \left(\begin{smallmatrix} q_1 \\ q_2 \\ q_3 \end{smallmatrix}\right)$. Entonces

$$\mathrm{d}(\mathbf{q}, \Pi) = \frac{|Aq_1 + Bq_2 + Cq_3 + D|}{\sqrt{A^2 + B^2 + C^2}}.$$

En efecto, podemos usar $\mathbf{n} = \left(\begin{smallmatrix} A \\ B \\ C \end{smallmatrix}\right)$ como vector normal para Π. Con ello la ecuación normal de Π será $\langle \mathbf{v} - \mathbf{p}, \mathbf{n} \rangle = 0$, donde $\mathbf{v} = \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right)$ es un punto que recorre Π. Manipulando esta ecuación normal se obtiene $\langle \mathbf{v}, \mathbf{n} \rangle - \langle \mathbf{p}, \mathbf{n} \rangle = 0$, ó $Ax_1 + Bx_2 + Cx_3 - \langle \mathbf{p}, \mathbf{n} \rangle = 0$, y por comparación con la ecuación cartesiana de Π, deberá tenerse que $-\langle \mathbf{p}, \mathbf{n} \rangle = D$.

Como $\langle \mathbf{q}, \mathbf{n} \rangle = Aq_1 + Bq_2 + Cq_3$, entonces

$$|\langle \mathbf{p} - \mathbf{q}, \mathbf{n} \rangle| = |\langle \mathbf{q}, \mathbf{n} \rangle - \langle \mathbf{p}, \mathbf{n} \rangle| = |Aq_1 + Bq_2 + Cq_3 + D|\,.$$

El resultado se sigue del cálculo de $\|\mathbf{n}\|$.

Cabe notar que hay otra manera, tal vez un poco menos eficiente, de calcular la proyección de $\mathbf{q}$ sobre Π, usando una ecuación vectorial para Π en vez de la normal: $\mathbf{v} = \mathbf{p} + \lambda \mathbf{d}_1 + \mu \mathbf{d}_2\,, \lambda, \mu \in \mathbb{R}$. Esto, junto a $\mathbf{v} = \mathbf{p} + t\mathbf{d}_1 \times \mathbf{d}_2$ que es la ecuación vectorial de L, entrega el sistema lineal $\mathbf{p} + \lambda \mathbf{d}_1 + \mu \mathbf{d}_2 = \mathbf{q} + t\mathbf{d}_1 \times \mathbf{d}_2$, o mejor $\left[\begin{array}{c|c|c} \mathbf{d}_1 & \mathbf{d}_2 & \mathbf{d}_1 \times \mathbf{d}_2 \end{array}\right] \left[\begin{smallmatrix} \lambda \\ \mu \\ t \end{smallmatrix}\right] = \mathbf{q} - \mathbf{p}$. Como $\mathbf{d}_1$ y $\mathbf{d}_2$ son no nulos y no paralelos, $\mathbf{d}_1$, $\mathbf{d}_2$ y $\mathbf{d}_1 \times \mathbf{d}_2$ no son coplanares con el origen, y entonces esta ecuación tiene una solución única para λ, μ y t, lo que entrega la proyección $\mathbf{r} = \mathbf{v}$.

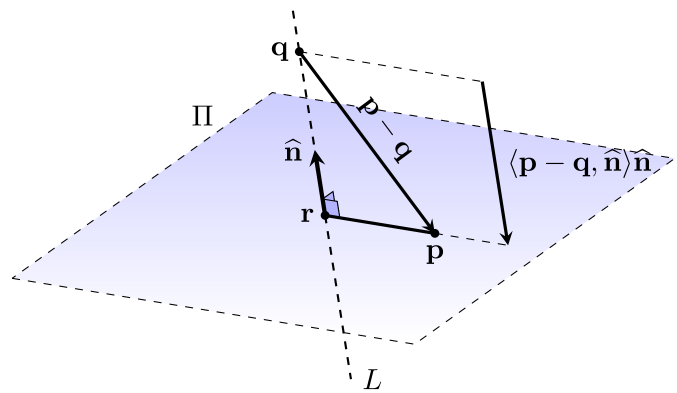
Una tercera manera de proceder, también algo mas larga que la primera, es utilizar la fórmula $\mathbf{r} = \mathbf{p} + \lambda \mathbf{d}_1 + \mu \mathbf{d}_2$ (que no dice más que el hecho de que la proyección $\mathbf{r}$ satisface la ecuación vectorial de Π pues es un elemento de este plano) y recordar que $\mathbf{r} - \mathbf{q} \perp \Pi$, es decir, $\mathbf{r} - \mathbf{q} \perp \mathbf{d}_1 \ \wedge \ \mathbf{r} - \mathbf{q} \perp \mathbf{d}_2$. Se obtiene el sistema de ecuaciones $\left\{\begin{array}{rcl} \langle \mathbf{d}_1, \mathbf{p} - \mathbf{q} + \lambda \mathbf{d}_1 + \mu \mathbf{d}_2 \rangle & = & 0 \\ \langle \mathbf{d}_2, \mathbf{p} - \mathbf{q} + \lambda \mathbf{d}_1 + \mu \mathbf{d}_2 \rangle & = & 0 \end{array}\right.$, o, en su forma matricial:

$$\begin{bmatrix} \langle \mathbf{d}_1, \mathbf{d}_1 \rangle & \langle \mathbf{d}_1, \mathbf{d}_2 \rangle \\ \langle \mathbf{d}_2, \mathbf{d}_1 \rangle & \langle \mathbf{d}_2, \mathbf{d}_2 \rangle \end{bmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \langle \mathbf{d}_1, \mathbf{q} - \mathbf{p} \rangle \\ \langle \mathbf{d}_2, \mathbf{q} - \mathbf{p} \rangle \end{pmatrix}.$$

Se puede probar que $\mathbf{d}_1$ **no** es paralelo a $\mathbf{d}_2$ **ssi** la matriz $\begin{bmatrix} \langle \mathbf{d}_1, \mathbf{d}_1 \rangle & \langle \mathbf{d}_1, \mathbf{d}_2 \rangle \\ \langle \mathbf{d}_2, \mathbf{d}_1 \rangle & \langle \mathbf{d}_2, \mathbf{d}_2 \rangle \end{bmatrix}$ es invertible, luego el sistema anterior tiene una solución única para λ y μ, lo que entrega la proyección $\mathbf{r}$.

Observación:

1. Notemos que la fórmula (2.8) era fácil de adivinar, pues si $\widehat{\mathbf{n}} = \frac{1}{\|\mathbf{n}\|}\mathbf{n}$ es el vector unitario en la dirección de $\mathbf{n}$, (2.8) se puede escribir como $\mathbf{r} = \mathbf{q} + \langle \mathbf{p} - \mathbf{q}, \widehat{\mathbf{n}} \rangle \, \widehat{\mathbf{n}}$



   y recordando que $\langle \mathbf{p} - \mathbf{q}, \widehat{\mathbf{n}} \rangle \, \widehat{\mathbf{n}}$ es la proyección (como vector) de $\mathbf{p} - \mathbf{q}$ en la dirección de $\widehat{\mathbf{n}}$, el punto $\mathbf{q} + \langle \mathbf{p} - \mathbf{q}, \widehat{\mathbf{n}} \rangle \, \widehat{\mathbf{n}}$ debería ser el lugar donde la recta que parte de $\mathbf{q}$ y va en la dirección de $\widehat{\mathbf{n}}$ encuentra al plano Π.

2. En muchos problemas se puede argumentar de manera similar a la observación 1) (es decir, proyectando sobre una recta vectores que conectan un par de puntos dados, etc.) para llegar rápidamente a una respuesta.

Guía de Ejercicios

1. Demuestre las siguientes propiedades relacionadas con el produto cruz:

(a) $\left(\forall \mathbf{x} = \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right) \in \mathbb{R}^3\right)$ $\mathbf{x} = x_1 \widehat{\mathbf{i}} + x_2 \widehat{\mathbf{j}} + x_3 \widehat{\mathbf{k}}$, y esta escritura es única, es decir, si $x_1 \widehat{\mathbf{i}} + x_2 \widehat{\mathbf{j}} + x_3 \widehat{\mathbf{k}} = y_1 \widehat{\mathbf{i}} + y_2 \widehat{\mathbf{j}} + y_3 \widehat{\mathbf{k}}$, entonces $x_1 = y_1 \wedge x_2 = y_2 \wedge x_3 = y_3$.

(b) $(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^3)$ $\mathbf{x} \times \mathbf{y} \perp \mathbf{x} \wedge \mathbf{x} \times \mathbf{y} \perp \mathbf{y}$.

(c) $(\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3)$ $\mathbf{x} \times (\mathbf{y} + \mathbf{z}) = \mathbf{x} \times \mathbf{y} + \mathbf{x} \times \mathbf{z} \wedge (\mathbf{x} + \mathbf{y}) \times \mathbf{z} = \mathbf{x} \times \mathbf{z} + \mathbf{y} \times \mathbf{z}$ ($\times$ distribuye sobre $+$).

(d) $(\forall \mathbf{x} \in \mathbb{R}^3)$ $\mathbf{x} \times \mathbf{x} = \mathbf{0}$.

(e) $(\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3)$ $(\mathbf{x} \times \mathbf{y}) \times \mathbf{z} = \langle \mathbf{x}, \mathbf{z} \rangle \mathbf{y} - \langle \mathbf{y}, \mathbf{z} \rangle \mathbf{x}$.

2. El objetivo de este ejercicio es probar:

> Sean $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ vectores no paralelos, y sea $\mathbf{z} \in \mathbb{R}^3$. Entonces
>
> *a)* $\mathbf{z} \perp \mathbf{x} \wedge \mathbf{z} \perp \mathbf{y} \implies (\exists \lambda \in \mathbb{R})\ \mathbf{z} = \lambda \mathbf{x} \times \mathbf{y}$.
>
> *b)* $\mathbf{z} \perp \mathbf{x} \times \mathbf{y} \implies (\exists s, t \in \mathbb{R})\ \mathbf{z} = s\mathbf{x} + t\mathbf{y}$.

Para ello:

(a) Sean $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ tres vectores no coplanares con el origen, es decir, que no existe un plano que pasa por el origen y que contiene a los tres vectores $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{z}$. Pruebe que esto equivale a:

$$\begin{array}{rrcl} (\forall s, t \in \mathbb{R}) & \mathbf{x} & \neq & \mathbf{0} + s\mathbf{y} + t\mathbf{z} \\ \wedge & \mathbf{y} & \neq & \mathbf{0} + s\mathbf{x} + t\mathbf{z} \\ \wedge & \mathbf{z} & \neq & \mathbf{0} + s\mathbf{x} + t\mathbf{y} \end{array}$$

(b) Pruebe que la condición anterior es equivalente a

$$(\forall \alpha, \beta, \gamma \in \mathbb{R})\ \alpha\mathbf{x} + \beta\mathbf{y} + \gamma\mathbf{z} = \mathbf{0} \implies \alpha = \beta = \gamma = 0. \tag{2.9}$$

(c) Sea $A \in M_{33}(\mathbb{R})$ la matriz descrita por columnas como $A = \left[\ \mathbf{x} \ \middle|\ \mathbf{y} \ \middle|\ \mathbf{z}\ \right]$. Pruebe que la propiedad 2.9 equivale a que el sistema homogéneo $A \left[\begin{smallmatrix} \alpha \\ \beta \\ \gamma \end{smallmatrix}\right] = \mathbf{0}$ tiene como solución única $\left[\begin{smallmatrix} \alpha \\ \beta \\ \gamma \end{smallmatrix}\right] = \mathbf{0}$, es decir, A es invertible.

(d) Concluya que tres vectores $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ no son coplanares con el origen ssi la matriz $A = \left[\ \mathbf{x} \ \middle|\ \mathbf{y} \ \middle|\ \mathbf{z}\ \right]$ es invertible.

(e) Considere ahora $\mathbf{x}$ e $\mathbf{y}$ no nulos y no paralelos. Pruebe que los vectores $\mathbf{x}, \mathbf{y}$ y $\mathbf{x} \times \mathbf{y} \neq 0$ son no coplanares con el origen.
Indicación: Verifique la condición 2.9.

(f) Use la conclusión de la parte 2c para probar que

$$\left(\forall \mathbf{z} \in \mathbb{R}^3\right)\ (\exists s, t, \lambda \in \mathbb{R})\ \mathbf{z} = s\mathbf{x} + t\mathbf{y} + \lambda \mathbf{x} \times \mathbf{y}. \tag{2.10}$$

(g) Pruebe, usando la propiedad 2.10, la proposición original.

3. Estudiar la proyección ortogonal de un punto $\mathbf{q} \in \mathbb{R}^2$ sobre una recta $L \subset \mathbb{R}^2$, usando ecuaciones normales.

Guía de Problemas

P1. Se definen las rectas

$$L_1\colon \begin{pmatrix}-1\\2\\1\end{pmatrix} + t\begin{pmatrix}1\\2\\-2\end{pmatrix},\ t \in \mathbb{R} \quad \text{y} \quad L_2\colon \begin{pmatrix}3\\1\\-1\end{pmatrix} + s\begin{pmatrix}0\\1\\-2\end{pmatrix},\ s \in \mathbb{R}.$$

(a) Verifique que L_1 y L_2 no se intersectan.

(b) Encuentre laecuación normal del plano Π que contiene a la recta L_1 y es paralelo a L_2.

(c) El punto $P = \begin{pmatrix}3\\1\\-1\end{pmatrix}$ pertenece a L_2. Encuentre la proyección ortogonal de P sobre el plano Π de la parte (b).

(d) Dé la ecuación del plano paralelo a Π que está a la misma distancia de L_1 y L_2 (Π es el plano de la parte (b)).

P2. **(a)** Sean P y Q puntos distintos en $\mathbb{R}^3$. Demuestre que el conjunto $A = \{x \in \mathbb{R}^3 : ||x - P|| = ||x - Q||\}$ es un plano. Encuentre un punto que pertenezca a A y encuentre un vector normal al plano A.

(b) En $\mathbb{R}^3$ considere las rectas

$$L_1\colon \begin{pmatrix}0\\5\\1\end{pmatrix} + t\begin{pmatrix}-1\\-3\\2\end{pmatrix},\ t \in \mathbb{R} \quad \text{y} \quad L_2\colon \begin{pmatrix}1\\2\\3\end{pmatrix} + s\begin{pmatrix}2\\4\\1\end{pmatrix},\ s \in \mathbb{R}.$$

(1) Demuestre que L_1 y L_2 se intersectan y encuentre la intersección.

(2) Encuentre el sistema de ecuaciones cartesianas que representan a la recta L que pasa por $P = \begin{pmatrix}1\\1\\1\end{pmatrix}$ y que es perpendicular al plano que contiene a L_1 y L_2.

(3) Determine las ecuaciones de los planos que son paralelos al plano que contiene a L_1 y L_2, y que se encuentra a distancia 1 del punto en que L_1 y L_2 se intersectan.

P3. (a) Sean $L_1 \subseteq \mathbb{R}^3$ el eje de las x y $L_2 \subseteq \mathbb{R}^3$ larecta que pasa por los puntos $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} -1 \\ 3 \\ -1 \end{pmatrix}$. Encuentre una tercera recta $L_3 \subseteq \mathbb{R}^3$ que sea perpendicular a L_1 y a L_2, y que además intersecte estas dos rectas.

(b) Sea Π el plano de $\mathbb{R}^3$ de ecuación cartesiana $x + z + y = 1$ y sea P el origen en $\mathbb{R}^3$. Calcular el punto $Q \in \mathbb{R}^3$ **simétrico** de P, con respecto a Π.

Ingeniería Matemática
FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Espacios vectoriales

3.1 Definiciones básicas

Las estructuras que hasta ahora se han estudiado están dotadas de operaciones que actuan internamente sobre el conjunto donde están definidas. Es decir, a un par de elementos del conjunto, se le asocia otro elemento del mismo conjunto. Veremos ahora un tipo diferente de estructura algebraica, donde la "multiplicación" actua sobre un par de elementos de conjuntos distintos. Esta estructura que estudiaremos está inspirada en la estructura lineal de $\mathbb{R}^n$. Consideremos $(V, +)$ un grupo Abeliano, es decir

- $+$ es ley de composición **interna**
- $+$ es asociativa y conmutativa.
- Existe un neutro, $0 \in V$, tal que: $\forall x \in V, x + 0 = 0 + x = x$.
- $\forall x \in V$, existe un inverso aditivo, $-x \in V$, tal que

$$x + (-x) = (-x) + x = 0.$$

Sea $\mathbb{K}$ un cuerpo y definamos una ley de composición denominada **externa**:

$$\begin{aligned} \mathbb{K} \times V &\to V \\ (\lambda, v) &\mapsto \lambda v \in V. \end{aligned}$$

Definición (Espacio vectorial) Dado un grupo abeliano $(V, +)$ y un cuerpo $(\mathbb{K}, +, \cdot)$, con una ley de composición externa. Diremos que V es un **espacio vectorial** sobre $\mathbb{K}$ si y sólo si la ley de composición externa satisface $\forall \lambda, \beta \in K, x, y \in V$:

(EV1) $(\lambda + \beta)x = \lambda x + \beta x$.

(EV2) $\lambda(x + y) = \lambda x + \lambda y$.

(EV3) $\lambda(\beta x) = (\lambda\beta)x$.

(EV4) $1\ x = x$, donde 1 es el neutro multiplicativo del cuerpo $\mathbb{K}$.

En tal caso, los elementos de V se denominan **vectores** y los de $\mathbb{K}$, **escalares**.

Observación: Señalemos que en la definición anterior participan dos estructuras algebraicas diferentes; un grupo Abeliano, $(V, +)$, y un cuerpo, $\mathbb{K}$. En tal sentido, dos espacios vectoriales son **iguales** si y sólo si ambos están definidos sobre los mismos grupos y cuerpos y, además con una ley de composición externa idéntica.

Ejemplos:

1. Es directo de la definición anterior que $\mathbb{K}^n$ es un espacio vectorial sobre $\mathbb{K}$.

2. Sea $(\mathcal{M}_{mn}(\mathbb{K}), +)$ el grupo Abeliano de las matrices de m filas y n columnas con elementos en $\mathbb{K}$. Definamos la ley de composición externa:

$$\begin{aligned} \mathbb{K} \times \mathcal{M}_{mn}(\mathbb{K}) &\to \mathcal{M}_{mn}(\mathbb{K}) \\ (\lambda, A) &\mapsto \lambda A = (\lambda a_{ij}) \end{aligned}$$

Es fácil verificar que se cumplen las propiedades (EV1) a (EV4). Por ejemplo:

$$\lambda(A+B) = \lambda \left\{ \begin{pmatrix} a_{11} \dots a_{1n} \\ \vdots \\ a_{m1} \dots a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} \dots b_{1n} \\ \vdots \\ b_{m1} \dots b_{mn} \end{pmatrix} \right\}$$

$$= \lambda \begin{pmatrix} a_{11} + b_{11} \dots a_{1n} + b_{1n} \\ \vdots \qquad\qquad \vdots \\ a_{m1} + b_{m1} \dots a_{mn} + b_{mn} \end{pmatrix} = \begin{pmatrix} \lambda(a_{11} + b_{11}) \dots \lambda(a_{1n} + b_{1n}) \\ \vdots \qquad\qquad \vdots \\ \lambda(a_{m1} + b_{m1}) \dots \lambda(a_{mn} + b_{mn}) \end{pmatrix}$$

$$= \begin{pmatrix} \lambda a_{11} \dots \lambda a_{1n} \\ \vdots \\ \lambda a_{m1} \dots \lambda a_{mn} \end{pmatrix} + \begin{pmatrix} \lambda b_{11} \dots \lambda b_{1n} \\ \vdots \quad \vdots \\ \lambda b_{m1} \dots \lambda b_{mn} \end{pmatrix} = \lambda A + \lambda B.$$

Concluímos entonces que $\mathcal{M}_{mn}(\mathbb{K})$ es un espacio vectorial sobre $\mathbb{K}$.

3. Sea $P_n(\mathbb{R})$ el conjunto de los polinomios de grado menor o igual a n, con la suma de funciones reales y la multiplicación por escalar definida por:

$$\begin{aligned} &\forall p, q \in P_n(\mathbb{R}),\ (p+q)(x) = p(x) + q(x) \\ &\forall \lambda \in \mathbb{R},\ (\lambda p)(x) = \lambda p(x) \end{aligned}$$

o bien, si $p(x) = \sum\limits_{i=0}^{n} a_i x^i, q(x) = \sum\limits_{i=0}^{n} b_i x^i$

$$(p+q)(x) = \sum_{i=0}^{n} (a_i + b_i) x^i, \quad (\lambda p)(x) = \sum_{i=0}^{n} (\lambda a_i) x^i.$$

Es directo verificar que $P_n(\mathbb{R})$ es espacio vectorial sobre $\mathbb{R}$. Además,

$$P_n(\mathbb{R}) \subseteq \mathcal{F}(\mathbb{R}, \mathbb{R}),$$

lo cual, nos llevará más adelante, a definir la noción de subespacio vectorial.

4. Tomemos ahora un cuerpo $(\mathbb{K},+,\cdot)$ arbitrario. Claramente $(\mathbb{K},+)$ es un grupo Abeliano. Además, definamos la ley de composición:

$$\begin{aligned} \mathbb{K}\times\mathbb{K} &\to \mathbb{K} \\ (\lambda, v) &\mapsto \lambda\cdot v \end{aligned}$$

donde "$\cdot$" es la multiplicación en el cuerpo $\mathbb{K}$. Como $\mathbb{K}$ es cuerpo, se verifican las propiedades . Concluímos entonces que todo cuerpo, $\mathbb{K}$, es espacio vectorial sobre si mismo. Dos ejemplos importantes son $\mathbb{R}$ y $\mathbb{C}$.

5. Sea el grupo Abeliano $(\mathbb{C},+)$ con la ley de composición externa:

$$\begin{aligned} \mathbb{R}\times\mathbb{C} &\to \mathbb{C} \\ (\lambda, a+bi) &\to \lambda(a+bi)=\lambda a+\lambda bi \end{aligned}$$

Es claro que $\mathbb{C}$ es un espacio vectorial sobre $\mathbb{R}$ pero no es el mismo espacio vectorial definido sobre $\mathbb{C}$. Es decir, el espacio vectorial $\mathbb{C}$ sobre $\mathbb{C}$ es **distinto** al espacio vectorial $\mathbb{C}$ sobre $\mathbb{R}$, pues los cuerpos sobre los cuales están definidos son diferentes.

Ejercicio 3.1: Sea V un e.v. sobre $\mathbb{K}$. Probar que:

1. $0v = v0 = 0,\ \forall v\in V, 0\in\mathbb{K}$, elemento neutro aditivo de $\mathbb{K}$.
2. $\lambda 0 = 0\lambda = 0, \forall\lambda\in\mathbb{K}, 0\in V$, elemento neutro aditivo de V.
3. $\lambda v = 0 \Rightarrow \lambda = 0 \vee v = 0, \lambda\in\mathbb{K}, v\in V$.

3.2 Subespacios vectoriales

Definición (Subespacio vectorial) Sea un espacio vectorial V sobre un cuerpo $\mathbb{K}$. Diremos que $U\neq\phi$, es un subespacio vectorial (s.e.v) de V si y sólo si:

1. $\forall u, v\in U,\ u+v\in U$
2. $\forall\lambda\in\mathbb{K}, \forall u\in U,\ \lambda u\in U$.

Es decir, ambas operaciones, la interna y la externa, son cerradas en U.

La siguiente proposición nos entrega una forma más compacta de caracterizar a un s.e.v.

Proposición 3.1. *Sea un espacio vectorial* V *sobre un cuerpo* $\mathbb{K}$. $U \neq \phi$, *es subespacio vectorial (s.e.v) de* V *si y sólo si:*

$$\forall \lambda_1, \lambda_2 \in \mathbb{K}, \ \forall u_1, u_2 \in U, \ \lambda_1 u_1 + \lambda_2 u_2 \in U.$$

Demostración. En efecto, si se verifica la primera definición entonces, de la propiedad 2, $\lambda_1 u_1 \in U$ y $\lambda_2 u_2 \in U$. De la propiedad 1 se concluye $\lambda_1 u_1 + \lambda_2 u_2 \in U$.

En el otro sentido, basta tomar $\lambda_1 = \lambda_2 = 1$ para verificar la primera propiedad y $\lambda_2 = 0$ para verificar la segunda. □

Observación: Si U es un s.e.v. de V entonces $0 \in U$ (basta tomar $\lambda = 0$). Luego el subespacio vectorial más pequeño de V es $\{0\}$. El más grande es, obviamente, V.

Ejemplos:

1. El conjunto de los polinomios de grado menor o igual a n, $P_n(\mathbb{R})$, es subespacio de $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

2. $U = \{(x, y) \mid ax + by = 0\}$ es un subespacio vectorial de $\mathbb{R}^2$. En efecto, sean $(x_1, y_1), (x_2, y_2) \in U, \lambda_1, \lambda_2 \in \mathbb{R}$:

$$\lambda_1(x_1, y_1) + \lambda_2(x_2, y_2) = (\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 y_1 + \lambda_2 y_2)$$

   Además, $a(\lambda_1 x_1 + \lambda_2 x_2) + b(\lambda_1 y_1 + \lambda_2 y_2) = \lambda_1(ax_1 + by_1) + \lambda_2(ax_2 + by_2) = \lambda_1 0 + \lambda_2 0 = 0$. Es decir, $\lambda_1(x_1, y_1) + \lambda_2(x_2, y_2) \in U$.
   El subespacio vectorial U está compuesto de todos los vectores de $\mathbb{R}^2$ contenidos en la recta $ax + by = 0$. Es claro que, dado un vector (x, y) de dicha recta, $\lambda(x, y)$, está en la misma recta así como la suma de dos vectores sobre ella. De esto concluímos que cualquiera recta que pasa por el origen es un s.e.v. de $\mathbb{R}^2$.

3. Sea la matriz $M \in M_{mn}(\mathbb{R})$, el conjunto $U = \{x \in \mathbb{R}^n \mid Mx = 0\}$ es un s.e.v. de $\mathbb{R}^n$. En efecto:
   Dados $\lambda_1, \lambda_2 \in \mathbb{R}, x, y \in \mathbb{R}^n$. $M(\lambda_1 x + \lambda_2 y) = \lambda_1 Mx + \lambda_2 My = \lambda_1 0 + \lambda_2 0 = 0$, luego, $\lambda_1 x + \lambda_2 y \in U$.

Supongamos ahora que, dado un espacio vectorial V sobre un cuerpo $\mathbb{K}$, se tienen U, W s.e.v. de V. Es directo que

$$U \cap W \text{ es también s.e.v de } V.$$

En efecto, si $x, y \in U \cap W, \lambda_1, \lambda_2 \in K$, entonces como U, W son s.e.v. de V, $\lambda_1 x + \lambda_2 y \in U$ y $\lambda_1 x + \lambda_2 y \in W$.

Sin embargo, la unión no es necesariamente un s.e.v de V.

Ejemplo 3.1.
Tomemos $U = \{(x,y) \in \mathbb{R}^2 / x - y = 0\}$, $W = \{(x,y) \in \mathbb{R}^2 / y - 2x = 0\}$. Es claro que ambos son s.e.v de $\mathbb{R}^2$. Sin embargo, $(1,1) \in U, (1,2) \in W$ pero $(1,1) + (1,2) = (2,3) \notin U \cup W$, luego no es s.e.v. (grafique para cerciorarse).

Esta constatación nos llevará, más adelante, a definir una operación entre espacios vectoriales (la "suma") tal que podamos hablar de "uniones" compatibles con la estructura de espacio vectorial.

3.3 Combinaciones lineales

Sea V un espacio vectorial sobre un cuerpo $\mathbb{K}$, y una colección de vectores $v_1, v_2, ..., v_n \in V$, y de escalares $\lambda_1, ..., \lambda_n \in K$. Denominamos **combinación lineal** a la suma ponderada de estos vectores:

$$\sum_{i=1}^{n} \lambda_i v_i = \lambda_1 v_1 + ... + \lambda_n v_n$$

Claramente, por inducción sobre n, se prueba que $\sum\limits_{i=1}^{n} \lambda_i v_i \in V$.

Dado un conjunto fijo, $v_1, ..., v_n \in V$ de vectores, definimos el conjunto de todas sus combinaciones lineales como sigue:

$$\langle \{v_1, ..., v_n\} \rangle = \{v \in V \mid v = \sum_{i=1}^{n} \lambda_i v_i, \lambda_i \in \mathbb{K}\}.$$

Proposición 3.2. *Sean V e.v. y $v_1, ..., v_n \in V$. Entonces $\langle \{v_1, \ldots, v_n\} \rangle$ es un subespacio vectorial de V. Además es el s.e.v. más pequeño que contiene los vectores $v_1, \ldots, v_n$. Es decir, si otro s.e.v U los contiene, entonces $\langle \{v_1, \ldots, v_n\} \rangle \subseteq U$.*

Por ende, $\langle \{v_1, \ldots, v_n\} \rangle$ es llamado **subespacio vectorial generado** *por $\{v_i\}_{i=1}^{n}$.*

Demostración. Para probar que es un subespacio vectorial de V, sean $u, v \in \langle \{v_1, \ldots, v_n\} \rangle, \alpha, \beta \in K$. Tenemos que $u = \sum\limits_{i=1}^{n} \lambda_i v_i,\ v = \sum\limits_{i=1}^{n} \lambda'_i v_i$, luego

$$\alpha u + \beta v = \sum_{i=1}^{n} (\alpha \lambda_i + \beta \lambda'_i) v_i \in \langle \{v_1, ..., v_n\} \rangle .$$

Para probar que es el más peque no, basta notar que como U es s.e.v. de V y contiene el conjunto $\{v_i\}_{i=1}^{n}$, entonces contiene todas sus combinaciones lineales. $\square$

Ejemplo 3.2.
Tomemos $v = (1,1) \in \mathbb{R}^2$, el subespacio $\langle\{v\}\rangle = \{\lambda(1,1)/\lambda \in \mathbb{R}\}$ corresponde a la recta que pasa por el orígen con dirección dada por el vector (1,1). Sean $v_1 = (1,0), v_2 = (0,1)$,

$$\langle\{v_1, v_2\}\rangle = \{\lambda(0,1) + \beta(1,0) \mid \lambda, \beta \in \mathbb{R}\} = \mathbb{R}^2.$$

En efecto, dado un vector arbitrario $(a,b) \in \mathbb{R}^2$, este puede escribirse como $(a,b) = a(1,0) + b(0,1)$.

3.4 Dependencia e independencia lineal

DEFINICIÓN Sea $\{v_i\}_{i=1}^n \subseteq V$, diremos que estos vectores son **linealmente dependientes** ($\ell.d.$) si y solo si: existen escalares $\{\lambda_1, ..., \lambda_n\}$, **no** todos nulos, tales que $\sum_{i=1}^n \lambda_i v_i = 0$. En caso contrario, es decir $\sum_{i=1}^n \lambda_i v_i = 0 \Rightarrow \lambda_i = 0 \ \forall i = 1, ..., n$, diremos que el conjunto de vectores $\{v_i\}_{i=1}^n$ es **linealmente independiente** ($\ell.i.$).

Ejemplos:

- $\{(1,1),(2,2)\}$ es un conjunto linealmente dependiente:

$$-2(1,1) + (2,2) = (0,0).$$

- Dado un espacio vectorial arbitrario V sobre $\mathbb{K}$, $\{0\}$ es linealmente dependiente pues: $\forall\lambda \ \lambda 0 = 0$.
- En general, cualquier conjunto de vectores, $\{v_i\}_{i=1}^n$ que contenga $0 \in V$ es linealmente dependiente. Basta tomar la combinación $0v_1 + 0v_2 + . \cdots + 0v_n + 1 \cdot 0 = 0$, con escalares no todos nulos.

De manera similar a lo anterior es fácil verificar que un conjunto dependiente sigue manteniendo esta propiedad al **agregarle** vectores. Además, un conjunto independiente, mantiene la independencia al **extraerle** vectores.

En efecto, sea $\{v_i\}_{i=1}^p \subseteq V$ un conjunto de vectores $\ell.d$, entonces existe una combinación lineal nula con escalares no todos nulos:

$$\sum_{i=1}^p \lambda_i v_i = 0.$$

Luego $\forall v \in V, \{v\} \cup \{v_i\}_{i=1}^p$ es $\ell.d.$, ya que basta tomar la combinación lineal:

$$0v + \sum_{i=1}^{n} \lambda_i v_i = 0.$$

Supongamos, ahora que el conjunto $\{v_i\}_{i=1}^p$ es $\ell.i.$, es directo de la de la definición que al sacar un vector, v_k, el conjunto $\{v_i\}_{i=1}^p \setminus \{v_k\}$ es $\ell.i.$

Ejemplos:

1. El conjunto $\{(1,1,1),(0,1,1),(1,0,0)\} \subseteq \mathbb{R}^3$ es $\ell.d.$ En efecto, $(1,1,1) = (0,1,1) + (1,0,0)$, es decir $(1,1,1) - (0,1,1) - (1,0,0) = (0,0,0)$ donde $\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = -1$.

2. En el espacio vectorial $\mathbb{P}_4(\mathbb{R})$, los polinomios $1, x$ son $\ell.i$ En efecto, tomemos una combinación lineal nula $\lambda 1 + \beta x = 0$, pero si $\beta \neq 0$ se tendrá que $\lambda + \beta x$ es un polinomio de grado 1, luego tiene a lo más una raíz y el polinomio 0 tiene infinitas raíces, luego $\lambda = \beta = 0$.

3. Los vectores en $\mathbb{R}^4, \{(1,1,0,1),(1,0,1,0),(0,0,1,1),(0,1,1,0)\}$, son $\ell.i.$ En efecto:

$$\lambda_1(1,1,0,1) + \lambda_2(1,0,1,0) + \lambda_3(0,0,1,1) + \lambda_4(0,1,1,0) = (0,0,0,0)$$

es equivalente al sistema de 4×4:

$$\begin{aligned} \lambda_1 + \lambda_2 &= 0 \\ \lambda_1 + \lambda_4 &= 0 \\ \lambda_2 + \lambda_3 + \lambda_4 &= 0 \\ \lambda_1 + \lambda_3 &= 0 \end{aligned}$$

en términos matriciales

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Aplicando Gauss a la matriz aumentada:

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

$$\rightarrow \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & -3 & 0 \end{pmatrix} \implies \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0.$$

En general en $\mathbb{K}^n$, dado un conjunto de m vectores $\{v_1, ..., v_m\}$, donde $v_i = (v_{1i}, ..., v_{ni})$, $i = 1, ..., m$, podemos estudiar su dependencia o independencia lineal a través de un sistema de n ecuaciones y m incógnitas. En efecto:

$$\lambda_1 v_1 + ... + \lambda_m v_m = 0$$

$$\iff \quad \lambda_1 \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix} + \lambda_2 \begin{pmatrix} v_{12} \\ v_{22} \\ \vdots \\ v_{n2} \end{pmatrix} + ... + \lambda_m \begin{pmatrix} v_{1m} \\ v_{2m} \\ \vdots \\ v_{nm} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}$$

$$\iff \quad \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1m} \\ v_{21} & v_{22} & \cdots & v_{2m} \\ \vdots & \vdots & & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nm} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{K}^m.$$

Definiendo $M = (v_{ij}) \in M_{nm}(\mathbb{K})$ (es decir poniendo los vectores originales como columnas) y $\bar{\lambda} \in \mathbb{K}^m$ como el vector de incógnitas, se tiene que para analizar la dependencia o independencia lineal basta estudiar las soluciones del sistema $M\bar{\lambda} = 0$.

Si existe $\bar{\lambda} \neq 0$ que satisface el sistema, entonces el conjunto de vectores $\{v_i\}_{i=1}^m$ es $\ell.d.$, en caso contrario es $\ell.i.$. Más aún, sabemos que si $m > n$ el sistema tiene más de una solución, en particular una no nula luego $\{v_i\}_{i=1}^m$ es $\ell.d.$.

Así se ha probado:

Teorema 3.1. *En $\mathbb{R}^n, m > n$ vectores son siempre linealmente dependientes.*

Observación: Conviene señalar que en $\mathbb{K}^n$ siempre es posible determinar n vectores independientes. Basta tomar el conjunto $\{e_i\}_{i=1}^n, e_i = (0, ..., 1, ..., 0)$, donde la componente no nula es la i-ésima.

Guía de Ejercicios

1. Sea V un e.v. sobre $\mathbb{K}$. Probar que:

(a) $0v = v0 = 0,\ \forall v \in V, 0 \in \mathbb{K}$, elemento neutro aditivo de $\mathbb{K}$.

(b) $\lambda 0 = 0\lambda = 0, \forall \lambda \in \mathbb{K}, 0 \in V$, elemento neutro aditivo de V.

(c) $\lambda v = 0 \Rightarrow \lambda = 0 \vee v = 0, \lambda \in \mathbb{K}, v \in V$.

2. Determinar cuáles de las siguientes estructuras son espacios vectoriales:

(a) $\mathbb{R}^2$ con la suma usual y la siguiente ponderación: $\alpha \in \mathbb{R},\ \alpha(a,b) = (\alpha a, b)$.

(b) $\mathbb{R}^2$ con la suma usual y la siguiente ponderación: $\alpha \in \mathbb{R},\ \alpha(a,b) = (\alpha a, 0)$.

(c) $\mathbb{R}^+$ con la "suma": $x \oplus y = x \times y$ y la ponderación: $\alpha x = x^\alpha$.

(d) El conjunto de las funciones de $[a,b]$ en $\mathbb{R}$ acotadas, con la suma y poderación (por escalares en $\mathbb{R}$) usuales.

(e) El conjunto de las funciones $f\colon \mathbb{R} \to \mathbb{R}$ tales que $f(-1) = f(1)$, con la suma y ponderación usuales.

3. Sea $P_n(\mathbb{R})$ el conjunto de los polinomios de grado menor o igual a n, con la suma de funciones reales y la multiplicación por escalar definida por:

$$\begin{aligned}&\forall p, q \in P_n(\mathbb{R}),\ (p+q)(x) = p(x) + q(x)\\ &\forall \lambda \in \mathbb{R},\ (\lambda p)(x) = \lambda p(x).\end{aligned}$$

Pruebe que $P_n(\mathbb{R})$ es espacio vectorial sobre $\mathbb{R}$.

4. Indique cuáles de los siguientes conjuntos son s.e.v. de $P_2(\mathbb{R})$ (considerado como en la parte anterior) y cuáles no lo son. Pruebe su respuesta.

(a) $U = \{p(x) = a(x^2 + x + 1) \mid a \in \mathbb{R}\}$.

(b) $U = \{p(x) = ax^2 + bx + b) \mid a, b \in \mathbb{R}\}$.

(c) $U = \{p(x) = x^2 + ax + b \mid a, b \in \mathbb{R}\}$.

(d) $U = \{p(x) = ax^2 + b \mid a, b \in \mathbb{R}\}$.

(e) $U = \{p(x) = ax^2 + b^2x + x + c \mid a, b, c \in \mathbb{R} \setminus \{0\}\}$.

5. Determine la dependencia o independencia lineal de cada uno de los siguientes conjuntos de vectores:

(a) $\left\{ \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 6 \\ -5 \end{pmatrix}, \begin{pmatrix} -7 \\ -5 \\ 5 \end{pmatrix} \right\} \subseteq \mathbb{R}^3$.

(b) $\{x^2 + 1, x^2 - 1, x^2 + x + 1\} \subseteq P_2(\mathbb{R})$.

(c) $\{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)\} \subseteq \mathcal{M}_{22}(\mathbb{R})$

6. **(a)** Sean a, b, c tres vectores de un espacio vectorial V. Determine si el conjunto $\{a + b, b + c, c + a\}$ es un conjunto l.i.

(b) Sea $\{u_1, u_2, \ldots, u_k\}$ un subconjunto l.i. de un espacio vectorial V. Pruebe que $\forall\alpha$ escalar, el conjunto:

$$\{u_1 + \alpha u_2, u_3, \ldots, u_k\} \quad \text{es l.i.}$$

Guía de Problemas

P1. Sea $P_3(\mathbb{R})$ el conjunto de polinomios de grado menor o igual a 3 con coeficientes reales. Sean $p_1, p_2, p_3 \in P_3(\mathbb{R})$ tales que

$$p_1(x) = 1 + 6x^2, \quad p_2(x) = 4x, \quad p_3(x) = 1 + 3 + 5x^2.$$

Demuestre que $P_2(\mathbb{R}) = \langle\{p_1, p_2, p_3\}\rangle$.

P2. Sean E, F e.v. sobre un cuerpo $\mathbb{K}$. Sea T una función $T : E \to F$, que satisface:

(1) $T(0_E) = 0_F$. En donde 0_E y 0_F son los neutros aditivos en cada e.v.

(2) $\forall x \in E,\ \forall\alpha \in \mathbb{K}$: $T(\alpha x) = \alpha T(x)$.

(3) $\forall x, y \in E$, $T(x + y) = T(x) + T(y)$.

Considere

$$T(E) = \{y \in F \mid y = T(x),\ x \in E\}.$$

(a) Muestre que $T(E)$ es un s.e.v. de F.

(b) Suponga además que T satisface que:

$$\forall x \in E,\ T(x) = 0 \implies x = 0.$$

Muestre que si $\{u_1, \ldots, u_n\} \subseteq E$ es l.i., entonces $\{T(u_1), \ldots, T(u_n)\} \subseteq F$ es l.i.

P3. Sean E y F espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$.

(a) Pruebe que dotando al conjunto $E \times F$ de las leyes

$$\forall(x, y), (u, v) \in E \times F, \quad (x, y) + (u, v) = (x + u, y + v)$$

$$\forall\alpha \in \mathbb{K}, \forall(x, y) \in E \times F, \quad \alpha(x, y) = (\alpha x, \alpha y),$$

éste resulta ser un e.v. sobre $\mathbb{K}$.

(b) Pruebe que $E \times \{0_F\}$ y $\{0_E\} \times F$ son subespacios vectoriales de $E \times F$. Determine $0_{E\times F}$ y pruebe que $(E \times \{0_F\}) \cap (\{0_E\} \times F) = \{0_{E\times F}\}$.

(c) Sea $\{e_1, e_2, \ldots, e_n\} \subseteq E$ un conjunto l.i. en E y $\{f_1, f_2, \ldots, f_m\} \subseteq F$ un conjunto l.i. en F. Determine un conjunto l.i. en $E \times F$, de tamaño $n + m$.

P4. Sea E el espacio vectorial sobre $\mathbb{R}$ de las funciones definidas sobre $\mathbb{Z}$ a valores reales:

$$E = \{f \mid f\colon \mathbb{Z} \to \mathbb{R} \text{ es función}\}.$$

(a) Sean $a_1, a_2 \in \mathbb{R}$ con $a_2 \neq 0$ y $\mathcal{F}_0$ el conjunto de las funciones $f \in E$ que verifican la condición:

$$\forall n \in \mathbb{Z},\ \ f(n) + a_1 f(n-1) + a_2 f(n-2) = 0.$$

Pruebe que $\mathcal{F}_0$ es un s.e.v. de E.

(b) Sea ahora $\mathcal{F}_1$ el conjunto de las funciones $f \in E$ que verifican la condición:

$$\forall n \in \mathbb{Z},\ \ f(n) + a_1 f(n-1) + a_2 f(n-2) = 1.$$

Pruebe que $\mathcal{F}_1$ **no** es un s.e.v. de E.

SEMANA 7:

Ingeniería Matemática
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Espacios vectoriales

3.5 Generadores de un espacio vectorial

Sea V un espacio vectorial sobre un cuerpo $\mathbb{K}$. Diremos que los vectores $\{v_1, ..., v_n\} \subseteq V$, **generan** V si y sólo sí:

$$\langle \{v_1, ..., v_n\} \rangle = V$$

o de manera equivalente:

$$\forall v \in V,\ \exists \{\lambda_i\}_{i=1}^n \subseteq \mathbb{K}, \quad \text{tal que} \quad v = \sum_{i=1}^n \lambda_i v_i.$$

Ejemplo 3.3.
Tomemos el conjunto $\{(1,0),(0,1),(1,1)\}$. Este genera $\mathbb{R}^2$. En efecto, $(x,y) = x(1,0) + y(0,1) + 0(1,1)$. Es decir $\langle \{(1,0),(0,1),(1,1)\} \rangle = \mathbb{R}^2$. Pero, en este caso, la forma de escribir (x,y) no es única:

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

En ambas combinaciones lineales "sobra" un vector. Más aún, constatamos que los tres vectores son l.d.:

$$(1,1) - (1,0) - (0,1) = (0,0)$$

Esto nos lleva a la noción de conjunto generador de $\mathbb{R}^2$ con un número mínimo de vectores linealmente independientes. En nuestro ejemplo anterior

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

¿Hay alguno más?.

Definición (Base) Dado un espacio vectorial V sobre $\mathbb{K}$, diremos que el conjunto de vectores $\{v_i\}_{i=1}^n$ es una **base** de V si y sólo sí:

(1) $\{v_i\}_{i=1}^n$ es un conjunto l.i.

(2) $V = \langle \{v_1, ..., v_n\} \rangle$.

Ejemplos:

1. En $\mathbb{K}^n$, el conjunto $\{e_i\}_{i=1}^n$, donde $e_i = (0, ..., 1, 0..,0)$ es base. En efecto, ya probamos que son l.i., además, dado $x = (x_1, ..., x_n) \in \mathbb{K}^n, x = \sum\limits_{i=1}^{n} x_i e_i$. Esta base de $\mathbb{K}^n$ se denomina **base canónica**.

2. Estudiemos, en $\mathbb{R}^3$, si el conjunto de vectores, $B = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$, es base. Veamos primero si B es generador. Sea $b = (b_1, b_2, b_3) \in \mathbb{R}^3$ y determinemos si existen escalares, $\{\lambda_1, \lambda_2, \lambda_3\}$, tales que: $(b_1, b_2, b_3) = \lambda_1(1, 1, 0) + \lambda_2(1, 0, 1) + \lambda_3(0, 1, 1)$. Esto es equivalente a estudiar el sistema de 3×3:

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

   Para resolverlo tomemos la matriz aumentada y escalaremos:

$$\begin{pmatrix} 1 & 1 & 0 & b_1 \\ 1 & 0 & 1 & b_2 \\ 0 & 1 & 1 & b_3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 & b_1 \\ 0 & -1 & 1 & b_2 - b_1 \\ 0 & 1 & 1 & b_3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 & b_1 \\ 0 & -1 & 1 & b_2 - b_1 \\ 0 & 0 & 2 & b_3 + b_2 - b_1 \end{pmatrix},$$

   de donde: $\lambda_3 = \frac{b_3+b_2-b_1}{2}, \lambda_2 = \frac{b_1+b_3-b_2}{2}, \lambda_1 = \frac{b_1+b_2-b_3}{2}$. Es decir, dado $b \in \mathbb{R}^3, b$ es combinación lineal de B:

$$(b_1, b_2, b_3) = \frac{b_1 + b_2 - b_3}{2}(1, 1, 0) + \frac{b_1 + b_3 - b_2}{2}(1, 0, 1) + \frac{b_2 + b_3 - b_1}{2}(0, 1, 1).$$

   Luego $\mathbb{R}^3 = \langle B \rangle$. Para determinar que el conjunto B es l.i. (y por lo tanto base)

   basta resolver el sistema anterior para $b = (0, 0, 0)$. En este caso se concluye $\lambda_1 = \lambda_3 = \lambda_3 = 0$, y por lo tanto son l.i.

3. En el espacio vectorial $P_n(\mathbb{R})$, el conjunto $B = \{1, x, x^2, ..., x^n\}$ es base. En efecto. Es claro que cualquier polinomio, q de grado $\leq n$ se escribe como combinación lineal de los vectores en B:

$$q(x) = \sum_{i=0}^{n} a_i x^i.$$

   para verificar que estos son $\ell.i$ tomemos la combinación lineal

$$\sum_{i=0}^{n} \lambda_i x^i = 0.$$

   Se tiene que, si el polinomio de la izquierda tiene algún coeficiente no nulo, entonces es un polinomio de grado entre 0 y n luego posee, a lo más, n raíces. Pero el polinomio de la derecha (polinomio nulo) tiene infinitas raíces. De esto se desprende que $\lambda_i = 0, \ i = 0, ..., n$.

Otra manera de verificar la independencia lineal es mediante derivación:

$$\frac{d}{dx}(\sum_{i=0}^{n} \lambda_i x^i) = \lambda_1 + 2\lambda_2 x + ... + n\lambda_n x^{n-1} = 0$$

$$\frac{d^2}{dx^2}(\sum_{i=0}^{n} \lambda_i x^i) = \ 2\lambda_2 + ... + n(n-1)\lambda_n x^{n-2} = 0$$

$$\vdots$$

$$\frac{d^n}{dx^n}(\sum_{i=0}^{n} \lambda_i x^i) = n!\lambda_n = 0$$

$$\Rightarrow \lambda_n = 0 \implies \lambda_{n-1} = 0 \implies ... \implies \lambda_1 = 0 \implies \lambda_0 = 0.$$

4. Existen espacios vectoriales en los cuales ningún conjunto finito de vectores es base, por ejemplo, sea S el conjunto de las sucesiones, es decir:

   $$S = \{(u_n)/u_n \in \mathbb{R}, \forall n \in \mathbb{N}\}.$$

   Claramente, con la suma $(u_n) + (v_n) = (u_n + v_n)$ y la multiplicación $\lambda(u_n) = (\lambda u_n), \lambda \in \mathbb{R}$, S es un espacio vectorial sobre $\mathbb{R}$.

   Vemos que el conjunto finito de sucesiones $B_n = \{x^{(1)}, \ldots, x^{(n)}\}$, donde $x_j^{(i)} = 1$ si y sólo si $i = j$ (0 en otro caso), es linealmente independiente. En efecto:

   $$\sum_{i=0}^{n} \lambda_i x^{(i)} = (\lambda_1, \lambda_2, ..., \lambda_n, 0, 0...) = (0, ..., 0, ...) \Rightarrow \lambda_i = 0 \quad \forall i = 1, ..., n.$$

   Pero, siempre es posible agrandar este conjunto. Basta tomar $B_{n+1} = B_n \cup \{x^{(n+1)}\}$. Además $\forall n \in \mathbb{N}, B_n$ no es generador pues dado n fijo, el vector $x^{(n+1)}$ es linealmente independiente de los vectores de B_n. Es decir, extendiendo adecuadamente las definiciones de l.i. y generador a conjuntos infinitos (lo cual se puede hacer, pero no lo haremos aquí), la "base" sería $B = \bigcup_{n\geq 1} B_n$, que es un conjunto infinito. Sin embargo, es posible probar que B tampoco puede generar a todas las sucesiones de S, es sólo un conjunto l.i.

5. Observemos que el conjunto de las progresiones aritméticas, PA es un s.e.v. de S, y este tiene una base finita. En efecto $\bar{\mathbf{1}} = (1, \ldots, 1, \ldots), \mathbf{e} = (1, 2, 3, \ldots, n, \ldots) \in PA$ y además son l.i.:

   $$\lambda\bar{\mathbf{1}} + \beta\mathbf{e} = (\lambda + \beta, \lambda + 2\beta, \lambda + 3\beta, ..., \lambda + u\beta, ...) = (0, ..., 0, ...)$$

   $$\Rightarrow (\lambda + \beta = 0) \wedge (\lambda + 2\beta = 0) \ \Rightarrow \ \lambda = \beta = 0.$$

   Además son generadores: Sea $(x_n)_n$ una progresión de diferencia d entonces se tiene $x_n = x_0 + nd$ y por lo tanto:

   $$(x_n) = x_0\bar{\mathbf{1}} + d\mathbf{e}.$$

Una caracterización de base, que nos será útil más adelante, es la siguiente:

Proposición 3.3. *Dado un espacio vectorial V, $B = \{v_i\}_{i=1}^n \subseteq V$ es una base si y sólo sí $\forall v \in V, v$ se escribe de* **manera única** *como combinación lineal de los vectores del conjunto B.*

Demostración. Demostremos esta propiedad:

$\Rightarrow$) Como B es base, $\langle B \rangle = V$, luego v se escribe como combinación lineal de los vectores en B. Supongamos que esto puede hacerse de dos maneras:

$$v = \sum_{i=1}^{n} \alpha_i v_i, \; v = \sum_{i=1}^{n} \beta_i v_i.$$

Se tiene entonces $\sum\limits_{i=1}^{n}(\alpha_i - \beta_i)v_i = 0$. Como B es un conjunto l.i., concluímos que $\alpha_i = \beta_i$ para todo i.

$\Leftarrow$) Por hipótesis, cualquier vector $v \in V$ es combinación lineal del conjunto B, luego B es generador. Supongamos que B es l.d., entonces $\sum\limits_{i=1}^{n} \lambda_i v_i = 0$ con escalares no todos nulos. Además $0v_1 + 0v + ... + 0v_n = 0$ luego el vector 0 se escribe de dos maneras distintas, lo cual es una contradicción. Concluímos entonces que B es base. $\square$

Observación: Por convención el conjunto vacío ϕ es la base del espacio vectorial $\{0\}$. Esto pues por definición diremos que ϕ es l.i. y además el subespacio más pequeño que lo contiene es $\{0\}$.

Otra propiedad importante es la siguiente:

Teorema 3.2. *Si $X = \{v_1, \ldots, v_n\} \subseteq V$ es un conjunto generador, entonces es posible extraer un subconjunto $B = \{v_{i_1}, \ldots, v_{i_s}\}$ que es base de V.*

Demostración. Procederemos de manera recursiva

1. Elijamos $v_{i_1} \in B$, el primer vector no nulo que encontremos (¿ Qué pasa si todos son cero?). Hagamos $B_1 = \{v_{i_1}\}$. Claramente B_1 es linealmente independiente.

2. Preguntamos ¿$X \setminus B_1 \subseteq \langle B_1 \rangle$?

   Si es cierto: Entonces B_1 es base. En efecto, $X \setminus B_1 \subseteq \langle B_1 \rangle \Rightarrow \forall j \neq i_1, v_j = \lambda_j v_{i_1}$. Como $V = \langle X \rangle \Rightarrow \forall v \in V$, $v = \sum\limits_{k=1}^{n} \alpha_k v_k = \sum\limits_{k \neq i_1} \alpha_k \lambda_k v_{i_1} + \alpha_{i_1} v_{i_1} = \lambda v_{i_1} \in \langle B_1 \rangle$, con $\lambda \in \mathbb{K}$ lo que implica que $V \subseteq \langle B_1 \rangle$. Es decir $V = \langle B_1 \rangle$, luego B_1 es l.i. y generador.

Si no es cierto: Entonces $\exists v_{i_2} \in X \setminus B_1, v_{i_2} \notin \langle B_1 \rangle$. Hacemos $B_2 = B_1 \cup \{v_{i_2}\}$ y repetimos el procedimiento anterior.

En el k-ésimo paso se habrá generado, a partir de $B_{k-1} = \{v_{i_1}, ..., v_{i_{k-1}}\}$, (que es linealmente independiente), el nuevo conjunto:

$$B_k = B_{k-1} \cup \{v_{i_k}\}, \quad v_{i_k} \notin \langle B_{k-1} \rangle .$$

B_k es linealmente independiente: En efecto, sea:

$$\sum_{j=1}^{k} \lambda_j v_{i_j} = 0.$$

Si $\lambda_k \neq 0$ entonces $v_{i_k} = -\frac{1}{\lambda_k} \sum_{j=1}^{k-1} \lambda_j v_{i_j} \in \langle B_{k-1} \rangle$ que es una contradicción. Luego $\lambda_k = 0$ y, como B_{k-1} es linealmente independiente, $\lambda_j = 0 \; \forall j = 1...k-1$.

Finalmente, como el conjunto de vectores X es finito, en un número finito de pasos determinamos un conjunto linealmente independiente $B_s = \{v_{i_1}, ..., v_{i_s}\}$, tal que $X \setminus B_s \subseteq \langle B_s \rangle$ y de manera análoga a lo ya hecho se prueba que es generador y por lo tanto base de V. $\square$

En la demostración anterior, elegimos el "próximo" vector, v_{i_k}, para "entrar" al conjunto B_{k-1}, exigiendo solamente: $v_{i_k} \notin \langle B_{k-1} \rangle$. Obviamente pueden existir varios vectores con tal propiedad. Esto permite afirmar que, en general, a partir de B podemos determinar varias bases (a menos claro que B mismo sea una base).

> **Ejemplo 3.4.**
> Si $B = \{(1,0),(0,1),(1,1)\}$, los subconjuntos $\{(1,0),(0,1)\}, \{(1,0),(1,1)\}$ y $\{(0,1),(1,1)\}$ son todos bases de $\mathbb{R}^2$.

Daremos ahora una generalización del Teorema 3.1.

Teorema 3.3. *Supongamos que tenemos $B = \{v_i\}_{i=1}^n$, base de V y un conjunto arbitrario $X = \{w_i\}_{i=1}^m \subseteq V$. Si $m > n$, entonces el conjunto X es l.d.*

Demostración. En efecto, como B es base: $w_i = \lambda_{1i} v_1 + \cdots + \lambda_{ni} v_n \quad 1 \leq i \leq m$. Tomemos una combinación lineal de los vectores de X:

$$\beta_1 \omega_1 + ... + \beta_m \omega_m = 0, \tag{3.1}$$

reemplazando cada w_i como combinación lineal de B, 3.1 equivale a:

$$\begin{aligned} & \sum_{i=1}^{m} \beta_i [\lambda_{1i} v_1 + ... + \lambda_{ni} v_n] = 0 \\ \iff & \sum_{i=1}^{m} \beta_i \sum_{j=1}^{n} \lambda_{ji} v_j = 0 \\ \iff & \sum_{j=1}^{n} v_j (\sum_{i=1}^{m} \beta_i \lambda_{ji}) = 0 \end{aligned}$$

y como $B = \{v_j\}_{j=1}^n$ es l.i., lo anterior es equivalente a

$$\sum_{i=1}^{m} \lambda_{ji}\beta_i = 0 \qquad 1 \leq j \leq n$$

o, de manera matricial $A\bar{\beta} = 0$, donde

$$A = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2m} \\ \vdots & \vdots & & \vdots \\ \lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nm} \end{pmatrix}, \ \bar{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}.$$

Del capítulo anterior sabemos que para $m > n$ el sistema anterior tiene más de una solución (infinitas), luego al menos una no nula. Es decir existen escalares $\beta_1, ..., \beta_m$, no todos nulos, solución de la ecuación 3.1. De donde concluímos que X es l.d. □

De este resultado se desprende directamente el corolario siguiente que es suma importancia:

Corolario 3.1. *Si $\{v_i\}_{i=1}^n$, y $\{u_i\}_{i=1}^m$ son bases de V, entonces $n = m$.*

Demostración. En efecto, si $n > m$, de la propiedad anterior concluímos $\{v_i\}_{i=1}^n$ es l.d., lo cual es una contradicción. □

Así, si existe una base **finita**, de cardinal n, entonces cualquier otra base contiene exactamente n vectores.

Esto nos lleva a definir la **dimensión** de un espacio vectorial.

Definición (Dimensión) Diremos que un espacio vectorial V sobre $\mathbb{K}$ es de dimensión n (finita) si admite una base de cardinalidad n. En caso que no exista una base finita, hablaremos de espacio vectorial de dimensión infinita.

Notaremos $\dim V$ al cardinal de una base ($\dim V = \infty$, si V no posee una base finita). En particular $\dim\{0\}=0$.

Ejemplos:

1. $\dim \mathbb{R}^n = n$.
2. $\dim S = \infty$.
3. $\dim PA = 2$.
4. $\dim M_{mn} = m \cdot n$.

Verifiquemos esta última. Tomemos el conjunto de matrices de 0 y 1:

$$B = \{\bar{E}[\alpha,\beta] = (e_{ij}^{\alpha\beta}), e_{ij}^{\alpha\beta} = 1 \Leftrightarrow (\alpha,\beta) = (i,j)\ \alpha \in \{1,...,m\}, \beta \in \{1,...,n\}\}.$$

Este conjunto corresponde a la base canónica del espacio de las matrices. Por ejemplo, para $m = 2$ y $n = 3$, el conjunto B está formado por las matrices:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.2}$$

Volvamos al caso general, claramente B es l.i.

$$\sum_{\alpha=1}^{m} \sum_{\beta=1}^{n} \lambda_{\alpha\beta} E[\alpha,\beta] = \begin{pmatrix} \lambda_{11} ... \lambda_{1n} \\ \vdots \\ \lambda_{m1} ... \lambda_{mn} \end{pmatrix} = \begin{pmatrix} 0 \ ... \ 0 \\ \vdots \\ 0 \ ... \ 0 \end{pmatrix}$$

$$\Longrightarrow \quad \lambda_{\alpha\beta} = 0 \quad \forall \alpha, \beta.$$

Además, dada $A = (a_{ij}) \in M_{mn}(\mathcal{K})$:

$$A = \sum_{\alpha=1}^{m} \sum_{\beta=1}^{n} a_{\alpha\beta} E[\alpha,\beta].$$

Luego, el conjunto $\{E[\alpha,\beta]\}_{\alpha,\beta}$ es base. Es decir $\dim M_{mn}(\mathcal{K}) = mn$.

Algunas propiedades de los espacios de dimensión finita son las siguientes:

Teorema 3.4.

1. *Sea* $\dim V = n$. *Si* $\{v_i\}_{i=1}^n$ *es l.i. entonces* $\{v_i\}_{i=1}^n$ *es base.*

2. *Sea* U *un s.e.v de* V, *luego* $\dim U \leq \dim V$ *más aún se tiene que* $\dim U = \dim V \Rightarrow U = V$.

Demostración.

1. Basta probar que $V = \langle\{v_1,...,v_n\}\rangle$. Supongamos $\exists u \notin \langle\{v_1,..,v_n\}\rangle$, luego $B = \{u, v_1,...,v_n\}$ es l.i. y $|B| = n+1 > n$. Lo que es una contradicción, pues todo conjunto de cardinal mayor que la dimensión debe ser l.d.

2. Supongamos $m = \dim U > n = \dim V$, entonces existe $\{u_i\}_{i=1}^m$ base de U, pero $\{u_i\}_{i=1}^m \subseteq V$ es l.i., de cardinal mayor que la dimensión, lo cual contradice el hecho que el máximo conjunto de vectores l.i. es de cardinalidad n.

   Para probar la última parte supongamos $U \neq V$. Sea $\{u_i\}_{i=1}^n$ una base de U. Como $U \neq V$ y $\langle\{u_1,...,u_n\}\rangle = U$,
   $\exists v \in V \setminus U \Rightarrow B = \{v\} \cup \{u_i\}_{i=1}^n$ es l.i. en V y $|B| > \dim V$, que al igual que antes da una contradicción. $\square$

Un resultado importante es la posibilidad de que, en un espacio vectorial V, dim $V = n$, un conjunto de vectores linealmente independiente, pueda completarse hasta formar una base de V.

Teorema 3.5 (Completación de base). *Dado V espacio vectorial sobre $\mathbb{K}$ con $\dim V = n$, y un conjunto de vectores l.i.; $X = \{v_1, \ldots, v_r\}$, $r < n$, entonces existen vectores $v_{r+1}, \ldots, v_n$, tales que el conjunto $\{v_1, \ldots, v_n\}$ es base de V.*

DEMOSTRACIÓN. Sea $U = \langle\{v_1, \ldots, v_r\}\rangle$, claramente $U \neq V$ (justifique). Luego $\exists v \in V \setminus U$. Definiendo $v_{r+1} = v$, el conjunto $\{v_1, ..., v_r, v_{r+1}\}$ es l.i. En efecto

$$\sum_{i=1}^{r+1} \lambda_i v_i = 0$$

Si $\lambda_{r+1} \neq 0$ entonces, se concluye que $v_{r+1} = -\frac{1}{\lambda_{r+1}} \sum_{i=1}^{r} \lambda_i v_i \in \langle\{v_1, ..., v_r\}\rangle$ lo cuál es una contradicción ya que $v_{r+1} \notin U$. Si $\lambda_{r+1} = 0 \Rightarrow \lambda_i = 0 \quad \forall i = 1, ..., r$ pues los vectores de X son l.i. Si $\langle\{v_1, ..., v_{r+1}\}\rangle = V$ hemos terminado, sino repetimos el procedimiento.

Ejemplo 3.5.
Consideremos, en $\mathbb{R}^4$ el conjunto

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

Claramente este conjunto es l.i. y además $(0, 0, 1, 0) \notin \langle\{(1, 1, 0, 0), (0, 1, 0, 0)\}\rangle$. Luego $\{(1, 1, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)\}$ es l.i.

Análogamente, el vector $(0, 0, 0, 1) \notin \langle\{(1, 1, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)\}\rangle$ luego el conjunto $B = \{(1, 1, 0, 0)(0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$ es l.i. y como contiene tantos vectores como dim $\mathbb{R}^4 = 4$, es base.

3.6 Suma de espacios vectoriales

Recordemos que la unión de subespacios de un espacio vectorial V, no es necesariamente un espacio vectorial. Definamos una operación entre subespacios que nos permita determinar uno nuevo, que contenga a ambos. Sean U, W subespacios de V, definimos la **suma de subespacios vectoriales** como sigue:

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

Es directo que $U + W$ es un s.e.v de V. En efecto, sean $x, y \in U + W, \lambda, \beta \in K$, luego $x = u_1 + w_1, y = u_2 + w_2$. Además, $\lambda x + \beta y = (\lambda u_1 + \beta u_2) + (\lambda w_1 + \beta w_2) \in U + W$ ya que U y W son s.e.v.

Ejemplo 3.6.
Tomemos en $\mathbb{R}^2$, los subespacios $U = \langle\{(1,1)\}\rangle\,, W = \langle\{(1,0)\}\rangle$

$$U + W = \{v = \lambda(1,1) + \beta(1,0) \mid \lambda, \beta \in \mathbb{R}\} = \mathbb{R}^2$$

En efecto, $(x,y) = y(1,1) + (x-y)(1,0)$

Sean, en $\mathbb{R}^3$, los subespacios $U = \langle\{(1,0,0)\}\rangle\,, W = \langle\{(0,1,0)\}\rangle$

$$U + W = \{(x,y,0) \mid x,y \in \mathbb{R}\} \neq \mathbb{R}^3.$$

En general, es directo que $U \subseteq U + W$ y $W \subseteq U + W$. Basta tomar los elementos de la forma $u + 0$ y $0 + w$, respectivamente.

Dado $V = U + W$, un vector no necesariamente admite una escritura única en términos de U y W.

Ejemplo 3.7.
Si

$$U = \langle\{(1,0),(0,1)\}\rangle\,,\ W = \langle\{(1,1)\}\rangle$$

luego $\mathbb{R}^2 = U + W$. Pero el vector $(1,1)$ se escribe de dos maneras:

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

donde $(1,1) \in U$ y $0 \in W$, pero además

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

donde esta vez $0 \in U$ y $(1,1) \in W$.

En el caso particular, en que la descomposición de un vector de V sea única en términos de los subespacios, diremos que V es **suma directa** de U y W,

Definición (Suma directa) Sea un espacio vectorial V y dos subespacios vectoriales U, W de V. Diremos que el subespacio $Z = U + W$ es **suma directa** de U y W, notado $U \oplus W = Z$, si $\forall v \in Z$, v se escribe de manera única como

$$v = u + w, \quad u \in U,\ w \in W.$$

En el caso en que V sea suma directa de U y W, diremos que estos últimos son **suplementarios.**

Ejemplo 3.8.
Sean, por ejemplo, los subespacios de $\mathbb{R}^2$, $V = \langle\{(1,0)\}\rangle\,, W = \langle\{(0,1)\}\rangle$
$\mathbb{R}^2 = U \oplus W,$ pues $(x,y) = x(1,0) + y(0,1)$ es única.

Una caracterización útil de la suma directa es la siguiente:

Proposición 3.4. *Dado V e.v. y U, W, Z s.e.v. de V, entonces*

$$Z = U \oplus W \iff (Z = U + W) \wedge (U \cap W = \{0\})$$

Demostración. En efecto:

$\Rightarrow$) Como Z es suma directa de U y W, $\forall v \in Z$, $\exists!\ u \in U, w \in W$ tal que $v = u + w \Rightarrow v \in U + W$, luego $Z = U + W$.

Supongamos ahora $v \in U \cap W$, entonces $v = v + 0 = 0 + v$. Es decir, v admite dos escrituras en $U + W$, luego necesariamente $v = 0$.

$\Leftarrow$) Como $Z = U + W, \forall v \in Z,\ v = u + w,\ u \in U, w \in W$. Veamos que esta descomposición es única. Supongamos $v = u + w$ y $v = u' + w'$, entonces $u - u' = w' - w$, donde $u - u' \in U$ y $w' - w \in W$, luego $u - u' \in U \cap W \Rightarrow u - u' = 0 \Rightarrow u = u' \Rightarrow w = w'$. □

Una propiedad interesante de la suma directa es:

Teorema 3.6. *Si $V = U \oplus W$ y V es de dimensión finita, entonces dim $V = \dim U + \dim W$.*

Demostración. En efecto, si U o W corresponde al subespacio nulo, $\{0\}$, el resultado es directo.

Supongamos entonces $U \neq \{0\} \neq W$. Sea $\{u_i\}_{i=1}^k$ base de $U, \{w_i\}_{i=1}^m$ base de W. Afirmamos que el conjunto $B = \{u_1, ..., u_k, w_1, ..., w_m\}$ es base de V:

1. B es linealmente independiente:

$$\sum_{i=1}^{k} \lambda_i u_i + \sum_{i=1}^{m} \beta_i w_i = 0$$

Como $\sum\limits_{i=1}^{k} \lambda_i u_i \in U, \sum\limits_{i=1}^{m} \beta_i w_i \in W$ y el vector 0 se escribe de manera única en $U + W$ como $0 = 0 + 0$, obtenemos:

$$\sum_{i=1}^{k} \lambda_i u_i = \sum_{i=1}^{m} \beta_i w_i = 0$$

y por independencia lineal de $\{u_i\}_{i=1}^k, \{w_i\}_{i=1}^m$ concluímos $\lambda_1 = ... = \lambda_k = \beta_1 = ... = \beta_k = 0$.

2. B es generador del e.v. V:

   Como $V = U \oplus W, \forall v \in V, v = u + w, u \in U, w \in W$. Pero $u \in U, w \in W$ se escriben como combinación lineal de sus bases: $u = \sum_{i=1}^{k} \alpha_i u_i$, $w = \sum_{i=1}^{m} \lambda_i w_i$, de donde,
   $v = \sum_{i=1}^{k} \alpha_i w_i + \sum_{i=1}^{m} \lambda_i w_i$, luego $\langle B \rangle = V$.

De (1) y (2) es directo que $\dim V = k + m = \dim U + \dim W$. □

Una pregunta importante es la siguiente: dado un s.e.v U de V , ¿ existe un suplementario de U?, es decir ¿ un s.e.v W tal que $V = U \oplus W$?

La respuesta es afirmativa, si $U = V$, basta tomar como suplementario $W = \{0\}$. Si $\dim U < \dim V$, tomemos una base de U : $X = \{u_i\}_{i=1}^{k}$. Por el teorema de completación de base, existen vectores $\{u_{k+1}, ..., u_n\}$ tales que $B = X \cup \{u_{k+1}, ..., u_n\}$ es base de V. Sea ahora
$W = \langle \{u_{k+1}, ..., u_n\} \rangle$. Afirmamos que $V = U \oplus W$. En efecto, como $\{u_1, ..., u_n\}$ es base de V, todo $v \in V$ puede escribirse, de manera única, como combinación lineal de esta base:

$$v = \sum_{i=1}^{k} \alpha_i u_i + \sum_{i=k+1}^{n} \alpha_i u_i$$

Como $u = \sum_{i=1}^{k} \alpha_i u_i \in U, w = \sum_{i=k+1}^{n} \alpha_i u_i \in W$, tenemos que un vector $v \in V$ se escribe de manera única como $v = u + w$.

Ejemplos: Tomemos $V = \mathbb{R}^3$ y los subespacios

$$U = \langle \{(1,0,0), (0,1,0)\} \rangle \,,\; W = \langle \{(0,1,0), (0,0,1)\} \rangle$$

$U \cap W = \langle \{(0,1,0)\} \rangle$. En efecto, si $v \in U \cap W, v = \alpha(1,0,0) + \beta(0,1,0) = \gamma(0,1,0) + \delta(0,0,1)$. De donde, $\alpha = \delta = 0, \beta = \gamma$. Luego $v = \gamma(0,1,0)$. Es decir $U \cap W = \langle \{(0,1,0)\} \rangle$, De esto concluímos que $U + W$ no es suma directa, sin embargo $\mathbb{R}^3 = U + W =$
$\langle \{(1,0,0), (0,1,0), (0,0,1)\} \rangle$.

En $\mathbb{R}^2$, cualquier par de rectas no colineales, que pasen por el origen, están asociados a subespacios suplementarios. Luego el suplementario de un s.e.v. U **no es único**.

El siguiente resultado, propuesto como ejercicio, extiende el Teorema 3.6 al caso en que la suma no es necesariamente directa.

Teorema 3.7. *Supongamos que* $V = U + W$, *entonces* $\dim V = \dim U + \dim W - \dim U \cap W$.

Demostración. Propuesta como ejercicio. Ver la guía para indicaciones. □

Guía de Ejercicios

1. Averigue si el conjunto de polinomios $\{1+x^3, 1-x^3, x-x^3, x+x^3\}$ constituye una base de $F = \{p \in \mathcal{P}(\mathbb{R}) \mid \mathrm{gr}(p) < 4\}$. Donde $\mathcal{P}(\mathbb{R})$ el es espacio de los polinomios a coeficientes reales.

2. Sea W es subespacio vectorial de $\mathbb{R}^4$ generado por el conjunto

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 2 \\ 3 \end{pmatrix} \right\}.$$

(a) Determine una base de W y su dimensión.

(b) Extienda la base encontrada antes a una base de $\mathbb{R}^4$.

3. Se desea probar el siguiente resultado:

> Supongamos que $V = U + W$, entonces $\dim V = \dim U + \dim W - \dim U \cap W$.

Para ello:

Considere $U \cap W$ como subespacio de U y W. Si se toma $\{z_1, ..., z_k\}$ base de $U \cap W$, se puede extender a

$$\begin{aligned} B_U &= \{z_1, ..., z_k, u_1, ..., u_l\} \text{ base de } U \text{ y} \\ B_W &= \{z_1, ..., z_k, w_1, ..., w_p\} \text{ base de } W. \end{aligned}$$

Se probará que $B = \{z_1, ..., z_k, u_1, ..., u_l, w_1, ..., w_p\}$ es base de V.

(a) Pruebe que B genera a V. Es decir, tome $v \in V$ y pruebe que es combinación lineal de los vectores de B. Use que $V = U + W$.

(b) Para probar que B es l.i., tome una combinación lineal de B igual a cero:

$$0 = \alpha_1 z_1 + ... + \alpha_k z_k + \gamma_1 u_1 + ... + \gamma_l u_l + \beta_1 w_1 + ... + \beta_p w_p$$

de donde

$$-(\beta_1 w_1 + ... + \beta_p w_p) = \alpha_1 z_1 + ... + \alpha_k z_k + \gamma_1 u_1 + ... + \gamma_l u_l. \tag{3.3}$$

Pruebe que existen $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ tales que

$$-(\beta_1 w_1 + ... + \beta_p w_p) = \lambda_1 z_1 + ... + \lambda_k z_k.$$

(c) Demuestre que $\beta_1 = \cdots = \beta_p = \lambda_1 = \cdots = \lambda_k = 0$.

(d) Reemplace lo anterior en 3.3 y use que B_U es base para probar que $\alpha_1 = \cdots = \alpha_k = \gamma_1 = \cdots = \gamma_l = 0$.

(e) Concluya que B es l.i. y por ende base de V.

(f) Concluya el resultado buscado.

Guía de Problemas

P1. Considere los conjuntos W y U definidos por:

$$W = \{M \in \mathcal{M}_{33}(\mathbb{R}) \mid M = \begin{pmatrix} a & b & 0 \\ b & e & c \\ 0 & c & i \end{pmatrix} \ \wedge\ a + e + i = 0,\ a, b, c, e, i \in \mathbb{R}\}.$$

$$U = \{M \in \mathcal{M}_{33}(\mathbb{R}) \mid M = \begin{pmatrix} 0 & r & s \\ r & 0 & 0 \\ s & 0 & 0 \end{pmatrix},\ r, s \in \mathbb{R}\}.$$

(a) Demuestre que W y U son subespacios vectoriales de $\mathcal{M}_{33}(\mathbb{R})$.

(b) Encuentre bases y dimensiones para los subespacios $U, W, U \cap W$ y $U + W$, y decida (justificando) si la suma $U + W$ es directa.

¿Cuántos vectores es preciso agregar a una base de $U + W$ para obtener una base de $S = \{M \in \mathcal{M}_{33}(R) \mid M \text{ simétrica}\}$? Justifique encontrando una base de S.

P2. Sea $m = 2n$ con $n > 0$ y considere el conjunto $\mathcal{P}_m(\mathbb{R})$ de los polinomios reales de grado menor o igual a m. Si cada $p \in \mathcal{P}_m(\mathbb{R})$ se escribe $p(x) = a_0 + a_1 x + \cdots + a_m x^m$, se define el conjunto

$$V = \{p \in \mathcal{P}_m(\mathbb{R}) \mid \forall i \in \{0, \ldots, m\}, a_i = a_{m-i}.\}$$

(a) Probar que V es subespacio vectorial de $\mathcal{P}_m(\mathbb{R})$ sobre los reales.

(b) Encontrar una base de V y deducir que su dimensión es $n + 1$.

(c) Probar que $\mathcal{P}_m(\mathbb{R}) = V \oplus \mathcal{P}_{n-1}(\mathbb{R})$.

(d) Se define

$$V' = \{p(x) = \sum_{i=0}^{m} a_i x^i \in \mathcal{P}_m(\mathbb{R}) \mid \forall i \in \{0, \ldots, m\},\ a_i = -a_{m-i}\}.$$

Probar que $\mathcal{P}_m(\mathbb{R}) = V \oplus V'$ (asuma que V' es un subespacio vectorial de $\mathcal{P}_m(\mathbb{R})$).

P3. Si $A \in \mathcal{M}_{nn}(\mathbb{R})$ y V es un subespacio vectorial de $\mathbb{R}^n$, se define:

$$A(V) = \{Ax \mid x \in V\}.$$

(a) **(1)** Pruebe que si $A \in \mathcal{M}_{nn}(\mathbb{R})$ y V es s.e.v. de $\mathbb{R}^n$ entonces $A(V)$ también es s.e.v. de $\mathbb{R}^n$.

(2) Sean V, W s.e.v. de $\mathbb{R}^n$ tales que $V \oplus W = \mathbb{R}^n$. Pruebe que si $A \in \mathcal{M}_{nn}(\mathbb{R})$ es invertible entonces $A(V) \oplus A(W) = \mathbb{R}^n$.

(3) Sean V, W s.e.v. de $\mathbb{R}^n$ tales que $V \oplus W = \mathbb{R}^n$. Pruebe que si $A(V) \oplus A(W) = \mathbb{R}^n$, entonces A es invertible.
Indicación: Pruebe que para todo $z \in \mathbb{R}^n$ el sistema $Ax = z$ tiene solución.

(b) **(1)** Sea W un s.e.v. de $\mathbb{R}^n$ y definamos $E = \{A \in \mathcal{M}_{nn}(\mathbb{R}) \mid A(\mathbb{R}^n) \subset W\}$. Muestre que E es un s.e.v. de $\mathcal{M}_{nn}(\mathbb{R})$.

(2) Sea $W = \{(t,t) \mid t \in \mathbb{R}\}$. Calcule la dimensión de $E = \{A \in \mathcal{M}_{22}(\mathbb{R}) \mid A(\mathbb{R}^2) \subset W\}$.

SEMANA 8:

Ingeniería Matemática
FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Transformaciones lineales

4.1 Introducción

Sea la matriz $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ y multipliquemos los vectores de $\mathbb{R}^2$ por esta matriz. Sea $v = \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)$, se tiene entonces $Av = \left(\begin{smallmatrix} x_2 \\ x_1 \end{smallmatrix}\right)$, es decir el vector que se obtiene de intercambiar las coordenadas de v.

Si tomamos $A = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, al multiplicar por el vector v obtenemos $Av = -\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)$, que es el vector simétrico con respecto al orígen.

Estos dos ejemplos nos dan funciones completamente distintas pero ellas comparten propiedades que son fundamentales: la imagen de la suma de dos vectores es la suma de las imagenes y la imagen de un vector ponderado por un real es la imagen del vector ponderada por el mismo escalar. Funciones que verifican estas propiedades son llamadas lineales y son el objeto de este capítulo. Veamos un caso más general.

En $\mathbb{R}^n$ consideremos T como la función:

$$\begin{aligned} T : \mathbb{R}^n &\to \mathbb{R}^m \\ v &\to Av \end{aligned}$$

donde $A \in \mathcal{M}_{mn}(\mathbb{R})$.

Este tipo de transformaciones tiene dos propiedades fundamentales:

1. $T(u + v) = T(u) + T(v) \qquad \forall u, v \in \mathbb{R}^n$
2. $T(\lambda v) = \lambda T(v) \qquad \forall \lambda \in \mathbb{R}, v \in \mathbb{R}^n$

En efecto, $T(u+v) = A(u+v)$. Como la multiplicación matricial distribuye con respecto a la suma:

$$T(u + v) = Au + Av = T(u) + T(v)$$

Por otra parte, $T(\lambda v) = A(\lambda v) = \lambda Av = \lambda T(v)$

4.2 Tranformaciones lineales

DEFINICIÓN (TANSFORMACIÓN LINEAL) Sean U, V dos espacios vectoriales sobre el *mismo* cuerpo $\mathbb{K}$. Diremos que una función $T : U \to V$ es una transformación (o función) **lineal** si y sólo si satisface:

$$1.\ \ \forall u_1, u_2 \in U,\ T(u_1 + u_2) = T(u_1) + T(u_2) \tag{4.1}$$

$$2.\ \ \forall u \in U, \lambda \in K,\ T(\lambda u) = \lambda T(u) \tag{4.2}$$

Señalemos que la propiedad 4.1 establece un homomorfismo entre los grupos $(U, +)$ y $(V, +)$.

Ejemplos:

1. Cualquier función $T : \mathbb{R}^n \to \mathbb{R}^m, X \mapsto AX$, es lineal.

2. El caso particular, $f : \mathbb{R} \to \mathbb{R}, x \to ax, a \in \mathbb{R}$ es lineal. En efecto, $f(x + y) = a(x + y) = ax + ay = f(x) + f(y)$. Además, $f(\lambda x) = a(\lambda x) = \lambda ax = \lambda f(x)$. Esta transformación corresponde a una recta (línea) que pasa por el orígen con pendiente a.

3. $f : \mathbb{R} \to \mathbb{R}, x \to x^2$, no es lineal. En efecto: $f(x+y) = (x+y)^2 = x^2+2xy+y^2 = f(x)+f(y)+2xy$, luego no es lineal. En general cualquier función real de grado superior a uno , trigonométrica, etcétera, no es lineal.

4. $f : P_3(\mathbb{R}) \to \mathbb{R}^4$, $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \to f(p) = (a_0, a_1, a_2, a_3)$, es lineal. En efecto; si $q(x) = b_0 + b_1x + b_2x^2 + b_3x^3$, obtenemos

$$\begin{aligned}
f(p + q) &= f((a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + (a_3 + b_3)x^3) = \\
&= (a_0 + b_0, a_1 + b_1, a_2 + b_2, a_3 + b_3) = (a_0, a_1, a_2, a_3) + (b_0, b_1, b_2, b_3) \\
&= f(p) + f(q) \\
f(\lambda p) &= f(\lambda a_0 + \lambda a_1x + \lambda a_2x^2 + \lambda a_3x^3) = (\lambda a_0, \lambda a_1, \lambda a_2, \lambda a_3) \\
&= \lambda(a_0, a_1, a_2, a_3) = \lambda f(p).
\end{aligned}$$

5. Sea $\mathcal{F}_d(\mathbb{R}, \mathbb{R})$ el conjunto de las funciones reales derivables. Es directo verificar que este conjunto es un s.e.v. de $\mathcal{F}(\mathbb{R}, \mathbb{R})$. Consideremos la transformación:

$$\begin{aligned}
T : \mathcal{F}_d(\mathbb{R}, \mathbb{R}) &\to \mathcal{F}(\mathbb{R}, \mathbb{R}) \\
f &\mapsto T(f)
\end{aligned}$$

donde $T(f)(x) = \frac{df}{dx}(x)$ es la función derivada. Es directo, de las propiedades de derivación que T es lineal.

6. Sea V e.v sobre $\mathbb{K}$, $V = V_1 \oplus V_2$ y sean:

$$P_1 : V \to V_1, \qquad P_2 : V \to V_2$$

$$v = v_1 + v_2 \mapsto P_1(v) = v_1 \qquad v = v_1 + v_2 \mapsto P_2(v) = v_2$$

Ambas funciones (¿porqué son funciones?) son lineales. Verifiquemos para P_1: sean $\lambda, \beta \in \mathbb{K}$ $v = v_1 + v_2, u = u_1 + u_2, v_i, u_i \in V_i, i = 1, 2$.

$$\begin{aligned} P_1(\lambda v + \beta u) &= P_1(\lambda v_1 + \beta u_1) + (\lambda v_2 + \beta u_2)) = \\ &= \lambda v_1 + \beta u_1 = \lambda P_1(v) + \beta P_1(u) \end{aligned}$$

La función, P_i, se denomina la proyección de V sobre V_i.

4.3 Propiedades de transformaciones lineales

Proposición 4.1. *Sea $T : U \to V$ una transformación lineal. Se tiene entonces:*

1. $T(0) = 0 \in V$.
2. $T(-u) = -T(u)$.
3. *T es lineal si y sólo si* $\forall \lambda_1, \lambda_2 \in K, \forall u_1, u_2 \in U$

$$T(\lambda_1 u_1 + \lambda_2 u_2) = \lambda_1 T(u_1) + \lambda_2 T(u_2).$$

Demostración. Demostremos estas propiedades:

1. $T(u) = T(u + 0) = T(u) + T(0)$
   como V es espacio vectorial $(V, +)$ es un grupo Abeliano, luego podemos cancelar $T(u)$; de donde $0 = T(0)$.

2. $T(-u) = T(-1u) = -1T(u) = -T(u)$.

3. $\Rightarrow)$ $T(\lambda_1 u_1 + \lambda_2 u_2) = T(\lambda_1 u_1) + T(\lambda_2 u_2) = \lambda_1 T(u_1) + \lambda_2 T(u_2)$.
   $\Leftarrow)$ Tomando $\lambda_1 = \lambda_2 = 1$ se verifica la propiedad (1) de linealidad. Considerando $\lambda_2 = 0$, se obtiene la propiedad (2).

Observación: Note que las demostraciones de las propiedades 1 y 2 podrían haber sido omitidas, ya que dichas propiedades son consecuencias de que T sea morfismo entre los grupos $(U, +)$ y $(V, +)$.

Dada una combinación lineal de vectores $\sum_{i=1}^{n} \lambda_i x_i \in U$ y la transformación lineal, $T : U \to V$. Se tiene, por inducción, que:

$$T(\sum_{i=1}^{n} \lambda_i x_i) = \sum_{i=1}^{n} \lambda_i T(x_i)$$

Si $\dim U = n$ y $\beta = \{u_i\}_{i=1}^{n}$ es una base de U, sabemos que $u \in U$ se escribe, de manera única, como $u = \sum_{i=1}^{n} \alpha_i u_i$, donde los escalares $\alpha = (\alpha_1, \ldots, \alpha_n)$ corresponden a las coordenadas de u con respecto a la base β. Aplicando a este vector u la transformación lineal T:

$$T(u) = T\left(\sum_{i=1}^{n} \alpha_i u_i\right) = \sum_{i=1}^{n} \alpha_i T(u_i) \tag{4.3}$$

De este cálculo se desprende que: basta definir la transformación T sobre una base de U, es decir calcular sólo $\{T(v_i)\}_{i=1}^{n}$. Para determinar el efecto de T sobre un vector $u \in V$ arbitrario, basta aplicar la ecuación 4.3.

Ejercicio 4.1: Pruebe que entonces dada una base $\beta = \{u_i\}_{i=1}^{n}$ de U y n vectores $X = \{v_i\}_{i=1}^{n}$ no necesariamente $\ell.i.$ de V existe una única transformación lineal T

$$\begin{aligned} T : U &\to V \\ u &\mapsto T(u) \end{aligned}$$

tal que $T(u_i) = v_i \; i = 1, ..., n$.

Consideremos ahora la transformación:

$$\begin{aligned} f : U &\to \mathbb{K}^n \\ u &\mapsto f(u) = \alpha = (\alpha_1, \ldots, \alpha_n) \end{aligned}$$

que a cada vector u le asocia las coordenadas con respecto a una base fija $\beta = \{u_i\}_{i=1}^{n}$.

Es claro que f es función (representación única de un vector con respecto a una base). Además, es lineal. Más aún f es biyectiva (verifíquelo). Esto nos lleva a la siguiente definición.

Definición (Isomorfismo) Sea $T : U \to V$ una transformación lineal, diremos que es un **isomorfísmo** si T es biyectiva.

Además diremos que U y V son isomorfos si existe un isomorfísmo entre U y V, en cuyo caso lo denotaremos como $U \cong V$.

En nuestro ejemplo anterior f es un isomorfismo y $U \cong \mathbb{K}^n$. Esto quiere decir que si U tiene dimensión finita n se "comporta" como si fuera $\mathbb{K}^n$. Esta isomorfía no es única, depende de la base elegida. Por el momento esto no debe preocuparnos. Calculemos además la imagen, a través del isomorfismo f, de la base $\{u_i\}_{i=1}^n$ del espacio vectorial U.

$$f(u_i) = f(0u_1 + \cdots + 1u_i + \cdots + 0u_n) = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = e_i \in \mathbb{K}^n$$

Luego, la base asociada a $\{u_i\}_{i=1}^n$ es la base canónica de $\mathbb{K}^n$.

4.4 Composición de funciones lineales

Consideremos U, V, W espacios vectoriales sobre el mismo cuerpo. Sean $T : U \to V$ y $L : V \to W$ dos transformaciones lineales, la composición $L \circ T : U \to W$ es fácil ver que es una función lineal. Si además L y T son biyectivas (isomorfísmos) entonces también lo es $L \circ T$.

Otra propiedad importante es que si $T : U \to V$ es un isomorfísmo tendremos que la función inversa, que existe pues T es biyectiva, también es un isomorfísmo. En efecto sólo resta probar que T^{-1} es lineal, pues es claro que es biyectiva (¿por qué?). Tomemos $v_1, v_2 \in V,\ \lambda, \beta \in \mathbb{K}$. Sean $T^{-1}(v_1) = u_1, T^{-1}(v_2) = u_2$ por lo tanto

$$T(\lambda u_1 + \beta u_2)) = \lambda T(u_1) + \beta T(u_2) = \lambda v_1 + \beta v_2,$$

esto quiere decir que

$$T^{-1}(\lambda v_1 + \beta v_2) = \lambda u_1 + \beta u_2 = \lambda T^{-1}(v_1) + \beta T^{-1}(v_2),$$

entonces T^{-1} es lineal.

Una transformación lineal importante es $id_U : U \to U$ que es la transformación identidad. Como sabemos que es biyectiva se tiene que id_U es un isomorfísmo.

Con todos estos ingredientes se puede concluir que la relación entre espacios vectoriales de ser isomorfos es una relación de equivalencia.

4.5 Subespacios Asociados a una transformación lineal

> Definición (Núcleo) Sea una transformación lineal $T : U \to V$. Definimos el **núcleo** de T como el conjunto:
>
> $$\operatorname{Ker} T = \{x \in U \mid T(x) = 0\}.$$

Claramente, $\operatorname{Ker} T \neq \phi$ ya que como $T(0) = 0$ tenemos que siempre $0 \in \operatorname{Ker} T$. Más aún, $\operatorname{Ker} T$ es un s.e.v. de U. En efecto, sean $x, y \in \operatorname{Ker} T, \lambda, \beta \in \mathbb{K}$;

$$T(\lambda x + \beta y) = \lambda T(x) + \beta T(y) = \lambda 0 + \beta 0 = 0$$

luego, $\lambda x + \beta y \in \operatorname{Ker} T$.

Otro conjunto importante en el estudio de transformaciones lineales es la **imagen**:

Definición (Imagen) Sea una transformación lineal $T : U \to V$. Definimos la **imagen** de T como el conjunto:

$$\operatorname{Im} T = T(U) = \{v \in V \mid \exists u \in U : \ v = f(u)\}$$

$\operatorname{Im} T$ es un s.e.v de V. En efecto, sean $v_1, v_2 \in \operatorname{Im} T$, existen entonces $u_1, u_2 \in U$ tales que $v_i = T(u_i), i = 1, 2$. Dados $\lambda, \beta \in \mathbb{K}$ se tiene:

$$\lambda v_1 + \beta v_2 = \lambda T(u_1) + \beta T(u_2) = T(\lambda u_1 + \beta u_2)$$

de donde concluímos $\lambda v_1 + \beta v_2 \in \operatorname{Im} T$.

Definición (Rango y nulidad) La dimensión de $\operatorname{Im} T$ se denomina el *rango* de la transformación T y se nota r. La dimensión del $\operatorname{Ker} T$ se llama **nulidad** y se suele denotar por la letra griega ν.

Ejemplo 4.1.
Desarrollemos un ejemplo. Sea la transformación lineal:

$$T : \mathbb{R}^4 \to \mathbb{R}^3$$
$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + x_2 \\ x_2 - x_3 \\ x_1 + x_3 \end{pmatrix}$$

o en términos matriciales:

$$T(x) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

Determinemos los subespacios $\operatorname{Ker} T$ y $\operatorname{Im} T$. Tenemos que $x \in \operatorname{Ker} T \Leftrightarrow T(x) = 0$,

lo que es equivalente a determinar la solución general del sistema homogéneo:

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Escalonando:

$$\begin{pmatrix}1&1&0&0\\0&1&-1&0\\1&0&1&0\end{pmatrix}\to\begin{pmatrix}1&1&0&0\\0&1&-1&0\\0&-1&1&0\end{pmatrix}\to\begin{pmatrix}1&1&0&0\\0&1&-1&0\\0&0&0&0\end{pmatrix}\tag{4.4}$$

De donde:

$$\begin{array}{rcl}x_1+x_2&=&0\\x_2-x_3&=&0\end{array}\quad\Longleftrightarrow\quad\begin{array}{rcl}x_1&=&-x_2\\x_2&=&x_3\\x_3&=&x_3\\x_4&=&x_4\end{array}$$

Lo cual permite escribir la solución general en función de dos parámetros:

$$\begin{pmatrix}x_1=-x_3\\x_2=x_3\\x_3=x_3\\x_4=x_4\end{pmatrix}\quad\Longleftrightarrow\quad\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=x_3\begin{pmatrix}-1\\1\\1\\0\end{pmatrix}+x_4\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

Luego, $\operatorname{Ker} T=\left\langle\left\{\left(\begin{smallmatrix}-1\\1\\1\\0\end{smallmatrix}\right),\left(\begin{smallmatrix}0\\0\\0\\1\end{smallmatrix}\right)\right\}\right\rangle$ es un s.e.v de $\mathbb{R}^4$ de dimensión dos. Determinemos ahora $\operatorname{Im} T$. Sea $\left(\begin{smallmatrix}y_1\\y_2\\y_3\end{smallmatrix}\right)\in\operatorname{Im} T$, es decir:

$$\begin{aligned}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}&=\begin{pmatrix}1&1&0&0\\0&1&-1&0\\1&0&1&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}\\&=x_1\begin{pmatrix}1\\0\\1\end{pmatrix}+x_2\begin{pmatrix}1\\1\\0\end{pmatrix}+x_3\begin{pmatrix}0\\-1\\1\end{pmatrix}\end{aligned}$$

Luego $\operatorname{Im} T=\left\langle\left\{\left(\begin{smallmatrix}1\\0\\1\end{smallmatrix}\right),\left(\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right),\left(\begin{smallmatrix}0\\-1\\1\end{smallmatrix}\right)\right\}\right\rangle$.

Pero del resultado de escalonar la matriz que tiene como columnas a los vectores $\left(\begin{smallmatrix}1\\0\\1\end{smallmatrix}\right),\left(\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right),\left(\begin{smallmatrix}0\\-1\\1\end{smallmatrix}\right)$ y $\left(\begin{smallmatrix}0\\0\\0\end{smallmatrix}\right)$, en 4.4 podemos extraer una base de $\left\langle\left\{\left(\begin{smallmatrix}1\\0\\1\end{smallmatrix}\right),\left(\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right),\left(\begin{smallmatrix}0\\-1\\1\end{smallmatrix}\right),\left(\begin{smallmatrix}0\\0\\0\end{smallmatrix}\right)\right\}\right\rangle=\left\langle\left\{\left(\begin{smallmatrix}1\\0\\1\end{smallmatrix}\right),\left(\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right),\left(\begin{smallmatrix}0\\-1\\1\end{smallmatrix}\right)\right\}\right\rangle$, tomando aquellos **vectores asociados a columnas con pivotes** de la matriz escalonada (Importante: ver guía de ejercicios y problemas). Es decir, $\left\{\left(\begin{smallmatrix}1\\0\\1\end{smallmatrix}\right),\left(\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right)\right\}$.

Concluímos entonces que $\operatorname{Im} T=\left\langle\left\{\left(\begin{smallmatrix}1\\0\\1\end{smallmatrix}\right),\left(\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right)\right\}\right\rangle$ es un s.e.v. de $\mathbb{R}^3$ de dimensión dos.

Observemos en este ejemplo que $4=\dim\mathbb{R}^4=\dim\operatorname{Ker} T+\dim\operatorname{Im} T$. Además T no es inyectiva ya que $\operatorname{Ker} T$ contiene más de un vector, es decir el vector $\left(\begin{smallmatrix}0\\0\\0\end{smallmatrix}\right)$ tiene más de una pre-imagen. Tampoco es epiyectiva ya que $\dim\operatorname{Im} T<3=\dim\mathbb{R}^3$.

Teorema 4.1. *Sea $T : U \to V$ una transformación lineal entonces*

$$T \text{ es inyectiva } \iff \operatorname{Ker} T = \{0\}$$

Demostración.

$\Longrightarrow$) Dado $u \in \operatorname{Ker} T,\ T(u) = 0 = T(0)$. Como T es inyectiva, se tiene $u = 0$.

$\Longleftarrow$) Sea $T(u_1) = T(u_2)$, por linealidad $T(u_1 - u_2) = 0$. Luego $u_1 - u_2 \in \operatorname{Ker} T = \{0\}$, de donde $u_1 = u_2$. □

De este resultado obtenemos el siguiente Corolario

Corolario 4.1. *Una transformación lineal $T : U \to V$, es un isomorfismo si y sólo si* $\operatorname{Ker} T = \{0\}$ *y* $\operatorname{Im} T = V$*, o equivalentemente* $\dim \operatorname{Im} T = \dim V$ *y* $\dim \operatorname{Ker} T = 0$.

Demostración. Para probar la biyectividad de T, basta notar que $\operatorname{Ker} T = \{0\}$ (inyectividad) y $\operatorname{Im} T = V$ (epiyectividad). La otra equivalencia sale de que $\operatorname{Im} T = V \Leftrightarrow \dim \operatorname{Im} T = \dim V$ y $\operatorname{Ker} T = \{0\} \Leftrightarrow \dim \operatorname{Ker} T = 0$. □

La inyectividad de aplicaciones lineales ($\operatorname{Ker} T = \{0\}$) tiene una consecuencia muy importante sobre los conjuntos linealmente independientes.

Teorema 4.2. *Si $T : U \to V$ es inyectiva, entonces $\{u_i\}_{i=1}^k$ es ℓ.i. en $U \Rightarrow \{T(u_i)\}_{i=1}^k$ es l.i. en V.*

Demostración. Sea la combinación lineal nula:

$$\sum_{i=1}^{n} \lambda_i T(u_i) = 0 \iff T(\sum_{i=1}^{n} \lambda_i u_i) = 0$$
$$\iff \sum_{i=1}^{n} \lambda_i u_i \in \operatorname{Ker} T = \{0\} \iff \sum_{i=1}^{n} \lambda_i u_i = 0$$

Como $\{u_i\}_{i=1}^n$ es $\ell.i.$, $\lambda_i = 0, i = 1, ..., n$. □

Se puede probar que en efcto esta propiedad caracteriza las aplicaciones lineales inyectivas (hacerlo!).

Finalizamos está sección mencionando un ejemplo interesante: $\mathbb{R}^{n+1} \cong P_n(\mathbb{R})$. En efecto, sea
$T : \mathbb{R}^{n+1} \to P_n(\mathbb{R})$ tal que:

$$\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto \sum_{i=0}^{n} a_i x^i \in P_n(\mathbb{R})$$

Es fácil verificar que T es lineal.

Guía de Ejercicios

1. El objetivo de este ejercicio es formalizar la técnica mencionada en la tutoría para extraer, de un conjunto generador de un cierto subespacio vectorial, una base. Esto en el contexto de $\mathbb{R}^n$.

Considere $v_1, \ldots, v_m \in \mathbb{R}^n$ y $U = \langle\{v_1, \ldots, v_m\}\rangle$. Sea además la matriz A, escrita por columnas como $A = (v_1, \ldots, v_m) \in \mathcal{M}_{nm}(\mathbb{R})$. Denotamos por v_{ij} a la coordenada i del vector j y, por ende, a la componente (i, j) de A.

Escalonando la matriz A, se obtiene:

$$\tilde{A} = \left(\begin{array}{cccccccccccccc}
\boxed{\tilde{v}_{11}} & \tilde{v}_{12} & \cdots & & & & & & & & & & & \\
0 & 0 & \cdots & 0 & \boxed{\tilde{v}_{2j_2}} & & & & & & & & & \\
& & & 0 & 0 & \ddots & & & & & & & & \\
& & & & & 0 & \boxed{\tilde{v}_{kj_k}} & \tilde{v}_{k(j_k+1)} & \cdots & & & & & \\
& & & & & 0 & 0 & \cdots & 0 & \boxed{\tilde{v}_{(k+1)j_{k+1}}} & & & & \\
& & & & & & & & 0 & 0 & \ddots & & & \\
& & & & & & & & & & 0 & \boxed{\tilde{v}_{lj_l}} & \cdots & \tilde{v}_{lm} \\
& & & & & & & & & & 0 & 0 & \cdots & 0 \\
& & & & & & & & & & \vdots & & & \vdots
\end{array}\right)$$

Los coeficientes encerrados en un rectángulo $\left(\boxed{\tilde{v}_{kj_k}}\right)$ son los pivotes de la matriz escalonada. Notamos por $\tilde{v}_j$ a la columna j de la matriz $\tilde{A}$.

(a) Demuestre que el conjunto de vectores $\{v_1, v_{j_2}, \ldots, v_{j_l}\}$, es decir las columnas correspondientes a los pivotes, es un conjunto l.i.
Indicación: Estudie qué sucede si en la matriz original A sólo hubiese puesto los vectores $v_1, v_{j_2}, \ldots, v_{j_l}$.

(b) Pruebe que las columnas de $\tilde{A}$ que no contienen pivotes, que están entre la del pivote k y la del $k+1$ (es decir $\tilde{v}_{j_k+1}, \ldots, \tilde{v}_{j_{k+1}-1}$), son combinación lineal de los vectores $\tilde{v}_1, \tilde{v}_{j_2}, \ldots, \tilde{v}_{j_k}$.
Indicación: Use inducción en k.

(c) Dado un conjunto de vectores $\{u_1, \ldots, u_m\} \in \mathbb{R}^n$ y una matriz **invertible** $E \in \mathcal{M}_{nn}(\mathbb{R})$, demuestre que:

$u \in \mathbb{R}^n$ es combinación lineal de $u_1, \ldots, u_m$ $\iff$ $Eu \in \mathbb{R}^n$ es combinación lineal de $Eu_1, \ldots, Eu_m$

(d) Use el resultado anterior para probar que, las columnas de A que no están asociadas a pivotes, entre la columna del pivote k y la del $k+1$ (es decir $v_{j_k+1}, \ldots, v_{j_{k+1}-1}$), son combinación lineal de los vectores $v_1, v_{j_2}, \ldots, v_{j_k}$.

(e) Concluya que $\{v_1, v_{j_2}, \ldots, v_{j_l}\}$ es base de U.

2. Sea un conjunto de vectores $v_1, \ldots, v_m \in \mathbb{R}^n$, con $U = \langle\{v_1, \ldots, v_m\}\rangle$ y $\dim(U) < n$. Muestre que, construyendo la matriz A del Ejercicio 1, "aumentándola" con la identidad en $\mathcal{M}_{nn}(\mathbb{R})$ (es decir $(A|I)$) y aplicando el mismo procedimiento, se consigue:

- Extraer una base de U, desde $\{v_1, \ldots, v_m\}$.
- Extender dicha base a una base de $\mathbb{R}^n$.

3. Sea U el subespacio vetorial de $\mathbb{R}^3$ generado por el conjunto $\left\{ \left(\begin{smallmatrix}1\\0\\1\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\-1\\2\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right) \right\}$. Extraiga una base de U y extiéndala a una base de $\mathbb{R}^3$.

Guía de Problemas

P1. Sea $\mathcal{P}_3(\mathbb{R})$ el conjunto de los polinomios de grado menor o igual a 3 con coeficientes reales. Se define la transformación $T\colon \mathcal{P}_3(\mathbb{R}) \to \mathcal{P}_3(\mathbb{R})$ por:

$$T(a_0 + a_1x + a_2x^2a_3x^3) = a_0 + (2a_1 + a_2)x + (2a_2 + a_3)x^2 + 2a_3x^3.$$

(a) Probar que T es una transformación lineal.

(b) Probar que T es una transformación biyectiva.

(c) Si id es la transformación identidad del espacio vectorial $\mathcal{P}_3(\mathbb{R})$ pruebe que $T - 2\text{id}$, $(T - 2\text{id})^2$, $(T - 2\text{id})^3$ y $T - \text{id}$ son transformaciones lineales.
Indicación: Recuerde que dada una función f, $f^k = \underbrace{f \circ f \circ \cdots \circ f}_{\text{k veces}}$.

(d) Encontrar bases y dimensión de $\text{Ker}(T - 2\text{id})$, $\text{Ker}(T - 2\text{id})^2$ y $\text{Ker}(T - 2\text{id})^3$.

(e) Probar que $\mathcal{P}_3(\mathbb{R}) = \text{Ker}(T - 2\text{id})^2 \oplus \text{Ker}(T - \text{id})$.

P2. **(a)** Considere la función $T : \mathbb{R}^4 \to \mathcal{M}_{22}(\mathbb{R})$ definida por

$$T\left(\begin{smallmatrix}x\\y\\z\\w\end{smallmatrix}\right) = \begin{pmatrix} x & y + w \\ 2x + 2y + z + w & x + y + z \end{pmatrix}.$$

(1) Demuestre que T es función lineal.

(2) Determine bases y dimensiones para $Ker(T)$ e $Im(T)$.

(b) Dada la transformación lineal $L : \mathbb{R}^3 \to \mathbb{R}^2$ se sabe que:

$$Ker(L) = \left\langle \left\{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix} \right\} \right\rangle \quad \text{y} \quad L\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}1\\-1\end{pmatrix}.$$

Encuentre en términos de x_1, x_2 y x_3, los valores de y_1 y y_2 que satisfacen $L\left(\begin{smallmatrix}x_1\\x_2\\x_3\end{smallmatrix}\right) = \left(\begin{smallmatrix}y_1\\y_2\end{smallmatrix}\right)$.

P3. Sean V, W e. v.'s sobre $\mathbb{R}$. En $V \times W$ se definen la suma y ponderación por escalar así:

$$\begin{aligned} (v, w) + (v', w') &= (v + v', w + w')\,, \quad \forall (v, w), (v', w') \in V \times W\,, \\ \lambda(v, w) &= (\lambda v, \lambda w)\,, \quad \forall (v, w) \in V \times W\,, \forall \lambda \in \mathbb{R}\,. \end{aligned}$$

Con estas operaciones $V \times W$ es un espacio vectorial sobre $\mathbb{R}$ (no lo pruebe).

Dada una función $f : V \to W$ se define su gráfico por

$$G_f = \{(v, w) \in V \times W : w = f(v)\}\,.$$

(a) Pruebe que f es lineal sí y sólo si G_f es subespacio vectorial de $V \times W$.

(b) Pruebe que $\{0_V\} \times W$ es sub–espacio vectorial de $V \times W$.
Nota: 0_V denota el neutro aditivo de V.

(c) Sea $f : V \to W$ lineal. Pruebe que $G_f \oplus (\{0_V\} \times W) = V \times W$.

Ingeniería Matemática
FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Transformaciones lineales

4.6 Teorema del Núcleo-Imagen (TNI)

El propósito de está sección es probar el siguiente teorema.

Teorema 4.3 (Núcleo-Imagen). *Sean U, V espacios vectoriales sobre un cuerpo $\mathbb{K}$, $T : U \to V$ una transformación lineal, $\dim U < \infty$. Entonces:*

$$\dim U = \dim \operatorname{Ker} T + \dim \operatorname{Im} T.$$

Demostración. Daremos una demostración constructiva de este teorema. Sea $\{u_1, ..., u_\nu\}$ una base de $\operatorname{Ker} T$. Completemos esta base a una base de U

$$\beta_U = \{u_1, ..., u_\nu, u_{\nu+1}, ..., u_n\}.$$

Probaremos que $\beta' = \{T(u_{\nu+1}), ..., T(u_n)\}$ es base de $\operatorname{Im} T$. Sea $v \in \operatorname{Im} T$ es decir $v \in V$ y existe $u \in U$ $T(u) = v$. Como β_U es base de U se tendrá que

$$u = \alpha_1 u_1 + ... + \alpha_\nu u_\nu + \alpha_{\nu+1} u_{\nu+1} + ... + \alpha_n u_n,$$

de donde

$$v = T(u) = \alpha_{\nu+1} T(u_{\nu+1}) + ... + \alpha_n T(u_n),$$

pues $T(u_1) = ... = T(u_\nu) = 0$. Es decir v es combinación de los elementos de β'. Probemos ahora que son l.i., para ello consideremos

$$\alpha_{\nu+1} T(u_{\nu+1}) + ... + \alpha_n T(u_n) = 0.$$

Como $T(\alpha_{\nu+1} u_{\nu+1} + ... + \alpha_n u_n) = \alpha_{\nu+1} T(u_{\nu+1}) + ... + \alpha_n u_n T(u_n) = 0$ se deduce que el vector $\alpha_{\nu+1} u_{\nu+1} + ... + \alpha_n \in \operatorname{Ker} T$ pero entonces existen escalares $\lambda_1, ..., \lambda_\nu$ tal que

$$\alpha_{\nu+1} u_{\nu+1} + ... + \alpha_n u_n = \lambda_1 u_1 + ... + \lambda_\nu u_\nu,$$

o equivalentemente

$$\lambda_1 u_1 + ... + \lambda_\nu u_\nu - \alpha_{\nu+1} u_{\nu+1} - ... - \alpha_n u_n = 0$$

Como los vectores $u_1, ..., u_n$ son $\ell.i.$ se deduce en particular que

$$\alpha_{\nu+1} = ... = \alpha_n = 0,$$

y por lo tanto $\beta' = \{T(u_{\nu+1}), ..., T(u_n)\}$ es base de $\operatorname{Im} T$. Pero entonces tenemos

$$\dim U = n = \nu + (n - \nu) = \dim \operatorname{Ker} T + \dim \operatorname{Im} T.$$

$\square$

Es directo del TNI que: si $\dim U = n$, y el rango de T es n se tiene $\dim \operatorname{Ker} T = 0$, es decir T es inyectiva. Otra aplicación directa es

Teorema 4.4. *Sea $T : U \to V$ aplicación lineal.*

1. *Si $\dim U = \dim V$ entonces*

$$T \text{ inyectiva} \iff T \text{ epiyectiva} \iff T \text{ biyectiva.}$$

2. *Si $\dim U > \dim V$ T no puede ser inyectiva.*
3. *Si $\dim U < \dim V$ T no puede ser epiyectiva.*
4. *Como conclusión de (2) y (3)*

$$U \text{ isomorfo a } V \iff \dim U = \dim V.$$

Demostración. Probemos sólo (2), el resto queda de ejercicio. Del TNI se tiene que

$$\dim U = \dim \operatorname{Ker} T + \dim \operatorname{Im} T < \dim V,$$

como $\dim \operatorname{Ker} T \geq 0$ se concluye que

$$\dim \operatorname{Im} T < \dim V,$$

y por lo tanto $\operatorname{Im} T$ es un s.e.v. estrictamente más pequeño que V, es decir T no puede ser epiyectiva.

Es importante señalar que en (1c), la implicancia $\Rightarrow$ es consecuencia de (2) y (3), mientras que $\Leftarrow$ se obtiene del hecho de que todo espacio de dimensión finita n es isomorfo a $\mathbb{K}^n$ (ver ejemplo anterior a la Definición 4.2). $\square$

Ejercicio 4.2: Probar que si $T : U \to V$ es lineal y W es un subespacio de U entonces

1. $T(W)$ es subespacio de V.
2. $\dim T(W) \leq \dim U$.
3. Si T es inyectiva $\dim T(W) = \dim W$.

4.7 Matriz Representante de una Transformación Lineal

Vimos anteriormente que, dado un e.v V sobre $\mathbb{K}, dimV = n$, se tenía $V \cong \mathbb{K}^n$. Esto nos permite "llevar" el trabajo de cualquier espacio de dimensión finita a un espacio de n-tuplas. Trataremos de extender esto para transformaciones lineales, reemplazando estas (en un sentido que precisaremos) por matrices.

Sean U, V espacios vectoriales sobre un cuerpo $\mathbb{K}, \dim U = p, \dim V = q$. Sean $\beta_U = \{u_1, ..., u_p\}, \beta_V = \{v_1, ..., v_q\}$ bases de U y V respectivamente. Consideremos finalmente una transformación lineal $T : U \to V$.

Sabemos que T queda completamente determinada por su acción sobre la base β_U

$$T(u_j) = \sum_{i=1}^{q} a_{ij} v_i \qquad 1 \leq j \leq p$$

o bien,

$$\begin{aligned} T(u_1) &= a_{11}v_1 + a_{21}v_2 + ... + a_{q1}v_q \\ T(u_2) &= a_{12}v_1 + a_{22}v_2 + ... + a_{q2}v_q \\ &\vdots \\ T(u_p) &= a_{1p}v_1 + a_{2p}v_2 + ... + a_{qp}v_q \end{aligned}$$

Es claro que la información importante que define T está contenida en las coordenadas, $(a_{1j}, a_{2j}, ..., a_{qj})$ de cada vector $T(u_j)$, con respecto a la base β_V, en el espacio de llegada. Esta información la guardamos en una matriz de q filas y p columnas, denominada **matriz representante de T, con respecto a las bases** β_U y β_V:

$$M_{\beta_U \beta_V}(T) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{q1} & a_{q2} & \dots & a_{qp} \end{pmatrix} \in M_{qp}(\mathbb{K})$$

donde, $q = \dim U,\ p = \dim V$.

En la columna j-ésima de esta matriz, aparecen las coordenadas del vector $T(u_j)$ con respecto a la base β_V:

$$A_{\bullet j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{qj} \end{pmatrix} \implies T(u_j) = \sum_{i=1}^{q} a_{ij} v_i$$

Observación: Una pregunta interesante es: ¿Qué sucede con esta matriz si cambiamos las bases?. Esto lo responderemos en breve.

Ejemplo 4.2.
Sea $T : \mathbb{R}^4 \to \mathbb{R}^3$, definida por

$$T(x) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

Consideremos $\beta = \beta_U = \{e_1, e_2, e_3, e_4\}$, base canónica de $\mathbb{R}^4$,

$$\beta' = \beta_V = \left\{ \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right) \right\}$$

base de $\mathbb{R}^3$. Se tiene entonces:

$$T(e_1) = T \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) = \frac{1}{2} \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right) + \frac{1}{2} \left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right) - \frac{1}{2} \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right)$$

$$T(e_2) = T \left(\begin{smallmatrix} 0 \\ 1 \\ 0 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right) = 1 \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right) + 0 \left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right) + 0 \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right)$$

$$T(e_3) = T \left(\begin{smallmatrix} 0 \\ 0 \\ 1 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right) = 0 \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right) + 0 \left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right) + 1 \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right)$$

$$T(e_4) = T \left(\begin{smallmatrix} 0 \\ 0 \\ 0 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right) = -\frac{1}{2} \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right) + \frac{1}{2} \left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right) + \frac{1}{2} \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right)$$

de donde:

$$M_{\beta\beta'}(T) = \begin{pmatrix} \frac{1}{2} & 1 & 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & 1 & \frac{1}{2} \end{pmatrix} \in M_{34}(\mathbb{R})$$

Tomemos ahora, en $\mathbb{R}^3$, la base canónica, $\beta' = \left\{ \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right) \right\}$:

$$T(e_1) = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) = 1e_1' + 0e_2' + 0e_3'$$

$$T(e_2) = \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right) = 1e_1' + 1e_2' + 0e_3'$$

$$T(e_3) = \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right) = 0e_1' + 1e_2' + 1e_3'$$

$$T(e_4) = \left(\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right) = 0e_1' + 0e_2' + 1e_3'$$

Luego,

$$M_{\beta\beta'}(T) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

¡Sorpresa!, esta matriz representante no sólo es distinta de la anterior, sino que corresponde a la matriz que define T en forma natural. Concluímos entonces que: la matriz representante *depende* de las bases elegidas y que la matriz representante con respecto a las bases canónicas es la matriz asociada a T. Esto último lo podemos verificar en el caso general:

$$\text{Consideremos} \qquad \begin{aligned} T_A : \mathbb{K}^p &\to \mathbb{K}^q \\ x &\mapsto Ax,\ A \in M_{qp}(\mathbb{K}) \end{aligned}$$

Sea $\beta = \{e_1, ..., e_p\}$ la base canónica de $\mathbb{K}^p$ y $\beta' = \{e'_1, ..., e'_q\}$ la base canónica de $\mathbb{K}^q$.

Se tiene entonces:

$$T_A(e_j) = Ae_j = \begin{pmatrix} a_{11}\ldots & a_{1j}\ldots & a_{1p} \\ \vdots & \vdots & \vdots \\ a_{q1}\ldots & a_{qj}\ldots & a_{qp} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{qj} \end{pmatrix}$$

o bien:

$$T_A(e_j) = a_{1j}e'_1 + ... + a_{qj}e'_m = \sum_{i=1}^{q} a_{ij}e'_i$$

de donde concluímos $M_{\beta\beta'}(T_A) = A$.

Veamos ahora como interactuan T y $M = M_{\beta\beta'}(T)$. Sea $u \in U$, entonces

$$u = \alpha_1 u_1 + ... + \alpha_p u_p,$$

y por lo tanto $T(u) = \sum_{j=1}^{p} \alpha_j T(u_j)$ pero $T(u_j) = \sum_{i=1}^{q} m_{ij} v_i$ y por lo tanto

$$T(u) = \sum_{j=1}^{p} \alpha_j \sum_{i=1}^{q} m_{ij} v_i = \sum_{i=1}^{q} (\sum_{j=1}^{p} \alpha_j m_{ij}) v_i.$$

Así las coordenadas de $T(u)$ con respecto a β' son

$$\sum_{j=1}^{p} \alpha_j m_{1j},\ \ldots,\ \sum_{j=1}^{p} \alpha_j m_{qj},$$

y por lo tanto si $\alpha = (\alpha_1, ..., \alpha_p)$ son las coordenadas de u entonces $M\alpha$ son las de $T(u)$. Esto permite trabajar con M en vez de T.

Ya tenemos una matriz que representa T, ahora deseamos “olvidarnos” de T y trabajar sólo con un representante matricial. Para hacer esto, en algebra acostumbramos establecer una isomorfía, en este caso, entre el espacio vectorial de todas estas matrices, $M_{qp}(\mathbb{K})$, y aquél de todas las transformaciones lineales. Para ello definamos primero este último espacio.

Definición Sea U, V espacios vectoriales sobre $\mathbb{K}$, de dimensiones $\dim U = p$ y $\dim V = q$ respectivamente.
Definimos

$$\mathcal{L}_{\mathbb{K}}(U, V) = \{T : U \to V \mid T \text{ es transformación lineal } \}$$

Es directo que $\mathcal{L}_{\mathbb{K}}(U, V)$, dotado de las operaciones:

$$(\lambda T)(x) = \lambda T(x), \ \forall \lambda \in \mathbb{K}, \forall T \in \mathcal{L}_{\mathbb{K}}(U, V)$$
$$(T + T')(x) = T(x) + T'(x), \ \forall T, T' \in \mathcal{L}_{\mathbb{K}}(U, V)$$

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

Tenemos entonces la propiedad buscada:

$$\mathcal{L}_{\mathbb{K}}(U, V) \cong M_{qp}(\mathbb{K})$$

En efecto, sean β_U, β_V bases de U y V respectivamente y definamos la función.

$$\begin{aligned} \varphi : \mathcal{L}_{\mathbb{K}}(U, V) &\to M_{qp}(\mathbb{K}) \\ T &\to \varphi(T) = M_{\beta_U \beta_V}(T) \end{aligned}$$

Es decir, a una transformación lineal arbitraria le asociamos su matriz representante con respecto a las bases β_U y β_V.

1. φ es una transformación lineal. Sean $\beta_U = \{u_1, ..., u_p\}, \beta_V = \{v_1, ..., v_q\}$, $T, L \in \mathcal{L}_K(U, V)$, $A = M_{\beta_U \beta_V}(T), B = M_{\beta_U \beta_V}(L)$.

   Calculemos $\varphi(T + L) = M_{\beta_U \beta_V}(T + L)$.

   $$\begin{aligned} (T + L)(u_j) &= T(u_j) + L(u_j) = \\ &= \sum_{i=1}^{q} a_{ij} v_i + \sum_{i=1}^{q} b_{ij} v_i = \sum_{i=1}^{q} (a_{ij} + b_{ij}) v_i \end{aligned}$$

   Luego, la j-ésima columna de $M_{\beta_U \beta_V}$ es el vector:

   $$\begin{pmatrix} a_{1j} + b_{1j} \\ \vdots \\ a_{qj} + b_{qj} \end{pmatrix} = A_{\bullet j} + B_{\bullet j}$$

Se concluye entonces:

$$\varphi(T+L) = M_{\beta_U\beta_V}(T+L) = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1j}+b_{1j} & \cdots & a_{1p}+b_{1p} \\ \vdots & & \vdots & & \vdots \\ a_{q1}+b_{q1} & \cdots & a_{qj}+b_{qj} & \cdots & a_{qp}+b_{qp} \end{pmatrix}$$

$$= A + B = M_{\beta_U\beta_V}(T) + M_{\beta_U\beta_V}(L) = \varphi(T) + \varphi(L)$$

De manera análoga se prueba que

$$\varphi(\lambda T) = M_{\beta_U\beta_V}(\lambda T).$$

2. φ es biyección. Sean $T \in \mathrm{Ker}(\varphi) \iff \varphi(T) = M_{\beta_U\beta_V}(T) = 0 \in M_{qp}(\mathbb{K})$

$$\iff T(u_j) = 0v_1 + ... + 0v_q = 0.$$

Por otra parte, $\forall u \in U, u = \sum\limits_{j=1}^{p} \alpha_j u_j$, de donde

$$\forall u \in U,\ T(u) = \sum_{j=1}^{p} \alpha_j T(u_j) = \sum_{j=1}^{p} \alpha_j 0 = 0,$$

concluyendo $T \equiv 0$, la transformación nula. Luego, φ es inyectiva.

Veamos la epiyectividad. Dada una matriz $A = (a_{ij}) \in M_{qp}(\mathbb{K})$, le asociamos la transformación lineal T, que a la base de U asocia: $T(u_j) = \sum\limits_{i=1}^{q} a_{ij}v_i$. Es directo que $\varphi(T) = M_{qp}(T) = A$, luego φ es epiyectiva. De (1) y (2) concluímos que φ es un isomorfismo.

Directamente, de este resultado se tiene: $\dim \mathcal{L}_{\mathbb{K}}(U,V) = \dim M_{qp}(\mathbb{K}) = pq = \dim U \dim V$.

Es decir, la dimensión del espacio de transformaciones lineales es *finita* y corresponde

al producto de las dimensiones de U y V.

En el contexto del resultado anterior ¿que relación existe entre la multiplicación de matrices y la composición de funciones lineales?.

Veamos que sucede al componer dos transformaciones lineales. Sean $T : U \to V$, y $L : V \to W$ transformaciones lineales, luego $L \circ T : U \to W$ es también lineal. Ahora bien,

sean $\dim U = p, \dim V = q, \dim W = r$ con bases $\beta_U = \{u_i\}_{i=1}^{p}, \beta_V = \{v_i\}_{i=1}^{q}, \beta_W = \{w_i\}_{i=1}^{r}$, respectivamente. Supongamos además que $M_{\beta_U\beta_V}(T) = B \in M_{qp}(\mathbb{K})$, $M_{\beta_U\beta_V}(L) = A \in M_{pr}(\mathbb{K})$. La pregunta es ¿cuál es la expresión de $M_{\beta_U\beta_W}(L \circ T)$? Los “datos” del problema son:

$$T(u_k) = \sum_{j=1}^{q} b_{jk} v_j \qquad 1 \le k \le p$$

$$L(v_j) = \sum_{i=1}^{r} a_{ij} w_i \qquad 1 \le j \le q$$

Calculemos ahora:

$$\begin{aligned}(L \circ T)(u_k) = L(T(u_k)) &= L(\sum_{j=1}^{q} b_{jk} v_j) = \\ &= \sum_{j=1}^{q} b_{jk} L(v_j) = \sum_{j=1}^{q} b_{jk} (\sum_{i=1}^{r} a_{ij} w_i) \\ &= \sum_{i=1}^{r} (\sum_{j=1}^{q} a_{ij} b_{jk}) w_i \qquad 1 \le k \le p\end{aligned}$$

obteniendo

$$M_{\beta_U \beta_W}(L \circ T) = \begin{pmatrix} \sum\limits_{j=1}^{q} a_{1j} b_{j1} & \cdots & \sum\limits_{j=1}^{q} a_{1j} b_{jp} \\ & \vdots & \\ \sum\limits_{j=1}^{q} a_{rj} b_{j1} & \cdots & \sum\limits_{j=1}^{q} a_{rj} b_{jp} \end{pmatrix}$$

$$= A \cdot B = M_{\beta_V \beta_W}(L) \cdot M_{\beta_U \beta_V}(T)$$

Hemos probado que la matriz representante de la composición de dos funciones lineales $L \circ T$, es el producto de las matrices representantes.

Una aplicación inmediata de este resultado es

Ejercicio 4.3: $T\colon U \to V$ es una transformación lineal y $\beta,\ \beta'$ son bases de U y V respectivamente entonces

1. T es invertible ssi $M_{\beta\beta'}(T)$ es una matriz invertible.
2. Si T es invertible,
$$(M_{\beta\beta'}(T))^{-1} = M_{\beta'\beta}(T^{-1}).$$

Esto nos permite, cuando se estudian transformaciones lineales sobre espacios vectoriales de dimensión finita, trabajar directamente en la estructura matricial correspondiente, sumando o multiplicando matrices.

En base a lo anterior se tiene el resultado siguiente:

Teorema 4.5. *Para toda matriz $A \in M_{nn}(\mathbb{K})$, las propiedades siguientes son equivalentes*

1. *A es invertible.*

2. *$T_A : \mathbb{K}^n \to \mathbb{K}^n, v \to T_A(v) = Av$ es un isomorfismo.*

3. *El conjunto $\{A_{\bullet j}\}_{j=1}^n$ es base de $\mathbb{K}^n$.*

Demostración. Demostremos primero (2)⇔(3). Para ello recordemos que A es la matriz representante de T_A con respecto a las bases canónicas: $T_A(e_j) = Ae_j = A\bullet j$.

$2 \Rightarrow 3$) Como T_A es isomorfismo, es en particular inyectiva, luego $\{T_A(e_j)\}_{j=1}^n$ es $\ell.i \Leftrightarrow \{A_{\bullet j}\}_{j=1}^n$ es $\ell.i$ en $\mathbb{K}^n$. Como $\dim \mathbb{K}^n = n$, este conjunto es base.

$3 \Rightarrow 2$) Sabemos $A_{\bullet j} = T_A(e_j)$ para $1 \leq j \leq n$, luego $\{A_{\bullet j}\}_{j=1}^n$ base, es decir $\{T_A(e_j)\}_{j=1}^n$ es base de $\mathbb{K}^n$, luego $\operatorname{Im} T_A = \langle \{T_A(e_j)\} \rangle = \mathbb{K}^n$. Del TNI se tiene:

$$\dim \mathbb{K}^n = n = \dim \operatorname{Im} T_A + dim \operatorname{Ker} T_A, \text{ luego } n = n + \dim \operatorname{Ker} T_A,$$

entonces $\dim \operatorname{Ker} T_A = 0$ y por lo tanto $\operatorname{Ker} T_A = \{0\}$.

Además las dimensiones de los espacios de partida y llegada son iguales, luego T_A es isomorfirmo.

(1)⇔(2) es consecuencia del Ejercicio 4.3, usando el hecho de que T_A tiene a A como matriz representante con respecto a las bases canónicas. □

4.8 Matriz de Pasaje o de Cambio de Base

Cuando definimos la matriz representante vimos que esta dependía de las bases elegidas. De esto surge el problema de la no unicidad de la representación: una misma transformación lineal tiene asociada más de una matriz, dependiendo de las bases escogidas.

Sea entonces V un espacio vectorial sobre $K, \dim V = n$, sean $\beta = \{v_i\}_{i=1}$ y $\beta' = \{v'_i\}_{i=1}^n$ dos bases de V. Claramente, los vectores de β' tienen una representación única en la base β:

$$\begin{aligned}
v'_1 &= p_{11}v_1 + p_{21}v_2 + ... + p_{n1}v_n \\
\vdots & \qquad \vdots \qquad\qquad\quad \vdots \\
v'_j &= p_{1j}v_1 + p_{2j}v_2 + ... + p_{nj}v_n \\
\vdots & \qquad \vdots \qquad\qquad\quad \vdots \\
v'_n &= p_{1n}v_1 + p_{2n}v_2 + ... + p_{nn}v_n
\end{aligned}$$

Denominamos *matriz de pasaje o de cambio de base de β a β'* a la matriz:

$$P = P_{\beta\beta'} = (p_{ij}) = \begin{pmatrix} p_{11} & \cdots & p_{1j} & \cdots & p_{1n} \\ \vdots & & \vdots & & \vdots \\ p_{n1} & \cdots & p_{nj} & \cdots & p_{nn} \end{pmatrix}$$

cuya columna j-ésima corresponde a las coordenadas del vector v'_j, de la "nueva" base β', con respecto a la "antigua" base β. Es fácil ver que está matriz es exactamente

$$M_{\beta'\beta}(id_V),$$

es decir la matriz representante de la identidad de V, colocando como base de partida β' y de llegada β.

Por otra parte, sabemos que a un vector $v \in V$ podemos asociarle de manera biunivoca un vector $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{K}^n$, correspondiente a sus coordenadas con respecto a la base β. De manera análoga le podemos asociar $\alpha' = (\alpha'_1, ..., \alpha'_n) \in \mathbb{K}^n$, sus coordenadas con respecto a la base β'.

Se tiene entonces:

$$v = \sum_{j=1}^{n} \alpha'_j v'_j = \sum_{j=1}^{n} \alpha'_j \left(\sum_{i=1}^{n} p_{ij} v_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{ij} \alpha'_j \right) v_i$$

Pero, además, $v = \sum\limits_{i=1}^{n} \alpha_i v_i$.

Como la representación es única, concluímos:

$$\alpha_i = \sum_{j=1}^{n} p_{ij} \alpha'_j \qquad 1 \leq i \leq n$$

matricialmente: $\alpha = P\alpha'$.

Esta ecuación establece la relación entre las coordenadas de un mismo vector $v \in V$ con respecto a ambas bases. ¿Es posible "despejar" α'?. Sí lo es, y esto es equivalente a probar que P es invertible. Pero por ser la matriz representante de un isomorfísmo es entonces invertible. Además P^{-1} es la matriz de cambio de base de β' a β: $M_{\beta\beta'}(id_V)$.

Observación: Cabe hacer notar que la notación no es muy afortunada, pues la matriz de cambio de base de β a β' es $M_{\beta'\beta}(id_V)$ es decir hay que expresar los vectores de β' en función de los de β.

Estudiaremos ahora la relación que existe entre las distintas matrices representantes de una misma transformación lineal. Esto lo deduciremos de la regla que tenemos para la matriz representante de una composición de aplicaciones lineales. Sea $T : U \to V$ lineal y consideremos cuatro bases:

$$\beta,\ \tilde{\beta} \text{ bases de } U$$
$$\beta',\ \tilde{\beta}' \text{ bases de } V.$$

La pregunta es ¿cómo se relacionan $M_{\beta\beta'}(T)$ y $M_{\tilde{\beta}\tilde{\beta}'}(T)$?.

Todo sale de la observación

$$id_V \circ T \circ id_U = T,$$

y por lo tanto

$$M_{\tilde{\beta}\tilde{\beta}'}(T) = M_{\tilde{\beta}\tilde{\beta}'}(id_V \circ T \circ id_U) = M_{\beta'\tilde{\beta}'}(id_V) M_{\beta\beta'}(T) M_{\tilde{\beta}\beta}(id_U). \tag{4.5}$$

Es decir la nueva matriz representante de T es la antigua multiplicada por matrices de cambio de base. La mejor forma de recordar esto es con un diagrama Recorremos este

$$\begin{tikzcd}
U, \overline{\beta} \arrow[r, "T"] \arrow[d, "id_U"'] & V, \overline{\beta}' \\
U, \beta \arrow[r, "T"] & V, \beta' \arrow[u, "id_V"']
\end{tikzcd}$$

diagrama de dos maneras distintas. para obtener la fórmula 4.5. Notemos que cuando hacemos el recorrido por “abajo” el orden que aparecen las matrices es reverso: la matriz más a la derecha en 4.5 es la que actua primero y así sucesivamente.

Las matrices $A = M_{\beta\beta'}(T)$ y $B = M_{\tilde{\beta}\tilde{\beta}'}(T)$ están relacionadas por un par de matrices invertibles que denotaremos por P, Q:

$$C = PAQ.$$

Diremos que entonces A y C son **semejantes**. Como ejercicio, probar que esta es una relación de equivalencia. Un caso importante es cuando $Q = P^{-1}$, es decir

$$C = PAP^{-1}.$$

En este caso especial diremos que A y C son **similares**. Esta también es una relación de equivalencia.

Se tiene que a diferencia de la semejanza no es fácil saber cuándo dos matrices son similares. Parte de este problema la estudiaremos en el próxima sección.

Guía de Ejercicios

1. Sea $T\colon U \to V$ una transformación lineal. Pruebe que

(a) Si $\dim U = \dim V$ entonces

$$T \text{ inyectiva} \iff T \text{ epiyectiva} \iff T \text{ biyectiva.}$$

(b) Si $\dim U < \dim V$ T no puede ser epiyectiva.

(c) U isomorfo a $V \iff \dim U = \dim V$.

2. Probar que si $T\colon U \to V$ es lineal y W es un subespacio de U entonces

(a) $T(W)$ es subespacio de V.

(b) $\dim T(W) \leq \dim U$.

(c) Si T es inyectiva $\dim T(W) = \dim W$.

3. $T\colon U \to V$ es una transformación lineal y $\beta,\ \beta'$ son bases de U y V respectivamente entonces

(a) T es invertible ssi $M_{\beta\beta'}(T)$ es una matriz invertible.

(b) Si T es invertible,

$$(M_{\beta\beta'}(T))^{-1} = M_{\beta'\beta}(T^{-1}).$$

Guía de Problemas

P1. Sean $\mathcal{M}_{22}(\mathbb{R})$ el espacio vectorial de las matrices de 2×2 a coeficientes reales sobre el cuerpo $\mathbb{R}$. Sea $\mathcal{P}_2(\mathbb{R})$ el espacio vectorial de los polinomios de grado menor o igual a 2 a coeficientes reales sobre el cuerpo $\mathbb{R}$. Sea

$$\begin{aligned} T : \mathcal{M}_{22}(\mathbb{R}) &\longrightarrow \mathcal{P}_2(\mathbb{R}) \\ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) &\longmapsto T(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)) = (a+d) + (b+c)x + (a+b+2c+d)x^2. \end{aligned}$$

(a) Demuestre que T es una transformación lineal.

(b) Sean la bases de $\mathcal{M}_{22}(\mathbb{R})$ y $\mathcal{P}_2(\mathbb{R})$ siguientes:

$$\beta_M = \{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)\}, \quad \beta_P = \{1, x, x^2\}.$$

Encuentre la matriz representante de T con respecto a las bases β_M y β_P.

(c) Sean la bases de $\mathcal{M}_{22}(\mathbb{R})$ y $\mathcal{P}_2(\mathbb{R})$ siguientes:

$$\beta'_M = \{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)\}, \quad \beta'_P = \{1+x, x+x^2, 1+x^2\}.$$

Encuentre la matriz representante de T con respecto a las bases β'_M y β'_P.

(d) Calule la dimensión del Núcleo y la dimensión de la Imagen de T.

P2. Considere las funciones lineales f y g tales que

$$f\colon \mathcal{M}_{22}(\mathbb{R}) \to \mathbb{R}^3 \quad \text{y} \quad \begin{aligned} g : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto g(x) = \left(\begin{smallmatrix} x_1+x_2-x_3 \\ 2x_1+x_3 \end{smallmatrix} \right). \end{aligned}$$

y las bases

$$B_1 = \left\{ \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 3 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 2 & 5 \\ 2 & -4 \end{pmatrix} \right\} \text{ de } \mathcal{M}_{22}(\mathbb{R})$$

y

$$B_2 = \left\{ \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\} \text{ de } \mathbb{R}^3$$

Sea $A = \left(\begin{smallmatrix} 1 & 1 & 1 & 1 \\ -2 & 0 & 0 & 1 \\ 2 & 1 & 1 & -1 \end{smallmatrix} \right)$ la matriz representante de la función f con respecto a las bases B_1 en $\mathcal{M}_{22}(\mathbb{R})$ y B_2 en $\mathbb{R}^3$.

(a) Encuentre la matriz representante de B de la función g con respecto a las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$ respectivamente.

(b) Obtenga la matriz representante de $g \circ f$ con respecto a la base B_1 en $\mathcal{M}_{22}(\mathbb{R})$ y a la base canónica de $\mathbb{R}^2$.

P3. Sea $T\colon V \to V$ una aplicación lineal con V espacio vectorial de dimensión finita. Demuestre que

$$V = Ker(T) \oplus Im(T) \iff Ker(T^2) = Ker(T).$$

P4. Sea $\beta = \{1, x, x^2\}$ la base canónica del espacio $\mathcal{P}_2(\mathbb{R})$. Considere

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

(a) Encuentre la base β' de $\mathcal{P}_2(\mathbb{R})$ tal que Q sea representante de la identidad de $\mathcal{P}_2(\mathbb{R})$ con β' en $\mathcal{P}_2(\mathbb{R})$con β. Si le sirve, si denotamos por $[p]_\beta$ el vector de coordenadas de p con respecto a la base β,

$$[p]_\beta = Q[p]_{\beta'} \quad \forall p \in \mathcal{P}_2(\mathbb{R}).$$

(b) Sea $T\colon \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}^3$ la transformación lineal cuya matriz representante con respecto a las bases β en $\mathcal{P}_2(\mathbb{R})$ y canónica en $\mathbb{R}^3$ es la matriz

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

Calcule la matriz representante de T con respecto a las bases β' en $\mathcal{P}_2(\mathbb{R})$ y canónica en $\mathbb{R}^3$, donde β' es la base encontrada en (a).

SEMANA 10:

Ingeniería Matemática
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Transformaciones lineales
Valores y Vectores Propios

5.1 Rango de una matriz y forma de Hermitte

Dada una matriz $A \in \mathcal{M}_{qp}(\mathbb{K})$ definimos el **rango** de A, como el rango de la transformación lineal $T(x) = Ax$. Denotaremos por $r(A)$ el rango de A. Así $r(A)$ es la dimensión de la imagen de A.

Como $Ax = x_1 A_{\bullet 1} + ... + x_p A_{\bullet p}$ si $x = (x_1, ..., x_p)$, entonces la imagen de A es generada por las columnas de A esto es:

$$\operatorname{Im} A = \langle \{A_{\bullet 1}, ..., A_{\bullet p}\} \rangle ,$$

por lo tanto el rango de una matriz es el número máximo de columnas l.i.

Ejercicio 5.1: Supongamos que A, B son semejantes es decir existen matrices invertibles P, Q tal que $A = PBQ$. Probar que entonces que A y B tienen el mismo rango:

$$r(A) = r(B).$$

Nuestro propósito es probar la recíproca de lo anterior, esto es si dos matrices de igual dimensión tienen el mismo rango entonces son semejantes. Para ello necesitaremos lo que se llama la **forma normal de Hermitte**.

Consideremos $A \in \mathcal{M}_{qp}(\mathbb{K})$, ella induce de forma natural una aplicación lineal T mediante: $T(x) = Ax$. Es claro que $A = M_{\beta\beta'}(T)$ donde β, β' son las bases canónicas de $\mathbb{K}^p$ y $\mathbb{K}^q$ respectivamente. Construiremos ahora dos nuevas bases. Comencemos por una base de $\operatorname{Ker} A$: $\{u_1, ..., u_\nu\}$, y extendámosla a una base de $\mathbb{K}^p$

$$\tilde{\beta} = \{w_1, ..., w_r, u_1, ..., u_\nu\}.$$

Notemos dos cosas: primero que hemos dado un orden especial a esta base donde los vectores que agregamos van al principio; y segundo que del teorema del núcleo imagen $p = \dim \mathbb{K}^p = \dim \operatorname{Im} A + \dim \operatorname{Ker} A$ y por lo tanto necesitamos agregar tanto vectores como rango de A es decir $r = r(A)$. Tomemos ahora $X = \{v_1, \ldots, v_r\}$ donde $v_i = Aw_i,\ i = 1, \ldots, r$. Probaremos que X es un conjunto l.i.. En efecto si

$$\lambda_1 v_1 + \ldots + \lambda_r v_r = 0,$$

se tendrá que

$$0 = \lambda_1 Aw_1 + \ldots + \lambda_r Aw_r = A(\lambda_1 w_1 + \ldots + \lambda_r w_r),$$

es decir el vector $\lambda_1 w_1 + \ldots + \lambda_r w_r$ está en el $\operatorname{Ker} A$. Como lo hemos hecho ya varias veces de aquí se deduce que los coeficientes deben ser cero. En efecto, por estar en el $\operatorname{Ker} A$ se debe tener que

$$\lambda_1 w_1 + \ldots + \lambda_r w_r = \alpha_1 u_1 + \ldots + \alpha_\nu u_\nu,$$

lo que implica que

$$\lambda_1 w_1 + \ldots + \lambda_r w_r - \alpha_1 u_1 - \ldots - \alpha_\nu u_\nu = 0,$$

y como los vectores son l.i. se tiene que $\lambda_1 = \cdots = \lambda_r = 0$.

Extendamos ahora X a una base de $\mathbb{K}^q$

$$\tilde{\beta}' = \{v_1, \ldots, v_r, v_{r+1}, \ldots, v_q\},$$

y calculemos la matriz representante de T con respecto a estas nuevas bases:

$$\begin{aligned}
&Aw_1 = v_1 = 1v_1 + 0v_2 + \ldots + 0v_r + 0v_{r+1} + \ldots + 0v_q \\
&Aw_2 = v_2 = 0v_1 + 1v_2 + \ldots + 0v_r + 0v_{r+1} + \ldots + 0v_q \\
&\vdots \\
&Aw_r = v_r = 0v_1 + 0v_2 + \ldots + 1v_r + 0v_{r+1} + \ldots + 0v_q \\
&Au_1 = 0 = 0v_1 + 0v_2 + \ldots + 0v_r + 0v_{r+1} + \ldots + 0v_q \\
&\vdots \\
&Au_\nu = 0 = 0v_1 + 0v_2 + \ldots + 0v_r + 0v_{r+1} + \ldots + 0v_q,
\end{aligned}$$

por lo tanto la matriz representante es:

$$M_{\tilde{\beta}\tilde{\beta}'}(T) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

donde I_r es una identidad de tamaño r y el resto son matrices de ceros de las dimensiones apropiadas. Usualmente esta matriz se denota por H y se le conoce como la forma normal de Hermitte. Usando el teorema de cambio de base se encuentra que existen matrices invertibles P, Q tal que

$$A = PHQ.$$

Por otro lado si $A, B \in \mathcal{M}_{qp}(\mathbb{K})$ tienen el mismo rango entonces se demuestra que existen matrices invertibles P, Q, R, S tal que

$$A = PHQ,\ B = RHS,\ \text{y por lo tanto } A = PR^{-1}BS^{-1}Q,$$

como $PR^{-1}, S^{-1}Q$ son invertibles se deduce que A, B son semejantes.

El rango de una matriz es el número máximo de columnas l.i., pero que hay del número máximo de filas l.i., o lo que es lo mismo, el rango de A^t. Veremos acontinuación que

$$r(A) = r(A^t).$$

Usando la forma normal de Hermite, se deduce que existen matrices invertibles P, Q tal que

$$A = PHQ \implies A^t = Q^t H^t P^t,$$

que es la forma normal de Hermite para A^t. Además es claro que $r(H) = r(H^t)$, pero entonces $r(A) = r(H) = r(H^t) = r(A^t)$. En particular se concluye que $r(A) \leq \min(p, q)$.

Un problema que se nos presenta es saber si hay una forma "fácil" de calcular el rango de una matriz. La respuesta es si, y se obtiene mediante escalonamiento. Notemos que la matriz escalonada de A que usualmente se denota $\tilde{A}$, tiene el mismo rango que A pues estas dos matrices son semejantes. En efecto $\tilde{A}$ se obtiene por premultiplicación de A por matrices elementales que son invertibles. Pero el rango de $\tilde{A}$ es fácil de calcular, pues tiene tantas filas l.i. como filas no nulas tenga, y esto debido a la forma triangular de $\tilde{A}$. Por lo tanto el rango de A es el número de filas no nulas en $\tilde{A}$.

Tomemos, por ejemplo, el escalonamiento siguiente:

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & -1 & 0 & 1 \\ 0 & -1 & -1 & -1 & 1 & 1 & 0 \\ -1 & 1 & -1 & 1 & 1 & 0 & 1 \end{pmatrix} \longrightarrow \tilde{A} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 2 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 & -2 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

Ejercicio 5.2: Probar que $r(AB) \leq r(A)$ y por lo tanto

$$r(AB) \leq \text{ mínimo } (r(A), r(B)).$$

5.2 Valores y vectores propios

En esta sección abordaremos el problema de saber cuándo para una aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^n$, es posible encontrar una base $\mathcal{B}$ con respecto a la cual la matriz representante de L sea diagonal. En forma equivalente, ¿cuándo una matriz A de $n \times n$ puede descomponerse de la forma

$$A = PDP^{-1} \text{ con } D \text{ diagonal ?} \tag{5.1}$$

Este problema de naturaleza puramente algebraica, tiene muchas aplicaciones en otras ramas de la Matemática, por ejemplo en Ecuaciones Diferenciales, Estadística, etc.

Tal vez el teorema más importante de este capítulo es el que dice que toda matriz simétrica puede representarse en la forma (5.1).

Comenzamos con las nociones de valores y vectores propios de una aplicación lineal $L : V \to V$ donde V es un espacio vectorial sobre $\mathbb{K}(\mathbb{K} = \mathbb{R} \text{ ó } \mathbb{C})$.

Definición (Vector y valor propio) Diremos que $x \in V$ es un **vector propio** de L si:

(i) $x \neq 0$

(ii) $\exists \lambda \in \mathbb{K}$ tal que $L(x) = \lambda x$.

En este caso el valor $\lambda \in \mathbb{K}$ que satisface (ii) se denomina **valor propio** asociado a x.

Diremos que $x \in V \setminus \{0\}$ es un vector propio de $A \in \mathcal{M}_{nn}$ si es vector propio de la aplicación lineal L dada por $L(x) = Ax$. Es decir

$$\exists \lambda \in \mathbb{K},\ Ax = \lambda x.$$

De la misma manera decimos que λ es valor propio de A.

Para la transformación lineal asociada a una matriz A mencionada anteriormente, se tiene:

Proposición 5.1. *Dada $A \in \mathcal{M}_{nn}(\mathbb{K})$, son equivalentes:*

(i) *$\exists x \neq 0\ Ax = \lambda x$.*

(ii) *$\exists x$ solución no trivial del sistema $(A - \lambda I)x = 0$.*

(iii) $\mathrm{Ker}(A - \lambda I) \neq \{0\}$

(iv) *$A - \lambda I$ no es invertible.*

Necesitamos una forma fácil de determinar qué valores de $\lambda \in \mathbb{K}$ son valores propios. Para ello sería útil que la condición (iv) pudiera expresarse como una ecuación en λ. Veremos que el cálculo de los valores propios se reduce a estudiar un polinomio de grado n cuyos coeficientes dependen de la matriz A.

Con este objetivo introduciremos una aplicación llamada **determinante**, que a cada matriz cuadrada A con coeficientes en $\mathbb{K}$ le asigna un elemento de $\mathbb{K}$.

Para ello primero definiremos A^{ij} la submatriz de A que se obtiene a partir de A eliminado la fila i y la columna j.

Ejemplo 5.1.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \qquad A^{11} = \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}$$

.

Con esto podemos definir el determinante de una matriz por un procedimiento inductivo de la siguiente forma.

> DEFINICIÓN (DETERMINANTE) Definimos el determinante de A como
>
> **(i)** Si A es de 1×1 $|A| = a$ donde $A = a$.
>
> **(ii)** Si A es de $n \times n$ $|A| = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} |A^{i1}|$.

Así, si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, entonces $|A| = a_{11}|A^{11}| - a_{21}|A^{21}| = ad - cb$.

Ejercicio 5.3: Obtener la fórmula para una matriz de 3×3.

A continuación daremos las propiedades más importantes de la función determinante, sin dar demostraciones, las cuales se presentan (por completitud) en el Apéndice.

Proposición 5.2. *Se tienen las siguiente propiedades:*

1. *El determinante es una función lineal de las filas de A es decir*

$$\forall t \in K, x, y \in \mathbb{K}^n : \left| \begin{pmatrix} A_{1\bullet} \\ \vdots \\ tx + y \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| = t \left| \begin{pmatrix} A_{1\bullet} \\ \vdots \\ x \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| + \left| \begin{pmatrix} A_{1\bullet} \\ \vdots \\ y \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right|$$

2. *Si B se obtiene de A permutando dos filas entonces $|A| = -|B|$.*

3. *Si A tiene dos filas iguales $|A| = 0$.*

4. *$|I| = 1$ donde I es la identidad.Si A es triangular superior entonces $|A| = \prod_{i=1}^{n} a_{ii}$.*

5. *Si a la fila i de la matriz A le sumamos la fila j ponderada por t entonces el determinante no cambia. Es decir*

$$\text{Si } B = \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{i-1\bullet} \\ A_{i\bullet} + tA_{j\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \text{ donde } i \neq j \text{ y } A = \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{i-1\bullet} \\ A_{i\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \text{ entonces } |B| = |A|.$$

6. *Si* $\tilde{A} = (\tilde{a}_{ij})$ *es una versión escalonada de* A *y* N_σ*, el número de permutaciones de filas realizadas para escalonar* A*. Entonces:*

$$|A| = (-1)^{N_\sigma} \cdot |\tilde{A}| = (-1)^{N_\sigma} \prod_{i=1}^{n} \tilde{a}_{ii}.$$

7. A *es invertible si y sólo si* $|A| \neq 0$.

8. *Una permutación del conjunto* $\{1, 2, ..., n\}$ *es una función* $\sigma \colon \{1, 2, ..., n\} \to \{1, 2, ..., n\}$*, biyectiva. Se suele usar la notación* $\sigma = \left(\begin{smallmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{smallmatrix}\right)$.

   Dada σ *una permutación del conjunto* $\{1, 2, ..., n\}$*, se define signo*$(\sigma) = \left|\begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}\right|$ *donde* $e_1, ..., e_n$ *son las filas de la matriz identidad. De (2) y (4) se obtiene que signo*$(\sigma) \in \{-1, 1\}$*. Se tiene además* $|A| = \sum\limits_{\sigma}$ *signo*(σ) $a_{1\sigma(1)} a_{2\sigma(2)} ... a_{n\sigma(n)}$ *donde la suma se extiende sobre todas las* $n!$ *permutaciones de* $\{1, ..., n\}$.

9. $|AB| = |A||B|$.

10. *Si* A *es invertible entonces* $|A^{-1}| = 1/|A|$.

11. $|A| = |A^t|$*. Además de (4) si* A *es triangular inferior entonces* $|A| = \prod\limits_{i=1}^{n} a_{ii}$.

12. *Para cualquiera* $i_0, j_0 \in \{1, \ldots, n\}$ *fijos, se tiene:*

$$|A| = \sum_{i=1}^{n} (-1)^{i+j_0} a_{ij_0} |A^{ij_0}| \quad \textit{(Cálculo de } |A|\textit{, usando la columna } j_0\textit{)}$$

$$|A| = \sum_{j=1}^{n} (-1)^{i_0+j} a_{i_0 j} |A^{i_0 j}| \quad \textit{(Cálculo de } |A|\textit{, usando la fila } i_0\textit{)}$$

La propiedad (7) nos dice que λ es valor propio de A si y sólo si $|A - \lambda I| = 0$. Esta ecuación es un polinomio de grado n en λ.

En efecto, usando la formula (8) se tiene

$$|A - \lambda I| = \sum_{\sigma} \text{ signo } (\sigma) \prod_{i=1}^{n} (A - \lambda I)_{i\sigma(i)}$$

si σ es la permutación identidad tendremos

$$\text{signo } (\sigma) \prod_{i=1}^{n} (A - \lambda I)_{i\sigma(i)} = \prod_{i=1}^{n} (a_{ii} - \lambda) = (-1)^n \lambda^n + q(\lambda)$$

donde $q(\lambda)$ es un polinomio de grado $\leq n - 1$.

Si σ es distinta de la permutación identidad se tendrá que existe un índice $i \in \{1,...,n\}$ tal que $i \neq \sigma(i)$. De donde $\ell(\lambda) = \text{signo}(\sigma)\prod_{i=1}^{n}(A-\lambda I)_{i\sigma(i)} = \text{signo } (\sigma)\prod_{i:\sigma(i)=i}(a_{ii}-\lambda)\prod_{i:\sigma(i)\neq i} a_{i\sigma(i)}$ es un polinomio de grado $\leq n-2$ (este último producto debe contener por lo menos dos términos ¿por qué?) y por lo tanto

$$|A-\lambda I| = (-1)^n\lambda^n + \alpha_{n-1}\lambda^{n-1} + ... + \alpha_1\lambda + \alpha_0,$$

donde $\alpha_{n-1},\dots,\alpha_0$ dependen de A. Por ejemplo evaluando en $\lambda = 0$ se tiene $|A| = \alpha_0$. ¿Cuánto vale el coeficiente de λ^{n-1}?.

Definición (Polinomio característico) $P(\lambda) = |A-\lambda I|$ es llamado el **polinomio característico de** A.

Ejemplo 5.2.

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

¿Cuáles son los valores propios de A?. Para ello estudiemos $A-\lambda I$.

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

pero entonces $0 = |A-\lambda I| = -(4-\lambda)(3+\lambda)+10 = -(12+\lambda-\lambda^2)+10 = \lambda^2-\lambda-2$. Los valores propios son $\lambda_1 = 2, \lambda_2 = -1$.

Ejemplo 5.3.

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad |A-\lambda I| = \left|\begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix}\right| = \lambda^2+1 = 0$$

no tiene solución en $\mathbb{R}$, y por tanto A no tiene valores propios reales. Sin embargo si el cuerpo donde estamos trabajando es $\mathbb{C}$ entonces A tiene dos valores propios.

Ejemplo 5.4.
Sea $D(\mathbb{R}) = \{$ espacio de funciones infinitamente diferenciables de $\mathbb{R} \to \mathbb{R}\}$ se prueba fácilmente que $D(\mathbb{R})$ es un espacio vectorial sobre $\mathbb{R}$.

Consideremos

$$\begin{aligned} L : D(\mathbb{R}) &\to D(\mathbb{R}) \\ f &\mapsto Lf = \frac{df}{dt}, \end{aligned}$$

es decir Lf es la función derivada de f.

Así por ejemplo si $f(t) = t^n$, $g(t) = cos(t)$ se tendrá $Lf = nt^{n-1}$; $Lg = -sen(t)$. Estudiemos si L tiene vectores propios:

$Lf = \lambda f$ es decir $\frac{df}{dt} = \lambda f$. Esta es una ecuación diferencial y una solución es: $f(t) = ce^{\lambda t}$.

Luego todo $\lambda \in \mathbb{R}$ es valor propio y un vector propio es $ce^{\lambda t}$, $c \neq 0$.

De la definición v es vector propio si $v \neq 0$ y existe $\lambda \in \mathbb{K}$ tal que $Av = \lambda v$. Así, λ es valor propio si y sólo si existe una solución no nula de la ecuación

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

Cualquier solución no nula de esta ecuación es entonces un vector propio de A con valor propio λ.

Notemos que si v es vector propio también lo es $10v$ y en general αv para cualquier $\alpha \in \mathbb{K}$, $\alpha \neq 0$.

Además, si v es vector propio, el valor propio asociado es único. Supongamos que $L(v) = \lambda_1 v = \lambda_2 v$ entonces $(\lambda_1 - \lambda_2)v = 0$. Como $v \neq 0$ se tendrá $\lambda_1 - \lambda_2 = 0$ y por lo tanto $\lambda_1 = \lambda_2$.

Para cada valor propio λ de A, introduciremos el **subespacio propio** W_λ que corresponde a $W_\lambda = \mathrm{Ker}(A - \lambda I)$. Notar que este subespacio debe contener vectores no nulos, pues λ es valor propio.

Si A y B son **similares** es decir si existe P invertible tal que $A = PBP^{-1}$, entonces los valores propios de A y B son los mismos y más aún, el polinomio característico es el mismo. En efecto:

$$A - \lambda I = PBP^{-1} - \lambda PP^{-1} = P(B - \lambda I)P^{-1}$$

Entonces

$$\begin{aligned} |(A - \lambda I)| &= |(P(B - \lambda I)P^{-1})| \\ &= |P||(B - \lambda I)||P^{-1}| \text{ pero } |P^{-1}| = 1/|P| \\ &= |B - \lambda I| \end{aligned}$$

luego A y B tienen el mismo polinomio característico, por lo tanto tienen los mismos valores propios.

Ejemplo 5.5.
Sea $A = \begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix}$ Sabemos que los valores propios son $\lambda_1 = 2, \lambda_2 = -1$. Calculemos los vectores propios asociados.

$\lambda_1 = 2$ La ecuación de vectores propios es

$$(A-2I)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & -5 \\ 2 & -5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

La solución general es $x_1 = \frac{5}{2}x_2$ y su espacio propio es:

$$W_2 = \operatorname{Ker}(A-2I) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1 = \frac{5}{2}x_2 \right\} = \left\langle \left\{ \begin{pmatrix} \frac{5}{2} \\ 1 \end{pmatrix} \right\} \right\rangle.$$

Así v es vector propio asociado a $\lambda_1 = 2$ si y solo si $v = \alpha \begin{pmatrix} \frac{5}{2} \\ 1 \end{pmatrix}$ $\alpha \neq 0$.

Para $\lambda_2 = -1$

$A-(-1)I = A+I = \begin{pmatrix} 5 & -5 \\ 2 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 5 & -5 \\ 0 & 0 \end{pmatrix}$, luego la solución general de la ecuación $(A+I)x = 0$ es: $x_1 = x_2$, y su espacio propio es

$$W_{-1} = \operatorname{Ker}(A-(-1)I) = \operatorname{Ker}(A+I) = \left\langle \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \right\rangle.$$

5.2 Apéndice

Hemos definido el determinante de una matriz A de $n \times n$ en forma recursiva de la siguiente manera $\mid A \mid = \sum_j (-1)^{j+1} a_{j1} \mid A^{j1} \mid$, donde $A^{k\ell}$ es la matriz de $n-1 \times n-1$, que se obtiene de A eliminándole la fila k y la columna ℓ. Probaremos ahora las propiedades fundamentales de esta función. Varias de las propiedades pueden probarse por inducción, dejaremos al lector la tarea de completar estas demostraciones.

1. Si

$$\tilde{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{k1}+b_{k1} & \cdots & a_{kn}+b_{kn} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

entonces

$$|\tilde{A}| = \left| \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{kn} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \right| + \left| \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ b_{k1} & \cdots & b_{kn} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \right| = |A| + |B|.$$

Por hipótesis de inducción (sobre el tamaño de la matriz) se tiene

$$\mid \tilde{A}^{j1} \mid = \mid A^{j1} \mid + \mid B^{j1} \mid .$$

Además, $\mid \tilde{A}^{k1} \mid = \mid A^{k1} \mid$, luego

$$\mid \tilde{A} \mid = \sum_{j=1}^{n} (-1)^{j+1} \tilde{a}_{j1} |\tilde{A}^{j1}| = \sum_{\substack{j=1 \\ j \neq k}}^{n} (-1)^{j+1} a_{j1} \{ \mid A^{j1} \mid + \mid B^{j1} \mid \} +$$

$$(-1)^{k+1}(a_{k+1} + b_{k+1}) \mid A^{k1} \mid = \mid A \mid + \mid B \mid$$

Análogamente se prueba que si

$$\tilde{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ ta_{k1} & \cdots & ta_{kn} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

entonces

$$\mid \tilde{A} \mid = t \left| \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{k1} & \cdots & a_{kn} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \right|.$$

Así visto como función de las filas de A el determinante tiene la propiedad lineal siguiente

$$\forall x, y \in \mathbb{K}^n, t \in K : \left| \begin{pmatrix} A_{1\bullet} \\ A_{2\bullet} \\ \vdots \\ tx + y \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| = t \left| \begin{pmatrix} A_{1\bullet} \\ A_{2\bullet} \\ \vdots \\ x \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| + \left| \begin{pmatrix} A_{1\bullet} \\ A_{2\bullet} \\ \vdots \\ y \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right|.$$

Es decir el determinante es lineal en cada una de las filas de la matriz. Si estamos considerando matrices de dimensión n entonces diremos que el determinante es **n-lineal**.

Probaremos simultáneamente por inducción las propiedades (2) y (3):

2. Si B se obtiene de A permutando dos filas entonces $|B| = -|A|$. Se dice que el determinante es una función **alternada**.

3. Si A tiene dos filas iguales entonces $|A| = 0$. Supongamos que A tiene las filas i y j

iguales, es decir:

$$A = \begin{array}{cl} \begin{pmatrix} A_{1\bullet} \\ \vdots \\ X \\ \vdots \\ X \\ \vdots \\ A_{n\bullet} \end{pmatrix} & \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array} \end{array}.$$

Cada matriz A^{k1} tiene dos filas iguales a menos que $k = i$ o $k = j$ de donde $\mid A^{k1} \mid = 0$ si $k \neq i \quad k \neq j$. Por lo tanto $\Rightarrow \mid A \mid = (-1)^{i+1} a_{i1} \mid A^{i1} \mid + (-1)^{j+1} a_{j1} \mid A^{j1} \mid$.

Las matrices A^{i1} y A^{j1} de dimensión $n-1 \times n-1$ son iguales excepto que la fila X esta en distinta posición. En A^{i1} esta en la fila $j-1$. En A^{j1} esta en la fila i. Así podemos llegar de la matriz A^{i1} a la matriz A^{j1} mediante $j-1-i$ permutaciones. Por cada permutación hay un cambio de signo en el determinante (de dimensión $n-1$). Luego $\mid A^{i1} \mid = (-1)^{j-1-i} \mid A^{j1} \mid$. Además $a_{i1} = a_{j1} \Rightarrow$

$$
\begin{aligned}
\mid A \mid &= a_{i1}\{(-1)^{i+1}(-1)^{j-1-i} \mid A^{j1} \mid +(-1)^{j+1} \mid A^{j1} \mid\} \\
&= \{(-1)^{j} - (-1)^{j}\} a_{i1} \mid A^{j1} \mid = 0
\end{aligned}
$$

Supongamos que $\tilde{A}$ se obtiene de A permutando las filas i y j:

$$
\tilde{A} = \begin{array}{c} \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{j\bullet} \\ \vdots \\ A_{i\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array}.
$$

Tomemos

$$
B = \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{i\bullet} + A_{j\bullet} \\ \vdots \\ A_{i\bullet} + A_{j\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array},
$$

como B tiene dos filas iguales se deduce que

$$
0 = \mid B \mid = \left| \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{i\bullet} \\ \vdots \\ A_{i\bullet} + A_{j\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| + \left| \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{j\bullet} \\ \vdots \\ A_{i\bullet} + A_{j\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| = \left| \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{i\bullet} \\ \vdots \\ A_{i\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| + \left| \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{i\bullet} \\ \vdots \\ A_{j\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| +
$$

$$
+ \left| \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{j\bullet} \\ \vdots \\ A_{i\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| + \left| \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{j\bullet} \\ \vdots \\ A_{j\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| = 0 + |A| + |\tilde{A}| + 0,
$$

y por lo tanto

$$
\mid \tilde{A} \mid = - \mid A \mid
$$

4. Sea I_n la identidad de $n \times n$ entonces $|I_n| = 1$. En efecto
$$|I_n| = 1|I_{n-1}| = 1.$$
De manera más general se prueba que si A es triangular superior entonces $|A| = \prod_{i=1}^{n} a_{ii}$, esto también es cierto para una triangular inferior pero lo probaremos más tarde.

5. Si a la fila i le sumamos la fila $j(i \neq j)$ amplificada por algún factor, entonces el determinante no cambia. Sea
$$\tilde{A} = \begin{array}{c} \left(\begin{array}{c} A_{1\bullet} \\ \vdots \\ A_{j\bullet} \\ \vdots \\ A_{i\bullet} + tA_{j\bullet} \\ \vdots \\ A_{n\bullet} \end{array}\right) \end{array} \begin{array}{l} \\ \\ \leftarrow j \\ \\ \leftarrow i \\ \\ \\ \end{array},$$
entonces
$$| \tilde{A} | = | A | + t \left| \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{j\bullet} \\ \vdots \\ A_{j\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| \begin{array}{l} \\ \\ \leftarrow j \\ \\ \leftarrow i \\ \\ \\ \end{array} = | A |$$

6. Las filas de A son l.d. $\Longleftrightarrow | A | = 0$. Supongamos que las filas de A son l.d. es decir $\exists \lambda_1, ..., \lambda_n$ no todos nulos tal que $\lambda_1 A_{1\bullet} + \lambda_2 A_{2\bullet} + ... + \lambda_n A_{n\bullet} = 0$ podemos suponer sin pérdida de generalidad que $\lambda_1 \neq 0$
$$\tilde{A} = \begin{pmatrix} \lambda_1 A_{1\bullet} + \lambda_2 A_{2\bullet} + \cdots + \lambda_n A_{n\bullet} \\ A_{2\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} =$$
$$= \begin{pmatrix} 0 \\ A_{2\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} = (-1)^{1+1} 0 \cdot | \tilde{A}^{11} | + \sum_{j=2}^{n} (-1)^{j+1} \tilde{a}_{j1} |\tilde{A}^{j1}|,$$
pero para $j \geq 2$ la primera fila de $\tilde{A}^{j1}$ es cero luego las filas de $\tilde{A}^{j1}$ son l.d., por tanto $| \tilde{A}^{j1} | = 0$. Así $| \tilde{A} | = 0$. Además
$$0 = \lambda_1 |A| + \lambda_2 \left| \begin{pmatrix} A_{2\bullet} \\ A_{2\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| + \cdots + \lambda_k \left| \begin{pmatrix} A_{k\bullet} \\ A_{2\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| + \cdots + \lambda_n \left| \begin{pmatrix} A_{n\bullet} \\ A_{2\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| = \lambda_1 |A|,$$

pero $\lambda_1 \neq 0$ y por lo tanto $\mid A \mid= 0$.

Si las filas son l.i. probaremos que $|A| \neq 0$. En efecto, mediante operaciones elementales sobre las filas de A (suma de filas y permutación de filas) se llega a

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \cdots & & \tilde{a}_{1n} \\ 0 & \tilde{a}_{22} & \cdots & & \tilde{a}_{2n} \\ \vdots & 0 & & & \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & 0 & 0 & \tilde{a}_{nn} \end{pmatrix}$$

pero entonces $\mid \tilde{A} \mid= \pm \mid A \mid$. Dado que $\mid \tilde{A} \mid= \prod_{i=1}^{n} \tilde{a}_{ii} \neq 0$ (ver (4)), se concluye que $\mid A \mid\neq 0$.

Falta probar que $|A| = |A^t|$. Esta proposición no es fácil, para ello necesitamos la siguiente fórmula:

7. $|A| = \sum_{\sigma}$ (signo σ) $a_{1\sigma(1)}\ a_{2\sigma(2)} \cdots a_{n\sigma(n)}$ donde la suma se extiende sobre todas las permutaciones de $\{1, ..., n\}$ y signo(σ) se define como

$$\text{signo } (\sigma) = \left| \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} \right|,$$

donde $\{e_1, ...e_n\}$ es la base canónica de $\mathbb{R}^n$

Ejemplo 5.6.

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ 2 & 1 & 3 & \cdots & n \end{pmatrix}$$

$$\begin{aligned} e_{\sigma(1)} &= e_2 \\ e_{\sigma(2)} &= e_1 \\ e_{\sigma(i)} &= e_i \quad i \geq 2 \end{aligned}$$

$$\text{signo}(\sigma) = \left| \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & 0 \\ & & & & 1 \end{pmatrix} \right| = (-1)^3 \mid I_{n-1} \mid = -1, \text{ luego signo } (\sigma) = -1$$

El término asociado a esta permutación σ en la fórmula (7) es:

$$-1\ a_{12}a_{21}a_{33}a_{44}...a_{nn}$$

Notemos que cada término de la fórmula (7) es el producto de una elemento por fila y columna.

Demostremos la fórmula. Dado que $A_{1\bullet} = a_{11}e_1 + a_{12}e_2 + ... + a_{1n}e_n$ se tendrá

$$| A |= \sum_{k=1}^{n} a_{1k} \left| \begin{pmatrix} e_k \\ A_{2\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right|.$$

Análogamente $A_{2\bullet} = \sum\limits_{\ell=1}^{n} a_{2\ell}e_\ell$, luego

$$\left| \begin{pmatrix} e_k \\ A_{2\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| = \sum_{\ell=1}^{n} \ a_{2\ell} \left| \begin{pmatrix} e_k \\ e_\ell \\ A_{3\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| = \sum_{\ell\neq k} a_{2\ell} \left| \begin{pmatrix} e_k \\ e_\ell \\ A_{3\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right|.$$

Por lo tanto

$$|A| = \sum_{k\neq\ell} a_{1k}a_{2\ell} \left| \begin{pmatrix} e_k \\ e_\ell \\ A_{j\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \right| = \sum_{\substack{k_1,\ldots,k_n \\ \text{todos} \\ \text{distintos}}} a_{1k_1}a_{2k_2}...a_{nk_n} \left| \begin{pmatrix} e_{k_1} \\ e_{k_2} \\ \vdots \\ e_{k_n} \end{pmatrix} \right|$$

Esta suma es la suma sobre todas las permutaciones

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix} \text{ y } \left| \begin{pmatrix} e_{k_1} \\ \vdots \\ e_{kn} \end{pmatrix} \right| = \text{ signo } (\sigma),$$

lo que prueba la fórmula deseada. Notemos que lo que se uso en probar esta fórmula es que el determinante es n-lineal y alternada. Por esto estudiaremos un poco como son este tipo de funciones.

Supongamos que tenemos una función F que asigna a cada matriz cuadrada de $n\times n$ un número de $\mathbb{K}$, de manera que es n-lineal y alternada en las filas de la matriz, entonces necesariamente:

$$F(A) = |A| \cdot F(I)$$

donde $F(I) \in \mathbb{K}$ es el valor que toma para la identidad de tamaño n. Este resultado dice que toda función n-lineal alternada debe ser un múltiplo del determinante. De igual forma como lo hicimos para el determinante se prueba que

$$F(A) = \sum_{\sigma} a_{1\sigma(1)}...a_{n\sigma(n)}F \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}.$$

Pero por ser F alternada se tendrá

$$F\begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = \pm F(I)$$

el signo depende de si se necesitan un número par o impar de permutaciones para llegar a la identidad pero

$$\left|\begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}\right|$$

es exactamente ± 1 dependiendo de la paridad del número de intercambios necesarios para llegar a la identidad. Luego, $F(A) = F(I)\sum\limits_{j}$ signo $(\sigma)a_{1\sigma(1)}...a_{n\sigma(n)} = F(I)|A|$.

8. $|AB| = |A||B|$

Sea

$$\begin{aligned} F : M_{nn}(\mathbb{K}) &\to \mathbb{K} \\ A &\mapsto |AB| \end{aligned}$$

probemos que F es n-lineal. En efecto

$$C_1 = \begin{pmatrix} A_{1\bullet} \\ \vdots \\ X \\ \vdots \\ A_{n\bullet} \end{pmatrix} \qquad C_2 = \begin{pmatrix} A_{1\bullet} \\ \vdots \\ Y \\ \vdots \\ A_{n\bullet} \end{pmatrix}$$

$$\tilde{A} = \begin{pmatrix} A_{1\bullet} \\ \vdots \\ X+Y \\ \vdots \\ A_{n\bullet} \end{pmatrix} \qquad \tilde{A}B = \begin{pmatrix} A_{1\bullet}B \\ \vdots \\ XB+YB \\ \vdots \\ A_{n\bullet}B \end{pmatrix},$$

y por lo tanto

$$F(\tilde{A}) = |\tilde{A}B| = \left|\begin{pmatrix} A_{1\bullet}B \\ \vdots \\ XB \\ \vdots \\ A_{n\bullet}B \end{pmatrix}\right| + \left|\begin{pmatrix} A_{1\bullet}B \\ \vdots \\ YB \\ \vdots \\ A_{n\bullet}B \end{pmatrix}\right| = |C_1B| + |C_2B| = F(C_1) + F(C_2).$$

De igual forma se prueba que

$$F\begin{pmatrix} A_{1\bullet} \\ \vdots \\ tA_{i\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} = tF\begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{i\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix}.$$

Si intercambiamos las filas i y j de A y definimos

$$\tilde{A} = \begin{pmatrix} A_{1\bullet} \\ \vdots \\ A_{j\bullet} \\ \vdots \\ A_{i\bullet} \\ \vdots \\ A_{n\bullet} \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array} \qquad \tilde{A}B = \begin{pmatrix} A_{1\bullet}B \\ \vdots \\ A_{j\bullet}B \\ \vdots \\ A_{i\bullet}B \\ \vdots \\ A_{n\bullet}B \end{pmatrix},$$

se tendrá $|\tilde{A}B| = -|AB|$ por lo tanto $F(\tilde{A}) = -F(A)$. Del resultado anterior concluimos que $F(A) = |A|F(I)$ pero $F(I) = |IB| = |B|$ por lo tanto $|AB| = |A||B|$

9. Si A es invertible entonces

$$|A^{-1}| = \frac{1}{|A|}$$

en efecto, $|A^{-1}A| = |I| = 1 = |A^{-1}| \cdot |A|$ luego $|A^{-1}| = \frac{1}{|A|}$.

Proposición 5.3. *Sean σ y σ' dos permutaciones entonces signo($\sigma \circ \sigma'$) = signo(σ) signo(σ').*

Demostración. Sean

$$A = \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} \qquad B = \begin{pmatrix} e_{\sigma'(1)} \\ \vdots \\ e_{\sigma'(n)} \end{pmatrix}$$

$$C = \begin{pmatrix} e_{\sigma\circ\sigma'(1)} \\ \vdots \\ e_{\sigma\circ\sigma'(n)} \end{pmatrix} = \begin{pmatrix} e_{\sigma(\sigma'(1))} \\ \vdots \\ e_{\sigma(\sigma'(n))} \end{pmatrix}.$$

Si escribimos A en términos de sus columnas, tendremos que el 1 de la primera columna se ubica en la posición k (fila k) que hace que $\sigma(k) = 1$, pero entonces $k = \sigma^{-1}(1)$, por lo tanto A descrita por sus columnas es: $A = (e_{\sigma^{-1}(1)} e_{\sigma^{-1}(2)} ... e_{\sigma^{-1}(n)})$.
Como
$e_j^t e_k = \begin{cases} 0 & j \neq k \\ 1 & j = k \end{cases}$, se tiene que la primera fila de BA tendrá la forma:

$$(BA)_{1j} = e^t_{\sigma'(1)} e_{\sigma^{-1}(j)} = \begin{cases} 1 & \sigma'(1) = \sigma^{-1}(j) \\ 0 & \text{si no} \end{cases},$$

es decir $(BA)_{1\bullet}$ es un vector de la base canónica, lo que debemos averiguar es en que posición tiene el "1", pero esto se deduce de la ecuación $\sigma'(1) = \sigma^{-1}(j)$ pues entonces $j = \sigma(\sigma'(1)) = \sigma \circ \sigma'(1)$, luego $(B \cdot A)_{1\bullet} = e_{\sigma o \sigma'(1)}$. Así se obtiene:

$$BA = \begin{pmatrix} e_{\sigma\circ\sigma'(1)} \\ e_{\sigma\circ\sigma'(2)} \\ \vdots \\ e_{\sigma\circ\sigma'(n)} \end{pmatrix}$$

Por lo tanto $C = B \cdot A$, de donde $|C| = |B||A|$ y entonces signo$(\sigma \circ \sigma')$= signo(σ) signo(σ'). En particular signo(σ^{-1})= signo(σ) pues

$$\text{signo}(\sigma \circ \sigma^{-1}) = \text{ signo}(id) = 1 = \text{ signo}(\sigma) \text{ signo}(\sigma^{-1}),$$

donde id es la permutación identidad. □

10. Probemos ahora que $|A| = |A^t|$. En efecto:

$$\begin{aligned} |A^t| &= \sum_\sigma \text{ signo } (\sigma)\ (A^t)_{1\sigma(1)}(A^t)_{2\sigma(2)} \cdots (A^t)_{n\sigma(n)} \\ &= \sum_\sigma (\text{ signo}(\sigma)) a_{\sigma(1)1}, a_{\sigma(2)} \cdots a_{\sigma(n)n}, \end{aligned}$$

pero $\prod_{i=1}^{n} a_{\sigma(i)i} = \prod_{i=1}^{n} a_{i\sigma^{-1}(i)}$ lo que se obtiene reordenando el producto de acuerdo a la primera coordenada. Luego

$$|A^t| = \sum_\sigma (\text{signo}(\sigma)) a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)}.$$

Como signo(σ) = signo(σ^{-1}) y la aplicación $\sigma \to \sigma^{-1}$ es una biyección se tiene:

$$|A^t| = \sum_\nu (\text{signo}(\nu)) a_{1\nu(1)} \cdots a_{2\nu(2)} = |A|$$

Fórmula para A^{-1}

Sea $F : M_{nn}(\mathbb{K}) \to \mathbb{K}$ $F(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} |A^{ij}|$ que corresponde a desarrollar el determinante de A por la columna j.

Se puede probar al igual que hicimos para el determinante que F es n-lineal y alternada luego $F(A) = |A| \cdot F(I)$ pero $F(I) = 1$, por tanto $F(A) = |A|$. Además si se usa $|A| = |A^t|$ se prueba que se puede calcular $|A|$ utilizando las filas de A esto es: $|A| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} |A^{ij}|$

Si $|A| \neq 0$ definamos B mediante:

$$b_{ij} = \frac{1}{|A|}(-1)^{i+j}|A^{ji}|.$$

Calculemos AB:

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{n} \frac{(-1)^{k+j}}{|A|} a_{ik}|A^{jk}|.$$

Analicemos ahora dos casos:

1. $j = i \quad (AB)_{ii} = \frac{1}{|A|}\sum_{k=1}^{n}(-1)^{k+i}a_{ik}|A^{ik}| = 1$

2. $j \neq i \quad (AB)_{ij} = \sum_{k=1}^{n}(-1)^{k+j}\frac{a_{ik}}{|A|}|A^{jk}|$, que corresponde al determinante de la siguiente matriz, calculado usando la fila j

$$\begin{array}{cc} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \vdots & \vdots \\ \frac{a_{i1}}{|A|} & \cdots & \frac{a_{in}}{|A|} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} & \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array} \end{array}$$

Como las filas de esta matriz son l.d. se tiene $\sum_{k=1}^{n}(-1)^{k+j}\frac{a_{jk}}{|A|}|A^{jk}| = 0$, y por tanto $AB = I$ lo que implica que

$$B = A^{-1}.$$

Ejemplo 5.7.
Sea

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

Se tendrá que

$$\begin{aligned} |A| &= ad - bc \\ |A^{11}| = d \quad |A^{12}| &= c \quad |A^{21}| = b \quad |A^{22}| = a \\ A^{-1} &= \frac{1}{|A|}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^{t} = \frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \end{aligned}$$

Guía de Ejercicios

1. Suponga que A, B son semejantes, es decir, existen matrices invertibles P, Q tal que $A = PBQ$. Pruebe que A y B tienen el mismo rango:

$$r(A) = r(B).$$

2. Probar que $r(AB) \leq r(A)$ y por lo tanto

$$r(AB) \leq \text{ mínimo } (r(A), r(B)).$$

3. Obtener la fórmula del determinante de una matriz de 3×3.

4. Demuestre que si $A, B \in M_{nxn}(\mathbb{R})$, entonces $P(A \cdot B) = P(B \cdot A)$ si A es invertible, y donde $P(A)$ denota "el polinomio característico de A"

5. Determinar los valores y los vectores propios correspondientes a las matrices siguientes:

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \begin{pmatrix} 3 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \quad \begin{pmatrix} 3 & 2 & 2 \\ -2 & 1 & -6 \\ 1 & -1 & 4 \end{pmatrix} \quad \begin{pmatrix} 3 & -2 & 1 \\ 2 & 1 & -1 \\ 2 & -6 & 4 \end{pmatrix}$$

Guía de Problemas

P1. (a) Pruebe que una matriz $A \in M_{mn}(\mathbb{K})$ tiene rango 1 si y sólo si existen $u \in \mathbb{K}^m$, $v \in \mathbb{K}^n$, ambos no nulos tales que $A = uv^t$.

(b) Si $\alpha_0, \alpha_1, ..., \alpha_{n-1} \in \mathbb{R}$ y $n \geqslant 2$, probar por inducción que el polinomio característico de la matriz

$$A = \begin{pmatrix} 0 & 0 & ... & 0 & -\alpha_0 \\ 1 & 0 & ... & 0 & -\alpha_1 \\ 0 & 1 & ... & 0 & -\alpha_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & ... & 1 & -\alpha_{n-1} \end{pmatrix}$$

es igual a: $(-1)^n (\sum\limits_{k=0}^{n-1} \alpha_k \lambda^k + \lambda^n)$.

P2. Sean $A, B \in M_{n,n}(\mathbb{R})$ invertibles y $\alpha \in \mathbb{R}$, $0 < \alpha < 1$.

a) Sea $P_{A^{-1}B}(\cdot)$ el polinomio característico de $A^{-1}B$. Pruebe que

$$|(\alpha A + (1-\alpha)B)| = (1-\alpha)^n |A| P_{A^{-1}B}\left(\frac{\alpha}{\alpha - 1}\right).$$

b) Demueste que existe $0 < \alpha < 1$ tal que $(\alpha A + (1-\alpha)B)$ es invertible.

P3. Sean A, B matrices de $n \times n$ con coeficientes reales tales que $AB = BA$

a) Pruebe que si $Bv \neq 0$ y v es vector propio de A asociado a λ entonces Bv también lo es.

b) Muestre que si v es vector propio de A perteneciente a un espacio propio de dimensión 1, entonces v es vector propio de B.

Valores y Vectores Propios

Sea $L : \mathbb{K}^n \to \mathbb{K}^n$ lineal y sea A la matriz representante de L con respecto a la base canónica B. Supongamos que existe $B' = \{v_1, \ldots, v_n\}$ base de vectores propios, es decir B' es una base de $\mathbb{K}^n$ y

$$\forall i, \exists \lambda_i \in \mathbb{K} : Av_i = \lambda_i v_i.$$

¿Cuál es la matriz representante de L con respecto a B'?

$$\begin{aligned}
Av_1 &= \lambda_1 v_1 = \lambda_1 v_1 + 0v_2 + \cdots + 0v_n \\
Av_2 &= \lambda_1 v_2 = 0v_1 + \lambda_2 v_2 + \cdots + 0v_n \\
&\vdots \\
Av_n &= \lambda_n v_n = 0v_1 + 0v_2 + \cdots + \lambda_n v_n
\end{aligned}$$

$$M_{B'B'}(L) = \begin{pmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & & \lambda_n \end{pmatrix} = D \text{ que es diagonal.}$$

Además se tiene el diagrama

$$\begin{tikzcd}
B \arrow[r, "A"] \arrow[d, "P^{-1}"'] & B \\
B' \arrow[r, "D"] & B' \arrow[u, "P"']
\end{tikzcd}$$

Por lo tanto $A = PDP^{-1}$

Algunas ventajas de conocer D:

(1) $r(A) = r(D) =$ número de valores propios no nulos.

(2) $|A| = |PDP^{-1}| = |P||D||P^{-1}| = |D| = \prod_{i=1}^{n} \lambda_i$.

(3) $|A| \neq 0$ entonces $\forall i$, $\lambda_i \neq 0$ y $A^{-1} = PD^{-1}P^{-1}$ donde

$$D^{-1} = \begin{pmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{-1} \end{pmatrix}.$$

En efecto,
$$A^{-1} = (PDP^{-1})^{-1} = (P^{-1})^{-1}D^{-1}P^{-1} = PD^{-1}P^{-1},$$
pero $D^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\lambda_n} \end{pmatrix}$

(4)

$$\begin{aligned} A^m = (PDP^{-1})^m &= (PDP^{-1})(PDP^{-1})\cdots(PDP^{-1}) \\ &= PD^mP^{-1} = P\begin{pmatrix} \lambda_1^m & & 0 \\ & \ddots & \\ 0 & & \lambda_n^m \end{pmatrix} P^{-1} \end{aligned}$$

Como hemos visto $A = PDP^{-1}$ donde $P = M_{B'B}(id_{\mathbb{K}^n})$ es decir tenemos que expresar los vectores propios en términos de la base canónica, y por lo tanto $P = (v_1, v_2, \ldots, v_n)$, donde los vectores v_i son las columnas de P.

Definición (Matriz diagonalizable) Diremos que $A \in M_{nn}(\mathbb{K})$ es **diagonalizable** si $\mathbb{K}^n$ admite una base de vectores propios de A.

Teorema 5.1. *A es diagonalizable si y sólo si A es similar a una matriz diagonal.*

Demostración.

($\Rightarrow$) Hemos visto que si A es diagonalizable entonces

$A = PDP^{-1}$ donde $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$

y $P = (v_1, \ldots, v_n)$. Luego A es similar a una diagonal.

($\Leftarrow$) Supongamos que A es similar a una diagonal, es decir existe P invertible tal que $A = PDP^{-1}$. De manera equivalente:
$$AP = PD.$$
Si suponemos que $P = (v_1, \ldots, v_n)$ (notación por columnas) y $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$, entoces la igualdad anterior se escribe:
$$A(v_1, \ldots, v_n) = (v_1, \ldots, v_n)\begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}.$$

Y por propiedades de la multiplcación de matrices esto equivale a:

$$(Av_1, \ldots, Av_n) = (d_1 v_1, \ldots, d_n v_n).$$

Finalmente, la igualdad vista por columnas nos dice que $\forall i \in \{1, \ldots, n\}$ $Av_i = d_i v_i$. Es decir, los vectores $v_1, \ldots, v_n$ son vectores propios de A (son no nulos pues P es invertibles), con valores propios respectivos $d_1 \ldots, d_n$.

Además, son claramente una base de $\mathbb{K}^n$, pues son las columnas de P que es invertible. $\square$

Teorema 5.2. *Sea $A \in M_{nn}(\mathbb{K})$ si $\{\lambda_i\}_{i=1,\ldots,k}$ son valores propios de A distintos y $\{v_i\}_{i=1\ldots,k}$ son vectores propios de A tal que $Av_i = \lambda_i v_i$, entonces $\{v_i\}_{i=1,\ldots,k}$ es un conjunto l.i.*

Demostración. Por inducción sobre k. Para $k = 1$, la propiedad es obviamente cierta. Supongamos

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0, \tag{5.2}$$

donde $Av_i = \lambda_i v_i$ y λ_i $i = 1, \ldots, k$ son todos distintos. Entonces

$$\alpha_1 \lambda_1 v_1 + \cdots + \alpha_k \lambda_k v_k = A(\alpha_1 v_1 + \cdots + \alpha_k v_k) = A0 = 0.$$

Multiplicado (5.2) por λ_k se tiene $\alpha_1 \lambda_k v_1 + \cdots + \alpha_k \lambda_k v_k = 0$. Luego

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

como $\{v_1 \ldots, v_{k-1}\}$ son l.i. se tiene $\alpha_i(\lambda_i - \lambda_k) = 0$ $i = 1, \ldots, k-1$. Como $\lambda_i \neq \lambda_k$ se debe tener que $\alpha_i = 0$ $i = 1, \ldots, k-1$ y de (5.2) se concluye que $\alpha_k = 0$. $\square$

Antes de seguir, veremos la extensión natural del concepto de suma directa de más de dos subespacios vectoriales. Definamos primero, dados $U_1, U_2, \ldots, U_k$ s.e.v. de un espacio vectorial V, la **suma** de ellos como:

$$\mathop{+}_{i=1}^{k} U_i = U_1 + U_2 + \cdots + U_k \triangleq \left\{ v = \sum_{i=1}^{k} u_i \mid \forall i \in \{1, \ldots, k\},\ u_i \in U_i \right\}.$$

Esto es claramente un s.e.v. de V.

Definición (Suma directa múltiple) Sea V espacio vectorial y $U_1, \ldots, U_k$ subespacios vectoriales de V. Decimos que el subespacio $Z = +_{i=1}^{k} U_i$ es **suma directa** de $U_1, \ldots, U_k$, notado $Z = \bigoplus_{i=1}^{k} U_i = U_1 \oplus U_2 \oplus \cdots \oplus U_k$, si para todo $v \in Z$, v se escribe de manera única como

$$v = \sum_{i=1}^{k} u_i, \quad \text{con } u_i \in U_i,\ \forall i \in \{1, \ldots, k\}.$$

Las siguientes propiedades se tienen de manera análoga al caso de dos subespacios:

Proposición 5.4. *Sea V espacio vectorial y $Z, U_1, \dots, U_k$ subespacios vectoriales de V, entonces:*

1. $Z = \bigoplus_{i=1}^{k} U_i \iff \left(Z = \mathop{+}_{i=1}^{k} U_i \ \wedge\ \forall j \in \{1, \dots, k\},\ U_j \cap \left(\mathop{+}_{\substack{i=1 \\ i \neq j}}^{k} U_i \right) = \{0\} \right).$

2. *Si* $Z = \mathop{+}_{i=1}^{k} U_i$ *y Z es de dimensión finita, entonces son equivalentes:*

   (a) $Z = \bigoplus_{i=1}^{k} U_i$.

   (b) $(\forall i \in \{1, \dots, k\})(\forall u_i \in U_i \setminus \{0\})\ \{u_1, \dots, u_k\}$ *es l.i.*

   (c) *La yuxtaposición de bases de los subespacios U_i es una base (y no sólo un generador) de Z.*

   (d) $\dim(Z) = \sum_{i=1}^{k} \dim(U_i)$.

Demostración. Probaremos $(a) \iff (b)$ en la parte 2. El resto de las demostraciones, quedan propuestas como ejercicio.

$(b) \implies (a)$ Sea $v \in Z$ y dos formas distintas de escribirlo:

$$v = \sum_{i=1}^{k} w_i = \sum_{i=1}^{k} w_i', \quad \text{con } w_i, w_i' \in U_i.$$

De aquí, suponiendo sin pérdida de generalidad que $\forall i \in \{1, \dots, k^*\},\ w_i - w_i' \neq 0$, sale que:

$$\sum_{i=1}^{k^*} (w_i - w_i') = 0,$$

lo cual es una combinación lineal nula, con escalares no nulos (todos 1), de los vectores $u_i = w_i - w_i' \in U_i \setminus \{0\}$. Pero esto es una contradicción con el hecho de que dichos vectores son l.i, por hipótesis.

Así se tiene que la escritura es única y por ende la suma es directa

$(a) \implies (b)$ Supongamos ahora que $Z = \bigoplus_{i=1}^{k} U_i$. Sea $\{u_1, \dots, u_k\}$, con $u_i \in U_i,\ \forall i \in \{1, \dots, k\}$. Estudiemos

$$\sum_{i=1}^{k} \alpha_i u_i = 0.$$

Como para cada $i \in \{1, \dots, k\}$, U_i es un s.e.v., luego $\alpha_i u_i \in U_i$. Así, hemos encontrado una forma de escribir al vector nulo como suma de elementos de los subespacios U_i.

Por unicidad de dicha escritura,

$$\forall i \in \{1, \dots, k\},\ \alpha_i u_i = 0.$$

Y esto implica, ya que $u_i \neq 0$, que $\alpha_i = 0$, $\forall i \in \{1, \ldots, k\}$. Es decir, $\{u_1, \ldots, u_k\}$ es l.i.
□

Teorema 5.3. *Sea $A \in M_{nn}(\mathbb{K})$, $\lambda_1, \ldots, \lambda_k$ los valores propios (distintos) de A y $W_{\lambda_i} = Ker(A - \lambda_i I)$, $i \in \{1, \ldots, k\}$. Si $W = W_{\lambda_1} + W_{\lambda_2} + \cdots + W_{\lambda_k}$, se tiene que*

$$W = W_{\lambda_1} \oplus W_{\lambda_2} \oplus \cdots \oplus W_{\lambda_k}.$$

En particular, A es diagonalizable si y sólo si

$$\mathbb{K}^n = W_{\lambda_1} \oplus W_{\lambda_2} \oplus \cdots \oplus W_{\lambda_k}.$$

Demostración. Directa de la Proposición 5.4. □

Corolario 5.1. *Supongamos que $A \in M_{nn}(\mathbb{K})$ y que $p(\lambda) = |A - \lambda I|$, el polinomio característico de A, tiene n raíces distintas en $\mathbb{K} : \lambda_1, \ldots, \lambda_n$, entonces*

1. *$W_{\lambda_i} = Ker(A - \lambda_i I)$ es de dimensión 1.*
2. *Sea $v_i \in W_{\lambda_i}$ con $v_i \neq 0$ entonces $\{v_1, \ldots, v_n\}$ es una base de vectores propios.*

Este teorema da un criterio simple para que $A \in M_{nn}(\mathbb{K})$ sea diagonizable, éste es que A tenga n valores propios distintos.

Ejemplo 5.8.
Sea

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

$$p(\lambda) = |A - \lambda I| = \left| \begin{pmatrix} 1-\lambda & -1 & -1 \\ -1 & 1-\lambda & -1 \\ -1 & -1 & 1-\lambda \end{pmatrix} \right|$$

$$\begin{aligned}
&= (1-\lambda)\{(1-\lambda)^2 - 1\} - 1\{(1-\lambda) + 1\} - \{1 + (1-\lambda)\} \\
&= (1-\lambda)^3 - (1-\lambda) - (1-\lambda) - 1 - 1 - (1-\lambda) \\
&= (1 - 3\lambda + 3\lambda^2 - \lambda^3) - 3 + 3\lambda - 2 \\
&= 3\lambda^2 - \lambda^3 - 4
\end{aligned}$$

Ahora $\lambda = 2$ es una raíz de $p(\lambda)$. Por otro lado $p(\lambda) \div \lambda - 2 = -\lambda^2 + \lambda + 2$

Así las otras raíces de p son $\lambda = -1$ y $\lambda = 2$. Por lo tanto $p(\lambda) = -(\lambda-2)(\lambda+1)(\lambda-2)$
Se puede probar que A tiene una base de vectores propios por lo que no es necesario que todos los valores propios sean distintos para que A sea diagonalizable.

Ejemplo 5.9.
Consideremos $A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ los valores propios son $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 2$ y $W_1 = \left\langle \left\{ \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \right\} \right\rangle$ $W_2 = \left\langle \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \right\rangle$. Así $\mathbb{R}^3 = W_1 \oplus W_2$, y A es diagonalizable similar a $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

Definición (Multiplicidad geométrica) Sean $A \in M_{nn}(\mathbb{K})$ y λ un valor propio de A. Definimos la **multiplicidad geométrica** de λ, $\gamma_A(\lambda)$, como la dimensión del espacio propio $W_\lambda = Ker(A - \lambda I)$.

Teorema 5.4. *Una matriz $A \in M_{nn}(\mathbb{K})$ es diagonalizable ssi la suma de las multiplicidades geométricas de sus valores propios es n.*

Demostración. Basta usar el Teorema 5.3 y el hecho de que si $\lambda_1, \ldots, \lambda_k$ son los v.p.'s de A,

$$\dim(W_{\lambda_1} \oplus \cdots \oplus W_{\lambda_k}) = \sum_{k=1}^{k} \gamma_A(\lambda_i).$$

□

Definición (Multiplicidad algebraica) Sean $A \in M_{nn}(\mathbb{K})$ y λ un valor propio de A. Definimos la **multiplicidad algebraica** de λ, $\alpha_A(\lambda)$, como la máxima potencia de $(x - \lambda)$ que divide al polinomio característico de A, $p_A(x) = |A - xI|$.

Proposición 5.5. *Sean $A \in M_{nn}(\mathbb{K})$ y λ_0 un valor propio de A, entonces:*

$$1 \leq \gamma_A(\lambda_0) \leq \alpha_A(\lambda_0) \leq n.$$

Ejercicio 5.4: Demuestre la Proposición 5.5. Para ello proceda de la siguiente manera:

1. Pruebe las desigualdades $1 \leq \gamma_A(\lambda_0)$ y $\alpha_A(\lambda_0) \leq n$.

2. Sea $\gamma = \gamma_A(\lambda_0)$. Considere una base $\beta_{\lambda_0} = \{v_1, \ldots, v_\gamma\}$ de $W_{\lambda_0} = Ker(A-\lambda_0 I)$.

3. Extienda dicha base a una base $\beta = \{v_1, \ldots, v_\gamma, v_{\gamma+1}, \ldots, v_n\}$ de $\mathbb{K}^n$.

4. Calcule la matriz representante de $T(x) = Ax$, con respecto a la base β. Es decir, $M_{\beta\beta}(T)$.

5. Pruebe que
$$p_A(\lambda) = |A - \lambda I| = (\lambda - \lambda_0)^\gamma q(\lambda).$$
En donde $q(\lambda)$ es un polinomio de grado $n - \gamma$.
Indicación: Pruebe y use que A y $M_{\beta\beta}(T)$ son similares.

6. Concluya el resultado.

Corolario 5.2. *Sea $A \in M_{nn}(\mathbb{K})$ y $p_A(\lambda)$ su polinomio característico. Se tiene entonces que A es diagonalizable si y sólo si $p_A(\lambda)$ se factoriza completamente en $\mathbb{K}$ en factores lineales. Es decir:*

$$p_A(\lambda) = c_A \cdot (\lambda - \lambda_1)^{\alpha_A(\lambda_1)}(\lambda - \lambda_2)^{\alpha_A(\lambda_2)} \ldots (\lambda - \lambda_k)^{\alpha_A(\lambda_k)},$$

y además, para todo valor propio λ de A, se tiene que $\gamma_A(\lambda) = \alpha_A(\lambda)$.

Demostración. Supongamos que A es diagonalizable y sean $\lambda_1, \ldots, \lambda_k$ sus valores propios. Entonces, gracias al Teorema 5.3 tenemos que $\mathbb{K}^n = \bigoplus_{i=1}^k W_{\lambda_i}$.

Sea entonces la fatorización en $\mathbb{K}$ del polinomio característico de A:

$$p_A(\lambda) = c_A \cdot (\lambda - \lambda_1)^{\alpha_A(\lambda_1)}(\lambda - \lambda_2)^{\alpha_A(\lambda_2)} \ldots (\lambda - \lambda_k)^{\alpha_A(\lambda_k)} q(\lambda),$$

con $q(\lambda)$ un polinomio.

Pero, gracias a la Proposición 5.5,

$$n = \dim(\mathbb{K}^n) = \dim(\bigoplus_{i=1}^k W_{\lambda_i}) = \sum_{i=1}^k \gamma_A(\lambda_i) \leq \sum_{i=1}^k \alpha_A(\lambda_i).$$

De donde $\sum_{i=1}^k \alpha_A(\lambda_i) = n$.

Ahora

$$\begin{aligned} n = \mathrm{gr}(p_A(\lambda)) &= \mathrm{gr}(c_A \cdot (\lambda - \lambda_1)^{\alpha_A(\lambda_1)}(\lambda - \lambda_2)^{\alpha_A(\lambda_2)} \ldots (\lambda - \lambda_k)^{\alpha_A(\lambda_k)}) + \mathrm{gr}(q(\lambda)) \\ &= \underbrace{\sum_{i=1}^k \alpha_A(\lambda_i)}_{n} + \mathrm{gr}(q(\lambda)). \end{aligned}$$

Así, $\mathrm{gr}(q(\lambda)) = 0$. Es decir, es una constante y se tiene la fatorización buscada.

Ahora, supongamos que $p_A(\lambda)$ se factoriza completamente en $\mathbb{K}$ de la siguiente manera:

$$p_A(\lambda) = c_A \cdot (\lambda - \lambda_1)^{\alpha_A(\lambda_1)}(\lambda - \lambda_2)^{\alpha_A(\lambda_2)} \ldots (\lambda - \lambda_k)^{\alpha_A(\lambda_k)}.$$

De donde $\sum_i^k \alpha_A(\lambda_i) = n$. Y, supongamos que A no es diagonalizable. Esto implica que

$$\dim(\bigoplus_{i=1}^{k} W_{\lambda_i}) = \sum_{i=1}^{k} \gamma_A(\lambda_i) < n = \sum_{i=1}^{k} \alpha_A(\lambda_i).$$

Lo cual contradice la hipótesis de igualdad de las multiplicidades. $\square$

Corolario 5.3. $A \in M_{nn}(\mathbb{C})$ *es diagonalizable si y sólo si para todo valor propio* λ *de* A*, se tiene que*

$$\gamma_A(\lambda) = \alpha_A(\lambda).$$

Guía de Ejercicios

1. Demostrar que dados V espacio vectorial y $Z, U_1, \ldots, U_k$ subespacios vectoriales de V, entonces:

(a) $Z = \bigoplus_{i=1}^{k} U_i \iff \left(Z = \mathop{+}_{i=1}^{k} U_i \ \wedge\ \forall j \in \{1, \ldots, k\},\ U_j \cap \left(\mathop{+}_{\substack{i=1 \\ i \neq j}}^{k} U_i \right) = \{0\} \right).$

(b) Si $Z = \mathop{+}_{i=1}^{k} U_i$ y Z es de dimensión finita, entonces son equivalentes:

(1) $Z = \bigoplus_{i=1}^{k} U_i$.

(2) $(\forall i \in \{1, \ldots, k\})(\forall u_i \in U_i \setminus \{0\})\ \{u_1, \ldots, u_k\}$ es l.i.

(3) La yuxtaposición de bases de los subespacios U_i es una base (y no sólo un generador) de Z.

(4) $\dim(Z) = \sum_{i=1}^{k} \dim(U_i)$.

2. Pruebe que, dados $A \in M_{nn}(\mathbb{K})$ y λ_0 un valor propio de A, entonces:

$$1 \leq \gamma_A(\lambda_0) \leq \alpha_A(\lambda_0) \leq n.$$

Para ello

(a) Pruebe las desigualdades $1 \leq \gamma_A(\lambda_0)$ y $\alpha_A(\lambda_0) \leq n$.

(b) Sea $\gamma = \gamma_A(\lambda_0)$. Considere una base $\beta_{\lambda_0} = \{v_1, \ldots, v_\gamma\}$ de $W_{\lambda_0} = Ker(A - \lambda_0 I)$.

(c) Extienda dicha base a una base $\beta = \{v_1, \ldots, v_\gamma, v_{\gamma+1}, \ldots, v_n\}$ de $\mathbb{K}^n$.

(d) Calcule la matriz representante de $T(x) = Ax$, con respecto a la base β. Es decir, $M_{\beta\beta}(T)$.

(e) Pruebe que

$$p_A(\lambda) = |A - \lambda I| = (\lambda - \lambda_0)^\gamma q(\lambda).$$

En donde $q(\lambda)$ es un polinomio de grado $n - \gamma$.
Indicación: Pruebe y use que A y $M_{\beta\beta}(T)$ son similares.

(f) Concluya el resultado.

Guía de Problemas

P1. **(a)** Sea $\mathcal{P}_2(\mathbb{R})$ el espacio vectorial de los polinomios a coeficientes reales de grado menor o igual a 2, y sea $T : \mathcal{P}_2(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$ la transformación lineal definida por

$$T(a_0 + a_1 x + a_2 x^2) = (a_0 + 2a_2) + a_1 x + (2a_0 + a_2)x^2.$$

(1) Verifique que la matriz representante de T con respecto a la base canónica $\{1, x, x^2\}$ de $\mathcal{P}_2(\mathbb{R})$ (en dominio y recorrido) es

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

(2) Calcule A^{2n} para cada $n \geq 1$ natural, diagonalizando A.

(3) Calcule $T^{2n}(1 + x + x^2)$, donde $T^{2n} = \underbrace{T \circ T \circ \cdots \circ T}_{2n}$.

Indicación: Recuerde que la matriz representante de T^{2n} es A^{2n}.

(b) Encuentre los valores reales de α y β de modo que la matriz

$$\begin{pmatrix} 2 & \alpha & 0 & 0 & 0 \\ 0 & 2 & \alpha & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & \beta \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

sea diagonalizable.

P2. Sea $a \in \mathbb{R}$. Se define una sucesión de números reales $(u_n)_{n\in\mathbb{N}}$ de la siguiente manera: $u_0 = 1$, $u_1 = a$, $(\forall n \geq 2)$ $u_n = au_{n-1} - u_{n-2}$.

Sean $x_0 = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, $(\forall n \geq 1)$ $x_n = \left(\begin{smallmatrix} u_n \\ u_{n-1} \end{smallmatrix}\right)$.

(a) Demuestre que para todo $n \geq 0$ se tiene que $x_n = A^n x_0$ con $A = \left(\begin{smallmatrix} a & -1 \\ 1 & 0 \end{smallmatrix}\right)$.

(b) Demuestre que si $|a| = 2$ entonces A no es diagonalizable.

(c) Demuestre que si $|a| > 2$ entonces A es diagonalizable.

(d) Asuma que $|a| > 2$ y denote $\lambda_1,\ \lambda_2$ los valores propios de A. Demuestre que

(1) $\left(\begin{smallmatrix} \lambda_1 \\ 1 \end{smallmatrix}\right)$ y $\left(\begin{smallmatrix} \lambda_2 \\ 1 \end{smallmatrix}\right)$ son vectores propios asociados a λ_1 y λ_2, respectivamente.

(2) Para todo $n \in \mathbb{N}$ se tiene que $u_n = \frac{\lambda_1^{n+1} - \lambda_2^{n+1}}{\lambda_1 - \lambda_2}$.

P3. Sean $R, S \in M_{nn}(\mathbb{R})$, con S invertible. Considere $A = RS$ y $B = SR$.

(a) Pruebe que $v \in \mathbb{R}^n$ es vector propio de A asociado al valor propio $\lambda \in R$ ssi Sv es vector propio de B asociado al mismo valor propio λ. Concluya que A y B tienen los mismos valores propios.

(b) Sean $W_\lambda(A) = \text{Ker}(A - \lambda I)$ el subespacio propio de A asociado al valor propio $\lambda \in \mathbb{R}$ y $W_\lambda(B)$ el subespacio propio de B asociado a λ. Pruebe que $\dim(W_\lambda(A)) = \dim(W_\lambda(B))$.

(c) Pruebe que A es diagonalizable ssi B es diagonalizable.

Ortogonalidad

6.1 Conjuntos ortogonales y ortonormales

Recordemos que la proyección de u sobre $v \neq 0$ está definida por $w = \frac{\langle u,v\rangle}{\langle v,v\rangle} v$. Este vector w es "el que está más cerca" de u en el espacio $\langle \{v\} \rangle$. Es decir

$$\min_{\lambda \in \mathbb{R}} \|u - \lambda v\| = \|u - w\|.$$

Notaremos por $u \perp v$ si $\langle u, v\rangle = 0$, es decir si u y v son ortogonales.

Definición (Conjunto ortogonal/ortonormal) Un conjunto $\{v_1, \ldots, v_k\}$ de $\mathbb{R}^n$ se dice **ortogonal** si $\langle v_i, v_j\rangle = 0$ para $i \neq j$. Si además $\|v_i\| = \langle v_i, v_i\rangle^{\frac{1}{2}} = 1$, diremos que el conjunto es **ortonormal**.

En el caso en que $\{v_1, \ldots, v_k\}$ es además base de un subespacio vectorial W, diremos que se trata de una **base ortonormal** de W.

Ejemplo 6.1.
La base canónica de $\mathbb{R}^n$ es una base ortonormal.

Veamos una propiedad útil de los conjuntos ortogonales:

Proposición 6.1. *Sea $\{v_1, \ldots, v_k\} \subseteq \mathbb{R}^n \setminus \{0\}$ ortogonal, entonces $\{v_1, \ldots, v_k\}$ es l.i.*

Demostración. Sea $\sum_{i=1}^k \alpha_i v_i$ una combinación lineal nula de los elementos de $\{v_1, \ldots, v_k\}$. Sea $j \in \{1, \ldots, k\}$, luego por propiedades del producto interno:

$$\left\langle \sum_{i=1}^k \alpha_i v_i, v_j \right\rangle = 0 \iff \sum_{i=1}^k \alpha_i \langle v_i, v_j\rangle = 0.$$

Pero dado que el conjunto es ortogonal, luego $\langle v_i, v_j\rangle = 0$ para $i \neq j$, de manera que

$$\alpha_j \langle v_j, v_j\rangle = \alpha_j \|v_j\|^2 = 0.$$

Y como $v_j \neq 0$, se deduce que $\alpha_j = 0$.

Como este procedimiento es válido para todo $j \in \{1, \ldots, k\}$, se concluye que $\{v_1, \ldots, v_k\}$ es l.i. $\square$

Ahora, en el caso de conjuntos ortonormales, hay propiedades importantes que motivan su búsqueda.

Proposición 6.2. *Sea W un subespacio vectorial de $\mathbb{R}^n$ y $\{u_1, \ldots, u_k\}$ una base ortonormal de W, entonces:*

1. *Si $x \in W$,*
$$x = \sum_{i=1}^{k} \langle x, u_i \rangle u_i.$$
2. *Si $x \in \mathbb{R}^n$ y definimos*
$$z = \sum_{i=1}^{k} \langle x, u_i \rangle u_i \in W,$$
entonces $(x - z) \perp w,\ \forall w \in W$. Y en particular $\mathrm{d}(x, z) = \min_{w \in W} \mathrm{d}(x, w)$, de hecho $\forall w \in W \setminus \{z\},\ \mathrm{d}(x, z) < \mathrm{d}(x, w)$.

Demostración.

1. Como $\{u_1, \ldots, u_k\}$ es una base de W y $x \in W$, luego
$$x = \sum_{i=0}^{k} \alpha_i u_i.$$
Tomemos $j \in \{1, \ldots, k\}$ y calculemos el producto interno entre x y u_j:
$$\begin{aligned} \langle x, u_j \rangle &= \left\langle \sum_{i=0}^{k} \alpha_i u_i, u_j \right\rangle \\ &= \sum_{i=0}^{k} \alpha_i \langle u_i, u_j \rangle . \end{aligned}$$
Como el conjunto $\{u_1, \ldots, u_k\}$ es ortogonal, se tiene que $\langle u_i, u_j \rangle = 0,\ \forall i \neq j$, de donde
$$\langle x, u_j \rangle = \alpha_j \langle u_j, u_j \rangle = \alpha_j \cdot 1.$$
Repitiendo este procedimiento para todo $j \in \{1, \ldots, k\}$, se concluye que $\alpha_j = \langle x, u_j \rangle$
$$x = \sum_{i=1}^{k} \langle x, u_i \rangle u_i.$$
2. Sea $x \in \mathbb{R}^n$, $z = \sum_{i=1}^{k} \langle x, u_i \rangle u_i$ y $w \in W$, veamos:
$$\langle x - z, w \rangle = \langle x, w \rangle - \langle z, w \rangle = \langle x, w \rangle - \left\langle \sum_{i=1}^{k} \langle x, u_i \rangle u_i, \sum_{j=1}^{k} \langle w, u_j \rangle u_j \right\rangle,$$

la última igualdad gracias a que $w \in W$. Así, usando las propiedades del producto interno:

$$\langle x - z, w \rangle = \langle x, w \rangle - \sum_{i=1}^{k} \sum_{j=1}^{k} \langle x, u_i \rangle \langle w, u_j \rangle \langle u_i, u_j \rangle$$

y como $\{u_1, \ldots, u_k\}$ es ortonormal, $\langle u_i, u_j \rangle = 0,\ \forall i \neq j$ y

$$\begin{aligned}
\langle x - z, w \rangle &= \langle x, w \rangle - \sum_{i=1}^{k} \langle x, u_i \rangle \langle w, u_i \rangle \langle u_i, u_i \rangle \\
&= \langle x, w \rangle - \sum_{i=1}^{k} \langle x, u_i \rangle \langle w, u_i \rangle \\
&= \langle x, w \rangle - \left\langle x, \underbrace{\sum_{i=1}^{k} \langle w, u_i \rangle u_i}_{w} \right\rangle = 0.
\end{aligned}$$

Para probar que $\mathrm{d}(x, z) = \min_{w \in W} \mathrm{d}(x, w)$, sea $w \in W$ distinto de z, luego por el Teorema de Pitágoras en $\mathbb{R}^n$, probado en el Ejercicio 2.4:

$$\|x - z\|^2 + \|w - z\|^2 = \|x - w\|^2 .$$

Pero como $\|w - z\|^2 > 0$, luego

$$\|x - z\|^2 < \|x - w\|^2 .$$

De esta manera

$$\forall w \in W,\ \mathrm{d}(x, z) = \|x - z\| < \|x - w\| = \mathrm{d}(x, w),$$

de donde se concluye. $\square$

6.2 Proyecciones Ortogonales

La Proposición 6.2 motiva la definición proyección ortogonal sobre un subespacio vectorial de $\mathbb{R}^n$. Asumiremos por ahora el siguiente resultado, el cual demostraremos en la Sección 6.4, que asegura la existencia de una base ortonormal de cualquier subespacio vectorial de $\mathbb{R}^n$.

Proposición 6.3. *Todo subespacio vectorial de $\mathbb{R}^n$ posee una base ortonormal.*

Definimos entonces:

Definición (Proyección ortogonal sobre un s.e.v.) Sea W un subespacio vectorial $\mathbb{R}^n$ y $\{u_1, \ldots, u_k\}$ una base ortonormal de W. Definimos la **proyección ortogonal sobre** W como la función

$$\begin{aligned} P : \mathbb{R}^n &\longrightarrow W \\ x &\longmapsto P(x) = \sum_{i=1}^{k} \langle x, u_i \rangle \, u_i \in W. \end{aligned}$$

Observación: Notemos que esta definición es correcta, es decir que P no depende de la base ortonormal considerada y que es función.

Gracias a la Proposición 6.2.2, sabemos que independiente de la base ortonormal considerada para definirla, la proyección es aquella que minimiza la distancia de x a W. Así, si hubiese dos proyecciones z_1, z_2, asociadas a bases distintas y fuesen distintas, luego:

$$\mathrm{d}(x, z_1) < \mathrm{d}(x, z_2),$$

lo cual contradice la propiedad que satisface z_2.

De la misma manera se prueba que P es función.

Veamos ahora las propiedades de la proyección ortogonal:

Proposición 6.4. *Sea W un subespacio vectorial de $\mathbb{R}^n$ y P su proyección ortogonal asociada, entonces:*

1. *P es una transformación lineal.*
2. *$\forall x \in \mathbb{R}^n, \forall w \in W, \quad (x - P(x)) \perp w$.*
3. *$\mathrm{d}(x, P(x)) = \min_{w \in W} \mathrm{d}(x, w)$.*

Demostración. 2 y 3 se deducen de la Proposición 6.2. La demostración de 1 queda de ejercicio .

6.3 Subespacio Ortogonal

Definición (Subespacio ortogonal) Sea W un subespacio de $\mathbb{R}^n$ definamos el **ortogonal** de W como:

$$W^{\perp} = \{u \in \mathbb{R}^n \mid \forall w \in W \quad \langle w, u \rangle = 0\}.$$

Proposición 6.5.

(i) $W^\perp$ *es un subespacio de* $\mathbb{R}^n$.

(ii) $\mathbb{R}^n = W \oplus W^\perp$.

(iii) $\dim(W^\perp) = n - \dim(W)$.

Demostración.

(i) Sea $u, v \in W^\perp$ y $\lambda \in \mathbb{R}$, debemos probar que $\lambda u + v \in W^\perp$. Sea $w \in W$:

$$\langle w, \lambda u + v \rangle = \lambda \langle w, u \rangle + \langle w, v \rangle = 0$$

lo que es válido para todo $w \in W$. Luego $\lambda u + v \in W^\perp$ por lo tanto $W^\perp$ es un subespacio vectorial de $\mathbb{R}^n$.

(ii) Es claro que $W + W^\perp \subseteq \mathbb{R}^n$. Para la otra inclusión, dado $x \in \mathbb{R}^n$, basta notar que

$$x = P(x) + (x - P(x)).$$

En donde P es la proyección ortogonal sobre W. Así, $P(x) \in W$ y $x - P(x) \in W^\perp$ (gracias a la Proposición 6.4), de donde $x \in W + W^\perp$.

Nos basta entonces probar que $W \cap W^\perp = \{0\}$. En efecto si $x \in W \cap W^\perp$, en particular es ortogonal a sí mismo, es decir $\langle x, x \rangle = 0$, de donde $x = 0$.

Se concluye entonces que $\mathbb{R}^n = W \oplus W^\perp$.

(iii) Se deduce directamente de (ii). □

6.4 Método de Gram-Schmidt

El objetivo de esta parte es asociar de forma "canónica" a un conjunto dado de $\mathbb{R}^n$, una base ortonormal del subespacio generado por él. Este método es conocido con el nombre de Método de Gram-Schmidt.

Probaremos el siguiente teorema:

Teorema 6.1 (Gram-Schmidt). *Dado un conjunto* $\{v_0, \dots, v_m\} \subseteq \mathbb{R}^n$, *existe un conjunto ortonormal* $\{u_0, \dots, u_k\}$ *tal que*

$$\langle \{v_0, \dots, v_m\} \rangle = \langle \{u_0, \dots, u_k\} \rangle .$$

Algoritmo 1 Método Gram-Schmidt

1: Partamos con $u_1 = \frac{v_1}{\|v_1\|}$

2: Luego, proyectamos v_2 sobre el subespacio generado por u_1:

$$z_2 = \langle v_2, v_1 \rangle \, u_1$$

y restamos esta proyección a v_2:

$$w_2 = v_2 - z_2$$

perpendicular a $\langle \{u_1\} \rangle$ y normalizamos, $u_2 = \frac{w_2}{\|w_2\|}$

3: Luego podemos pasar a trabajar con v_2

$$\begin{aligned} z_3 &= \langle v_3, u_1 \rangle \, u_1 + \langle v_3, u_1 \rangle \, u_2 \\ w_3 &= v_3 - z_3, \end{aligned}$$

perpendicular a $\langle \{u_1, u_2\} \rangle$ y normalizamos, $u_3 = \frac{w_3}{\|w_3\|}$

4: se sigue del mismo modo: una vez obtenidos $\{u_1, \ldots, u_k\}$, proyectamos v_{k+1} sobre $\langle \{u_1, \ldots, u_k\} \rangle$ y definimos

$$\begin{aligned} z_{k+1} &= \sum_{i=0}^{k} \langle v_{k+1}, u_i \rangle \, u_i \\ w_{k+1} &= v_{k+1} - z_{k+1}, \end{aligned}$$

perpendicular a $\langle \{u_1, \ldots, u_k\} \rangle$ y normalizamos, $u_{k+1} = \frac{w_{k+1}}{\|w_{k+1}\|}$

5: El proceso termina cuando hemos procesado todos los v_i.

Demostración. La demostración de este teorema es precisamente el método Gram-Schmidt, que entrega un procedimiento algorítmico para construir el conjunto buscado. Sea entonces $\{v_0, \dots, v_m\}$ un conjunto de vectores y definimos el conjunto $\{u_0, \dots, u_k\}$ mediante el Algoritmo 1.

Probemos que esta definición del $\{u_0, \dots, u_k\}$ cumple lo pedido. Lo haremos por inducción en m.

- **Caso base.** Para $m = 0$, si $v_0 = 0$ luego $w_0 = 0$ y el algoritmo termina sin generar vectores u_j. Esto es consistente con que la única base de $\{0\}$ es el conjunto vacío.

  Si $v_0 \neq 0$, por ende $w_0 \neq 0$, se tiene el resultado ya que
  $$u_0 = \frac{v_0}{\|v_0\|},$$
  que es claramente de norma 1 y además, como es ponderación de v_0, se tiene que $\langle \{u_0\} \rangle = \langle \{v_0\} \rangle$.

- **Paso inductivo.** Supongamos que dado el conjunto $\{v_0, \dots, v_{m-1}\}$, el conjunto $\{u_0, \dots, u_{k-1}\}$ es ortonormal y tal que
  $$\langle \{v_0, \dots, v_{m-1}\} \rangle = \langle \{u_0, \dots, u_{k-1}\} \rangle .$$
  Con $k = \dim(\langle \{v_0, \dots, v_{m-1}\} \rangle)$.

  Veamos los dos casos posibles:

  Si $w_k = 0$, se tiene que $v_m = \sum_{l=0}^{k-1} \langle v_m, u_l \rangle u_l$, de donde $v_k \in \langle \{u_0, \dots, u_{k-1}\} \rangle$ y por ende
  $$\langle \{v_0, \dots, v_m\} \rangle = \langle \{u_0, \dots, u_{k-1}\} \rangle .$$
  Y la base ortonormal generada es $\{u_0, \dots, u_{k-1}\}$.

  Ahora si $w_k \neq 0$, según el algoritmo, u_k estará definido por:
  $$u_k = \frac{w_k}{\|w_k\|} = \frac{1}{\|w_k\|} \left(v_m - \sum_{l=0}^{k-1} \langle v_m, u_l \rangle u_l \right) .$$

  Veamos entonces, para cualquier $i \in \{0, \dots, k-1\}$
  $$\begin{aligned} \langle u_k, u_i \rangle &= \left\langle \frac{1}{\|w_k\|} \left(v_m - \sum_{l=0}^{k-1} \langle v_m, u_l \rangle u_l \right), u_i \right\rangle \\ &= \frac{1}{\|w_k\|} \left(\langle v_k, u_i \rangle - \sum_{l=0}^{k-1} \langle v_m, u_l \rangle \langle u_l, u_i \rangle \right) \end{aligned}$$
  y por hipótesis de inducción, $\forall l \neq i,\ \langle u_l, u_i \rangle = 0$ y $\langle u_i, u_i \rangle = 1$, luego
  $$\begin{aligned} \langle u_k, u_i \rangle &= \frac{1}{\|w_k\|} (\langle v_m, u_i \rangle - \langle v_m, u_i \rangle \cdot 1) \\ &= 0. \end{aligned}$$

Es decir, el conjunto $\{u_0, \ldots, u_k\}$ es ortogonal, y como es claro que $\forall i \in \{0, \ldots, k\}$ $\|u_i\| = 1$, es además ortonormal.

Ahora, por construcción de u_k se tiene

$$v_m = \|w_k\| \, u_k + \sum_{l=0}^{k-1} \langle v_m, u_l \rangle \, u_l$$

es decir, v_m es combinación lineal de $\{u_0, \ldots, u_k\}$, lo que equivale a $v_m \in \langle \{u_0, \ldots, u_k\} \rangle$. Así, gracias a la hipótesis inductiva, se tiene que

$$\langle \{v_0, \ldots, v_m\} \rangle \subseteq \langle \{u_0, \ldots, u_k\} \rangle . \tag{6.1}$$

Por otra parte, u_k es combinación lineal de v_m y $u_0, \ldots, u_{k-1}$, ya que

$$u_k = \frac{1}{\|w_k\|} v_m - \sum_{l=0}^{k-1} \frac{\langle v_m, u_l \rangle}{\|w_k\|} u_l .$$

Pero gracias a (6.1), existen $\alpha_0, \ldots, \alpha_m \in \mathbb{R}$ tales que $\sum\limits_{l=0}^{k-1} \frac{\langle v_m, u_l \rangle}{\|w_k\|} u_l = \sum\limits_{l=0}^{m} \alpha_l u_l$. Así $u_k \in \langle \{v_0, \ldots, v_m\} \rangle$ y luego

$$\langle \{u_0, \ldots, u_k\} \rangle \subseteq \langle \{v_0, \ldots, v_m\} \rangle .$$

Obteniendo lo buscado.

Luego, por Principio de Inducción, se tiene el teorema. □

Observación:

- Notemos que como corolario del Teorema 6.1, se obtiene la Proposición 6.3.
- Además, el Teorema 6.1 puede ser aplicado en particular a una base $\{v_0, \ldots, v_{n-1}\}$ de $\mathbb{R}^n$.

Observación: Observar que si alguno de los vectores a normalizar, v_1 o los w_i, vale 0, significa que el v_i correspondiente no aportará un nuevo vector a la base, y simplemente se descarta, pasándose a trabajar con v_{i+1}.

Esto trae como consecuencia que el número de vectores u_i sea menor que el de los v_j si estos últimos son l.d.

En general, esto hará que al procesar el k-ésimo v_k, estemos produciendo un u_i, con $i < k$, pero no hemos introducido esto en la descripción del algoritmo para no confundir la notación.

Ejemplo 6.2.

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\} \text{ en } \mathbb{R}^3.$$

$u_0 = v_0/\langle v_0, v_0\rangle^{\frac{1}{2}}$, como $\langle v_0, v_0\rangle = 1$ se tiene $u_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

$$\langle w_1, w_1\rangle = 2 \qquad \Rightarrow u_1 = \frac{1}{\sqrt{2}}\, w_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{aligned} w_2 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} \cdot 2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

$w_2 = 0$, por ende ignoramos $\left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right)$ y redefinimos w_2.

$$\begin{aligned} w_2 &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}. \end{aligned}$$

Como $\langle w_2, w_2\rangle = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$, luego

$$u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

6.5 Producto Hermítico

Ahora, extendemos el producto interno de $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ a $\langle \cdot, \cdot \rangle_H : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$, de la siguiente forma.

DEFINICIÓN (PRODUCTO HERMÍTICO) Dados $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{C}^n$,

$$\langle u, v \rangle_H = \sum_{j=1}^{n} u_j \overline{v_j}.$$

Ejercicio 6.1: Probar que el producto hermítico satisface: $\forall \lambda \in \mathbb{C}, u, v, w \in \mathbb{C}^n$

1. $\langle u, v \rangle_H = \overline{\langle v, u \rangle}_H$.
2. $\langle \lambda u + v, w \rangle_H = \lambda \langle u, w \rangle_H + \langle v, w \rangle_H$.
3. $\langle u, u \rangle_H \in \mathbb{R}$, más aún $\langle u, u \rangle_H \geq 0$ y $\langle u, u \rangle_H = 0 \iff u = 0$.

De 2a y 2b, se deduce que

$$\langle u, \lambda v + w \rangle_H = \bar{\lambda} \langle u, v \rangle_H + \langle u, w \rangle_H .$$

Observación: Los conceptos de ortogonalidad, proyección ortogonal, subespacio ortogonal y el método de Gram-Schmidt, pueden desarrollarse también en $\mathbb{C}^n$, como espacio vectorial sobre $\mathbb{C}$.

Guía de Ejercicios

1. Dado W un s.e.v. de $\mathbb{R}^n$ y $P\colon \mathbb{R}^n \to W$ su proyección ortogonal asociada. Probar que P es lineal.

2. Demuestre las siguientes propiedades del producto hermítico:

(a) $\langle u, v\rangle_H = \overline{\langle v, u\rangle}_H$.

(b) $\langle \lambda u + v, w\rangle_H = \lambda \langle u, w\rangle_H + \langle v, w\rangle_H$.

(c) $\langle u, u\rangle_H \in \mathbb{R}$, más aún $\langle u, u\rangle_H \geq 0$ y $\langle u, u\rangle_H = 0 \iff u = 0$.

Guía de Problemas

P1. Sea W el subespacio vectorial de $\mathbb{R}^4$ generado por:

$$\left\{ \begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}, \begin{pmatrix}1\\3\\1\\3\end{pmatrix} \begin{pmatrix}2\\3\\2\\3\end{pmatrix} \right\}.$$

(a) Encuentre una base ortonormal de W y de $W^\perp$.

(b) Encuentre la descomposición de $v = \left(\begin{smallmatrix}1\\2\\3\\4\end{smallmatrix}\right)$ en $W + W^\perp$.

P2. Sea $A \in M_{nn}(\mathbb{R})$ una matriz invertible y U un s.e.v. de $\mathbb{R}^n$ de dimensión $0 < k < n$. Se define

$$W = \{y \in \mathbb{R}^n \mid \forall u \in U,\ \langle Au, Ay\rangle = 0\}.$$

(a) Probar que W es s.e.v. de $\mathbb{R}^n$ y probar que $W \cap U = \{0\}$.

(b) Sea $V = \{v \in \mathbb{R}^n \mid \exists u \in U,\ v = Au\}$. Pruebe que V es s.e.v. de $\mathbb{R}^n$.

(c) Sea $\{v_1, \ldots, v_k\}$ una base ortonormal de V. Probar que $\{u_1, \ldots, u_k\}$ donde $u_i = A^{-1}v_i$, $i \in \{1, \ldots, k\}$, es una base de U que verifica la propiedad: $\langle Au_i, Au_j\rangle = 0$ si $i \neq j$ y $\langle Au_i, Au_j\rangle = 1$ si $i = j$.

(d) Probar que

$$w \in W \iff \forall i \in \{1, \ldots, k\},\ \langle Aw, Au_i\rangle = 0.$$

Donde $\{u_1 \ldots, u_k\}$ es la base del punto anterior.

(e) Probar que si $v \in \mathbb{R}^n$ entonces $z = \sum_{i=1}^{k} \langle Av, Au_i\rangle u_i \in U$ y que $v - z \in W$.

(f) Deducir que $\mathbb{R}^n = U \oplus W$ y calcular la dimensión de W.

P3. **(a)** Sea $W = \left\langle \left\{ \begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}2\\1\\1\\0\end{pmatrix} \right\} \right\rangle$.

(a.1) Encuentre una base ortonormal de $W^{\perp}$.

(a.2) Sean $x_1, x_2, x_3, x_4 \in \mathbb{R}$. Encuentre (en términos de x_1, x_2, x_3, x_4) los valores y_1, y_2, y_3, y_4 que satisfacen $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$, donde $T\colon \mathbb{R}^4 \to \mathbb{R}^4$ es una transformación lineal para la cual $Ker(T) = W$ y, para todo $x \in W^{\perp}$, $T(x) = x$.

(b) Sean U, V subespacios vectoriales de $\mathbb{R}^n$. Demuestre que

(b.1) $(U + V)^{\perp} = U^{\perp} \cap V^{\perp}$.

(b.2) $(U^{\perp})^{\perp} = U$.

(b.3) $(U \cap V)^{\perp} = U^{\perp} + V^{\perp}$.

(b.4) $U \oplus V = \mathbb{R}^n \iff U^{\perp} \oplus V^{\perp} = \mathbb{R}^n$.

Ingeniería Matemática
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Ortogonalidad

6.6 Espectro de una matriz simétrica

Sea $A \in M_{nn}(\mathbb{R})$ una matriz simétrica entonces

$$\forall u, v \in \mathbb{R}^n : \langle Au, v\rangle = \langle u, Av\rangle. \tag{6.2}$$

En efecto, $\langle Au, v\rangle = (Au)^t v = u^t A^t v = u^t Av = \langle u, Av\rangle$

Probemos ahora que si $A \in M_{nn}(\mathbb{R})$ satisface 6.2 entonces A es simétrica. Esto se deduce del siguiente hecho fácil de probar: si $\{e_1, \ldots, e_n\}$ la base canónica de $\mathbb{R}^n$ y A es una matriz de $n \times n$ entonces

$$\langle Ae_i, e_j\rangle = a_{ji}. \tag{6.3}$$

Utlizando esta propiedad y 6.2 se obtiene que

$$a_{ij} = \langle Ae_j, e_i\rangle = \langle e_j, Ae_i\rangle = \langle Ae_i, e_j\rangle = a_{ji},$$

donde además se usa el hecho que el producto interno en $\mathbb{R}^n$ es simétrico.

Dada $A \in M_{nn}(\mathbb{C})$, definimos la matriz adjunta de A, A^*, de la siguiente forma:

$$\forall k, \ell : (A^*)_{k\ell} = \overline{a_{\ell k}}.$$

Ejemplo 6.3.

$$A = \begin{pmatrix} 1 & i \\ 2+i & 3 \end{pmatrix} \quad A^* = \begin{pmatrix} 1 & 2-i \\ -i & 3 \end{pmatrix}$$

Decimos que A es **hermítica** si $A = A^*$, notar que si $A \in M_{nn}(\mathbb{R})$ entonces $A = A^*$ si y sólo so $A = A^t$. De forma análoga al caso simétrico se prueba que

$$\forall u, v \in \mathbb{C}^n \;\; \langle Au, v\rangle_H = \langle u, Av\rangle_H$$

si y sólo si A es hermítica. Así, en particular, $\forall u, v \in \mathbb{C}^n \; \langle Au, v\rangle_H = \langle u, Av\rangle_H$ si $A \in M_{nn}(\mathbb{R})$ y A es simétrica.

Sea $A \in M_{nn}(\mathbb{C})$ entonces el polinomio característico $p(\lambda) = |A - \lambda I|$ es un polinomio de grado n con coeficientes complejos y por lo tanto tiene n raíces en $\mathbb{C}$ (con posibles repeticiones). Denotaremos por

$$\sigma(A) = \{\lambda \in \mathbb{C} \mid |A - \lambda I| = 0\}$$

al conjunto de las raices del polinomio característico. Este conjunto lo llamaremos el **espectro** de A. Como $M_{nn}(\mathbb{R}) \subseteq M_{nn}(\mathbb{C})$ podemos definir de igual manera el espectro de A, para $A \in M_{nn}(\mathbb{R})$.

En general el espectro de A es un subconjunto de $\mathbb{C}$, aún cuando $A \in M_{nn}(\mathbb{R})$, como lo muestra el próximo ejemplo.

Ejemplo 6.4.

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

$$B - \lambda I = \begin{pmatrix} -\lambda & 1 & 0 \\ -1 & -\lambda & 0 \\ 0 & 0 & 1-\lambda \end{pmatrix} \qquad |B - \lambda I| = (1-\lambda)\{\lambda^2 + 1\}$$

Las raíces son $\lambda = 1 \quad \lambda = \pm i$. Si B es considerada como matriz de coeficientes reales (es decir estamos trabajando en $\mathbb{R}^n$) B tiene un sólo valor propio que es 1. Si el cuerpo considerado es $\mathbb{C}$ entonces B tiene 3 valores propios $1, i, -i$. En cualquier caso el espectro de B es $\sigma(B) = \{1, i, -i\}$.

Teorema 6.2. *Sea $A \in M_{nn}(\mathbb{C})$ hermítica entonces el espectro de A es subconjunto de $\mathbb{R}$.*

Demostración. Sea $\lambda \in \sigma(A)$ (en general λ será complejo) entonces, $|A - \lambda I| = 0$, es decir $A - \lambda I$ no es invertible, luego debe existir $v \in \mathbb{C}^n, v \neq 0$ tal que $(A - \lambda I)v = 0$. Por lo tanto

$$\langle Av, v\rangle_H = \langle \lambda v, v\rangle_H = \lambda \langle v, v\rangle_H,$$

pero

$$\langle Av, v\rangle_H = \langle v, Av\rangle_H = \langle v, \lambda v\rangle_H = \overline{\lambda} \langle v, v\rangle_H$$

como $v \neq 0$, $\langle v, v\rangle_H \neq 0$ entonces $\lambda = \overline{\lambda}$ es decir $\lambda \in \mathbb{R}$. Se tiene que para A hermítica el polinomio característico $p(\lambda) = |A - \lambda I|$ tiene todas sus raíces reales. $\square$

Corolario 6.1. *Si $A \in M_{nn}(\mathbb{R})$ es simétrica entonces $\sigma(A) \subseteq \mathbb{R}$.*

Teorema 6.3. *Si $A \in M_{nn}(\mathbb{C})$ es hermítica y $\lambda_1 \neq \lambda_2$ son valores propios y v_1, v_2 son vectores propios asociados a λ_1 y λ_2 respectivamente entonces $\langle v_1, v_2\rangle_H = 0$.*

Demostración. Consideremos $\langle Av_1, v_2\rangle_H = \langle \lambda_1 v_1, v_2\rangle_H = \lambda_1 \langle v_1, v_2\rangle_H$, pero

$$\langle Av_1, v_2\rangle_H = \langle v_1, Av_2\rangle_H = \langle v_1, \lambda_2 v_2\rangle_H = \overline{\lambda_2} \langle v_1, v_2\rangle_H$$

pero $\lambda_2 \in \mathbb{R}$ y por tanto $(\lambda_1 - \lambda_2) \langle v_1, v_2\rangle_H = 0$, como $\lambda_1 \neq \lambda_2$, $\langle v_1, v_2\rangle_H = 0$. $\square$

Notemos que si $u, v \in \mathbb{R}^n \subseteq \mathbb{C}^n$, entonces $\langle u, v\rangle_H = \langle u, v\rangle$, por lo tanto se obtiene el siguiente corolario importante.

Corolario 6.2. *Vectores propios de $A \in M_{nn}(\mathbb{R})$, simétrica, asociados a valores propios distintos son ortogonales.*

Definición (Transformación simétrica) Sea W un subespacio vectorial de $\mathbb{R}^n$ y $L : W \to W$ una aplicación lineal, diremos que L es simétrica (en W) si

$$\forall u, v \in W : \langle L(u), v\rangle = \langle u, L(v)\rangle$$

Lema 2. *$L : \mathbb{R}^n \to \mathbb{R}^n$ es simétrica si y sólo si la matriz representante con respecto a la base canónica de $\mathbb{R}^n$ es simétrica.*

Demostración. Si A es la matriz representante de L con respecto a la base canónica entonces $L(u) = Au$ de donde

$$\langle L(u), v\rangle = \langle Au, v\rangle \;\; \langle u, L(v)\rangle = \langle u, Av\rangle,$$

luego L es simétrica si y sólo si A es simétrica. $\square$

Ejercicio 6.2: Sea W subespacio vectorial de $\mathbb{R}^n$ y $L\colon W \to W$ lineal, probar que:

1. Si $B = \{u_1, \ldots, u_n\}$ es una base **ortonormal** de W, entonces
$$L \text{ simétrica} \iff M_{BB}(L) \text{ simétrica.}$$
2. Si $B = \{u_1, \ldots, u_n\}$ es una base de W y
$$\begin{aligned} \varphi\colon W &\longrightarrow \mathbb{R}^n \\ x = \textstyle\sum_{i=0}^n \alpha_i u_i &\longmapsto \varphi(x) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}. \end{aligned}$$
Entonces x es vector propio de L asociado al v.p. λ si y sólo si $\varphi(x)$ es vector propio de $M_{BB}(L)$ asociado al v.p. λ.

Ahora estamos en posición de probar uno de los teoremas más importantes de la sección.

Teorema 6.4. *Sea W subespacio de $\mathbb{R}^n$, $\dim W \geq 1$ y $L : W \to W$ lineal, simétrica, entonces existe una base ortonormal de W de vectores propios de L.*

Demostración. Por inducción sobre la dimensión de W. Si $\dim W = 1$. Entonces $W = \langle \{v\} \rangle$, con $v \neq 0$. Notemos que

$$L(v) \in W = \langle \{v\} \rangle \, ,$$

luego existe $\lambda \in \mathbb{R}$ tal que $L(v) = \lambda v$. Es decir, λ es un valor propio de L y v vector propio de L.

Definiendo

$$\widetilde{v} = \frac{1}{\|v\|} v,$$

se tendrá $L\widetilde{v} = \lambda \widetilde{v}$ y además $\{\widetilde{v}\}$ es una base ortonormal de W. Esto prueba el caso en que la dimensión sea 1.

Supongamos que la propiedad es cierta para espacios de dimensión $k-1$, por demostrar que es cierta para espacios de dimensión k. Sea W subespacio de dimensión k y $L : W \to W$ lineal simétrica. Sea $B = \{w_1, \ldots, w_k\}$ base ortonormal de W, gracias al Ejercicio 6.2 la matriz $M_{BB}(L) \in \mathbb{R}^k$ es simétrica, por lo que tiene todos sus valores propios reales.

Sea λ_0 un valor propio cualquiera, por lo tanto existe un vector $z = (z_1, \ldots, z_k)^t \neq 0$, tal que $M_{BB}(L)z = \lambda_0 z$.

Sea $v = z_1 w_1 + z_2 w_2 + \cdots + z_k w_k \in W$. Como $v \neq 0$ (pues $z \neq 0$ y $\{w_1, \ldots, w_k\}$ es l.i.), luego gracias a la parte 2 del Ejercicio 6.2, es vector propio de L (ya que $\varphi(v) = z$).

Sea $\widetilde{W} = \{u \in W \mid \langle u, v \rangle = 0\}$ el ortogonal de $\langle \{v\} \rangle$ con respecto a W. Se tiene que $k = \dim \ W = 1 + \dim \widetilde{W}$ lo que implica que $\dim \ \widetilde{W} = k - 1$

Sea $u \in \widetilde{W} \Rightarrow \langle L(u), v \rangle = \langle u, L(v) \rangle = \langle u, \lambda_0 v \rangle = \lambda_0 \langle u, v \rangle = 0$, luego $L(\widetilde{W}) \subseteq \widetilde{W}$. Tomemos la restricción de L a $\widetilde{W}$, es decir,

$$\begin{aligned} \widetilde{L} : \widetilde{W} &\to \widetilde{W} \\ u &\mapsto \widetilde{L}_u = Lu, \end{aligned}$$

entonces $\widetilde{L}$ es lineal y simétrica. Verifiquemos que $\widetilde{L}$ es simétrica. Consideremos $u, w \in \widetilde{W}$

$$\left\langle \widetilde{L}(u), w \right\rangle = \langle L(u), w \rangle = \langle u, L(w) \rangle = \left\langle u, \widetilde{L}(w) \right\rangle ,$$

es decir $\widetilde{L}$ es simétrica en $\widetilde{W}$. Como $\dim \widetilde{W} = k - 1$ se tiene por hipótesis de inducción que existe una base ortonormal de $\widetilde{W}$ de vectores propios de $\widetilde{L}$: $\{\widetilde{w}_1, \ldots, \widetilde{w}_{k-1}\}$, es decir $L(w_i) = \widetilde{L}(\widetilde{w}_i) = \lambda_i \widetilde{w}_i$ luego $\widetilde{w}_1, \ldots, \widetilde{w}_{k-1}$ son vectores propios de L. Finalmente, consideremos

$$B' = \left\{ \frac{v}{\|v\|}, \widetilde{w}_1, \ldots, \widetilde{w}_{k-1} \right\}$$

es una base ortonormal de W, de vectores propios de L, y por lo tanto el resultado queda demostrado. $\square$

Corolario 6.3. *Sea $A \in M_{nn}(\mathbb{R})$, simétrica, entonces A es diagonalizable y más aún, $A = PDP^t$ donde $P = (v_1, \ldots, v_n)$ y $B' = \{v_1, \ldots, v_n\}$ es una base ortonormal de vectores propios de A. D es la matriz diagonal de los valores propios correspondientes.*

Demostración. Sea $L(x) = Ax$, entonces L es simétrica. Por el teorema anterior existe una base ortonormal $B' = \{v_1, \ldots, v_n\}$ de vectores propios. Luego A es diagonizable y más aún $A = PDP^{-1}$ donde $P = (v_1, v_2, \ldots, v_n)$ y $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \cdots & \\ 0 & & \lambda_n \end{pmatrix}$. Resta por demostrar que $P^{-1} = P^t$. Calculemos

$$(P^tP)_{ij} = v_i^t v_j = \langle v_i, v_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases},$$

por lo tanto $P^tP = I$. □

Una matriz que satisface $P^{-1} = P^t$ se denomina **ortogonal** o **unitaria**.

Ejercicio 6.3: Verificar que P es ortogonal si y sólo si

$$\forall u, v \in \mathbb{R}^n \qquad \langle Pu, Pv \rangle = \langle u, v \rangle .$$

Entonces si P es ortogonal $\|Pu\|^2 = \langle Pu, Pu \rangle = \langle u, u \rangle = \|u\|^2$, es decir, P preserva la norma.

Notemos que no hay unicidad en la base de vectores propios y por lo tanto una matriz diagonalizable A puede descomponerse de la forma "PDP^{-1}" de muchas maneras. Lo que probamos para una matriz simétrica es que se puede tomar una base ortonormal de vectores propios y que para esta base se tiene la descomposición "PDP^t".

Ejercicio 6.4: Probar que $P_\theta = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}$ es unitaria en $\mathbb{R}^2$.

Ejemplo 6.5.

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

los valores propios son $\lambda = 2$ y $\lambda = 1$ en efecto:

$$|A - \lambda I| = \left| \begin{pmatrix} 2-\lambda & 0 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{pmatrix} \right| = (2-\lambda)^2(1-\lambda).$$

Calculemos el espacio propio de $\lambda = 2$.

$$0 = (A - 2I)u = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

luego $W_2 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid z = 0 \right\}$. Determinemos W_1:

$$0 = (A - I)u = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

es decir $W_1 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x = y = 0 \right\}$. Por lo tanto una base de vectores propios es:

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

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

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

Pero A también posee una base ortonormal de vectores propios:

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

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

Observación: Cabe señalar que en el caso de que si no hubiese sido sencillo obtener a simple vista una base ortonormal del subespacio propio asociado a $\lambda = 2$, podemos utilizar el método de Gram-Schmidt en dicho subespacio. En general, la base ortonormal de vectores propios de $\mathbb{R}^n$ (que sabemos existe en el caso de A simétrica) puede obtenerse utilizando Gram-Schmidt en cada subespacio propio.

El Teorema 6.4 también es cierto para $L(x) = Ax$, con A hermítica, y la demostración es análoga.

Ejemplo 6.6.
Veamos el caso de las proyecciones ortogonales.

Dado W subespacio de $\mathbb{R}^n$ consideremos $W^\perp$ el ortogonal de W. Entonces $\mathbb{R}^n = W \oplus W^\perp$. Sea además $P\colon \mathbb{R}^n \to \mathbb{R}^n$, la proyección ortogonal sobre W.

Esta funcion P verifica

1. $P^2 = P$**.**

   En efecto: $P(u) = w \in W$ luego $w = w + 0$ es la única descomposición, de donde
   $$PP(u) = P(w) = w.$$
   Se tiene además $P(I - P) = 0$.

2. $\mathrm{Im}(P) = W$ **y** $\mathrm{Ker}(P) = W^\perp$**.**

   En efecto, es claro que $Im(P) \subseteq W$. Por otro lado, si $w \in W$ $P(w) = w$, luego $\mathrm{Im}(P) = W$. Ahora si $v \in W^\perp$ $v = 0 + v$ es la descomposición única y por tanto $P(v) = 0$, así $W^\perp \subseteq \mathrm{Ker}(P)$. Sea $v \in \mathrm{Ker}(P)$, luego $P(v) = 0$ y por tanto $v = 0 + v$ es decir $v \in W^\perp$.

3. P **es simétrica.**

   En efecto sean $u_1 = w_1 + v_1, u_2 = w_2 + v_2$, luego
   $$\langle P(u_1), u_2\rangle = \langle w_1, w_2 + v_2\rangle = \langle w_1, w_2\rangle + \langle w_1, v_2\rangle = \langle w_1, w_2\rangle =$$
   $$\langle w_1 + v_1, w_2\rangle = \langle u_1, P(u_2)\rangle$$

4. **Los valores propios de** P **son 0 ó 1.**

   Sea $P(u) = \lambda u,\ u \neq 0$, entonces como $P^2 = P$ se tiene $P(u) = P^2(u) = P(P(u)) = P(\lambda u) = \lambda P(u) = \lambda^2 u$. De donde $(\lambda^2 - \lambda)u = 0$. Como $u \neq 0$ $\lambda^2 - \lambda = 0$ entonces $\lambda =$ es 0 ó 1.

Identificaremos P con su matriz representante respecto a la base canónica. Dado que P es simétrica $P = \Gamma D \Gamma^t$ con Γ ortogonal y D satisface:

$$D = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

Como P y D son similares tienen el mismo rango, así $r = \mathrm{r}(P) = \dim(Im(P)) = \dim(W)$. Γ puede construirse de la siguiente forma: $\Gamma = (\Gamma_1, \ldots, \Gamma_n)$. Escogemos $\{\Gamma_1, \ldots, \Gamma_r\}$ una base ortonormal de W, la cual completamos a una base ortonormal de $\mathbb{R}^n$ $\{\Gamma_1, \ldots, \Gamma_n\}$.

Puede comprobarse directamente que $I - P$ es la proyección ortogonal sobre $W^\perp$.

Guía de Ejercicios

1. Sea W subespacio vectorial de $\mathbb{R}^n$ y $L\colon W \to W$ lineal, probar que:

(a) Si $B = \{u_1, \ldots, u_n\}$ es una base **ortonormal** de W, entonces

$$L \text{ simétrica} \iff M_{BB}(L) \text{ simétrica.}$$

(b) Si $B = \{u_1, \ldots, u_n\}$ es una base de W y

$$\begin{aligned} \varphi\colon W &\longrightarrow \mathbb{R}^n \\ x = \textstyle\sum_{i=0}^{k} \alpha_i u_i &\longmapsto \varphi(x) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix}. \end{aligned}$$

Entonces x es vector propio de L asociado al v.p. λ si y sólo si $\varphi(x)$ es vector propio de $M_{BB}(L)$ asociado al v.p. λ.

2. Verificar que P es ortogonal si y sólo si

$$\forall u, v \in \mathbb{R}^n \qquad \langle Pu, Pv \rangle = \langle u, v \rangle .$$

3. Probar que $P_\theta = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}$ es ortogonal en $\mathbb{R}^2$.

Guía de Problemas

P1. Considere la matriz $A \in \mathcal{M}_{33}(\mathbb{R})$ siguiente,

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

(a) Calcular los valores propios de A.

(b) Calcular una base de vectores propios de A.

(c) Encontrar una matriz $P \in \mathcal{M}_{33}(\mathbb{R})$ y $D \in \mathcal{M}_{33}(\mathbb{R})$, D diagonal, tal que $A = PDP^t$.

(d) ¿Es A una matriz invertible?, justifique su respuesta.

P2. **(a)** Sea $A \in \mathcal{M}_{33}(\mathbb{R})$. Se sabe que A es simétrica y que $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ es vector propio de A asociado al valor propio 2, y además la dimensión del $Ker(T)$ es igual a 2. Calcular A.

(b) Sea $A = \begin{pmatrix} 1 & a & 0 \\ 0 & 1 & a \\ 0 & 0 & 3 \end{pmatrix}$, donde a es un parámetro real. Calcule los valores de $a \in \mathbb{R}$ para los cuales la matriz A es diagonalizable.

P3. Sean $v_1, \dots, v_k \in \mathbb{R}^n$ ortogonales y de norma 1, sean $\alpha_1, \dots, \alpha_k \in \mathbb{R} \setminus \{0\}$, y

$$A = \alpha_1 \cdot v_1 v_1^t + \cdots + \alpha_k \cdot v_k v_k^t \in \mathcal{M}_{nn}(\mathbb{R}).$$

(a) Pruebe que $\mathrm{Im}(A) = \langle \{v_1, \dots, v_k\} \rangle$, que $\mathrm{Ker}(A) = \langle \{v_1, \dots, v_k\} \rangle^{\perp}$, y determine $r(A)$, es decir el rango de A.

(b) Sea $\{w_1, \dots w_{n-k}\}$ base ortonormal del $Ker(A)$. Demuestre que $\{v_1, \dots, v_k, w_1, \dots w_{n-k}\}$ es una base de $\mathbb{R}^n$ de vectores propios de A. Especifique cuál es el valor propio correspondiente a cada uno de estos vectores propios.

(c) Sea $A = \begin{pmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{pmatrix}$. Determine $\{v_1, v_2\}$ base ortonormal de $\mathbb{R}^2$ y $\alpha_1, \alpha_2 \in \mathbb{R}$ tales que

$$A = \alpha_1 \cdot v_1 v_1^t + \alpha_2 \cdot v_2 v_2^t.$$

Ingeniería Matemática
FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Formas cuadráticas

7.1 Formas cuadráticas y matrices definidas positivas

En esta sección estudiaremos las formas cuadráticas inducidas por una matriz simétrica. Las formas cuadráticas aparecen en varias aplicaciones en matemáticas. Una de ellas es la caracterización de mínimos o máximos para funciones de varias variables. Nosotros nos interesaremos en la caracterización de matrices definidas positivas. Al final del capítulo estudiaremos las cónicas en $\mathbb{R}^2$.

DEFINICIÓN (FORMA CUADRÁTICA) Dada $A \in M_{nn}(\mathbb{R})$ **simétrica**, definimos:

$$\begin{aligned} q : \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto q(x) = x^t A x. \end{aligned}$$

q es llamada una **forma cuadrática** en $\mathbb{R}^n$.

Notemos que una forma cuadrática en $\mathbb{R}^n$ es **homogénea de grado 2**, es decir, $\forall x \in \mathbb{R}^n$ $q(\lambda x) = \lambda^2 q(x)$. En efecto, $q(\lambda x) = (\lambda x)^t A(\lambda x) = \lambda x^t A \lambda x = \lambda^2 x^t A x = \lambda^2 q(x)$.

Ejemplo 7.1.
$q : \mathbb{R}^2 \to \mathbb{R},\ q(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2 = (x_1, x_2) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$

Se tiene $q(1, -1) = 0$, en realidad $q(x_1, -x_1) = 0$.

DEFINICIÓN ((SEMI)DEFINIDA POSITIVA/NEGATIVA) Sea $A \in M_{nn}(\mathbb{R})$, simétrica diremos que

- A es definida positiva si $\forall x \neq 0 \quad x^t A x > 0$.
- A es semidefinida positiva si $\forall x\ x^t A x \geq 0$.
- A es definida negativa $\forall x \neq 0\ x^t A x < 0$.
- A es semidefinida negativa $\forall x\ x^t A x \leq 0$.

Notar que A es definida (semidefinida) positiva ssi $-A$ es definida (semidefinida) negativa.

Ejemplo 7.2.

$$
\begin{aligned}
q(x) &= 2x_1^2 + 2x_1x_2 + 2x_2^2 = (x_1+x_2)^2 + x_1^2 + x_2^2 \\
&= (x_1, x_2)\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}
\end{aligned}
$$

$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ es definida positiva, en efecto, si $q(x) = 0$, entonces $x_1^2 = 0, x_2^2 = 0$, luego $x_1 = x_2 = 0$.

Teorema 7.1. *Supongamos que $A \in M_{nn}(\mathbb{R})$ es simétrica. Entonces las siguientes proposiciones son equivalentes.*

1. $\forall x \neq 0 \quad x^{\,t}Ax > 0$ *(A es definida positiva)*

2. *Los valores propios de A son positivos.*

3. *Sea*

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

$$
A^{(1)} = [a_{11}] \quad A^{(2)} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad A^{(3)} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & A_{33} \end{pmatrix} \ldots
$$

$$
\ldots\ A^{(n)} = A
$$

entonces $|A^{(1)}| = a_{11} > 0 \quad , |A^{(2)}| > 0, \cdots |A^{(i)}| > 0, \cdots, |A^{(n)}| = |A| > 0$

4. *El método de Gauss utilizando sólo operaciones elementales del tipo $E_{pq}(\alpha, 1)$, $p < q$, permite escalonar A y además los pivotes son siempre positivos.*

Demostración. (1) $\Rightarrow$ (2).

Sea λ un valor propio de A y $v \neq 0$ un vector propio correspondiente al valor propio λ.

$$
0 < v^tAv = v^t(\lambda v) = \lambda v^t v = \lambda\|v\|^2 \implies \lambda > 0
$$

(2) $\Rightarrow$ (1). Sabemos que por ser A simétrica tiene la descomposición $A = PDP^t$, donde las columnas de P son una base ortonormal de vectores propios y D es la diagonal de los valores propios respectivos. Así

$$x^t Ax = x^t PDP^t x = (P^t x)^t DP^t x,$$

entonces si definimos $z = P^t x$ se tendrá que en términos de estas nuevas variables la forma cuadrática queda

$$x^t Ax = z^t Dz = \lambda_1 z_1^2 + \cdots + \lambda_n z_n^2 \geq 0,$$

pues los valores propios son positivos. Más aún la única forma en que esta suma de cuadrados sea cero es que $z_1 = \cdots = z_n = 0$ es decir $z = 0$, pero entonces $P^t x = 0 \Rightarrow x = P0 = 0$ (recordemos que $P^{-1} = P^t$).

(1) $\Rightarrow$ (3). Sea $a \neq 0$ y definamos $x = \begin{pmatrix} a \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, entonces

$$\begin{aligned} 0 < \ x^t Ax &= \sum_{i=1}^{n} x_i (Ax)_i = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij} x_j \\ &= a^2 a_{11}, \end{aligned}$$

luego $a_{11} > 0$. Sea $v = (v_1, \ldots, v_k)^t \in \mathbb{R}^k$, tomemos $x = (v_1, \ldots, v_k, 0, \ldots, 0)^t \in \mathbb{R}^n$

$$x^t Ax = \sum_{i,j=1}^{n} x_i x_j a_{ij} = \sum_{i,j=1}^{k} v_i v_j a_{ij} = v^t A^{(k)} v$$

si $v \neq 0$ se tiene $x \neq 0$, luego $v^t A^{(k)} v = x^t Ax > 0$ es decir $A^{(k)}$ es definida positiva.

De donde $A^{(1)}, A^{(2)}, \ldots, A^{(n)} = A$ son todas definidas positivas. Pero una matriz definida positiva tiene todos sus valores propios positivos (por lo ya demostrado). Como el determinante es el producto de los valores propios entonces el determinante es positivo. Luego
$|A^{(1)}| > 0, \ldots, |A^{(n)}| > 0$.

(3) $\Rightarrow$ (4).

Por inducción sobre i $\quad a_{11} > 0$ es el primer pivote. En la etapa i tendremos

$$\tilde{A}_i = \begin{pmatrix} a_{11} & \tilde{a}_{12} & & \cdots & \tilde{a}_{1i} & \cdots & \tilde{a}_{1n} \\ 0 & \tilde{a}_{22} & & \vdots & & & \vdots \\ \vdots & & \ddots & \tilde{a}_{i-1i-1} & & & \vdots \\ 0 & \cdots & \cdots & 0 & \tilde{a}_{ii} & & \vdots \\ 0 & \cdots & \cdots & 0 & \tilde{a}_{ii+1} & & \vdots \\ \vdots & & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \tilde{a}_{in} & & \tilde{a}_{nn} \end{pmatrix} = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}.$$

Donde $B_1 : i \times i, B_2 : i \times n-i, B_3 : n-i \times i, B_4 : n-i \times n-i$. Por hipótesis de inducción, se obtuvo $\tilde{A}_i$ mediante el algoritmo de Gauss, sin utilizar permutaciones de filas pues todos los pivotes $\tilde{a}_{11}, \ldots, \tilde{a}_{i-1i-1}$ son positivos. Así, $\tilde{A}_i = (\prod_k E_k)A$ donde E_k son matrices simples del tipo $E_{p,q}(\alpha, 1)$ que son invertibles y triangulares inferiores con diagonal 1. Por tanto $A = C \cdot \tilde{A}_i$ donde $C = (\prod_k E_k)^{-1}$ es también triangular inferior con unos en la diagonal. Así si $C = \left(\begin{smallmatrix} C_1 & C_2 \\ C_3 & C_4 \end{smallmatrix}\right)$ donde $C_1 : i \times i, C_2 : i \times (n-i), C_3 : (n-i) \times i, C_4 : (n-i) \times (n-i)$, se tendrá $C_2 = 0$ y además C_1 es triangular inferior con diagonal 1, y por lo tanto $|C_1| = 1$. Luego:

$$A = \begin{pmatrix} C_1 & 0 \\ \\ C_2 & C_3 \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ \\ B_3 & B_4 \end{pmatrix} = \begin{pmatrix} C_1 B_1 & C_1 B_2 \\ \\ C_2 B_1 + C_3 B_3 & C_2 B_2 + C_3 B_4 \end{pmatrix}.$$

Por lo tanto $A^{(i)} = C_1 B_1$, de donde $0 < |A^{(i)}| = |C_1 B_1| = |C_1||B_1| = |B_1|$. Por otro lado B_1 es triangular superior, luego $|B_1| = \prod_{j=1}^{i} \tilde{a}_{jj}$, pero $\tilde{a}_{11}, \tilde{a}_{22}, \ldots, \tilde{a}_{i-1i-1} > 0$ se concluye, $\tilde{a}_{ii} > 0$. Así el método de Gauss no necesita intercambiar filas y además todos los pivotes son > 0.

(4) $\Rightarrow$ (1).

Como todos los pivotes son positivos y no hay que hacer intercambios de filas se tendra que $A = LDU$, con L triangular inferior D diagonal y U triangular superior. U y L tienen 1 en la diagonal y $d_{ii} = \tilde{a}_{ii}$. Como A es simétrica $U = L^t$ entonces

$$A = LDL^t = (L\sqrt{D})\ (\sqrt{D}L^t) = RR^t, \text{ con } R = L\sqrt{D}$$

y

$$\sqrt{D} = \begin{pmatrix} \sqrt{\tilde{a}_{11}} & & & 0 \\ & \sqrt{\tilde{a}_{22}} & & \\ & & \ddots & \\ 0 & & & \sqrt{\tilde{a}_{nn}} \end{pmatrix}.$$

Dado que $|A| = |\tilde{A}| = \tilde{a}_{11}\tilde{a}_{22}, \ldots, \tilde{a}_{nn} > 0$ entonces $0 < |A| = |RR^t| = |R|^2$ luego $|R| \neq 0$, y por lo tanto R es invertible y así su traspuesta.

$$x^t A x = x^t (RR^t) x = (R^t x)^t R^t x = \|R^t x\|^2.$$

Si $x \neq 0 \Rightarrow R^t x \neq 0$ pues R^t es invertible luego $\|R^t x\|^2 > 0$ de donde $x^t A x > 0$ si $x \neq 0$ es decir A es definida positiva. $\square$

Hemos probado de paso la que toda matriz definida positiva tiene la siguiente descomposición.

Definición (Descomposición de Cholesky) Decimos que $A \in \mathcal{M}_{nn}(\mathbb{R})$ admite una descomposición de Cholesky si existe R matriz triangular inferior, con diagonal estrictamente positiva tal que

$$A = RR^t.$$

Notemos que A es definida positiva ssi existe R triangular inferior con diagonal no nula tal que:

$$A = RR^t.$$

Observación: Hay que hacer una precisión sobre la propiedad 4 del teorema anterior. Es fácil ver que la matriz

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

es definida positiva, pero si al escalonar utilizamos $E_{12}(\frac{1}{2}, -1)$ obtenemos la matriz escalonada

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

y los pivotes no son todos positivos!. Sin embargo si utilizamos $E_{12}(-\frac{1}{2}, 1)$ obtenemos

$$\tilde{A} = \begin{pmatrix} 2 & 1 \\ 0 & \frac{3}{2} \end{pmatrix}.$$

Por lo tanto es importante recordar que para estudiar si una matriz es definida positiva las operaciones elementales permitidas son $E_{pq}(\alpha, 1)$.

7.2 Formas canónicas

Sea $q(x) = x^t Ax$ una forma cuadrática. Como A es simétrica, entonces $A = PDP^t$ donde $P = (v_1, \cdots, v_n)$ con $\{v_1, \cdots, v_n\}$ una base ortonormal de vectores propios de A y

$$D = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix},$$

siendo $\lambda_1, \ldots, \lambda_n$ los valores propios de A.

Sea $y = P^t x$ entonces

$$x^t Ax = x^t PDP^t x = y^t Dy = \sum_{i=1}^{n} \lambda_i y_i^2$$

Así haciendo este cambio de variables la forma cuadrática tiene la expresión

$$\tilde{q}(y) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

Ejemplo 7.3.

$$q(x) = x_1^2 + 2x_2^2 + \sqrt{3}x_1x_2 = x^t \begin{pmatrix} 1 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 2 \end{pmatrix} x.$$

Consideremos A la matriz que define esta forma cuadrática.

$$A = \begin{pmatrix} 1 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 2 \end{pmatrix},$$

Calculemos los valores y vectores propios:

$$|A - \lambda I| = \left| \begin{pmatrix} 1-\lambda & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 2-\lambda \end{pmatrix} \right| = (1-\lambda)(2-\lambda) - \frac{3}{4} = 0.$$

Luego el polinomio característico es

$$\lambda^2 - 3\lambda + \frac{5}{4} = 0,$$

cuyas raices son

$$\lambda = \frac{3 \pm \sqrt{9-5}}{2} = \frac{3 \pm 2}{2} = \begin{cases} \lambda_1 = \frac{5}{2} \\ \lambda_2 = \frac{1}{2} \end{cases}$$

Calculemos los vectores propios:

$\lambda = \frac{5}{2}$

$$A - \frac{5}{2}I = \begin{pmatrix} 1-\frac{5}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 2-\frac{5}{2} \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

las filas son $\ell.d.$ (¿ por qué?) luego

$$(A - \frac{5}{2}I)v = 0 \iff \quad -\frac{3}{2}v_1 + \frac{\sqrt{3}}{2}v_2 = 0 \iff v_2 = \sqrt{3}v_1.$$

Sea $v_1 = 1, v_2 = \sqrt{3}$, un vector propio colineal de largo uno es:

$$w_1 = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}.$$

$\lambda = \frac{1}{2}$

$$A - \frac{1}{2}I = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{3}{2} \end{pmatrix}$$

y por lo tanto

$$(A - \frac{1}{2}I)v = 0 \qquad \iff \qquad \frac{1}{2}v_1 + \frac{\sqrt{3}}{2}v_2 = 0 \Leftrightarrow v_1 = -\sqrt{3}v_2.$$

Un vector propio de largo uno es:

$$w_2 = \frac{1}{2}\begin{pmatrix} -\sqrt{3} \\ 1 \end{pmatrix},$$

y una base de ortonormal de vectores propios está formada por $\{w_1, w_2\}$. Se tiene que

$$A = PDP^t = \begin{pmatrix} 1 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} \frac{5}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}.$$

Consideremos el cambio de variables:

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = P^t x = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

La forma cuadrática en términos de y es $\tilde{q}(y) = \frac{5}{2}y_1^2 + \frac{1}{2}y_2^2$. El cambio de coordenadas de x a y corresponde a una rotación (¿en qué ángulo?).

Volvamos al caso general. En el sistema $y = P^t x$, la forma cuadrática se escribe como

$$\tilde{q}(y) = \sum_{i=1}^{n} \lambda_i y_i^2 \quad = \sum_{i=1}^{p} \lambda_i y_i^2 + \sum_{i=p+1}^{r} \lambda_i y_i^2$$

donde:

$$\lambda_1, \ldots, \lambda_p \text{ son } > 0, \quad \lambda_{p+1}, \ldots, \lambda_r \text{ son } < 0, \quad \lambda_{r+1} = \cdots = \lambda_n = 0.$$

Definamos

$$\begin{aligned} z_i &= \sqrt{\lambda_i} y_i & i &= 1, \ldots, p \\ z_i &= \sqrt{-\lambda_i} y_i & i &= p+1, \ldots, r \\ z_i &= y_i & i &= r+1, \ldots, n \end{aligned}$$

es decir,

$$z = \begin{pmatrix} \sqrt{\lambda_1} & & & & & & & & 0 \\ & \sqrt{\lambda_2} & & & & & & & \\ & & \ddots & & & & & & \\ & & & \sqrt{\lambda_p} & & & & & \\ & & & & \sqrt{-\lambda_{p+1}} & & & & \\ & & & & & \ddots & & & \\ & & & & & & \sqrt{-\lambda_r} & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ 0 & & & & & & & & 1 \end{pmatrix} y = Qy.$$

luego tenemos $z = QP^t x$. En las nuevas variables z, es la forma cuadrática:

$$\hat{q}(z) = z_1^2 + z_2^2 + \cdots + z_p^2 - z_{p+1}^2 - z_{p+2}^2 \cdots - z_r^2,$$

donde r corresponde al rango de A. El número $2p - r$ = número de valores propios positivos menos el número de valores propios negativos, es llamada la signatura.

El cambio de coordenadas $y = P^t x$ corresponde a una rotación de sistemas de coordenadas. El cambio $z = Qy$ corresponde a dilatar y contraer ciertas direcciones (homotecia).

En nuestro ejemplo $z_1 = \sqrt{\frac{5}{2}} y_1 \qquad z_2 = \sqrt{\frac{1}{2}} y_2$

$$z_1^2 + z_2^2 = (\sqrt{\frac{5}{2}} y_1)^2 + (\sqrt{\frac{1}{2}} y_2)^2 = \frac{5}{2} y_1^2 + \frac{1}{2} y_2^2$$

Para ubicar el nuevo sistema, con respecto al inicial, consideremos el problema en dos etapas. En el cambio $y = P^t x$, el eje correspondiente a "y_1" está en la dirección "x" tal que

$$P^t x = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ luego } \; x = P \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = v_1,$$

la primera columma de P que corresponde al primer vector propio. De igual forma el i-ésimo eje del sistema de coordenadas "y" es el i-ésimo vector propio v_i. De igual forma el sistema de coordenadas "z" tiene sus ejes en las direcciones $v_1, \ldots, v_n$, solamente que algunos ejes se han contraído y otros expandido.

Hemos probado el siguiente resultado

Teorema 7.2. *Sea $x^t A x$ una forma cuadrática, existe L invertible tal que si $z = Lx$, entonces en términos de las variables z la forma cuadrática se expresa como $\hat{q}(z) = z_1^2 + \cdots + z_p^2 - (z_{p+1}^2 + \cdots + z_r^2)$, donde r=rango A = número de valores propios $\neq 0$, p = número de valores propios > 0.*

7.3 Cónicas en $\mathbb{R}^2$

Una cónica en $\mathbb{R}^2$ es el conjunto solución de una ecuación del tipo

$$ax^2 + by^2 + 2cxy + dy + fx = e$$

o en forma matricial

$$(x, y) \begin{pmatrix} a & c \\ c & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + (d, f) \begin{pmatrix} x \\ y \end{pmatrix} = e = v^t A v + g^t v = e,$$

donde

$$A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}, \ v = \begin{pmatrix} x \\ y \end{pmatrix}, \ g = \begin{pmatrix} d \\ f \end{pmatrix}$$

Supongamos que $A = PDP^t$, con $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ y $P = (v_1 v_2)$. Entonces la ecuación de la cónica es:

$$e = v^t PDP^t v + g^t v = v^t PDP^t v + gPP^t v.$$

Sea $u = P^t v$. En este nuevo sistema de coordenadas la ecuación de la cónica es:

$$e = u^t Du + \tilde{g}^t u \ \text{ con } \ \tilde{g} = P^t g = \begin{pmatrix} \tilde{f} \\ \tilde{d} \end{pmatrix}.$$

La expresión que toma la cónica es entonces

$$e = \lambda_1 u_1^2 + \lambda_2 u_2^2 + \tilde{f} u_1 + \tilde{d} u_2$$

1. $\lambda_1 = \lambda_2 = 0$ ($\iff a = b = c = 0$)

   El conjunto solución es $\{(x, y) \mid dy + fx = e\}$ que en general es una recta (pudiendo ser vacío si $d = f = 0$ y $e \neq 0$).

2. El caso interesante es cuando $\lambda_1 \neq 0$ o $\lambda_2 \neq 0$, llamemos

   $$u_1' = u_1 - \alpha$$
   $$u_2' = u_2 - \beta$$

   $$\lambda_1(u_1' + \alpha)^2 + \lambda_2(u_2' + \beta)^2 + \tilde{f}(u_1' + \alpha) + \tilde{d}(u_2' + \beta) = e$$

   $$\lambda_1(u_1')^2 + \lambda_2(u_2')^2 + u_1'\{2\lambda_1\alpha + \tilde{f}\} + u_2'\{2\beta\lambda_2 + \tilde{d}\} = \tilde{e}$$

   con $\tilde{e} = e - (\lambda_1\alpha^2 + \lambda_2\beta^2 + \tilde{f}\alpha + \tilde{d}\beta)$

   Si $\lambda_1 = 0$ ($\lambda_2 \neq 0$) tomemos: $\beta = -\frac{\tilde{d}}{2\lambda_2}$, $\alpha = 0$ se llega a la ecuación

   $$\lambda_2(u_2')^2 + \tilde{f} u_1' = \tilde{e}.$$

   Si $\tilde{f} = 0 \quad (u_2')^2 = \frac{\tilde{e}}{\lambda_2}$. Así, el conjunto solución serán dos rectas paralelas: $u_2' = \pm\sqrt{\frac{\tilde{e}}{\lambda_2}}$ si $\frac{\tilde{e}}{\lambda_2} \geq 0$ (que en realidad es una sola si $\tilde{e} = 0$) o será vacío si $\frac{\tilde{e}}{\lambda_2} < 0$. Si $\tilde{f} \neq 0$
   $u_1' = \frac{\tilde{e} - \lambda_2(u_2')^2}{\tilde{f}}$, que corresponde a una parábola, en las coordenada u_1', u_2'. El caso $\lambda_1 \neq 0$ $\lambda_2 = 0$ es totalmente análogo.

   Falta ver el caso $\lambda_1 \neq 0$ y $\lambda_2 \neq 0$, tomamos $\alpha = -\frac{\tilde{f}}{2\lambda_1}$ y $\beta = -\frac{\tilde{d}}{2\lambda_2}$. En el sistema u_1', u_2' tenemos la ecuación

   $$\lambda_1(u_1')^2 + \lambda_2(u_2')^2 = \tilde{e}$$

3. $\lambda_1 > 0, \lambda_2 > 0, \tilde{e} \geq 0$; o $\lambda_1 < 0, \lambda_2 < 0, \tilde{e} \leq 0$ la solución es una elipse (una circunferencia si $\lambda_1 = \lambda_2$) en el sistema u_1', u_2' Si $\lambda_1 > 0, \lambda_2 > 0, \quad \tilde{e} < 0$; o $\lambda_1 < 0, \lambda_2 < 0, \tilde{e} > 0$ no hay solución.

4. $\lambda_1 > 0$ y $\lambda_2 < 0$ podemos suponer que $\tilde{e} \geq 0$ de otra manera multiplicamos por -1

   $\lambda_1 (u_1')^2 - |\lambda_2| (u_2')^2 = \tilde{e}$, que corresponde a una hipérbola con eje de simetría el eje u_1'. El caso $\lambda_1 < 0, \lambda_2 > 0, \quad \tilde{e} \geq 0$ el conjunto solución es una hipérbola con eje de simetría u_2'.

Notemos que en general el sistema (u_1, u_2) corresponde a una rotación con respecto al eje (x, y) y sus ejes están en la dirección de los vectores propio. El sistema (u_1', u_2') corresponde a una traslación o cambio de origen del sistema (u_1, u_2).

Guía de Ejercicios

1. Escriba en forma x^tAx las siguientes formas cuadráticas, en donde $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$:

(a) $q(x) = -2x_1^2 - \frac{1}{2}x_1x_2 + 5x_2^2$.

(b) $q(x) = x_1(2x_1 - x_2) + x_2(3x_1 + x_2)$.

(c) $q(x) = x_1^2 + y_2^2 - y_3^2 - x_1x_2 + x_1x_3$.

(d) $q(x) = (x_1 + x_2)^2 - (x_1 + x_3)^2$.

2. Señale si las matrices siguientes son definidas positivas/negativas, o ninguna de las dos:

(a) $\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$.

(b) $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$.

(c) $\begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix}$.

(d) $\begin{pmatrix} -1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$.

3. Sea $A \in \mathcal{M}_{nn}(\mathbb{R})$ una matriz simétrica.

(a) Probar que $v \in \mathbb{R}^n$ es vector propio de A si y sólo si es vector propio de $I - A$. Si $\lambda \in \mathbb{R}$ es el valor propio de A asociado a v, ¿cuál es el valor propio de $I - A$ asociado a v?

(b) Probar que la matriz $I - A$ es definida positiva si y sólo si todos los valores propios de A son menores estrictos que uno.

Guía de Problemas

P1. Considere la cónica de ecuación $5x^2 + 5y^2 + 6xy + 16x + 16y = -15$. Realice el cambio de variables que permite escribir la cónica de manera centrada y escriba la nueva expresión (escribir explícitamente el cambio de variables). Identifique la cónica resultante y dibújela.

P2. Sea $A \in \mathcal{M}_{22}(\mathbb{R})$ simétrica definida postiva, $b \in \mathbb{R}^2$, $c \in \mathbb{R}$ y $f\colon \mathbb{R}^2 \to \mathbb{R}$ tal que

$$f(x) = x^tAx - b^tx + c.$$

Se quiere minimizar f.

(a) Sea $x_0 = \frac{1}{2}A^{-1}b$. Pruebe que $f(x) = (x - x_0)^tA(x - x_0) - x_0^tAx_0 + c$.

(b) Concluya que el único mínimo de f se alcanza en x_0, i.e., que $f(x) \geq f(x_0)$ para todo $x \in \mathbb{R}^2$ y que la igualdad se alcanza solamente cuando $x = x_0$-

P3. Sea $\alpha \in \mathbb{R}$ y considere la ecuación

$$\alpha x^2 + \alpha y^2 + 2(\alpha - 1)xy - \sqrt{2}x + \sqrt{2}y = 0.$$

Determine todos los valores de α tales que la ecuación corresponde a:

(a) Circunferencia.

(b) Elipse.

(c) Recta o rectas.

(d) Parábola.

(e) Hipérbola.

(f) Conjunto vacío.

(g) Un punto.

Forma de Jordan

Ingeniería Matemática
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

8.1 Definición

Estudiaremos aquí un procedimiento que permite obtener, a partir de una matriz A, otra *similar* con gran cantidad de ceros. Obviamente, el ideal es obtener una matriz diagonal, pero esto no siempre es posible. Nos conformaremos con una forma algo más compleja que la matriz diagonal, pero con la ventaja de que ésta puede obtenerse para matrices arbitarias.

La idea general de reducir una matriz consiste en determinar una base de modo que la transformación lineal asociada a la matriz inicial tenga una representación simple en la nueva base.

Partamos con un ejemplo. Consideremos la transformación lineal $T_A : \mathbb{C}^9 \to \mathbb{C}^9$ dada por $x \mapsto Ax$, cuya matriz representante J, con respecto a una base β_J es:

$$J = \begin{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} & & & & \bigcirc \\ & \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} & & & \\ \bigcirc & & (2) & & \\ & & & \begin{pmatrix} 5 & 1 \\ 0 & 5 \end{pmatrix} & \\ & & & & (6) \end{pmatrix}$$

Es claro que $\sigma(A) = \{2, 5, 6\}$, con multiplicidades algebraicas $6, 2, 1$ respectivamente.

La base β_J la notaremos como sigue:

$$\beta_J = \{v_{11}^1, v_{12}^1, v_{13}^1, v_{21}^1, v_{22}^1, v_{31}^1, v_{11}^2, v_{12}^2, v_{11}^3\}$$

De la definición de matriz representante, obtenemos:

$$
\begin{aligned}
T_A(v_{11}^1) &= 2v_{11}^1 \\
T_A(v_{12}^1) &= v_{11}^1 + 2v_{12}^1 \\
T_A(v_{13}^1) &= v_{12}^1 + 2v_{13}^1 \\
T_A(v_{21}^1) &= 2v_{21}^1 \\
T_A(v_{22}^1) &= v_{21}^1 + 2v_{22}^1 \\
T_A(v_{31}^1) &= 2v_{31}^1 \\
T_A(v_{11}^2) &= 5v_{11}^2 \\
T_A(v_{12}^2) &= v_{11}^2 + 5v_{11}^2 \\
T_A(v_{11}^3) &= 6v_{11}^3
\end{aligned}
$$

Vemos que v_{11}^1 es vector propio (cabeza de serie) asociado al valor propio $\lambda_1 = 2, v_{12}^1$ es una "cola" asociada a v_{11}^1 y v_{13}^1 es "cola" asociada a v_{12}^1. Posteriormente obtenemos otro vector propio asociado a $\lambda_1 = 2$ con una cola y finalmente un tercer vector propio asociado a $\lambda_1 = 2$. Luego se repite el mismo esquema: un vector propio y una cola asociados a $\lambda_2 = 5$ y finalmente un vector propio asociado a $\lambda_3 = 6$.

En general la matriz J que nos interesa en este párrafo tiene la forma:

$$
\begin{pmatrix}
\begin{pmatrix} \lambda_1 & 1 & \\ & \ddots & 1 \\ 0 & & \lambda_1 \end{pmatrix} & & & & & & \\
& \ddots & & & & & \\
& & \begin{pmatrix} \lambda_1 & 1 & \\ & \ddots & 1 \\ 0 & & \lambda_1 \end{pmatrix} & & \bigcirc & & \\
& & & \ddots & & & \\
\bigcirc & & & & \begin{pmatrix} \lambda_r & 1 & \\ & \ddots & 1 \\ & & \lambda_r \end{pmatrix} & & \\
& & & & & \ddots & \\
& & & & & & \begin{pmatrix} \lambda_r & 1 & \\ & \ddots & 1 \\ & & \lambda_r \end{pmatrix}
\end{pmatrix}
$$

donde los bloques son de tamaño $s(1,1), \ldots, s(1,p_1), \ldots, s(r,1), \ldots, s(r,p_r)$, y el espectro es $\sigma(A) = \{\lambda_1, \ldots, \lambda_r\}$, con multiplicidades $\sum_{j=1}^{p_1} s(1,j), \ldots, \sum_{j=1}^{p_r} s(r,j)$, respectivamente. La base se escribe:

$$
\beta_J = \{v_{11}^1, \ldots, v_{1s(1,1)}^1, v_{21}^1, \ldots, v_{2s(1,2)}^1, \ldots, v_{p_1 1}^1, \ldots, v_{p_1 s(1,p_1)}^1, \ldots, v_{p_r 1}^r, \ldots, v_{p_r s(r,p_r)}^r\}
$$

o mediante una tabla de doble entrada:

$$
\begin{array}{ccccccccc}
 & \downarrow & & \downarrow & & \downarrow & \downarrow & & \downarrow \\
v_{11}^{1} & v_{21}^{1} & \cdots & v_{p_1 1}^{1} & \cdots & v_{11}^{r} & v_{21}^{r} & \cdots & v_{p_r 1}^{r} \\
v_{12}^{1} & v_{22}^{1} & \cdots & v_{p_1 2}^{1} & \cdots & v_{12}^{r} & v_{22}^{r} & \cdots & v_{p_r 2}^{r} \\
\vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\
v_{1s(1,1)}^{1} & v_{2s(2,1)}^{1} & \cdots & v_{p_1 s(1,p_1)}^{1} & \cdots & v_{1s(r,1)}^{r} & v_{2s(r,2)}^{r} & \cdots & v_{p_r s(r,p_r)}^{r} \\
\downarrow & \downarrow & & \downarrow & & \downarrow & \downarrow & &
\end{array} \tag{8.1}
$$

vectores de la base asociados a λ_1 $\quad\cdots\quad$ vectores de la base asociados a λ_r

las flechas indican el encadenamiento de los vectores de base. Por ejemplo al vector $v_{1s(1,1)}^{1}$ le sigue v_{21}^{1}.

El primer elemento de cada columna corresponde a un vector propio. Bajo un vector propio figuran las colas asociadas.

Además:

$$
T_A(v_{ij}^k) = A v_{ij}^k = \begin{cases} \lambda_k v_{ij}^k & \text{si } j = 1 \quad (1^{\text{era}} \text{ fila de la tabla}) \\ v_{ij-1}^k + \lambda_k v_{ij}^k & \text{si } j > 1 \end{cases} \tag{8.2}
$$

Una base β_J, con estas características se denomina *base de Jordan* asociada a la transformación lineal T y su matriz representante es justamente J.

También es claro que los subespacios propios son:

$$
\begin{aligned}
V_{\lambda_1} &= \left\langle \left\{ v_{11}^1, \ldots, v_{p_1 1}^1 \right\} \right\rangle, \quad \dim\ V_{\lambda_1} = p_1 \\
V_{\lambda_2} &= \left\langle \left\{ v_{11}^2, \ldots, v_{p_2 1}^2 \right\} \right\rangle, \quad \dim\ V_{\lambda_2} = p_2 \\
&\vdots \\
V_{\lambda_r} &= \left\langle \left\{ v_{11}^r, \ldots, v_{p_r 1}^r \right\} \right\rangle, \quad \dim\ V_{\lambda_r} = p_r
\end{aligned}
$$

Los bloques se notan:

$$
J_i = \begin{pmatrix} \lambda_i & 1 & & \bigcirc \\ & \ddots & \ddots & \\ \bigcirc & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}
$$

Definición (Forma de Jordan) Una transformación lineal

$$
\begin{aligned}
T_A : \mathbb{C}^n &\longrightarrow \mathbb{C}^n \\
x &\longmapsto T_A(x) = Ax
\end{aligned}
$$

se reduce a la **forma de Jordan** si y sólo si es posible determinar una base β_J con las características definidas en (8.1) y (8.2).

Veamos otro ejemplo. Sea la matriz:

$$J = \begin{pmatrix} 8 & 1 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} J_1 & & \bigcirc \\ & J_2 & \\ \bigcirc & & J_3 \end{pmatrix}$$

donde $J_1 = \begin{pmatrix} 8 & 1 \\ 0 & 8 \end{pmatrix}$, $J_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $J_3 = (0)$.

Tenemos $\sigma(J) = \{\lambda_1 = 8, \lambda_2 = 0\}$ con multiplicidades 2 y 3 respectivamente.

$$V_{\lambda_1} = V_8 = \langle \{e_1\} \rangle\,, \quad V_{\lambda_2} = V_0 = \langle \{e_3, e_5\} \rangle$$

donde $\{e_i\}_{i=1}^5$ es la base canónica de $\mathbb{C}^5$. De acuerdo a nuestra notación, la base de Jordan, β_J es:

$$\begin{array}{ccc} v_{11}^1 & v_{11}^2 & v_{21}^2 \\ & & \\ v_{12}^1 & v_{12}^2 & \\ & & \\ \text{asociados a} & \text{asociados a} & \\ V_{\lambda_1} & V_{\lambda_2} & \end{array}$$

Estudiemos el esquema de similitud entre la matriz A (representante de T con respecto a la base canónica) y β_J:

$$\begin{array}{cccccccc} & & A & & & & & \\ T: & \mathbb{C}^5 & \rightarrow & \mathbb{C}^5 & & & & \\ & \beta & & \beta & & \beta & : & \text{base canónica} \\ P & \uparrow & & \uparrow & P & \beta_J & : & \text{base de Jordan} \\ & \mathbb{C}^5 & \rightarrow & \mathbb{C}^5 & & & & \\ & \beta_J & J & \beta_J & & & & \end{array}$$

Luego $J = P^{-1}AP$ o equivalentemente $AP = PJ$. Pero P es la matriz de pasaje de la base canónica a $\beta_J = \{v_{11}^1, v_{12}^1, v_{11}^2, v_{21}^2, v_{21}^2\}$. Es directo que:

$$P = (v_{11}^1, v_{12}^1, v_{11}^2, v_{12}^2, v_{21}^2)$$

y se tiene:

$$Av_{11}^1 = PJ_{\bullet 1} = P\begin{pmatrix}8\\0\\0\\0\\0\end{pmatrix} = 8v_{11}^1, \qquad Av_{12}^1 = PJ_{\bullet 2} = P\begin{pmatrix}1\\8\\0\\0\\0\end{pmatrix} = v_{11}^1 + 8v_{12}^1,$$

$$Av_{11}^2 = PJ_{\bullet 3} = P\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix} = 0v_{11}^2, \qquad Av_{12}^2 = PJ_{\bullet 4} = P\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} = v_{11}^2 + 0v_{12}^2,$$

$$Av_{21}^2 = PJ_{\bullet 5} = P\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix} = 0v_{21}^2$$

recuperándose la relación (8.2).

Simplifiquemos por un momento la notación. Si escribimos $P = (v^1, \ldots, v^n)$ tal que $\beta = \{v^1, \ldots, v^n\}$ es base de Jordan, se verifica:

$$\forall i = 1, \ldots, n \qquad Av^i = \lambda_i v^i \text{ o bien } Av^i = v^{i-1} + \lambda v^i.$$

En el primer caso v^i es un vector propio y en el segundo lo denominaremos vector propio generalizado (existe una “cola”).

En estas condiciones se tiene el teorema siguiente:

Teorema 8.1. *Una transformación lineal arbitraria, $T_A = \mathbb{C}^n \to \mathbb{C}^n,\ x \to Ax$, es reductible a la forma de Jordan, o de manera equivalente:*

$$\forall A \in \mathcal{M}_{nn}(\mathbb{C}),\ \exists J \in \mathcal{M}_{nn}(\mathbb{C})$$

matriz de Jordan, similar a la matriz A:

$$A = PJP^{-1}$$

donde las columnas de P son los vectores de la base asociada a β_J.

Demostración. La demostración la omitimos aquí y no se cubrirá en clases, sin embargo se presenta en el apéndice de la semana, **sólo por completitud**. $\square$

8.2 Aplicaciones de la forma de Jordan

Potencias de una matriz

Vimos anteriormente que la diagonalización de matrices nos entregaba una forma fácil de calcular las potencias de una matriz. Sin embargo, esto sólo es aplicable a matrices que admitan una diagonalización.

A continuación veremos que la forma de Jordan tiene también utilidad para el cálculo de potencias, con la ventaja que cualquier matriz cuadrada admite una forma de Jordan. Veamos primero que la forma de Jordan nos da una descomposición especial, para cualquier matriz cuadrada.

Si J es la matriz de Jordan de A, cada bloque de J es de la forma:

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda_i \end{pmatrix} = \lambda_i I_i + \begin{pmatrix} 0 & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & 0 \end{pmatrix},$$

en donde I_i es la identidad de la dimensión correspondiente. Llamemos N_i a la segunda matriz de la suma. Esta matriz tiene la siguiente propiedad, propuesta como ejercicio:

> **Ejercicio 8.1:** Suponga que $N_i \in \mathcal{M}_{ss}(\mathbb{C})$, pruebe entonces por inducción que $\forall m \in \mathbb{N}$, $\forall p, q \in \{1, \ldots, s\}$,
>
> $$(N_i^m)_{pq} = \begin{cases} 1 & \text{si } q = p + m, \\ \\ 0 & \text{si } q \neq p + m. \end{cases}$$
>
> Además deduzca que $N_i^s = \mathbf{0}$. Una matriz con esta propiedad, se dice **nilpotente**.

Luego, reescribiendola matriz A por bloques:

$$A = P \left[\underbrace{\begin{pmatrix} \lambda_1 I_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \lambda_2 I_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \lambda_r I_r \end{pmatrix}}_{D} + \underbrace{\begin{pmatrix} N_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & N_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & N_r \end{pmatrix}}_{N} \right] P^{-1}.$$

Notar que los valores propios $\lambda_1, \ldots, \lambda_r$ son los valores propios de A con posible repetición.

Se tiene que la matriz N es nilpotente, gracias al siguiente ejercicio:

> **Ejercicio 8.2:** Sea B una matriz diagonal por bloques, es decir
>
> $$B = \begin{pmatrix} B_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & B_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & B_r \end{pmatrix}$$
>
> en donde $B_1, \dots, B_r$ son matrices cuadradas y el resto son matrices cero.
>
> Sea $m \in \mathbb{N}$, pruebe que entonces
>
> $$B^m = \begin{pmatrix} B_1^m & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & B_2^m & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & B_r^m \end{pmatrix}.$$
>
> Concluya que la matriz N anterior es nilpotente.

Hemos probado entonces la siguiente propiedad:

Proposición 8.1. *Dada $A \in \mathcal{M}_{nn}(\mathbb{C})$, existen una matriz invertible P, una matriz diagonal D y una matriz nilpotente N, tales que*

$$A = P(D+N)P^{-1}.$$

Usaremos lo anterior para calcular potencias de una matriz. Sea $A \in \mathcal{M}_{nn}(\mathbb{C})$ y $m \in \mathbb{N}$. Queremos calcular A^m. Por lo visto en la parte de diagonalización:

$$A^m = P(D+N)^m P^t.$$

Debemos entonces estudiar $(D+N)^m$, que gracias al Ejercicio 8.2 es equivalente a estudiar $(\lambda_i I_i + N_i)^m$, es decir cada bloque de la forma de Jordan.

Se puede probar que el Teorema del Binomio de Newton es también válido para matrices, **si la multiplicación de éstas conmuta**. Este es precisamente el caso de $\lambda_i I_i$ y N_i, que conmutan pues $\lambda_i I_i$ es ponderación de la identidad. Luego

$$(\lambda_i I_i + N_i)^m = \sum_{k=0}^{m} \binom{m}{k} (\lambda_i I_i)^{m-k} N_i^k.$$

Para $k \in \{0, \dots, m\}$, $(\lambda_i I_i)^{m-k} = \lambda_i^{m-k} I_i$. Así, gracias al Ejercicio 8.1:

$$\binom{m}{k}(\lambda_i I_i)^{m-k} N_i^k = \binom{m}{k} \lambda_i^{m-k} N_i^k = \left(\begin{array}{ccccc} 0 & \dots & \binom{m}{k}\lambda_i^{m-k} & & \mathbf{0} \\ & \ddots & & \ddots & \\ & & & & \binom{m}{k}\lambda_i^{m-k} \\ & & & \ddots & \vdots \\ \mathbf{0} & & & & 0 \end{array}\right)$$

En donde los términos $\binom{m}{k}\lambda_i^{m-k}$ están ubicados desde la posición $(1,k)$ hasta la (k,n). Notar además que, si $\lambda_i + N_i \in \mathcal{M}_{ss}(\mathbb{C})$, la matriz anterior es nula para $k \geq s$.

Finalmente, sumando sobre k, se obtiene:

$$(\lambda_i I_i + N_i)^m = \begin{pmatrix} \lambda_i^m & \binom{m}{1}\lambda_i^{m-1} & \binom{m}{2}\lambda_i^{m-2} & \dots & \binom{m}{s-1}\lambda_i^{m-(s-1)} \\ 0 & \lambda_i^m & \binom{m}{1}\lambda_i^{m-1} & \binom{m}{2}\lambda_i^{m-2} & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \dots & 0 & \lambda_i^m & \binom{m}{1}\lambda_i^{m-1} \\ 0 & \dots & \dots & 0 & \lambda_i^m \end{pmatrix}. \tag{8.3}$$

Y tenemos entonces que

$$A^m = P \begin{pmatrix} (\lambda_1 I_1 + N_1)^m & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & (\lambda_2 I_2 + N_2)^m & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & (\lambda_r I_r + N_r)^m \end{pmatrix} P^{-1},$$

con los bloques $(\lambda_i I_i + N_i)^m$ como en (8.3).

Matriz exponencial

Otra aplicación importante está relacionada con calcular la exponencial de una matrix, la generalización de la función exponencial real para matrices. Dicha matriz exponencial tiene utilidad en ciertos métodos de resolución de ecuaciones diferenciales (que verás en cursos más adelante).

Veamos primero la definición de la matriz exponencial.

Definición (Matriz exponencial) Sea $A \in \mathcal{M}_{nn}(\mathbb{C})$. La **exponencial** de A, denotada por e^A, es la matriz en $\mathcal{M}_{nn}(\mathbb{C})$ dada por la serie de potencias

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Esta serie siempre converge, por lo que e^A está bien definida.

Observación: El último punto tratado en la definición, acerca de la convergencia, lo aceptaremos aquí. Esto sale de los tópicos a tratar en el curso.

Veamos primero algunas propiedades de la exponencial, las cuales no probaremos aquí:

Proposición 8.2. *Dadas matrices $B, C \in \mathcal{M}_{nn}(\mathbb{C})$, se tiene*

1. *Si C es invertible, $e^{CBC^{-1}} = Ce^BC^{-1}$.*
2. *Si $BC = CB$, entonces $e^{B+C} = e^Be^C = e^Ce^B$.*
3. *Si B es diagonal, $B = \operatorname{diag}(b_1, \ldots, b_n)$, entonces*
$$e^B = \operatorname{diag}(e^{b_1}, \ldots, e^{b_n}).$$

Sea entonces $A \in \mathcal{M}_{nn}(\mathbb{C})$, que gracias a la Proposición 8.1 se escribe como $A = P(D+N)P^{-1}$, con D matriz diagonal con los vectores propios de A y N una matriz nilpotente.

Calculamos entonces la exponencial de A, utilizando la Proposición 8.2 y el hecho de que D y N conmutan (ya se sus bloques conmutan):

$$e^A = Pe^{(D+N)}P^{-1} = P(e^De^N)P^{-1}.$$

Veamos e^N,

$$e^N = \sum_{k=0}^{\infty} \frac{1}{k!} N^k.$$

Gracias al Ejercicio 8.2 y sumando en k, se tiene que (este paso se puede formalizar, pero no lo haremos aquí)

$$e^N = \begin{pmatrix} e^{N_1} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & e^{N_2} & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & e^{N_r} \end{pmatrix}.$$

Además, por la Proposición 8.2,

$$e^D = \begin{pmatrix} e^{\lambda_1} I_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & e^{\lambda_2} I_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & e^{\lambda_r} I_r \end{pmatrix},$$

el i-ésimo bloque de e^De^N es:

$$e^{\lambda_i} I_i e^{N_i} = e^{\lambda_i} \sum_{k=0}^{\infty} \frac{1}{k!} N_i^k = e^{\lambda_i} \sum_{k=0}^{s-1} \frac{1}{k!} N_i^k.$$

Ya que si suponemos $N_i \in \mathcal{M}_{ss}(\mathbb{C})$, como N_i es nilpotente, para $k \geq s$ se tiene $N_i^k = \mathbf{0}$. Así:

$$e^{\lambda_i} e^{N_i} = \begin{pmatrix} e^{\lambda_i} & \frac{1}{1!} e^{\lambda_i} & \frac{1}{2!} e^{\lambda_i} & \ldots & \frac{1}{(s-1)!} e^{\lambda_i} \\ 0 & e^{\lambda_i} & \frac{1}{1!} e^{\lambda_i} & \frac{1}{2!} e^{\lambda_i} & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ldots & 0 & e^{\lambda_i} & \frac{1}{1!} e^{\lambda_i} \\ 0 & \ldots & \ldots & 0 & e^{\lambda_i} \end{pmatrix} \tag{8.4}$$

Finalmente

$$e^A = P \begin{pmatrix} e^{\lambda_1}e^{N_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & e^{\lambda_2}e^{N_2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & e^{\lambda_r}e^{N_r} \end{pmatrix} P^{-1},$$

con cada bloque como en (8.4).

Teorema de Cayley-Hamilton

El Teorema de Cayley-Hamilton señala que una matriz es siempre raiz de su polinomio característico. Para comprender a qué nos referimos con esto, notemos que dado un polinomio

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \in \mathcal{P}_n(\mathbb{C}),$$

podemos asociarle naturalmente una función de $\mathcal{M}_{nn}(\mathbb{C})$ a $\mathcal{M}_{nn}(\mathbb{C})$. Dada $X \in \mathcal{M}_{nn}(\mathbb{C})$:

$$p(X) = a_0I + a_1X + a_2X^2 + \cdots + a_nX^n.$$

Enunciemos entonces el teorema:

Teorema 8.2 (Cayley-Hamilton). *Dada una matriz $A \in \mathcal{M}_{nn}(\mathbb{C})$ y $p_A(x)$ su polinomio característico, entonces*

$$p_A(A) = \mathbf{0}.$$

Demostración. Consideremos $A = PJP^{-1}$, la forma de Jordan de A. Sabemos que dos matrices similares tienen el mismo polinomio característico, luego

$$p_A(x) = p_J(x).$$

Ahora, como J es triangular superior, $p_J(x)$ puede ser calculado como:

$$p_J(x) = |J - xI| = (x - \lambda_1)^{s_1}(x - \lambda_2)^{s_2} \dots (x - \lambda_r)^{s_r},$$

en donde cada término $(x - \lambda_i)^{s_i}$ corresponde al bloque de Jordan J_i.

Calculemos entonces $p_A(A)$. Gracias a lo visto en la parte de diagonalización, no es difícil probar que:

$$p_A(A) = p_J(A) = P \cdot p_J(J) \cdot P^{-1} = P(J - \lambda_1 I)^{s_1}(J - \lambda_2 I)^{s_2} \dots (J - \lambda_r I)^{s_r} P^{-1}.$$

Ahora, mirando el bloque i-ésimo del término i de este producto (de dimensiones $s \times s$), tenemos que:

$$(J_i - \lambda_i I_i)^s = N_i^s = \mathbf{0},$$

ya que gracias al Ejercicio 8.1, la matriz N_i es nilpotente.

Luego, usando el Ejercicio 8.2, se concluye. □

Traza de una matriz

Una última utilidad de la forma de Jordan, que veremos, es permitir caracterizar la **traza** de una matriz.

DEFINICIÓN (TRAZA) La **traza** es la función $\mathrm{tr}\colon \mathcal{M}_{nn}(\mathbb{C}) \to \mathbb{C}$, definida por:

$$\forall A = (a_{ij}) \in \mathcal{M}_{nn}(\mathbb{C}), \quad \mathrm{tr}(A) = \sum_{k=0}^{n} a_{ii}.$$

Es decir, es la suma de los elementos de la diagonal de la matriz.

Las siguientes propiedades quedan de ejercicio:

Proposición 8.3. *Se tienen las siguientes propiedades:*

1. *La función tr es lineal.*
2. $\forall A, B \in \mathcal{M}_{nn}(\mathbb{C})$, $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

DEMOSTRACIÓN. Propuesta como ejercicio. □

Así, podemos probar el siguiente resultado:

Proposición 8.4. *Sea $A \in \mathcal{M}_{nn}(\mathbb{C})$, con valores propios **distintos** $\lambda_1, \ldots, \lambda_r$. Entonces*

$$\mathrm{tr}(A) = \sum_{i=1}^{r} \alpha_A(\lambda_i) \cdot \lambda_i.$$

En donde recordemos que $\alpha_A(\lambda_i)$ es la multiplicidad algebraica de λ_i.

DEMOSTRACIÓN. Sea $A = PJP^{-1}$, la forma de Jordan de A, luego

$$\mathrm{tr}(A) = \mathrm{tr}(PJP^{-1}) = \mathrm{tr}(P^{-1}PJ) = \mathrm{tr}(J).$$

En donde ocupamos la Proposición 8.3, con PJ y P^{-1}.

Ahora, en la diagonal de J están precisamente los valores propios de A, repetidos tantas veces como su mutiplicidad algebraica. Así,

$$\mathrm{tr}(A) = \mathrm{tr}(J) = \sum_{i=1}^{r} \alpha_A(\lambda_i) \cdot \lambda_i.$$

□

Forma de Jordan

Probaremos el siguiente teorema:

Teorema 8.3. *Una transformación lineal arbitraria, $T_A = \mathbb{C}^n \to \mathbb{C}^n$, $x \to Ax$, es reductible a la forma de Jordan, o de manera equivalente:*

$$\forall A \in \mathcal{M}_{nn}(\mathbb{C}),\ \exists J \in \mathcal{M}_{nn}(\mathbb{C})$$

matriz de Jordan, similar a la matriz A:

$$A = PJP^{-1}$$

donde las columnas de P son los vectores de la base asociada a β_J.

Demostración. Por inducción sobre la dimensión, n, del espacio. Es trivialmente cierto para $n = 1$ (verifique). Supongamos que el resultado se cumple $\forall k \leq n-1$. Veremos primero:

Caso 1: La matriz A es singular ($det A = 0$).

En este caso $r(A) = r < n$, es decir $\dim Im(T_A) = r(A) = r < n$. Sea:

$$T' = T\ |_{ImT_A} \colon ImT_A \to ImT_A$$

Por hipótesis, como $\dim ImT_A < n$, existe la descomposición de Jordan. Existe entonces una base de ImT_A que verifica las hipótesis. Más explícitamente: Si $\sigma(T') = \{\lambda_1, \ldots, \lambda_m\}$ y además los subespacios propios:

$$V_{\lambda_1} = \left\langle \{v_{11}^1, \ldots, v_{p_1 1}^1\} \right\rangle,\ \ldots,\ V_{\lambda_m} = \left\langle \{v_{11}^m, \ldots, v_{p_m 1}^m\} \right\rangle$$

obteniéndose la base β, de Jordan para T':

$$\begin{array}{ccccc}
v_{11}^1 & \cdots & v_{p_1 1}^1 & \to & \text{vectores propios de } V_{\lambda_1} \\
 & & & & \\
v_{12}^1 & & v_{p_1 2}^1 & & \\
\vdots & & \vdots & & \text{colas asociadas} \\
v_{1s(1,1)}^1 & & v_{ps(1,p_1)}^1 & & \text{a los vectores propios de } V_{\lambda_1} \\
\vdots & & \vdots & & \vdots \\
v_{11}^m & \cdots & v_{p_m 1}^m & \to & \text{vectores propios de} V_{\lambda_m} \\
 & & & & \\
v_{12}^m & \cdots & v_{p_m 2}^m & & \\
\vdots & & \vdots & & \text{colas asociadas} \\
v_{1s(m,1)}^m & \cdots & v_{p_m s(m,p_m)}^m & &
\end{array}$$

donde $r = \dim ImT_A = \sum_{i,j} s(i,j)$ y se verifica:

$$T'(v_{ij}^k) = \begin{cases} \lambda_k v_{ij}^k & \text{si } j = 1 \\ v_{ij-1}^k + \lambda_k v_{ij}^k & \text{si } j > 1 \end{cases}$$

Donde:

$$J = \begin{pmatrix} J_{11} & & & & & & & & & \\ & \ddots & & & & & \bigcirc & & & \\ & & J_{p1} & & & & & & & \\ & & & J_{12} & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & J_{p_2 2} & & & & \\ & & & & & & \ddots & & & \\ & & \bigcirc & & & & & J_{1m} & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & J_{p_m m} \end{pmatrix}$$

y cada bloque:

$$J_{ik} = \begin{pmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \bigcirc \\ \bigcirc & & \ddots & 1 \\ & & & \lambda_k \end{pmatrix}$$

es de dimensión $s(k,i) \times s(k,i)$, asociado a los vectores de base $v_{i1}^k, \ldots, v_{is(k,i)}^k$.

Subcaso 1: $KerT_A \cap ImT_A = \{0\}$.

De acuerdo al Teorema del Núcleo-Imagen (TNI), $\dim KerT_A + \dim ImT_A = n$, luego $\mathbb{C}^n = KerT_A \oplus ImT_A$, de donde la base de Jordan será:

$$\beta_J = \{w^1, \ldots, w^r\} \cup \{z^1, \ldots, z^{n-r}\}$$

donde $\{w^i\}_{i=1}^r$ es base de Jordan de ImT_A y $\{z^i\}_{i=1}^{n-r}$ es una base del núcleo de T_A.

La matriz asociada es:

$$J = \begin{pmatrix} J_{11} & & & & & \\ & \ddots & & & \bigcirc & \\ & & J_{p_m m} & & & \\ & & & 0 & & \\ & \bigcirc & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

$n - r$ bloques de 1×1 con el valor propio 0

que verifica claramente:

$$T_A w^i = \lambda_i w^i \quad \text{ó} \quad T_A w^i = w^{i-1} + \lambda_i w^i \qquad 1 \leq i \leq r$$

y además, $T_A z^i = 0 z^i \quad 1 \leq i \leq n-r$.

Subcaso 2: $KerT_A \cap ImT_A = \langle \{u^1, \ldots, u^p\} \rangle$, subespacio de dimensión $p \geq 1$.

Es decir $u^i \in KerT_A$ y $u^i \in ImT_A$. Estos vectores son propios $T_A(u^i) = 0u^i$. Además, como $u^i \in ImT_A$ estos son vectores propios de $T' = T|_{ImT_A}$. Por hipótesis de Jordan existen, para cada uno de los vectores u^i, un bloque (eventualmente de 1×1). Sin pérdida de generalidad, supongamos que en la base de Jordan asociada a T', se tiene $\lambda_1 = 0$. Luego obtenemos los bloques siguientes:

$$\begin{array}{cccc} v^1_{11} & v^1_{21} & \cdots & v^1_{p1} \\ v^1_{12} & v^1_{22} & & \vdots \\ \vdots & \vdots & & \vdots \\ v^1_{1s(1,1)} & v^1_{2s(1,2)} & \cdots & v^1_{ps(1,p)} \end{array} \quad \text{vectores propios asociados a } \lambda_1 = 0$$

que verifican la recurrencia:

$$\begin{aligned} & T_A v^1_{11} = 0 \cdot v^1_{11} \\ & T_A v^1_{12} = v^1_{11} + 0 \cdot v^1_{12} \\ & \vdots \\ & T_A v^1_{1s(1,1)} = v^1_{1s(1,1)-1} + 0 \cdot v^1_{1s(1,1)} \\ & \vdots \\ & T_A v^1_{p1} = 0 \cdot v^1_{p1} \\ & \vdots \\ & T_A v^1_{ps(1,p)} = v^1_{ps(1,p)-1} + 0 \cdot v^1_{ps(1,p)} \end{aligned} \tag{8.5}$$

Para "completar" la base de Jordan asociada a ImT_A a una base de $\mathbb{C}^n$, elegimos el conjunto de vectores $\beta_2 = \{y^j\}_{j=1}^p$ tal que:

$$\begin{aligned} Ay^1 &= v^1_{1s(1,1)} \\ Ay^2 &= v^1_{2s(1,2)} \\ &\vdots \\ Ay^p &= v^1_{ps(1,p)} \end{aligned}$$

Es decir, los vectores $\{y^j\}_{j=1}^p$ son soluciones de los sistemas asociados al último vector de cada bloque del vector $\lambda_1 = 0$. Estos sistemas tienen solución ya que, por hipótesis, $v^1_{js(1,j)} \in ImT_A, \quad \forall j = 1, \ldots, p$.

Además, son vectores linealmente independientes. En efecto, si $\sum_{j=1}^p c_j y^j = 0$ entonces

$$\sum_{j=1}^p c_j T_A y^j = \sum_{j=1}^p c_j A y^j = \sum_{j=1}^p c_j v^1_{js(1,j)} = 0$$

como, por hipótesis el conjunto $\{v^1_{js(1,j)}\}_{j=1}^p$ es linealmente independiente, concluímos que $c_j = 0, \quad \forall j = 1, \ldots, n$.

Más aún, los vectores $\{y^j\}_{j=1}^p$ satisfacen:

$$T_A(y^j) = Ay^j = v_{js(1,j)} + 0y^j \qquad 1 \leq j \leq p.$$

luego podemos redefinir el orden de los vectores en el conjunto $\beta_2 \cup \beta_J$ para obtener una descomposición en bloques de Jordan. Por ejemplo, veamos la inclusión del vector y^1.

Teníamos (ver Ecuación (8.5)):

$$\begin{aligned} T_A v^1_{11} &= 0v^1_{11} \\ &\vdots \\ T_A v^1_{1s(1,1)} &= v^1_{1s(1,1)-1} + 0v^1_{1s(1,1)} \end{aligned}$$

definiendo $v^1_{1s(1,1)+1} = y^1$, se tiene que

$$T_A v^1_{1s(1,1)+1} = v^1_{1s(1,1)} + 0v^1_{1s(1,1)+1}$$

y redefinimos el primer bloque:

$$v^1_{11}, \ldots, v^1_{1s(1,1)}, v^1_{1s(1,1)+1} = y^1$$

que satisface la recurrencia de Jordan.

En general, para cada bloque asociado a λ_1, agregamos el vector $v^1_{js(1,j)+1} = y^j$, lo cual satisface la recurrencia de Jordan. Los nuevos bloques tienen la forma:

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

Debemos verificar que el conjunto $\beta_2 \cup \beta_J$ es $\ell.i.$: Consideremos φ una combinación lineal nula, escrita por bloques:

$$\begin{aligned} \varphi = {} & [\alpha^1_{11} v^1_{11} + \alpha^1_{12} v^1_{12} + \cdots + \alpha^1_{1s(1,1)} v^1_{1s(1,1)}] + \ldots \\ & + \cdots + [\alpha^1_{p1} v^1_{p1} + \alpha^1_{p2} v^1_{p2} + \cdots + \alpha^1_{ps(1,p)} v^1_{1s(1,p)}] + \ldots \\ & + [\alpha^k_{11} v^k_{11} + \cdots + \alpha^k_{1s(k,1)} v^k_{1s(k,1)}] + \ldots \\ & + [\alpha^k_{p_k 1} v_{p_k s(1,p_k)} + \cdots + \alpha^k_{p_k s(k,p_k)} v^k_{p_k s(k,p_k)}] + \ldots \\ & + \beta_1 y^1 + \beta_2 y^2 + \cdots + \beta_p y^p = 0 \end{aligned} \tag{8.6}$$

Aplicando la transformación T_A a esta ecuación obtenemos:

$$\begin{aligned}
T_A(\varphi) = A\varphi = [\alpha_{11}^1 A v_{11}^1 + \alpha_{12}^1 A_2 v_{12}^1 + \cdots + \alpha_{1s(1,1)}^1 A v_{1s(1,1)}^1] + \dots \\
\cdots + [\alpha_{p1}^1 A v_{p1}^1 + \alpha_{p2}^1 A v_{p2}^1 + \cdots + \alpha_{ps(1,p)}^1 A v_{1s(1,p)}^1] + \\
\cdots + [\alpha_{11}^k A v_{11}^k + \cdots + \alpha_{1s(k,1)}^k A v_{1s(k,1)}^k] + \\
\cdots + [\alpha_{p_k 1} A v_{p_k s(1,p_k)} + \cdots + \alpha_{p_k s(k,p_k)}^k A v_{p_k s(k,p_k)}^k] + \dots \\
\cdots + \beta_1 A y^1 + \cdots + \beta_p A y^p = 0
\end{aligned}$$

Reemplazando según la recurrencia de Jordan y recordando que $\lambda_1 = 0$:

$$\begin{aligned}
A\varphi = [\alpha_{12}^1 v_{11}^1 + \alpha_{13}^1 v_{12}^1 + \cdots + \alpha_{1s(1,1)}^1 v_{1s(1,1)-1}^1] + \dots \\
+ [\alpha_{p2}^1 v_{p1}^1 + \alpha_{p3}^1 v_{p2}^1 + \cdots + \alpha_{ps(1,p)}^1 v_{1s(1,p)-1}^1] + \dots \\
+ [\alpha_{11}^k \lambda_k v_{11}^k + \cdots + \alpha_{1s(k,1)}^k (\lambda_k v_{pts(k,1)}^k + v_{1s(k,1)-1}^k)] \\
\cdots + [\alpha_{p_k 1}^k \lambda_k v_{p_k 1}^k + \cdots + \alpha_{p_k s(k,p_k)}^k (\lambda_k v_{p_k s(k,p_k)}^k + v_{p_k s(k,p_k)-1}^k)] + \dots \\
+ \beta_1 v_{1s(1,1)}^1 + \beta_2 v_{2s(1,2)}^1 + \cdots + \beta_p v_{ps(1,p)}^1 = 0
\end{aligned}$$

Reordenando, agregamos cada vector asociado a los escalares β_j a su respectivo bloque:

$$\begin{aligned}
T_A(\varphi) = [\alpha_{12}^1 v_{11}^1 + \cdots + \alpha_{1s(1,1)}^1 v_{1s(1,1)-1}^1 + \beta^1 v_{1s(1,1)}^1] + \dots \\
\cdots + \{(\alpha_{11}^k \lambda_k + \alpha_{12}^k) v_{11}^k + (\alpha_{12}^k \lambda_k + \alpha_{13}^k) v_{12}^k + \dots \\
\cdots + (\alpha_{1s(k,1)-1}^k \lambda_k + \alpha_{1s(k,1)}^k) v_{1s(k,1)-1}^k + \alpha_{1s(k,1)}^k \lambda_k v_{1s(k,1)}^k\} + \cdots = 0
\end{aligned}$$

como los vectores son $\ell.i$ se tiene que

$$\begin{aligned}
\alpha_{12}^1 = \alpha_{13}^1 = \cdots = \alpha_{1s(1,1)}^1 = \beta_1 = 0 \\
\vdots \\
\alpha_{p2}^1 = \alpha_{p3}^1 = \cdots = \alpha_{ps(1,p)}^1 = \beta_p = 0
\end{aligned}$$

y $\forall k \geq 1$ (bloques asociados a λ_k):

$$\begin{aligned}
\alpha_{11}^k \lambda_k + \alpha_{12}^k = 0 \\
\alpha_{12}^k \lambda_k + \alpha_{13}^k = 0 \\
\vdots \\
\alpha_{1s(k,1)-1}^k \lambda_k + \alpha_{1s(k,1)}^k = 0 \\
\alpha_{1s(k,1)}^k \lambda_k = 0
\end{aligned}$$

como $\lambda_k \neq 0$, reemplazando desde la última a la primera ecuación, concluímos $\alpha_{ij}^k = 0 \quad \forall k, \forall i, j$. Recordemos que, al aplicar T_A, desaparecieron los coeficientes $\alpha_{11}^1, \dots, \alpha_{p1}^1$, pero, como el resto de los coeficientes α_{ij}^k, β_j, es nulo se tiene en (8.6):

$$\alpha_{11}^1 v_{11}^1 + \cdots + \alpha_{p1}^1 v_{p1}^1 = 0$$

Como $\{v_{j1}^1\}_{j=1}^p$ es $\ell.i$, concluímos $\alpha_{11}^1 = \cdots = \alpha_{p1}^1 = 0$, luego el conjunto β_2 o β_J es $\ell.i$.

El conjunto anterior $\beta_2 \cup \beta_J$ tiene $r + p$ vectores, donde $r = \dim ImT_A$, es decir $| \beta_2 \cup \beta_J | = p + r \leq n$. Nos falta entonces considerar $KerT_A \setminus ImT_A \cap KerT_A$. Pero esto es fácil, sea:

$$KerT_A = \left\langle \left\{ u^1, \ldots, u^p \right\} \right\rangle \oplus \langle \{ z_1, \ldots, z_q \} \rangle$$

donde $\langle \{ u^1, \ldots, u^p \} \rangle = ImT_A \cap KerT_A$ y $\langle \{ z_1, \ldots, z_q \} \rangle$ es un suplementario en $KerT_A$.

Afirmamos que la base de Jordan de $\mathbb{C}^n$ está dada por:

$$\beta'_J = \beta_J \cup \beta_2 \cup \beta_3, \quad \beta_3 = \{ z_i \}_{i=1}^{q}.$$

Claramente estos n vectores son $\ell.i.$, pues si

$$\sum_{j=1}^{r} \alpha_j w_j + \sum_{j=1}^{p} \beta_j y_j + \sum_{j=1}^{q} \gamma_j z_j = 0$$

aplicando T_A y recordando que $z_j \in KerT_A$ obtenemos:

$$\sum_{j=1}^{r} \alpha_j T_A w_j + \sum_{j=1}^{p} \beta_j T_A y_j = 0$$

que corresponde al análisis de la independencia lineal de $\beta_2 \cup \beta_J$ que hicimos previamente, luego $\alpha_j = \beta_j = 0$. Sólo queda entonces $\sum\limits_{j=1}^{q} \gamma_j z_j = 0$ de donde $\gamma_j = 0$ pués $\{ z_j \}$ es $\ell.i.$

La matriz de Jordan, agregando los $\{ z_i \}_{i=1}^{q}$ al final de la base, es la siguiente:

$$J = \begin{pmatrix}
\tilde{J}_{11} & & & & & & & & & & & & \\
& \ddots & & & & & & & & & & & \\
& & \tilde{J}_{p1} & & & & & & \bigcirc & & & & \\
& & & J_{12} & & & & & & & & & \\
& & & & \ddots & & & & & & & & \\
& & & & & J_{p_2 2} & & & & & & & \\
& & & & & & \ddots & & & & & & \\
& & & & & & & J_{1m} & & & & & \\
& & & & & & & & \ddots & & & & \\
& & \bigcirc & & & & & & & J_{p_m m} & & & \\
& & & & & & & & & & 0 & & \\
& & & & & & & & & & & \cdots & \\
& & & & & & & & & & & & 0
\end{pmatrix}$$

donde $\tilde{J}_{11}, \ldots, \tilde{J}_{p1}$ son los bloques asociados a $\lambda_1 = 0$ con la "cola" agregada $\{ y_i \}$. El último grupo de q bloques de 0's corresponde a la base del suplementario de $ImT_A \cap KerT_A$ en $KerT_A$.

Con esto hemos demostrado que, si A es regular (no invertible), esta es reductible a la forma de Jordan.

Veamos ahora el caso invertible: Como $A \in M_{nn}(\mathbb{C}), \exists \lambda \in \mathbb{C}$ valor propio asociado a A. Consideremos la matriz singular $A' = A - \lambda I$. Por el procedimiento desarrollado anteriormente existen matrices J' y P invertible tal que:

$$
\begin{aligned}
J' = P^{-1}A'P &\Leftrightarrow J' = P^{-1}(A - \lambda I)P \\
\Leftrightarrow \quad PJ'P^{-1} = A - \lambda I &\Leftrightarrow \quad A = PJ'P^{-1} + \lambda I = \\
&= PJ'P^{-1} + \lambda PP^{-1} = P(J' + \lambda I)P^{-1}
\end{aligned}
$$

luego la matriz de Jordan asociada a T_A es $J = J' + \lambda I$, donde λ aparece en la diagonal principal. $\square$

Ejercicio 8.3: La forma de Jordan es única, salvo permutación del orden de la base.

Guía de Ejercicios

1. Sea $N_i \in \mathcal{M}_{ss}(\mathbb{C})$, definida por

$$\begin{pmatrix} 0 & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & 0 \end{pmatrix}.$$

Pruebe por inducción que $\forall m \in \mathbb{N}$, $\forall p, q \in \{1, \ldots, s\}$,

$$(N_i^m)_{pq} = \begin{cases} 1 & \text{si } q = p + m, \\ \\ 0 & \text{si } q \neq p + m. \end{cases}$$

Además deduzca que $N_i^s = \mathbf{0}$. Una matriz con esta propiedad, se dice **nilpotente**.

2. Sea B una matriz diagonal por bloques, es decir

$$B = \begin{pmatrix} B_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & B_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & B_r \end{pmatrix}$$

en donde $B_1, \ldots, B_r$ son matrices cuadradas y el resto son matrices cero.

Sea $m \in \mathbb{N}$, pruebe que entonces

$$B^m = \begin{pmatrix} B_1^m & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & B_2^m & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & B_r^m \end{pmatrix}.$$

3. Pruebe las siguientes propiedades:

(a) La función $\mathrm{tr}\colon \mathcal{M}_{nn}(\mathbb{C}) \to \mathbb{C}$ es lineal.

(b) $\forall A \in \mathcal{M}_{mn}(\mathbb{C}),\ \forall B \in \mathcal{M}_{nm}(\mathbb{C})$, $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

4. $A \in \mathcal{M}_{nn}(\mathbb{C})$, una matriz no necesariamente diagonalizable tal que $\sigma(A) \subseteq \mathbb{R}$. Pruebe que

$$|e^A| = e^{\mathrm{tr}(A)}.$$

5. Calcule e^A, A^3 y A^5, para las siguientes matrices:

(a) $\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$.

(b) $\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$.

(c) $\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

Guía de Problemas

P1. Se define la siguiente recurrencia de números reales

$$\begin{cases} x_{n+1} = 2x_n - x_{n-1}, & \forall n \geq 1 \\ x_1 = 1, x_0 = 0. \end{cases}$$

(a) Para $n \in \mathbb{N}$, defina $v_n = \left(\begin{smallmatrix} x_n \\ x_{n-1} \end{smallmatrix} \right) \in \mathbb{R}^2$. Encuentre $A \in \mathcal{M}_{22}(\mathbb{R})$ tal que

$$v_{n+1} = Av_n.$$

(b) Pruebe que $v_n = A^n v_0$, $\forall n \in \mathbb{N}$.

(c) Use lo anterior para resolver la recurrencia, obteniendo un valor explícito para x_n.
Indicación: Use la forma de Jordan.

P2. **(a)** Sea $\lambda \in \mathbb{R} \setminus \{0\}$. Pruebe por inducción que

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} \quad \forall n \geq 1.$$

(b) Considere

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

Encuentre P invertible y J en forma de Jordan tal que $A = PJP^{-1}$.

(c) Para la matriz A de la parte anterior y $n \geq 1$ cualquiera calcule A^n explícitamente, encontrando expresiones para los coeficientes de A^n en función de n.