Apuntes Álgebra Lineal

MIGUEL MONTES LORENZO

May 2022

Índice general

Capítulo 1

Espacios Vectoriales

1.1. Espacio vectorial

Un espacio vectorrial (E) es una estructura algebraica formada por un conjunto (C) no vacío, una operación interna llamada suma $(+)$ y una operación externa llamada producto por escalar $(\cdot)$. A los elementos de un espacio vectorial se los llama vectores.

Las operaciones $(+, \cdot)$ de un espacio vectorial (E) cumplen las siguientes 2 propiedades :

1. Suma $\longrightarrow \quad \forall \bar{a}, \bar{b} \in E \qquad \bar{a} + \bar{b} = \bar{c} \quad tq. \ \bar{c} \in E$
2. Producto por escalar $\longrightarrow \quad \forall \bar{a} \in E \ \& \ \lambda \in \mathbb{R} \qquad \lambda \cdot \bar{a} = \bar{d} \quad tq. \ \bar{d} \in E$

1.2. Axiomas de un espacio vectorial

Para que la terna $(C, +, \cdot)$ constituya un espacio vectorial, esta debe cumplir una serie de propiedades internas llamadas axiomas. (Si se cumplen estos axiomas también se cumplirán las propiedades de la suma y el producto por escalar previamente mencionadas).

1.2.1. Axiomas de la suma $(+ : E \times E \longmapsto E)$

$\forall \bar{x}, \bar{y}, \bar{z} \in E$

- Asociativa $\longrightarrow$ $\bar{x} + (\bar{y} + \bar{z}) = (\bar{x} + \bar{y}) + \bar{z}$
- Conmuntativa $\longrightarrow$ $\bar{x} + \bar{y} = \bar{y} + \bar{x}$
- $\exists$ elemento neutro $\longrightarrow$ $\exists \bar{0} \in E$ tq. $\bar{x} + \bar{0} = \bar{x}$
- $\exists$ elemento opuesto $\longrightarrow$ $\exists (-\bar{x}) \in E$ tq. $\bar{x} + (-\bar{x}) = \bar{0}$

1.2.2. Axiomas del producto por escalar $(\cdot : \mathbb{R} \times E \longmapsto E)$

$\forall \bar{x}, \bar{y} \in E$ y $\forall \lambda, \mu \in \mathbb{R}$

- Asociativa $\longrightarrow$ $\lambda(\mu \cdot \bar{x}) = (\lambda \cdot \mu)\bar{x}$
- Distributiva en vectores $\longrightarrow$ $\lambda(\bar{x} + \bar{y}) = \lambda\bar{x} + \lambda\bar{y}$
- Distributiva en escalares $\longrightarrow$ $(\lambda + \mu)\bar{x} = \lambda\bar{x} + \mu\bar{x}$
- $\exists$ elemento identidad $\longrightarrow$ $\exists 1$ tq. $1 \cdot \bar{x} = \bar{x}$

Ejemplo

Demostrar que la terna formada por el conjunto $\mathbb{R}^n$ y las siguientes dos operaciones de suma $\boxplus$ y producto por escalar $\boxdot$ forman una estructura de espacio vectorial:

$$(x_1, \ldots, x_n) \boxplus (y_1, \ldots, y_n) = \left(\sqrt[3]{x_1^3 + y_1^3}, \ldots, \sqrt[3]{x_n^3 + y_n^3}\right)$$

$$\lambda \boxdot (x_1, \ldots, x_n) = \left(\sqrt[3]{\lambda}x_1, \ldots, \sqrt[3]{\lambda}x_n\right)$$

Axiomas de la suma

- Asociativa

$$\bar{x} \boxplus (\bar{y} \boxplus \bar{z}) = \bar{x} \boxplus \left(\left\{\sqrt{y_i^3 + z_i^3}\right\}_{i=1}^n\right) = \left(\left\{\sqrt[3]{x_i^3 + \left(\sqrt{y_i^3 + z_i^3}\right)^3}\right\}_{i=1}^n\right) =$$

$$= \left(\left\{\sqrt[3]{x_i^3 + y_i^3 + z_i^3}\right\}_{i=1}^n\right) = \left(\left\{\sqrt[3]{\left(\sqrt{x_i^3 + y_i^3}\right)^3 + z_i^3}\right\}_{i=1}^n\right) = \left(\left\{\sqrt[3]{x_i^3 + y_i^3}\right\}_{i=1}^n \boxplus \bar{z}\right) =$$

$$= (\bar{x} \boxplus \bar{y}) \boxplus \bar{z}$$

- Conmutativa

$$\bar{x} \boxplus \bar{y} = \left(\left\{\sqrt[3]{x_i^3 + y_i^3}\right\}_{i=1}^n\right) = \left(\left\{\sqrt[3]{y_i^3 + x_i^3}\right\}_{i=1}^n\right) = \bar{y} \boxplus \bar{x}$$

- $\exists$ elemento neutro

$$\bar{x} \boxplus \bar{0} = \left(\left\{\sqrt[3]{x_i^3 + 0}\right\}_{i=1}^n\right) = \left(\left\{\sqrt[3]{x_i^3}\right\}_{i=1}^n\right) = (\{x_i\}_{i=1}^n) = \bar{x}$$

- $\exists$ elemento opuesto

$$\bar{x} \boxplus (-\bar{x}) = \left(\left\{\sqrt[3]{x_i^3 + (-\bar{x})_i^3}\right\}_{i=1}^n\right) = \left(\left\{\sqrt[3]{x_i^3 - x_i^3}\right\}_{i=1}^n\right) = (\{0\}_{i=1}^n) = \bar{0}$$

Axiomas del producto

- Asociativa

$$\lambda \boxdot (\mu \boxdot \bar{x}) = \lambda \boxdot (\{\sqrt[3]{\mu}x_i\}_{i=1}^n) = \left(\left\{\sqrt[3]{\lambda}\sqrt[3]{\mu}x_i\right\}_{i=1}^n\right) = \left(\left\{\sqrt[3]{\mu}\sqrt[3]{\lambda}x_i\right\}_{i=1}^n\right) =$$

$$= \mu \boxdot \left(\left\{\sqrt[3]{\lambda}x_i\right\}_{i=1}^n\right) = \mu \boxdot (\lambda \boxdot \bar{x})$$

- Distributiva en vectores

$$\lambda \boxdot (\bar{x} \boxplus \bar{y}) = \lambda \boxdot \left(\left\{\sqrt[3]{x_i^3 + y_i^3}\right\}_{i=1}^n\right) = \left(\left\{\sqrt[3]{\lambda}\sqrt[3]{x_i^3 + y_i^3}\right\}_{i=1}^n\right) = \left(\left\{\sqrt[3]{\lambda(x_i^3 + y_i^3)}\right\}_{i=1}^n\right) =$$

$$= \left(\left\{\sqrt[3]{\lambda x_i^3 + \lambda y_i^3}\right\}_{i=1}^n\right) = (\lambda \boxdot \bar{x}) \boxplus (\lambda \boxdot \bar{y})$$

- Distributiva en escalares

$$(\lambda + \mu) \boxdot \bar{x} = \left(\left\{\sqrt[3]{(\lambda + \mu)x_i}\right\}_{i=1}^n\right) = \left(\left\{\sqrt[3]{(\lambda + \mu)x_i^3}\right\}_{i=1}^n\right) = \left(\left\{\sqrt[3]{\lambda x_i^3 + \mu x_i^3}\right\}_{i=1}^n\right) =$$

$$= (\lambda \boxdot \bar{x}) + (\mu \boxdot \bar{x})$$

- $\exists$ elemento identidad

$$1 \boxdot \bar{x} = \left(\left\{\sqrt[3]{1}x_i\right\}_{i=1}^n\right) = \left(\left\{1 \cdot x_i^3\right\}_{i=1}^n\right) = \left(\left\{x_i^3\right\}_{i=1}^n\right) = \bar{x}$$

$\square$

1.3. Propiedades de un espacio vectorial

De la verificación de los axiomas de espacio vectorial se desprenden otra serie de propiedades que siempre se cumplen dentro de este tipo de estructura algebraica.

- EXISTENCIA DE UN ÚNICO ELEMENTO NEUTRO ($\bar{0}$) :
  ($\forall \bar{x}_i \in E \;\; \exists! \bar{0} \;\mid\; \bar{x}_i + \bar{0} = \bar{x}_i$) Demostración por reducción al absurdo

  Suponemos que $\exists\, \bar{0}_1, \bar{0}_2 \in E \quad \text{tq.} \quad \forall \bar{x} \in E \;\; \bar{0}_1 + \bar{x} = \bar{x}$ y $\bar{0}_2 + \bar{x} = \bar{x}$

  $$\bar{0}_1 \Leftarrow \bar{0}_1 + \bar{0}_2 \Rightarrow \bar{0}_2 \quad !! \qquad \square$$

- PARA CADA $\bar{x}_i \in E$ EXISTE UN ÚNICO ELEMENTO OPUESTO DE LA SUMA
  ($\forall \bar{x}_i \in E \;\; \exists! x_j \;\mid\; \bar{x}_i + \bar{x}_j = \bar{0}$) Demostración por reducción al absurdo

  Suponemos que $\exists\, \bar{x}, \bar{x}_1, \bar{x}_2 \in V \quad \text{tq.} \quad \bar{x} + \bar{x}_1 = \bar{0}$ y $\bar{x} + \bar{x}_2 = \bar{0}$

  $$\bar{x}_1 + \bar{x} + \bar{x}_2 = (\bar{x}_1 + \bar{x}) + \bar{x}_2 = (\bar{x} + \bar{x}_1) + \bar{x}_2 = \bar{0} + \bar{x}_2 = \bar{x}_2 \; !!$$

  $$\bar{x}_1 + \bar{x} + \bar{x}_2 = \bar{x}_1 + (\bar{x} + \bar{x}_2) = (\bar{x} + \bar{x}_2) + \bar{x}_1 = \bar{0} + \bar{x}_1 = \bar{x}_1 \; !!$$

  $$\square$$

- PARA TODO $\bar{x}$ CONTENIDO EN E, $\bar{x} \cdot \bar{0} = \bar{0}$:

  $$\begin{cases} 0 \cdot \bar{x} = 0 \cdot \bar{x} + \bar{0} \\ 0 \cdot \bar{x} = (0+0) \cdot \bar{x} = 0 \cdot \bar{x} + 0 \cdot \bar{x} \end{cases} \qquad \square$$

  $$\bar{0} \Leftarrow \cancel{0 \cdot \bar{x}} + \bar{0} = \cancel{0 \cdot \bar{x}} + 0 \cdot \bar{x} \Rightarrow 0 \cdot \bar{x}$$

- PARA TODO λ CONTENIDO EN $\mathbb{R}$, $\lambda \cdot \bar{0} = \bar{0}$:

  $$\begin{cases} \lambda \cdot \bar{x} = \lambda \cdot \bar{0} + \bar{0} \\ 0 \cdot \bar{x} = \lambda \cdot (\bar{0} + \bar{0}) = \lambda \cdot \bar{0} + \lambda \cdot \bar{0} \end{cases} \qquad \square$$

  $$\bar{0} \Leftarrow \cancel{\lambda \cdot \bar{0}} + \bar{0} = \cancel{\lambda \cdot \bar{0}} + \lambda \cdot \bar{0} \Rightarrow \lambda \cdot \bar{0}$$

- SI ($\lambda \cdot \bar{x} = \bar{0}$), O BIEN ($\lambda = 0$), O BIEN($\bar{x} = \bar{0}$):

  - Caso 1 $\quad \longrightarrow \quad \lambda = 0 \quad \longrightarrow \quad \lambda \cdot \bar{x} = 0 \cdot \bar{x} = \bar{0}$

  - Caso 2 $\quad \longrightarrow \quad \lambda \neq 0 \quad \longrightarrow \quad \forall \lambda \in \mathbb{R} \;\; \exists \lambda^{-1}$ tq. $\lambda \cdot \lambda^{-1} = 1$
    $\lambda \cdot \bar{x} = \bar{0} \quad \longrightarrow \quad \lambda^{-1} \cdot (\lambda \cdot \bar{x}) = \lambda^{-1} \cdot (\bar{0}) \quad \longrightarrow \quad (\lambda^{-1} \cdot \lambda) \cdot \bar{x} = \lambda^{-1} \cdot \bar{0} \quad \longrightarrow \quad 1 \cdot \bar{x} = \bar{0} \quad \longrightarrow \quad \bar{x} = \bar{0}$

  $$\square$$

- PARA TODO $\lambda \in \mathbb{R}$ SE CUMPLE QUE $(-\lambda) \cdot \bar{x} = -\lambda \cdot \bar{x} = \lambda \cdot (-\bar{x})$:
  Veamos que los 3 elementos son el (único) opuesto de $\lambda \cdot \bar{x}$

  - $(-\lambda) \cdot \bar{x} + \lambda \cdot \bar{x} = (-\lambda + \lambda) \cdot \bar{x} = 0 \cdot \bar{x} = \bar{0}$
  - $-(\lambda \cdot \bar{x}) + \lambda \cdot \bar{x} = \bar{0}$ $\qquad \square$
  - $\lambda \cdot (-\bar{x}) + \lambda \cdot \bar{x} = \lambda(-\bar{x} + \bar{x}) = \lambda \cdot 0 = \bar{0}$

1.4. Subespacio vectorial

Un subespacio vectorial es un espacio vectorial contenido dentro de otro de dimensión mayor (Que contiene a un mayor número de vectores) con el que comparte las mismas operaciones de suma y producto por escalar.

Ejemplos

$\mathbb{R}$ es subespacio de $\mathbb{R}^2 \quad \longrightarrow \quad x \in \mathbb{R} \quad \bar{x} = (x_1) \xrightarrow{\mathbb{R} \to \mathbb{R}^2} \bar{x} = (x_1, 0)$

$\mathbb{R}^2$ es subespacio de $\mathbb{R}^n \quad \longrightarrow \quad x \in \mathbb{R}^2 \quad \bar{x} = (x_1, x_2) \xrightarrow{\mathbb{R}^2 \to \mathbb{R}^n} \bar{x} = (x_1, x_2, 0, \ldots, 0)$

S_2 es subespacio de M_2

$\longrightarrow \quad \bar{x} \in S_2 \quad \bar{x} = \left(\begin{smallmatrix} x_1 & x_2 \\ x_2 & x_3 \end{smallmatrix}\right) = (x_1, x_2, x_3)_{B_c S_2} \xrightarrow{S_2 \to M_2} \bar{x} = (x_1, x_2, x_2, x_3)_{B_c M_2}$

Nota. S_2 denota el espaio de las matrices simétricas 2×2.

1.5. Teorema de caracterización

Cuando tenemos un subconjunto V que comparte unas mismas operaciones de suma y producto por escalar con un espacio vectorial E, para probar que V es también un espacio vectorial (subespacio vectorial de E) basta con probar que cumple las condiciones del teorema de caracterización. Si no las cumple, tampoco cumplirá los axiomas de espacio vectorial.

Las condiciones del teorema de caracterización son las siguientes:

$$\forall \bar{x}, \bar{y} \in V \qquad \bar{x} + \bar{y} \in V$$
$$\forall \bar{x} \in V \ \mathbf{y}\ \forall \lambda \in \mathbb{R} \qquad \lambda \cdot \bar{x} \in V$$

Lo que es equivalente a:

$$\forall \bar{x}, \bar{y} \in V \ \mathbf{y}\ \forall \lambda, \mu \in \mathbb{R} \qquad \lambda \bar{x} + \mu \bar{y} \in V$$

Ejemplo
Demostrar que S_2 (Matrices simétricas 2×2) es subespacio de M_2

$$A = \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix} \qquad B = \begin{pmatrix} b_1 & b_2 \\ b_2 & b_3 \end{pmatrix} \qquad A, B \in S_2 = \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix}$$

- Suma $\longrightarrow$ Tomamos que $a_n + b_n = c_n$

$$A + B = \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix} + \begin{pmatrix} b_1 & b_2 \\ b_2 & b_3 \end{pmatrix} = \begin{pmatrix} a_1+b_1 & a_2+b_2 \\ a_2+b_2 & a_3+b_3 \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \\ c_2 & c_3 \end{pmatrix} \in S_2$$

- Producto por escalar $\longrightarrow$ Tomamos que $\lambda \cdot a_n = d_n$

$$\lambda \cdot A = \lambda \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix} = \begin{pmatrix} \lambda a_1 & \lambda a_2 \\ \lambda a_2 & \lambda a_3 \end{pmatrix} = \begin{pmatrix} d_1 & d_2 \\ d_2 & d_3 \end{pmatrix} \in S_2$$

Esto demuestra que S_2 es subespacio vectorial de cualquier espacio vectorial E (es decir $S_2 \in E$) que tenga las mismas operaciones de suma y producto $(+, \cdot)$ que S_2. En este caso, solamente M_2.

1.6. Sistemas libres y sistemas ligados

Un sistema $S = \{\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n\}$ se dice **libre**, o lo que es lo mismo, que todos sus vectores son linealmente independientes entre sí, cuando:

$$\lambda_1\bar{x}_1 + \lambda_2\bar{x}_2 + \cdots + \lambda_n\bar{x}_n = \bar{0} \iff \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$$

En caso contrario, si existe alguna combinación lineal de los vectores del sistema que con algún coeficiente de los vectores del sistema distinto de 0 proporcione el vector $\bar{0}$, entonces se dice que el sistema S es **ligado**, o lo que es lo mismo, que sus vectores son linealmente dependientes entre sí.

1.6.1. Propiedades de sistemas libres y sistemas ligados

- Si un sistema $S = \{\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n\}$ es un sistema **libre**, cualquier sistema S' extraido de S es **libre** también.
- Si $\bar{x} \neq \bar{0} \quad \longrightarrow \quad S = \{\bar{x}\}$ es un sistema **libre**.
- Un sistema S es **ligado** si y solo si al menos uno de sus vectores se puede expresar como combinación lineal de los demás.

  DEMOSTRACIÓN

  $S = \{\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n\} \quad \longrightarrow \quad$ ligado

  Es decir, $\{\lambda_i\}_{i=1}^n = \{\lambda_1, \lambda_2, \ldots, \lambda_n\} \quad (\lambda_i \in \mathbb{R} \;\&\; \{\lambda_i\}_{i=1}^n \neq \{0\}_{i=1}^n)$ tq. $\lambda_1\bar{x}_1 + \lambda_2\bar{x}_2 + \cdots + \lambda_n\bar{x}_n = \bar{0}$

  Dividiendo entre λ_j

  $$\frac{\lambda_1}{\lambda_j}\bar{x}_1 + \frac{\lambda_2}{\lambda_j}\bar{x}_2 + \cdots + 1 \cdot \bar{x}_j + \cdots + \frac{\lambda_{n-1}}{\lambda_j}\bar{x}_{n-1} + \cdots + \frac{\lambda_n}{\lambda_j}\bar{x}_n = \bar{0}$$

  Despejando $\bar{x}_j$

  $$-\left(\frac{\lambda_1}{\lambda_j}\bar{x}_1 + \frac{\lambda_2}{\lambda_j}\bar{x}_2 + \cdots + \frac{\lambda_{n-1}}{\lambda_j}\bar{x}_{n-1} + \cdots + \frac{\lambda_n}{\lambda_j}\bar{x}_n\right) = \bar{x}_j$$

  $\square$

- Si un sistema S contiene al vector $\bar{0}$ es ligado.
- Si un sistema $S = \{\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n\}$ es un sistema **ligado**, cualquier sistema S' que contenga a S es **ligado** también.
- Si un sistema S es libre y el sistema $S' = S \cup \bar{x}$ es un sistema **ligado**, entonces $\bar{x}$ es combinación lineal de los vectores de S.

1.6.2. Definición de $L\{S\}$

Dado un sistema S, se define el sistema generador de S, que denotaremos por $L\{S\}$, como el conjunto de todas las combinaciones lineales que se pueden obtener con los elementos de S.

$$S = \{\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n\}$$
$$L\{S\} = \{\bar{v} = \lambda_1\bar{x}_1 + \lambda_2\bar{x}_2, \lambda_3\bar{x}_3 + \cdots + \lambda_n\bar{x}_n \quad | \quad \lambda_i \in \mathbb{R}\}$$

Dos sistemas S_1 y S_2 son **equivalentes** si $L\{S_1\}$ y $L\{S_2\}$ generan el mismo conjunto de combinaciones lineales.

Aplicar las siguientes operaciones sobre un sistema lo mantiene siempre equivalente:

- Cambiar el orden de 2 vectores
- Sumar un vector del sistema a otro vector del sistema
- Multiplicar un vector por $\lambda \; (0 \neq \lambda \in \mathbb{R})$
- Añadir / Eliminar una combinación lineal del sitema

1.7. Base de un espacio vectorial

Una base de un espacio vectorial E es un sistema generador **libre** y **ordenado** de E.

Es decir, una base de E es un sistema generador que utiliza el mínimo número de vectores linealmente independientes necesario para generar cualquier vector contenido en dicho espacio E. Además, el orden en el que se referencian (Con las coordenadas del nuevo vector) los vectores de la base debe ser invariable (El mismo para todos los vectores que se quieran describir).

Nota. Si un espacio vectorial admite un número finito de vectores en su base se dice que es finitamente generado. En caso contrario se dice que es infinitamente generado.

1.7.1. Coordenadas de una base

Sea $B = \{\bar{e}_1, \bar{e}_2, \dots, \bar{e}_n\}$ una base de E y $\bar{x} \in E$ un vector tal que:

$$\bar{x} = x_1\bar{e}_1 + x_2\bar{e}_2 + \cdots + x_n\bar{e}_n$$

A $\{x_1, x_2, \dots, x_n\}$ se las llama **coordenadas de $\bar{x}$ en base** B, de forma que $\bar{x} = (x_1, x_2, \dots, x_n)_B$.

Ejemplos

$$(1,2,3) = (1,2,3)_{B_c\mathbb{R}^3} \quad \longrightarrow \quad B_c\mathbb{R}^3 = \{(1,0,0),(0,1,0),(0,0,1)\}$$
$$(1,2,3) = (1,1,1)_{B'} \quad \longrightarrow \quad B' = \{(1,1,1),(0,1,1),(0,0,1)\}$$

Cuando trabajamos con espacios vectoriales que describen elementos no vectoriales $(\mathbb{R}_{n[x]}, M_n, \dots)$, los elementos de la base también son elementos no vectoriales. La única condición que se debe cumplir es que los elementos del espacio se puedan describir como una combinación lineal (Indicada a través de las coordenadas de un vector) de los elementos de dicha base.

Ejemplos

$$3 + 4x + 5x^2 = (3,4,5)_{B_c\mathbb{R}_{2[x]}} \quad \longrightarrow \quad B_c\mathbb{R}_{2[x]} = \{1, x, x^2\}$$
$$\left(\begin{smallmatrix}1 & 2\\ 3 & 4\end{smallmatrix}\right) = (1,2,3,4)_{B_cM_2} \quad \longrightarrow \quad B_cM_2 = \{\left(\begin{smallmatrix}1 & 0\\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & 1\\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & 0\\ 1 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & 0\\ 0 & 1\end{smallmatrix}\right)\}$$

1.7.2. Base canónica de un espacio vectorial

Todo espacio vectorial, salvo el $E = \{\bar{0}\}$, admite al menos una base. La base canónica es la base con tantos vectores como coordenadas, y formada por vectores de una unidad en una coordenada y ceros en el resto.

Ejemplo

Expresar $\mathbb{R}^3$ en función de varios parámetros multiplicados por los vectores de su base canónica:
$B_c(\mathbb{R}^3) = \{(1,0,0),(0,1,0),(0,0,1)\}$

Por tanto:
$L\{B(\mathbb{R}^3)\} = \{\alpha(1,0,0) + \beta(0,1,0) + \gamma(0,0,1)\} = (\alpha, \beta, \gamma)$

Ejemplo

Expresar $\mathbb{R}_{2[x]}$ en función de varios parámetros multiplicados por los elementos de su base canónica:
$B_c(\mathbb{R}_{2[x]}) = \{1, x, x^2\} = \{(1,0,0)_{B\mathbb{R}_{2[x]}}, (0,1,0)_{B\mathbb{R}_{2[x]}}, (0,0,1)_{B\mathbb{R}_{2[x]}}\}$

Por tanto:
$L\{B(\mathbb{R}^3)\} = \{\alpha(0,0,1)_{B\mathbb{R}_{2[x]}} + \beta(0,0,1)_{B\mathbb{R}_{2[x]}} + \gamma(0,0,1)_{B\mathbb{R}_{2[x]}}\} = (\alpha,\beta,\gamma)_{B\mathbb{R}_{2[x]}} = \alpha + \beta x + \gamma x^2$

Ejemplo

Expresar M_2 en función de varios parámetros multiplicados por los elementos de su base canónica:
$B_c(M_2) = \{\left(\begin{smallmatrix}1 & 0\\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & 1\\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & 0\\ 1 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & 0\\ 0 & 1\end{smallmatrix}\right)\} = \{(1,0,0,0)_{BM_2}, (0,1,0,0)_{BM_2}, (0,0,1,0)_{BM_2}, (0,0,0,1)_{BM_2}\}$

Por tanto: $L\{B(M_2)\} = \{\alpha(1,0,0,0)_{BM_2} + \beta(0,1,0,0)_{BM_2} + \gamma(0,0,1,0)_{BM_2} + \delta(0,0,0,1)_{BM_2}\} = (\alpha,\beta,\gamma,\delta)_{BM_2} = \left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$

1.7.3. Pasar de sistema de vectores a base

Para asegurar que un sistema de vectores es una base además de un sistema de generadores dicho sistema debe ser libre, ya que la propia definición de base pide que ningún vector sea combinación lineal del resto.

Ejemplo

Hallar una base de $V = L\{(1,0,3),(2,3,1),(-3,-6,1)\}$

Triangulemos los vectores del sistema:

$$\begin{pmatrix}1 & 0 & 3\\ 2 & 3 & 1\\ -3 & -6 & 1\end{pmatrix} \sim \begin{pmatrix}1 & 0 & 3\\ 0 & 3 & -5\\ 0 & -6 & 10\end{pmatrix} \sim \begin{pmatrix}1 & 0 & 3\\ 0 & 3 & -5\\ 0 & 0 & 0\end{pmatrix} \sim \begin{pmatrix}1 & 0 & 3\\ 0 & 3 & -5\end{pmatrix}$$

$$B(V) = \{(1,0,3),(1,3,-5)\}$$

1.7.4. Dimensión de un espacio vectorial

Dado un espacio vectorial E, se define como dimensión de E, que denotaremos por $dim(E)$, al **número de vectores que conforman su base**.

Ejemplo

$$dim(\mathbb{R}) = 1 \qquad dim(\mathbb{R}^2) = 2 \qquad dim(\mathbb{R}^3) = 3 \qquad dim(\mathbb{R}^n) = n$$
$$dim(M_2) = 4 \qquad dim(M_3) = 9 \qquad dim(M_4) = 16 \qquad dim(M_n) = n^2$$
$$dim(\mathbb{R}_{1[x]}) = 2 \qquad dim(\mathbb{R}_{2[x]}) = 3 \qquad dim(\mathbb{R}_{3[x]}) = 4 \qquad dim(\mathbb{R}_{n[x]}) = n+1$$

De su propia definición se desprende que un **subespacio vectorial** siempre tiene dimensión menor que el espacio vectorial de dimensión máxima que lo contiene. Es decir, que su base siempre tiene menos vectores que coordenadas.

Ejemplo

Determinar la dimensión del subespacio vectorial de los polinominos de grado 2 sin término independiente $(0 + \alpha x + \beta x^2) \in \mathbb{R}_{2[x]}$.

$$V = L\{0, x, x^2\} = L\{(0,0,0)_{B_c\mathbb{R}_{2[x]}}, (0,1,0)_{B_c\mathbb{R}_{2[x]}}, (0,0,1)_{B_c\mathbb{R}_{2[x]}}\}$$
$$B(V) = \{(0,1,0)_{B_c\mathbb{R}_{2[x]}}, (0,0,1)_{B_c\mathbb{R}_{2[x]}}\} \quad \longrightarrow \quad dim(V) = 2$$

Ejemplo

Determinar la dimensión del subespacio vectorial generado por $L\{(1,0,2),(-1,3,2),(3,-3,2)\} \in \mathbb{R}^3$.
$V = L\{(1,0,2),(-1,3,2),(3,-3,2)\}$

Triangulemos los vectores del sistema:

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

$$B(V) = \{(1,0,2),(0,3,4)\} \quad \longrightarrow \quad dim(V) = 2$$

Nota. para cualquier valor de $n \in \mathbb{N}$, los conjuntos de las matrices simétricas S_n y las matrices antisimétricas A_n cumplen que:

$$dim(S_n) = \frac{n(n+1)}{2} \qquad dim(A_n) = \frac{n(n-1)}{2}$$

1.8. Ecuaciones paramétricas e implícitas de un subespacio vectorial

1.8.1. Ecuaciones paramétricas

Las ecuaciones paramétricas de un subespacio vectorial $V = L\{\bar{e}_1, \dots, \bar{e}_k\}$ describen, en función de unos ciertos parámetros variables $\{\alpha_1, \dots, \alpha_k\}$, las posibles coordenadas que puede poseer un vector $\bar{x}$ de E (tal que $V \in E$ y $dim(E) = \text{máx}$) si está contenido también en V.

$$\forall \bar{x} \in V \quad \text{tq.} \quad \bar{x} = (x_1, \dots, x_n) \quad B_V = \{\bar{e}_1, \dots, \bar{e}_k\} \quad \text{Con } k < n$$

$$\bar{x} = \alpha_1 \cdot \bar{e}_1 + \cdots + \alpha_k \cdot \bar{e}_k \qquad \begin{cases} x_1 = e_{1,1} \cdot \alpha_1 \cdot + \cdots + e_{k,1} \cdot \alpha_k \\ \vdots \\ x_n = e_{1,n} \cdot \alpha_1 + \cdots + e_{k,n} \cdot \alpha_k \end{cases}$$

Nota. Todo vector $\bar{x} \in V$ es una combinación lineal de los k vectores de la base de V (Y no de los n vectores de la base de E)

Nota. A diferencia de las ecuaciones implícitas, las ecuaciones paramétricas no solo son útiles para describir los vectores de un subespacio vectorial, sino que también nos pueden servir para llevar a cabo cambios de base sin necesidad de emplear matrices.

Ejemplo
Obtener las ecuaciones paramétricas del subespacio vectorial V generado por $L\{(1,0,-1),(0,-1,2),(1,1,-3)\}$.

1. Obtenemos una base de V.

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

$$B(V) = \{(1,0,1),(0,-1,2)\}$$

2. Expresamos las coordenadas del vector genérico del subespacio en función de parámetros.

$$\{B(V)\} = \alpha(1,0,1) + \beta(0,-1,2) = (\alpha, -\beta, \alpha + 2\beta)$$

3. Reescribimos las coordenades del vector genérico como un sistema de ecuaciones.

$$\text{Ecuaciones paramétricas:} \quad \begin{cases} x_1 &= \alpha + 0 \\ x_2 &= 0 - \beta \\ x_3 &= \alpha + 2\beta \end{cases}$$

1.8.2. Ecuaciones implícitas

Las ecuaciones implícitas de un subespacio vectorial V establecen las relaciones que se tienen que cumplir entre las coordenadas los vectores de E (tal que $V \in E$ & $dim(E) = \text{máx}$) para que dichos vectores estén contenidos en V. Por tanto, si queremos comprobar si un vector de E está en V, basta con ver si sus coordenadas satisfacen las ecuaciones implícitas de V.

Para que las ecuaciones de un sistema puedan ser tomadas como las ecuaciones implícitas de un subespacio vectorial, dicho sistema debe ser **homogéneo**, es decir, que los términos indepencientes de todas las ecuaciones sean iguales a 0. Esto se puede deducir de la propiedad que dice que un espacio vectorial debe contener el elementeo neutro de la suma, que es siempre el vector $\bar{0}$ del espacio. Si un sistema no es homogéneo, al sustituir las coordenadas por las del vector $\bar{0}$, dejarán de verificarse las igualdades del sistema.

$$\forall \bar{v} \in V \quad \text{tq.} \quad \bar{v} = (v_1, \dots, v_n) \qquad V = \begin{cases} \lambda_{1,1}v_1 + \cdots + \lambda_{1,n}v_n &= 0 \\ \vdots & \vdots \\ \lambda_{m,1}v_1 + \cdots + \lambda_{m,n}v_n &= 0 \end{cases} \quad \text{Con } m < n$$

Ejemplo
El subespacio vectorial de los polinomios de grado menor o igual que 3 sin término indemendiente y con el término de grado 1 el doble que el de grado 3.

$$V = \{(x_1, x_2, x_3, x_4)_{\mathbb{R}_{2[x]}} \;\mid\; x_1 = 0,\; x_2 = 2x_4\}$$

$$\text{Ecuaciones implícitas:} \quad \begin{cases} x_1 = 0 \\ x_2 - 2x_4 = 0 \end{cases}$$

Nota. No existen unas únicas ecuaciones ímplícitas para un determinado subespacio. En realidad, si se realizan combinaciones lineales entre ellas el sistema de ecuaciones resultante seguirá describiendo el mismo subespacio vectorial.

Si no estamos seguros, por la forma en la que las hemos obtenido, de que las ecuaciones implícitas de las que disponemos sean las mínimas necesarias, podemos simplificarlas (Triangular sus coeficientes).

Ejemplo

$$\text{Simplificar las ecuaciones implícitas:} \quad \begin{cases} x_1 + x_2 + x_3 = 0 \\ 2x_4 - 2x_2 = 0 \\ -x_3 - x_1 - x_4 = 0 \end{cases}$$

Triangular sus coeficientes.

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 0 & -2 & 0 & 2 & 0 \\ -1 & 0 & -1 & -1 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\text{Ecuaciones implícitas simplificadas:} \quad \begin{cases} x_1 + x_2 + x_3 = 0 \\ x_2 - x_4 = 0 \end{cases}$$

1.8.3. Transición de paramétricas a implícitas

Para obtener las ecuaciones implícitas de un subespacio V a partir de sus ecuaciones paraméticas debemos obtener todas las ecuaciones no paramétricas (Que no dependan de los parámetros) y homogéneas (Ya que, como en las ecuaciones paramétricas no hay términos independientes, si se eliminan los parámetros la ecuación resultante será homogénea) posibles que respeten las condiciones de las ecuaciones paramétricas. Es decir, tenemos que obtener todas las combinaciones posibles de las ecuaciones paramétricas que simplifiquen todos los parámetros. Para ello, lo más sencillo que podemos hacer es triangular los parámetros de las ecuaciones paramétricas.

Ejemplo
Obtener las ecuaciones implícitas del subespacio vectorial V de base $B(V) = \{(1,0,-1,0),(0,-1,1,-1)\}$.

1. Obtenemos sus ecuaciones paramétricas.

$$\text{Ecuaciones paramétricas:} \quad \begin{cases} x_1 &= \alpha \\ x_2 &= -\beta \\ x_3 &= -\alpha + \beta \\ x_4 &= -\beta \end{cases} \longrightarrow \left(\begin{array}{c|cc} x_1 & 1 & 0 \\ x_2 & 0 & -1 \\ x_3 & -1 & 1 \\ x_4 & 0 & -1 \end{array}\right)$$

2. Triangulamos los parámetros.

$$\left(\begin{array}{c|cc} x_1 & 1 & 0 \\ x_2 & 0 & -1 \\ x_3 & -1 & 1 \\ x_4 & 0 & -1 \end{array}\right) \sim \left(\begin{array}{r|cc} x_1 & 1 & 0 \\ x_2 & 0 & -1 \\ x_1 + x_3 & 0 & 1 \\ x_4 & 0 & -1 \end{array}\right) \sim \left(\begin{array}{r|cc} x_1 & 1 & 0 \\ x_2 & 0 & -1 \\ x_1 + x_2 + x_3 & 0 & 0 \\ (-x_2) + x_4 & 0 & 0 \end{array}\right)$$

3. Extraemos las ecuaciones implícitas.

$$\text{Ecuaciones implícitas:} \quad \begin{cases} x_1 + x_2 + x_3 = 0 \\ x_2 - x_4 = 0 \end{cases}$$

Ejemplo
Obtener las ecuaciones implícitas del subespacio vectorial V generado por $L\{(2,3,0),(0,1,1),(2,2,-1)\}$.

1. Obtenemos las ecuaciones paramétricas de su generador (No hace falta obtener las de su base ya que al triangular sus parámetros ya lo estamos simplificando).

$$\text{Ecuaciones paramétricas:} \quad \begin{cases} x_1 &= 2\alpha + 2\gamma \\ x_2 &= 3\alpha + \beta + 2\gamma \\ x_3 &= + \beta - \gamma \end{cases} \longrightarrow \left(\begin{array}{c|ccc} x_1 & 2 & 0 & 2 \\ x_2 & 3 & 1 & 2 \\ x_3 & 0 & 1 & -1 \end{array}\right)$$

2. Triangulamos los parámetros.

$$\left(\begin{array}{c|ccc} x_1 & 2 & 0 & 2 \\ x_2 & 3 & 1 & 2 \\ x_3 & 0 & 1 & -1 \end{array}\right) \sim \left(\begin{array}{r|ccc} 3x_1 & 6 & 0 & 6 \\ 2x_2 & 6 & 2 & 4 \\ x_3 & 0 & 1 & -1 \end{array}\right) \sim \left(\begin{array}{r|ccc} 3x_1 & 6 & 0 & 6 \\ (-3x_1) + 2x_2 & 0 & 2 & -2 \\ 2x_3 & 0 & 2 & -2 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{r|ccc} 3x_1 & 6 & 0 & 6 \\ (-3x_1) + 2x_2 & 0 & 2 & -2 \\ 3x_1 - 2x_2 + 2x_3 & 0 & 0 & 0 \end{array}\right)$$

3. Extraemos las ecuaciones implícitas.

$$\text{Ecuaciones implícitas:} \quad \{\; 3x_1 - 2x_2 + 2x_3 = 0$$

1.8.4. Transición de implícitas a paraméticas

Para obtener las ecuaciones paramétricas de un subespacio V a partir de sus ecuaciones implícitas debemos extraer de las implícitas las relaciones entre parámetros que nos permitan simplificar las ecuaciones paramétricas del espacio de dimensión máxima.

Nota. Esta forma de calcular las implícitas de un subespacio vectorial V es equivalente a calcular la intersección de V con E (Espacio vectorial de dimensión máxima) expresado en la base canónica.

Ejemplo
Obtener las ecuaciones paramétricas del subespacio vectorial $V = \{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 = 0\}$.

1. Obtenemos la relación entre parámetros que nos da la ecuación implícita.

$$\{\; x_1 + x_2 + x_3 = 0 \quad \longrightarrow \quad \alpha + \beta + \gamma = 0 \quad \longrightarrow \quad \alpha = -\beta - \gamma$$

2. Sustituimos la relación entre parámetros que hemos obtenido en las ecuaciones paramétricas de dimensión máxima.

$$\begin{cases} x_1 &= \alpha \\ x_2 &= \beta \\ x_3 &= \gamma \end{cases} \xrightarrow{\alpha = -\beta - \gamma} \begin{cases} x_1 &= -\beta - \gamma \\ x_2 &= \beta \\ x_3 &= \gamma \end{cases}$$

Ejemplo
Obtener una base del subespacio vectorial $V = \{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 = 0,\; x_1 - x_2 = 0,\; x_1 + 2x_4 = 0\}$.

1. Obtenemos la relación entre parámetros que nos da la primera ecuación implícita.

$$\{\; x_1 + x_2 + x_3 = 0 \quad \rightarrow \quad \alpha + \beta + \gamma = 0 \quad \rightarrow \quad \alpha = -\beta - \gamma$$

2. Sustituimos la relación entre parámetros que hemos obtenido en las ecuaciones paramétricas de dimensión máxima.

$$\begin{cases} x_1 &= \alpha \\ x_2 &= \beta \\ x_3 &= \gamma \\ x_4 &= \delta \end{cases} \xrightarrow{\alpha = -\beta - \gamma} \begin{cases} x_1 &= -\beta - \gamma \\ x_2 &= +\beta \\ x_3 &= +\gamma \\ x_4 &= +\delta \end{cases}$$

3. Repetimos los pasos 1 y 2 utilizando la relación entre parámetros que nos da la segunda ecuación implícita.

$$\{\; x_1 - x_2 = 0 \quad \rightarrow \quad \alpha - \beta = 0 \quad \rightarrow \quad \alpha = \beta \quad \rightarrow \quad -\beta - \gamma = \beta \quad \rightarrow \quad \gamma = -2\beta$$

$$\begin{cases} x_1 &= -\beta - \gamma \\ x_2 &= +\beta \\ x_3 &= +\gamma \\ x_4 &= +\delta \end{cases} \xrightarrow{\gamma = -2\beta} \begin{cases} x_1 &= +\beta \\ x_2 &= +\beta \\ x_3 &= -2\beta \\ x_4 &= +\delta \end{cases}$$

4. Repetimos los pasos 1 y 2 utilizando la relación entre parámetros que nos da la tercera ecuación implícita.

$$\{\; x_1 + 2x_4 = 0 \;\rightarrow\; \alpha + 2\delta = 0 \;\rightarrow\; \alpha = -2\delta \;\rightarrow\; -(\beta + \gamma) = -2\delta \;\rightarrow\; \delta = \tfrac{1}{2}(\beta + \gamma) \;\rightarrow\; \delta = -\tfrac{1}{2}(\beta)$$

$$\begin{cases} x_1 &= -\beta - \gamma \\ x_2 &= +\beta \\ x_3 &= +\gamma \\ x_4 &= +\delta \end{cases} \xrightarrow{\delta = -\frac{1}{2}(\beta)} \begin{cases} x_1 &= +\beta \\ x_2 &= +\beta \\ x_3 &= -2\beta \\ x_4 &= -\frac{1}{2}(\beta) \end{cases} \sim \begin{cases} x_1 &= +2\beta \\ x_2 &= +2\beta \\ x_3 &= -4\beta \\ x_4 &= -\beta \end{cases}$$

5. Tomamos un vector que cumpla las ecuaciones paramétricas resultantes y formamos una base.

$$B(V) = \{(2,2,-4,-1)\}$$

1.8.5. Ecuaciones paramétricas e implícitas de un espacio vectorial de dimensión máxima.

Cada vector del generador de V añade una dimensión al subespacio, por tanto cada parámetro de las ecuaciones paramétricas hace lo propio. Por el contrario cada ecuación de las implícitas de V le impone una condición que reduce la dimensión del subespacio en una unidad respecto a la dimensión de E. De esta forma, si el generador de V tiene tantos vectores como coordenadas o 0 ecuaciones implícitas, entonces $V = E$.

1.8.6. Teorema de las dimensiones

Sean V y E tales que $V \in E$ y $dim(E) = máx$, siempre se cumple la siguiente relación entre sus coordenadas:

$$dim(E) = dim(V) + \text{nº ec(s). imp. V}$$

$$\text{nº coord. } E = \text{nº parám. } V + \text{nº ec(s). imp. V}$$

1.9. Suma de subespacios vectoriales

La suma de 2 subespacios vectoriales, que denotamos por $V \cup W$ o $V + W$, da como resultado el espacio vectorial generado por el sistema de todos los vectores de cada una de las bases de los subespacios. Si queremos expresar la suma como un nuevo espacio vectorial debemos asegurarnos de que el sistema de vectores obtenido tiene todos sus vectores linealmente independientes entre sí (Formando así una base).

Ejemplo

Hallar la suma de los subespacios vectoriales V y W.
$V = L\{(1,0,0,0),(1,1,0,2),(0,0,0,1)\}$
$W = L\{(1,1,0,0),(0,1,3,0)\}$

1. Triangulemos el sitema de vectores formado por los vectores de ambas bases.

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

2. Obtenemos la base de la suma de los subespacios V y W.

$B(V+W) = \{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\} = B_c\mathbb{R}^4$
$(V+W) = L\{B(V+W)\} = \mathbb{R}^4$

1.10. Intersección de subespacios vectoriales

La intersección de 2 espacios vectoriales, que denotamos por $V \cap W$, es el espacio vectorial que contiene a los vectores contenidos en cada uno de los espacios originales. La intersección se obtiene como el espacio de ecuaciones implícitas resultantes de la unión y la simplificación de las ecuaciones implícitas de ambos espacios vectoriales (V y W).

Ejemplo

Hallar la intersección de los subespacios vectoriales V y W.
$V = \{(x_1, x_2, x_3) \mid x_1 - x_2 + x_3 = 0\}$
$W = \{(x_1, x_2, x_3) \mid x_1 - x_2 = 0,\ x_3 = 0\}$

1. Juntar todas las ecuaciones implícitas

$$\text{Ecuaciones implícitas de } V \text{ y } W\text{:} \quad \left\{\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ x_1 - x_2 &= 0 \\ x_3 &= 0 \end{aligned}\right. \quad \to \quad \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

2. Simplificamos las ecuaciones implícitas triangulando la matriz de sus coeficientes.

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

3. Obtenemos las nuevas implícitas, que nos dan la intersección de V y W.

$$\text{Ecuaciones implícitas de } V \cap W\text{:} \quad \left\{\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ x_3 &= 0 \end{aligned}\right.$$

Cuando no disponemos de los 2 espacios vectoriales expresados en sus ecuaciones implícitas, otro procedimiento que podemos utilizar para obtener su intersección, consiste en introducir las relaciones entre parámetros provenientes de las implícitas de uno de ellos en las ecuaciones paramétricas del otro.

Ejemplo

Hallar una base de la intersección de los subespacios vectoriales V y W.
$V = \{(1, 0, 0, 0), (1, 1, 0, 2)\}$
$W = \{(1, 0, 0, 2), (1, 1, 0, 0)\}$

1. Obtenemos las ecuaciones implícitas de uno de los subespacios y sacamos las relaciones entre sus parámetros. En este caso, tomamos las implícitas de V.

   Nota. Si podemos, conviene tomar la base del espacio de menor dimensión, para partir del menor número posible de ecuaciones implícitas.

$$\text{Param. } V\text{:} \left\{\begin{array}{lcl} x_1 & = & +\alpha + \beta \\ x_2 & = & \quad\ +\beta \\ x_3 & = & 0 \\ x_4 & = & \quad\ +2\beta \end{array}\right. \longrightarrow \quad \text{Imp. } V\text{:} \left\{\begin{aligned} 2x_2 - x_4 &= 0 \\ x_3 &= 0 \end{aligned}\right.$$

$$\left\{\begin{aligned} 2x_2 - x_4 &= 0 \\ x_3 &= 0 \end{aligned}\right. \to \left\{\begin{aligned} 2(\beta) - (2\alpha) &= 0 \\ 0 &= 0 \end{aligned}\right. \to \left\{\ \beta = \alpha\right.$$

2. Sustituimos las equivalencias de las ecuaciones paramétricas de W en las ecuaciones implícitas de V.

$$\text{Param. } W\text{:} \left\{\begin{array}{lcl} x_1 & = & +\alpha + \beta \\ x_2 & = & \quad\ +\beta \\ x_3 & = & 0 \\ x_4 & = & +2\alpha \end{array}\right.$$

3. Utilizamos las relaciones entre parámetros obtenidas para simplificar las ecuaciones paramétricas de W.

$$\left\{\begin{array}{lcl} x_1 & = & +\alpha + \beta \\ x_2 & = & \quad\ +\beta \\ x_3 & = & 0 \\ x_4 & = & +2\alpha \end{array}\right. \xrightarrow{\ \beta = \alpha\ } \left\{\begin{array}{lcl} x_1 & = & +2\alpha \\ x_2 & = & +\alpha \\ x_3 & = & 0 \\ x_4 & = & +2\alpha \end{array}\right.$$

4. Obtenemos una base de la intersección de V y W.

$$\text{B}\ (V \cap W) = \{(2, 1, 0, 2)\}:$$

1.11. Espacios vectoriales complementarios o suplementarios

Dos subespacios vectoriales V y W se dicen **complementarios o suplementarios** de un espacio vectorial E cuando $V + W = E$.

1.12. Cambios de base

Cuando hablamos de cambios de base nos referimos al proceso que consiste en pasar de expresar un vector o subespacio vectorial en una base B_1 a expresarlo en otra base B_2. (Ambas bases de espacios vectoriales de dimensión máxima).

Para hacer cambios de base de vectores (El cambio de base de un subespacio vectorial es el cambio de base de los vectores de su base), generalmente usamos las matrices de cambio de base, aunque también podemos hacerlo mediante un sistema de ecuaciones.

1.12.1. Mediante un sistema de ecuaciones paramétricas

Para llevar a cabo un cambio de base de un vector expresado en una base B_1 sin recurrir a una matriz de cambio de base debemos primero obtener el vector en la base canónica y posteriormente recurrir a un sistema de ecuaciones paramétricas para expresarlo en la base B_2.

$$B_1 \longrightarrow Bc \longrightarrow B_2$$

DESARROLLO

Sea $\bar{x}_{B_1} = (x'_1, \ldots, x'_n)_{B_1}$ un vector expresado en la base $B_1 = \{\bar{e}'_1, \ldots, \bar{e}'_n\}$, para obtener este mismo vector expresado en la base $B_2 = \{\bar{e}''_1, \ldots, \bar{e}''_n\}$ debemos seguir el siguiente desarrollo:

1. Obtenemos el vector $\bar{x}_{B_1}$, expresado en la base B_1, en la base canónica $B_c = \{\bar{e}_1, \ldots, \bar{e}_n\}$ multiplicando cada coordenada por su correspondiente vector de la base B_1:

$$\bar{x}_{B_1} = (x'_1, \ldots, x'_n)_{B_1}$$
$$\bar{x}_{B_c} = (x'_1 \cdot \bar{e}'_1 + \cdots + x'_n \cdot \bar{e}'_n) = (x_1, \ldots, x_n)_{B_c}$$

2. Utilizamos un sistema de ecuaciones paramétricas para obtener el vector $\bar{x}_{B_c}$, expresado en la base canónica, en la base $B_2 = \{\bar{e}''_1, \ldots, \bar{e}''_n\}$:

$$\bar{x}_{B_c} = (x_1, \ldots, x_n)_{B_c}$$
$$\begin{cases} x_1 = e''_{1,1} \cdot x''_1 + \cdots + e''_{n,1} \cdot x''_n \\ \vdots \qquad\qquad \vdots \\ x_n = e''_{1,n} \cdot x''_1 + \cdots + e''_{n,n} \cdot x''_n \end{cases}$$

Al tratarse de un sistema compatible y determinado, ya que tiene el mismo número de ecuaciones que de incógnitas, tiene exactamente una solución (una combinación de $\{x''_1, \ldots, x''_n\}$ distinta de $\bar{0}$ que satisface el sistema).

$$\bar{x}_{B_2} = (x''_1, \ldots, x''_n)_{B_2}$$

Ejemplo

Hallar el vector $\bar{v} = (1, 2)$, expresado en la base B_1, en la base B_2.

$B_1 = \{(-1, 1), (2, 1)\}$
$B_2 = \{(1, 3), (0, 1)\}$

1. Obtenemos el vector $\bar{v}_{B_1}$ en base canónica.

$$\bar{v}_{B_c} = (1 \cdot (-1, 1) + 2 \cdot (2, 1)) = (3, 3)_{B_c}$$

2. Utilizamos un sistema de ecuaciones paramétricas para obtener el vector $\bar{v}_{B_c}$ en la base B_2.

$$\begin{cases} 3 = 1 \cdot x''_1 + 0 \cdot x''_2 \\ 3 = 3 \cdot x''_1 + 1 \cdot x''_2 \end{cases} \longrightarrow \begin{matrix} x''_1 = 3 \\ x''_2 = (-6) \end{matrix} \longrightarrow \bar{x}_{B_2} = (3, -6)_{B_2}$$

1.12.2. Mediante una matriz de cambio de base

La matriz de cambio de la base B_1 a la base B_2 se denota de la siguiente forma ($M_{B_2 \leftarrow B_1}$), y cumple la siguiente igualdad:

$$\begin{pmatrix} x''_1 \\ x''_2 \\ x''_3 \\ \vdots \end{pmatrix}_{B_2} = M_{B_2 \leftarrow B_1} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ \vdots \end{pmatrix}_{B_1}$$

La matriz de cambio de base B_1 a la base B_2 se construye colocando de forma vertical los vectores de la base B_1 expresados en la base B_2.

Ejemplo

Hallar la matriz de cambio de base de la B_1 a la base B_2.

$B_1 = \{(1, 0, 1), (0, 3, 1), (1, 2, 1)\}$
$B_2 = \{(1, -1, 1), (0, 1, 1), (0, 0, -1)\}$

La matriz $M_{B_2 \leftarrow B_1}$ es una matriz con los vectores de B_1 en base B_2 colocados en columnas. Para obtener cada uno de los vectores de B_1 en base $B2$ tenemos que expresar dicho vector como combinación lineal de los vectores de B_2.

$$[(1, 0, 1), (0, 3, 1), (1, 2, 1)] = [L\{B_2\}_1, L\{B_2\}_2, L\{B_2\}_3]$$
$$L\{B_2\} = \alpha(1, -1, 1) + \beta(0, 1, 1) + \gamma(0, 0, -1)$$

1. Obtenemos el vector $(1, 0, 1)$ en base B_2 .

$$\begin{cases} \alpha = 1 \\ -\alpha + \beta = 0 \\ \alpha + \beta + \gamma = 1 \end{cases} \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & 1 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right) \rightarrow (1, 1, 1)_{B_2}$$

2. Obtenemos el vector $(0, 3, 1)$ en base B_2 .

$$\begin{cases} \alpha = 0 \\ -\alpha + \beta = 3 \\ \alpha + \beta + \gamma = 1 \end{cases} \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 3 \\ 1 & 1 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 1 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right) \rightarrow (0, 3, 2)_{B_2}$$

3. Obtenemos el vector $(1, 2, 1)$ en base B_2 .

$$\begin{cases} \alpha = 1 \\ -\alpha + \beta = 2 \\ \alpha + \beta + \gamma = 1 \end{cases} \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 2 \\ 1 & 1 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 1 & -1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 3 \end{array}\right) \rightarrow (1, 3, 3)_{B_2}$$

4. Formamos la matriz de cambio de base con los vectores que hemos obtenido.

$$M_{B_2 \leftarrow B_1} = \left(\begin{array}{c|c|c} 1 & 0 & 1 \\ 1 & 3 & 3 \\ 1 & 2 & 3 \end{array}\right)$$

1.12.3. Caso específico de cambio de base B a B_c

En el caso concreto de cambio de una base cualquiera a la base canónica, al hacer el sistema para expresar $\bar{e}_i = (e_{i1}, e_{i2}, e_{i3}) \in B$ en la base canónica siempre vamos a obtener igualdades de la forma $(e_{i1}, e_{i2}, e_{i3}) = (\alpha, \beta, \gamma)$. Por tanto, tomaremos la matriz $M_{B_c \leftarrow B}$ como la matriz resultante de colocar los vectores de la base B directamente en vertical.

Ejemplo

Calcular la matriz de cambio de base de una base B cualquera de $\mathbb{R}^3$ a la base canónica $B_c\mathbb{R}^3$.

$$B = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\} = \{\bar{e}_i\} \qquad B_c = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

$$\begin{cases} \alpha = e_{i1} \\ \beta = e_{i2} \\ \gamma = e_{i3} \end{cases} \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & e_{i1} \\ 0 & 1 & 0 & e_{i2} \\ 0 & 0 & 1 & e_{i3} \end{array}\right) \rightarrow (e_{i1}, e_{i2}, e_{i3})_{B_c}$$

$$M_{Bc \leftarrow B} = \left(\begin{array}{c|c|c} \bar{e}_1 & \bar{e}_2 & \bar{e}_3 \end{array}\right)$$

1.12.4. Relación entre matrices de cambios de base opuestos

Dadas dos bases B_1 y B_2. Las matrices $M_{B_2 \leftarrow B_1}$ y $M_{B_1 \leftarrow B_2}$ guardan entre sí una relación inversa.

Ejemplo

Conociendo la matriz de cambio de base B_1 a base B_2, $M_{B_2 \leftarrow B_1} = \left(\begin{smallmatrix} 1 & 2 & 0 \\ 3 & 6 & 1 \\ 0 & 1 & 0 \end{smallmatrix}\right)$, calcular la matriz $M_{B_1 \leftarrow B_2}$ de cambio de base B_2 a B_1.

$$M_{B_1 \leftarrow B_2} = M^{-1}_{B_2 \leftarrow B_1} = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 6 & 1 \\ 0 & 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 1 \\ -3 & 1 & 0 \end{pmatrix}$$

1.12.5. Concatenación de matrices de cambio de base $(B_1 \longrightarrow B_2 \longrightarrow B_3)$

Sean B_1, B_2 y B_3 tres bases distintas, sus matrices de cambio de base cumplen la siguiente propiedad:

$$M_{B_3 \leftarrow B_2 \leftarrow B_1} = M_{B_3 \leftarrow B_2} \cdot M_{B_2 \leftarrow B_1} = M_{B_3 \leftarrow B_1}$$

Esta propiedad nos permite aprovecharnos de la simplicidad de las matrices de cambio a la base canónica para construir cambios de base entre dos bases no canónicas, evitando así tener que resolver el sistema de ecuaciones.

$$M_{B_2 \leftarrow B_1} = M_{B_2 \leftarrow B_c} \cdot M_{B_c \leftarrow B_1} = M^{-1}_{B_c \leftarrow B_2} \cdot M_{B_c \leftarrow B_1}$$

Ejemplo

Calcular la matriz $M_{B_2 \leftarrow B_1}$ de cambio de base entre las bases $B_1 = \{(0, 1, 3), (5, 4, 0), (2, 3, 4)\}$ y $B_2 = \{(1, 1, 1), (3, 4, 3), (3, 3, 4)\}$.

1. Obtenemos las matrices de cambio de base entre ambas bases y la base canónica:

$$M_{B_c \leftarrow B_1} = \begin{pmatrix} 0 & 5 & 2 \\ 1 & 4 & 3 \\ 3 & 0 & 4 \end{pmatrix} \qquad M_{B_2 \leftarrow B_c} = M^{-1}_{B_c \leftarrow B_2} = \begin{pmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{pmatrix}^{-1} = \begin{pmatrix} 7 & -3 & 3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

2. Multiplicamos las dos matrices anteriores para obtener la matriz de cambio de base B_1 a B_2:

$$M_{B_2 \leftarrow B_1} = M_{B_2 \leftarrow B_c} \cdot M_{B_c \leftarrow B_1} = \begin{pmatrix} 7 & -3 & 3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 5 & 2 \\ 1 & 4 & 3 \\ 3 & 0 & 4 \end{pmatrix} = \begin{pmatrix} -12 & 23 & -7 \\ 1 & -1 & 1 \\ 3 & -5 & 2 \end{pmatrix}$$

Capítulo 2

Aplicaciones Lineales

2.1. Aplicación Lineal

Una aplicación es una transformación que toma elementos de uno o varios conjuntos (C_i) dando lugar a elementos de otro u otros conjuntos (G_i).

$$a : C_1 \times \cdots \times C_n \longmapsto G_1 \times \cdots \times G_n$$

Una aplicación lineal u homomorfismo es una aplicación que toma un vector de un espacio vectorial y lo transforma en un vector de otro espacio vectorial cumpliendo una serie de características de linealidad.

$$f : E \longmapsto F$$

2.1.1. Teorema de caracterización de aplicaciones lineales

Sean E y F dos espacios vectoriales. Una función o aplicación $f : E \mapsto F$ es una aplicación lineal u homomorfismo si y sólo si cumple las siguientes características:

$$\begin{array}{ll} \forall \bar{x}, \bar{y} \in E & f(\bar{x} + \bar{y}) = f(\bar{x}) + f(\bar{y}) \\ \forall \bar{x} \in E \quad \forall \lambda \in \mathbb{R} & f(\lambda \bar{x}) = \lambda f(\bar{x}) \end{array}$$

Lo que es equivalente a:

$$\forall \bar{x}, \bar{y} \in E \quad \forall \lambda, \mu \in \mathbb{R} \qquad f(\lambda \bar{x} + \mu \bar{y}) = \lambda f(\bar{x}) + \mu f(\bar{y})$$

Nota. Si una aplicación es lineal, entonces $f(\bar{0}) = \bar{0}$, ya que: $f(\bar{0}) = f(0 \cdot \bar{0}) = 0 f(\bar{0}) = \bar{0}$.
Nota. Si una aplicación es lineal, entonces $f(-\bar{x}) = -f(\bar{x})$, ya que: $f(-\bar{x}) = f((-1) \cdot \bar{x}) = (-1) f(\bar{x}) = -f(\bar{x})$.

Ejemplo
Demostración de que la aplicación $f(x, y, z) = (2x, y)$ es lineal.

MÉTODO 1 - Demostrando las propiedades por separado:

1. Demostrar que $f(\bar{x} + \bar{y}) = f(\bar{x} + \bar{y}) \qquad \forall \bar{x}, \bar{y} \in E$:

$$f(\bar{x} + \bar{y}) = f((x_1, x_2, x_3) + (y_1, y_2, y_3)) = f((x_1 + y_1, x_2 + y_2, x_3 + y_3) = (2(x_1 + y_1), x_2 + y_2)$$
$$f(\bar{x}) + f(\bar{y}) = f(x_1, x_2, x_3) + f(y_1, y_2, y_3) = (2x_1, x_2) + (2y_1, y_2) = (2(x_1 + y_1), x_2 + y_2)$$

2. Demostrar que $f(\lambda \bar{x}) = \lambda f(\bar{x}) \qquad \forall \bar{x} \in E \quad \forall \lambda \in \mathbb{R}$:

$$f(\lambda \bar{x}) = f(\lambda (x_1, x_2, x_3)) = f(\lambda x_1, \lambda x_2, \lambda x_3) = (2\lambda x_1, \lambda x_2)$$
$$\lambda f(\bar{x}) = \lambda f(x_1, x_2, x_3) = \lambda (2x_1, x_2) = (2\lambda x_1, \lambda x_2)$$

MÉTODO 2 - Demostrando las 2 propiedades al mismo tiempo:

$$\forall \bar{x}, \bar{y} \in E \qquad \forall \lambda, \mu \in \mathbb{R}$$
$$f(\lambda \bar{x} + \mu \bar{y}) = f(\lambda (x_1, x_2, x_3) + \mu (y_1, y_2, y_3)) = f(\lambda x_1 + \mu y_1, \lambda x_2 + \mu y_2, \lambda x_3 + \mu y_3) =$$
$$= (2(\lambda x_1 + \mu y_1), \lambda x_2 + \mu y_2) = (2\lambda x_1, \lambda x_2) + (2\mu y_1, \mu y_2) = \lambda (2x_1, x_2) + \mu (2y_1, y_2) =$$
$$= \lambda f(x_1, x_2, x_3) + \mu f(y_1, y_2, y_3) = \lambda f(\bar{x}) + \mu f(\bar{y})$$

Ejemplo
Demostración de que la aplicación $f(x, y, z) = (x - 3y, 2z + 1)$ no es lineal.

Basta con probar que $f(\bar{0}) \neq \bar{0}$:

$$f(\bar{0}) = f(0, 0, 0) = (0 - 3 \cdot 0, 2 \cdot 0 + 1) = (0, 1) \neq \bar{0}$$

Ejemplo
Demostración de que la aplicación $f(M) = tr(M) \quad \forall M \in M_n$ es lineal.

Sean: $A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} \qquad B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \cdots & b_{n,n} \end{pmatrix} \qquad \lambda A + \mu B = \begin{pmatrix} \lambda a_{1,1} + \mu b_{1,1} & \cdots & \lambda a_{1,n} + \mu b_{1,n} \\ \vdots & \ddots & \vdots \\ \lambda a_{n,1} + \mu b_{n,1} & \cdots & \lambda a_{n,n} + \mu b_{n,n} \end{pmatrix}$

Tenemos que: $\forall A, B \in M_n \qquad \forall \lambda, \mu \in \mathbb{R}$

$$f(\lambda A + \mu B) = tr(\lambda A + \mu B) = \sum_{i=1}^{n} (\lambda a_{i,i} + \mu b_{i,i}) = \sum_{i=1}^{n} \lambda \cdot a_{i,i} + \sum_{i=1}^{n} \mu \cdot b_{i,i} = \lambda \sum_{i=1}^{n} a_{i,i} + \mu \sum_{i=1}^{n} b_{i,i} =$$
$$= \lambda \cdot tr(A) + \mu \cdot tr(B) = \lambda f(A) + \mu f(B)$$

2.1.2. Clasificación de aplicaciones lineales

Sean E y F dos conjuntos y $f : E \mapsto F$ una aplicación (no necesariamente lineal):

- f se define **inyectiva** si dados dos elementos cualquiera distintos de su dominio su imagen también es distinta:
$$\forall \bar{x}_1, \bar{x}_2 \in E \text{ tq. } \bar{x}_1 \neq \bar{x}_2 \implies f(\bar{x}_1) \neq f(\bar{x}_2)$$
- f se define **sobreyectiva** si todo elemento de la imagen tiene una preimagen en el dominio:
$$\forall \bar{y} \in F, \ \exists \bar{x} \in E \quad \text{tq.} \quad f(\bar{x}) = \bar{y}$$
- f se define **biyectiva** si es al mismo tiempo inyectiva y sobreyectiva.

Sea $f : E \mapsto F$ una aplicación lineal u homomorfismo, se hace la siguiente clasificación según su tipo:

- **Endomorfismo**: si los espacio vectoriales del dominio y de la imagen del homomorfismo f son iguales ($E = F$).
- **Monomorfismo**: si el homomorfismo de f es **inyectivo**.
- **Epimorfismo**: si el homomorfismo de f es **sobreyectivo**.
- **Isomorfismo**: si el homomorfismo de f es **biyectivo** (inyectivo y sobreyectivo).
- **Automorfismo**: si el homomorfismo de f es al mismo tiempo un **endomorfimo** y **un isomorfismo**.

2.2. Imagen de una aplicación lineal

Sean E y F dos espacios vectoriales de igual dimensión, y $f : E \longmapsto F$ una aplicación lineal que toma vectores de E dando lugar a vectores de F.

Sea V un subespacio vectorial de E, el conjunto $f(V)$ (conjunto de las transformaciones por f de todos los vectores contenidos en V) es un subespacio vectorial de F. A este subespacio se lo llama "imagen de V respecto de f", y se lo denota por $f(V)$.

$$\forall\, V \in E \qquad f(V) = W \in F$$

Dada una base $B(V)$ del subespacio vectorial V, una posible base del subespacio vectorial $f(V)$ es el sistema formado por las evaluaciones de los vectores de $B(V)$ en f. La dimensión de base del subespacio imagen no tiene porque ser la misma que la de V si la aplicación f no es inyectiva.

Ejemplo
Dados el subespacio vectorial $V \in \mathbb{R}^3$ y la aplicación $f : \mathbb{R}^3 \mapsto \mathbb{R}^3$:

$V = L\{(3, 7, 1), (1, 1, 0)\}$
$f(x, y, z) = (z, z, x + y - z)$

Una base de la imagen de V respecto de f es:

1. Calculamos la imagen de los vectores de la base de V respecto de la aplicación lineal f:

$$f(3, 7, 1) = (1, 1, 9)$$
$$f(1, 1, 0) = (0, 0, 2)$$

2. La imagen de V respecto de f es el generador formado por la imagen de los vectores de su base:

$$f(V) = L\{(1, 1, 9), (0, 0, 2)\}$$

En particular, si $f(V) = f(E)$, entonces $[f(V) = F(E) = F]$ recibe el nombre de subespacio imagen de f, y se denota por Imf.

Nota. f es sobreyectiva si y solo si E y F tienen la misma dimensión.

2.2.1. Preimagen de una aplicación lineal

Sean E y F dos espacios vectoriales de igual dimensión, y $f : E \longmapsto F$ una aplicación lineal que toma vectores de E dando lugar a vectores de F.

Sea W un subespacio vectorial de F, se denomina preimagen de W al subespacio de E:

$$f^{-1}(W) = \{\bar{x} \in E \mid f(\bar{x}) \in W\}$$

Y se lo denota por $f^{-1}(W)$.

Ejemplo
Dados el subespacio vectorial $V \in \mathbb{R}^3$ y la aplicación $f : \mathbb{R}^3 \mapsto \mathbb{R}^3$:

$V = \{(x_1, x_2, x_3) | x_1 + x_2 - x_3 = 0\}$
$f(x, y, z) = (x, x, y + z)$

Una base de la preimagen de V respecto de f es:

1. Obtenemos las ecuaciones implícitas de $f^{-1}(V)$:

$$f^{-1}(V) = \{(x_1', x_2', x_3') \mid f(x_1', x_2', x_3') \in V\}$$
$$f^{-1}(V) = \{(x_1', x_2', x_3') \mid (x_1', x_1', x_2' + x_3') \in V\}$$
$$f^{-1}(V) = \{(x_1', x_2', x_3') \mid (x_1') + (x_1') - (x_2' + x_3') = 0\}$$
$$f^{-1}(V) = \{(x_1', x_2', x_3') \mid 2x_1' - x_2' - x_3' = 0\}$$

2. Obtenemos una base de $f^{-1}(V)$ a través de sus ecuaciones paramétricas:

$$2x_1' + x_2' - x_3' = 0 \quad \rightarrow \quad x_3' = 2x_1' - x_2' \quad \rightarrow \quad \gamma = 2\alpha - \beta$$

$$\begin{cases} x_1 = \alpha \\ x_2 = \beta \\ x_3 = \gamma \end{cases} \xrightarrow{\gamma = 2\alpha - \beta} \begin{cases} x_1 = \alpha \\ x_2 = \beta \\ x_3 = 2\alpha - \beta \end{cases} \qquad f^{-1}(V) = \{(1, 0, 2), (0, 1, -1)\}$$

2.3. Núcleo o kernel de una aplicación lineal

Sean E y F dos espacios vectoriales, se llama núcleo o kernel de una aplicación $f : E \longmapsto F$, que se denota por $kerf$, al subespacio vectorial V de E tal que:

$$f(V) = \bar{0}_F$$

O lo que es lo mismo, el subespacio de E: $f^{-1}(\bar{0}_F)$

Ejemplo
Dada la aplicación $f : \mathbb{R}^3 \mapsto \mathbb{R}^3$:

$f(x, y, z) = (x + y + z, x + y, z)$

Una base de su kernel o núcleo $kerf$ es:

1. Obtenemos las ecuaciones implícitas de $kerf$:

$$f^{-1}(V) = \{(x'_1, x'_2, x'_3) \ | \ f(x'_1, x'_2, x'_3) = \bar{0}\}$$
$$f^{-1}(V) = \{(x'_1, x'_2, x'_3) \ | \ (x'_1 + x'_2 + x'_3, x'_1 + x'_2, x'_3) = \bar{0}\}$$
$$f^{-1}(V) = \{(x'_1, x'_2, x'_3) \ | \ x'_1 + x'_2 + x'_3 = 0; \ \ x'_1 + x'_2 = 0; \ \ x'_3 = 0\}$$
$$\text{Ecuaciones implícitas de } kerf\text{:} \begin{cases} x'_1 + x'_2 + x'_3 = 0 \\ x'_1 + x'_2 = 0 \\ x'_3 = 0 \end{cases}$$

2. Simplificamos el sistema de ecuaciones implícitas (triangulando sus coeficientes):

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{cases} x'_1 + x'_2 + x'_3 = 0 \\ x'_3 = 0 \end{cases}$$

3. Obtenemos una base de $kerf$ a través de sus ecuaciones paramétricas:

$$\begin{cases} x'_1 + x'_2 + x'_3 = 0 \quad \to \quad x'_3 = -x'_1 - x'_2 & \to \quad \beta = -\alpha - \gamma \\ x'_3 = 0 \quad \to \quad x'_3 = 0 & \to \quad \gamma = 0 \end{cases}$$

$$\begin{cases} x_1 = \alpha \\ x_2 = \beta \\ x_3 = \gamma \end{cases} \xrightarrow{\beta = -\alpha - \gamma} \begin{cases} x_1 = \alpha \\ x_2 = -\alpha - \gamma \\ x_3 = \gamma \end{cases} \xrightarrow{\gamma = 0} \begin{cases} x_1 = \alpha \\ x_2 = -\alpha \\ x_3 = 0 \end{cases}$$

$$B_{kerf} = \{(1, -1, 0)\}$$

2.3.1. Propiedades del núcleo de una aplicación lineal

1. $Kerf$ es un subespacio vectorial de E.

2. f es una aplicación lineal inyectiva (monomorfismo) si y solo si $Kerf = \{\bar{0}_E\}$.

3. $Kerf = \{\bar{0}_E\} \Longleftrightarrow$ las imágenes de los vectores de un sistema generador libre S de E constituyen un sistema generador libre S' de F. Es decir, que E y F deben tener la misma dimensión.

2.4. Rango de una aplicación lineal

El rango de una aplicación lineal f es la dimensión del espacio $f(E) = Imf$. Si los espacios E y $f(E)$ son de dimensión finita, se tiene que:

$$dim(E) = dim\ kerf + rg(f) = dim\ kerf + dim\ Imf$$

Nota. El rango de una aplicación lineal f es igual al rango de su matriz asociada $M(f)$.

2.5. Matriz de una aplicación lineal

Sean E y F dos espacios vectoriales y $f : E \longmapsto F$ una aplicación lineal, se define como matriz asociada a (f), que denotamos por $M_{B_F \leftarrow B_E}(f)$, la matriz que $\forall \bar{x} \in E$ y $\forall \bar{y} \in F$ cumple la siguiente relación:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{B_F} = M_{B_F \leftarrow B_E}(f) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{B_E}$$

La matriz asociada a (f) se obtiene colocando en columnas las imágenes $f(\bar{e}_i)$ de los vectores $\bar{e}_i$ de la base B_E expresados en la propia base B_E. Veamos que, como una base está conformada por vectores linealmente independientes, la única forma de expresar un vector de una base en esa misma base es él mismo, es decir, $B_E = \{\bar{e}_1, \ldots, \bar{e}_n\}_{Bc} = \{(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\}_{B_E}$. Por lo que, en la práctica, para construir $M_{B_F \leftarrow B_E}(f)$, emplearemos las imágenes en f de vectores de la misma forma que los de la base canónica.

Nota. Cuando una aplicación $f : E \longmapsto E$ lleva los vectores de E al mismo espacio E en la misma base B_E, se suele denotar su matriz por $M_{B_E}(f)$.

Ejemplo
Dada una aplicación lineal $f : E \longmapsto F$ en su forma analítica y expresada en la base B_c :

$f(x_1, x_2, x_3) = (2x_2 + 2x_3, 3x_1)$
$B_E = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\} = \{(1, -2, -4), (1, 0, 0), (1, 0, 3)\}$
$B_F = \{(2, -1), (0, 1)\}$

Hallar la matriz asociada a la aplicación (f) expresada entre las bases B_E y B_F.

1. Para construir la matriz pedida debemos primero construir la matriz de la aplicación en base canónica y después concatenarla con las dos de cambio de base:

$$B_E \longrightarrow B_c \xrightarrow{f} B_c \longrightarrow B_F$$

2. Obtenemos las imágenes respecto de f de los vectores de la base de B_E expresados en la propia base B_E (vectores canónicos):

$$f(\bar{e}_1) = f((1, 0, 0)_{B_E}) = (1, 0)_{B_F}$$
$$f(\bar{e}_2) = f((0, 1, 0)_{B_E}) = (0, 1)_{B_F}$$
$$f(\bar{e}_3) = f((0, 0, 1)_{B_E}) = (2, 0)_{B_F}$$

3. Construimos la matriz de la aplicación lineal en la base canónica:

$$M_{B_F \leftarrow B_E}(f) = \left(\begin{array}{c|c|c} f(1,0,0) & f(0,1,0) & f(0,0,1) \end{array} \right) = \left(\begin{array}{c|c|c} 2 & 2 & 0 \\ 0 & 0 & 3 \end{array} \right)$$

4. Obtenemos las matrices de cambio de base entre ambas bases y la base canónica:

$$M_{B_c \leftarrow B_E} = \begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 0 \\ -4 & 0 & 3 \end{pmatrix} \qquad M_{B_2 \leftarrow B_c} = M^{-1}_{B_c \leftarrow B_2} = \begin{pmatrix} 2 & 3 \\ -1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & -3 \\ 1 & 2 \end{pmatrix}$$

5. Construimos la matriz de la aplicación lineal entre las bases indicadas:

$$M_{B_c \leftarrow B_E} = M_{B_2 \leftarrow B_c} M(f)_{B_c} M_{B_c \leftarrow B_3}$$

$$M_{B_c \leftarrow B_E} = \begin{pmatrix} -1 & -3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 0 \\ -4 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 38 & -2 & -29 \\ -26 & 2 & 20 \end{pmatrix}$$

Ejemplo
Dada una aplicación lineal $f : E \longmapsto F$ expresada directamente entre las bases B_1 y B_2 :

$f((x_1, x_2, x_3)_{B_E}) = (x_1 + 2x_3, x_2)_{B_F}$
$B_E = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\} = \{(1, 1, 1), (0, 1, 2), (0, 2, 1)\}$
$B_F = \{(1, 0), (1, 1)\}$

Hallar la matriz asociada a la aplicación (f) expresada entre las bases B_E y B_F.

1. Las bases no tienen ninguna relevancia, ya que, como nos dan la aplicación f directamente expresada entre las bases B_E y B_F, no es necesario hacer ningún cambio de base:

$$B_E \xrightarrow{f} B_F$$

2. Obtenemos las imágenes respecto de f de los vectores de la base de B_E expresados en la propia base B_E (vectores canónicos):

$$f(\bar{e}_1) = f((1, 0, 0)_{B_E}) = (0, 3)_{B_F}$$
$$f(\bar{e}_2) = f((0, 1, 0)_{B_E}) = (2, 0)_{B_F}$$
$$f(\bar{e}_3) = f((0, 0, 1)_{B_E}) = (2, 0)_{B_F}$$

3. Construimos la matriz de la aplicación lineal directamente entre las bases indicadas:

$$M_{B_F \leftarrow B_E}(f) = \left(\begin{array}{c|c|c} f(1,0,0) & f(0,1,0) & f(0,0,1) \end{array} \right) = \left(\begin{array}{c|c|c} 1 & 0 & 2 \\ 0 & 1 & 0 \end{array} \right)$$

Ejemplo
Se consideran los espacios vectoriales $\mathbb{R}^2$ y $\mathbb{R}^3$. Se define entre ellos la familia de homomorfismos $f_a : \mathbb{R}^3 \mapsto \mathbb{R}^2$ que verifica:

- Las ecuaciones implícitas de $ker f_a$ son $\left\{ \begin{array}{lcl} x + ay + z & = & 0 \\ ax + y + z & = & 0 \end{array} \right.$
- $f_a(0, 0, a - 1) = (0, a - 1)$
- $f_a(1, 0, 1) = (2, 1)$

Se pide calcular la matriz de los homomorfismos f_a expresada entre bases canónicas:

La matriz de una aplicación lineal se construye colocando ordenadamente y en vertical las imágenes de los vectores canónicos. En este caso, se obtendrá una matriz con elementos dependientes de a.

1. Combinar linealmente las dos preimágenes que ofrece el problema (aplicando las propiedades de aplicaciones lineales que nos da el teorema de caracterización) para obtener tantos vectores canónicos como sea posible:

$$\text{Propiedades de aplicaciones lineales:} \quad \left\{ \begin{array}{c} f(\bar{x}) + f(\bar{y}) = f(\bar{x} + \bar{y}) \\ \lambda \cdot f(\bar{x}) = f(\lambda \bar{x}) \end{array} \right.$$

$$f(0, 0, 1) = f(\frac{1}{a-1}((0, 0, a - 1)) = \frac{1}{a-1} f(0, 0, a - 1) = \frac{1}{a-1}(0, a - 1) = (0, 1)$$
$$f(1, 0, 0) = f((1, 0, 1) - (0, 0, 1)) = f(1, 0, 1) - f(0, 0, 1) = (2, 1) - (0, 1) = (2, 0)$$

$$f(1, 0, 0) = (2, 0) \qquad f(0, 0, 1) = (0, 1)$$

2. Obtener un generador del kernel de la aplicación f mediante sus las ecuaciones paramétricas:

$$\begin{array}{l} x + ay + z = 0 \longrightarrow z = -(x + ay) \longrightarrow \gamma = -(\alpha + a\beta) \\ ax + y + z = 0 \longrightarrow z = -(ax + a) \longrightarrow \gamma = -(a\alpha + \beta) \end{array}$$

$$-(\alpha + a\beta) = -(\alpha + a\beta) \longrightarrow \alpha(1 - a) = \beta(1 - a) \iff \alpha = \beta$$

$$\left\{ \begin{array}{lcl} x_1 & = & \alpha \\ x_2 & = & \quad \beta \\ x_3 & = & \qquad \gamma \end{array} \right. \rightarrow \left\{ \begin{array}{lcl} x_1 & = & \alpha \\ x_2 & = & \qquad \beta \\ x_3 & = & -\alpha \; -a\beta \end{array} \right. \rightarrow \left\{ \begin{array}{lcl} x_1 & = & \alpha \\ x_2 & = & \alpha \\ x_3 & = & -\alpha \; -a\beta \end{array} \right. \rightarrow ker f_a = L\{(1, 1, -(a + 1)\}$$

3. Combinar linealmente las preimágenes que ya tenemos con la que acabamos de calcular del kernel para obtener el vector canónico que nos falta:

$$f(1, 1, -a - 1) + f(0, 0, a - 1) = f(1, 1, -2) = (0, 0) - (0, -a - 1) = (0, a - 1) \longrightarrow$$
$$\longrightarrow f(1, 1, -2) + 2 \cdot f(0, 0, 1) = f(1, 1, 0) = (0, a - 1) + (0, 2) = (0, a + 1) \longrightarrow$$
$$\longrightarrow f(1, 1, 0) - f(1, 0, 0) = f(0, 1, 0) = (0, a + 1) - (2, 0) = (-2, a + 1)$$
$$f(0, 1, 0) = (-2, a + 1)$$

4. Colocar ordenadamente las imágenes de los tres vectores canónicos para obtener la matriz de los homomorfismos f_a

$$M(f_a) = \left(\begin{array}{c|c|c} f(1,0,0) & f(0,1,0) & f(0,0,1) \end{array} \right) = \left(\begin{array}{c|c|c} 2 & -2 & 0 \\ 0 & a + 1 & 1 \end{array} \right)$$

2.5.1. Teorema de unicidad de aplicaciones lineales

El teorema de unicidad nos dice que una aplicación lineal queda perfectamente definida por las imágenes de los elementos de su dominio (Imágenes de vectores del espacio origen), es decir, que queda perfectamente definida por su matriz asociada (Ya que esta se forma con las imágenes de la base del espacio origen).

Esto significa que, aunque una aplicación lineal se pueda expresar de muchas formas distintas, si la matriz asociada a todas ellas es la misma, la aplicación es la misma.

DEMOSTRACIÓN

Veamos que se puede descomponer una aplicación en una combinación lineal de las imágenes de los elementos de su base.

$$\text{Sea} \quad \bar{x} \in E \quad \text{con} \quad B_E = \{\bar{e}_1, \ldots, \bar{e}_n\} \quad \text{tq.} \quad \bar{x} = (\alpha\bar{e}_1 + \cdots + \gamma\bar{e}_n)$$

$$f(\bar{x}) = f(\alpha\bar{e}_1 + \cdots + \gamma\bar{e}_n) = f(\alpha\bar{e}_1) + \cdots + f(\gamma\bar{e}_n) = \alpha f(\bar{e}_1) + \cdots + \gamma f(\bar{e}_n) = \alpha f_1 + \cdots + \gamma f_n$$

$\square$

Esto quiere decir que la aplicación (f) viene determinada por el conjunto $\{f_1, \ldots, f_n\}$ de imágenes de la base B_E (ya que el conjunto de coeficientes $\{\alpha, \ldots, \gamma\}$ dependen del vector $\bar{x}$, no de la aplicación f). Cualquier otra aplicación que comparta $\{f_1, \ldots, f_n\}$ será exactamente la misma.

2.5.2. Operaciones con matrices de aplicación lineal

Sean E, F, y G espacios vectoriales y B_E, B_F, B_G sus respectivas bases. Sean f y g dos aplicaciones lineales de E en F se cumple que:

- La **suma** de aplicaciones, $f + g$, definida por:

$$(f + g)(\bar{x}) = f(\bar{x}) + g(\bar{x}), \quad \forall \bar{x} \in E$$

es también lineal y cumple que:

$$M(f + g)_{B_F \leftarrow B_E} = M(f)_{B_F \leftarrow B_E} + M(g)_{B_F \leftarrow B_E}$$

Ejemplo
Dadas dos aplicaciones lineales f y g de E en F y un vector $\bar{x}$:

$M(f)_{Bc} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 2 \\ 3 & 1 & 0 \end{pmatrix} \qquad M(g)_{Bc} = \begin{pmatrix} 0 & 2 & -1 \\ 1 & -4 & 2 \\ -2 & 1 & 1 \end{pmatrix} \qquad \bar{x} = (1, 1, 3)$

Veamos que $M(f + g)_{B_F \leftarrow B_E} = M(f)_{B_F \leftarrow B_E} + M(g)_{B_F \leftarrow B_E}$:

1. $M(f)_{B_F \leftarrow B_E}(\bar{x}) + M(g)_{B_F \leftarrow B_E}(\bar{x})$:

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 2 \\ 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ 4 \end{pmatrix} \qquad \begin{pmatrix} 0 & 2 & -1 \\ 1 & -4 & 2 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix} \qquad \begin{pmatrix} 4 \\ 7 \\ 4 \end{pmatrix} + \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 10 \\ 6 \end{pmatrix}$$

2. $(M(f)_{B_F \leftarrow B_E} + M(g)_{B_F \leftarrow B_E})(\bar{x})$:

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 2 \\ 3 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 2 & -1 \\ 1 & -4 & 2 \\ -2 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & -4 & 4 \\ 1 & 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 0 \\ 2 & -4 & 4 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 10 \\ 6 \end{pmatrix}$$

- El **producto por escalar** de una aplicación, $\lambda \cdot f$, definido por:

$$(\lambda f)(\bar{x}) = \lambda \cdot f(\bar{x}), \quad \forall \bar{x} \in E \ \forall \lambda \in \mathbb{R}$$

es también lineal y cumple que:

$$M(\lambda f)_{B_F \leftarrow B_E} = \lambda \cdot M(f)_{B_F \leftarrow B_E}$$

Ejemplo
Dada una aplicación lineal f, un vector $\bar{x}$ y un escalar λ:

$M(f)_{Bc} = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \qquad \bar{x} = (2, 1, 0) \qquad \lambda = 2$

Veamos que $M(\lambda \cdot f)_{B_F \leftarrow B_E} = \lambda \cdot \lambda \cdot M(f)_{B_F \leftarrow B_E}$:

1. $(\lambda \cdot M(f)_{B_F \leftarrow B_E})(\bar{x})$:

$$\left(2 \cdot \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \right) = \begin{pmatrix} 2 & 0 & 2 \\ 4 & 2 & 4 \\ 0 & 2 & 2 \end{pmatrix} \qquad \begin{pmatrix} 2 & 0 & 2 \\ 4 & 2 & 4 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 10 \\ 2 \end{pmatrix}$$

2. $\lambda \cdot (M(f)_{B_F \leftarrow B_E}(\bar{x}))$:

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix} \qquad \left(2 \cdot \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 4 \\ 10 \\ 2 \end{pmatrix}$$

- La **composición** de dos aplicaciones lineales, $h \circ f$ (sea h una aplicación de F en G), definida por:

$$(h \circ f)(\bar{x}) = h(f(\bar{x})), \quad \forall \bar{x} \in E$$

es también lineal y cumple que:

$$M(h \circ f)_{B_G \leftarrow B_E} = M(h)_{B_G \leftarrow B_F} \cdot M(f)_{B_F \leftarrow B_E}$$

DIAGRAMA

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[rr, bend right, "h \circ f"'] & F \arrow[r, "h"] & G \end{tikzcd}$$

Ejemplo
Dadas, una aplicación lineal f de E en F, una aplicación lineal h de F en G, y un vector $\bar{x}$:

$M(f)_{Bc} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix} \qquad M(h)_{Bc} = \begin{pmatrix} 0 & 1 & 1 \\ 2 & 0 & 2 \\ 3 & 2 & 0 \end{pmatrix} \qquad \bar{x} = (1, 1, 0)$

Veamos que $M(h \circ f)_{B_G \leftarrow B_E} = M(h)_{B_F \leftarrow B_E} \cdot M(h)_{B_F \leftarrow B_E}$:

1. $M(h)_{B_G \leftarrow B_F}(M(f)_{B_F \leftarrow B_E}(\bar{x}))$:

$$\begin{pmatrix} 0 & 1 & 1 \\ 2 & 0 & 2 \\ 3 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 2 \\ 0 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 12 \\ 9 \\ 11 \end{pmatrix}$$

2. $(M(h)_{B_G \leftarrow B_F} \cdot M(f)_{B_F \leftarrow B_E})(\bar{x})$:

$$\begin{pmatrix} 0 & 1 & 2 \\ 0 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 2 & 0 & 2 \\ 3 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 8 & 4 & 2 \\ 7 & 2 & 4 \\ 7 & 4 & 6 \end{pmatrix} \qquad \begin{pmatrix} 8 & 4 & 2 \\ 7 & 2 & 4 \\ 7 & 4 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 12 \\ 9 \\ 11 \end{pmatrix}$$

2.5.3. Cambio de base de una aplicación lineal

Una aplicación lineal $f : E \longmapsto F$ lleva vectores del espacio E con una determinada base B_1 a vectores del espacio F en una determinada base B_2. Si se quiere llevar vectores del espacio E en una determinada base B_1' a vectores del espacio F en una determinada base B_2', se deben aplicar cambios de base sobre la entrada y la salida de la aplicación lineal. El producto (la composición) de matrices de cambio de base y matrices de aplicación lineal cumple la siguiente relación:

$$M(f)_{B_2' \leftarrow B_1'} = M_{B_2' \leftarrow B_2} \cdot M(f)_{B_2 \leftarrow B_1} \cdot M_{B_1 \leftarrow B_1'}$$

DIAGRAMA

$$\begin{tikzcd} B_1 \arrow[r, "f"] & B_2 \arrow[d, "M_{B_2' \leftarrow B_2}"] \\ B_1' \arrow[u, "M_{B_1 \leftarrow B_1'}"] & B_2' \end{tikzcd}$$

Ejemplo
Se tiene la matriz $M(f)_{Bc}$ de f en la base canónica de $\mathbb{R}^3$, y dos bases B_1 y B_2 de $\mathbb{R}^3$:

$M(f)_{Bc} = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 2 \\ 3 & 1 & 0 \end{pmatrix}$

$B_1 = \{(1, 2, 3), (2, 1, 1), (1, 0, 0)\}$

$B_2 = \{(3, 0, -2), (-6, 1, 4), (-1, 0, 1)\}$

La matriz $M(f)_{B_2 \leftarrow B_1}$ de f expresada entre las bases B_1 y B_2 es:

1. Veamos que $M(f)_{B_2 \leftarrow B_1} = M_{B_2 \leftarrow Bc} \cdot M(f)_{B_c} \cdot M_{Bc \leftarrow B_1}$. Por tanto, para obtener la aplicación $M(f)_{B_2 \leftarrow B_1}$ primero necesitamos obtener las matrices de cambio de base $M_{B_2 \leftarrow Bc}$ y $M_{Bc \leftarrow B_1}$:

$$\begin{tikzcd} Bc \arrow[r, "f"] & Bc \arrow[d, "M_{B_2 \leftarrow Bc}"] \\ B_1 \arrow[u, "M_{Bc \leftarrow B_1}"] & B_2 \end{tikzcd}$$

2. La matriz $M_{Bc \leftarrow B_1}$ se obtiene simplemente colocando los vectores de la base de B_1 en columnas:

$$M_{Bc \leftarrow B_1} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 3 & 1 & 0 \end{pmatrix}$$

3. La matriz $M_{B_2 \leftarrow Bc}$ se obtiene invirtiendo la matriz $M_{Bc \leftarrow B_2}$. La matriz $M_{Bc \leftarrow B_2}$ se obtiene colocando los vectores de la base de B_2 en columnas:

$$M_{Bc \leftarrow B_1} = \begin{pmatrix} 3 & -6 & -1 \\ 0 & 1 & 0 \\ -2 & 4 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 3 \end{pmatrix}$$

4. Obtenemos la matriz de f expresada entre las bases B_1 y B_2 aplicando $M(f)_{B_2 \leftarrow B_1} = M_{B_2 \leftarrow Bc} \cdot M(f)_{B_c} \cdot M_{Bc \leftarrow B_1}$:

$$M(f)_{B_2 \leftarrow B_1} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 2 \\ 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 3 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 27 & 20 & 6 \\ 7 & 4 & 1 \\ 31 & 31 & 11 \end{pmatrix}$$

2.5.4. Matriz de cambio de base como un caso particular de aplicación lineal

Un cambio de base se define como una aplicación lineal que permite relacionar entre sí las coordenadas de un espacio vectorial expresadas respecto a dos bases distintas. Es decir, cuando tenemos una aplicación $f : E \longmapsto E$ que toma vectores en una base B_1 de E y los lleva a otra base B_2 de E es indistinto entenderla como una aplicación, como un cambio de base, o como la aplicación Id entre las bases B_1 y B_2.

DIAGRAMA

$$B_1 \xrightarrow{M_{B_2 \leftarrow B_1}} B_2 \quad \sim \quad \begin{tikzcd} Bc \arrow[r, "Id"] & Bc \arrow[d, "M_{B_2 \leftarrow Bc}"] \\ B_1 \arrow[u, "M_{B_c \leftarrow B_1}"] & B_2 \end{tikzcd} \quad \sim \quad B_1 \xrightarrow{M(f)} B_2$$

Nota. La aplicación $Id : E \longmapsto E$ es la aplicación que toma vectores de E en una determinada base B y los lleva a vectores de E en esa misma base, es decir, no hace ningún cambio sobre el vector. Su matriz de aplicación asociada es la matriz diagonal de 1(s).

2.6. Espacio vectorial de homomorfismos (aplicaciones lineales)

Sean (E y F) dos espacios vectoriales y (f y g) dos aplicaciones lineales de E en F:

1. Si la aplicación definida por $(f+g)(\bar{x}) \quad \forall \bar{x} \in E$ es lineal, entonces $f(\bar{x}) + g((\bar{x})$ es también lineal y $M_{B_F \leftarrow B_E}(f+g) = M_{B_F \leftarrow B_E}(f) + M_{B_F \leftarrow B_E}(g)$

2. la aplicación definida por $f(\lambda f)\bar{x} \quad \forall \bar{x} \in E$ y $\forall \lambda \in \mathbb{R}$ es lineal, entonces $\lambda(f(\bar{x}))$ es también lineal y $\lambda \cdot M_{B_F \leftarrow B_E}(f) = M_{B_F \leftarrow B_E}(\lambda f)$

El conjunto de los homomorfismos de E en F, al que denotamos por $Hom(E,F)$, con las operaciones de suma y producto por escalares anteriores da lugar un espacio vectorial y $dim\,(Hom(E,F)) = dim(E) \cdot dim(F)$

Nota 1. $Hom(E,F)$ es espacio vectorial solo si E y F son de dimensión finita

Nota 2. Al ser $Hom(E,F)$ un espacio vectorial se pueden aplicar todas las propiedades y axiomas de espacios vectoriales

Ejemplo
Espacio vectorial de los endomorfismos de $\mathbb{R}^2$ $[End(\mathbb{R}^2) = Hom(\mathbb{R}^2, \mathbb{R}^2)]$.

$B_{End(\mathbb{R}^2)} = \{\left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right), \left(\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right)\} = \{f_1, f_2, f_3, f_4\} = \{f_i\}_{i=1}^{n}$

$\forall f \in End(\mathbb{R}^2)$

$$f = \sum_{i=1}^{4} \lambda_i f_i = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 + \lambda_4 f_4 = \lambda_1 \left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right) + \lambda_2 \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right) + \lambda_3 \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right) + \lambda_4 \left(\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right)$$

2.7. Espacio dual

El espacio dual asociado a un espacio vectorial (E) es el homomorfismo de las funciones que llevan a los vectores de dicho espacio vectorial a $(\mathbb{R})$. Se denota por E^*.

$$E^* = Hom(E, \mathbb{R})$$

Los vectores de un espacio dual, al tratarse de aplicaciones lineales $f : E \times \mathbb{R}$ que transforman vectores n-dimensionales en escalares, tienen siempre dimensión $1 \times n$.

Nota. Se denomina forma lineal a una aplicación lineal que lleva vectores de E a $\mathbb{R}$ $(Hom(E, \mathbb{R}))$

Ejemplo
La forma lineal $(1, 0, 2) \in \mathbb{R}^{3*}$ toma vectores de $\mathbb{R}^3$ y los transforma en escalares.

$$(1,0,2)\begin{pmatrix}2\\1\\4\end{pmatrix} = (1 \cdot 2 + 0 \cdot 1 + 2 \cdot 4) = 10$$

2.7.1. Noción de evaluador del espacio dual

Sea E^* el espacio dual de E, cada vector $\bar{e}^* \in E^*$ define una forma lineal de E en $\mathbb{R}$ que se denomina "evaluador en $\bar{e}^*$" y que se denota por $ev[\bar{e}^*]$. Es decir, tomando como fijo un vector cualquiera $\bar{e}^* \in E^*$, obtenemos una aplicación $ev[\bar{e}^*]$ que toma vectores de E y los multiplica por el vector $\bar{e}^*$.

$$\text{Sean:} \quad \bar{e}^* \text{ (fijo) } \in E^* \quad y \quad \bar{v} \text{ (genérico) } \in E$$
$$ev[\bar{e}^*] : E \longmapsto \mathbb{R}$$
$$ev[\bar{e}^*](\bar{v}) = \bar{e}^* \cdot \bar{v} = \lambda$$

DIAGRAMA

$$\bar{e} \in E \xrightarrow[\bar{e}^* \in E^*]{ev[\bar{e}^*](\bar{e}) \in E^*} \mathbb{R}$$

2.7.2. Base del espacio dual

Aunque cualquier base que tomemos para el espacio dual puede generar vectores que sirvan como $Hom(E, \mathbb{R})$, esta debe cumplir una serie de condiciones. La base del espacio dual, que se denota por B^*, es el conjunto de formas lineales que llevan a su correspondiente vector de la base de E a $1 \in \mathbb{R}$ y al resto de vectores de la base canónica de E a $0 \in \mathbb{R}$.

$$B_{E^*} = B^* = \{\bar{e}_1^*, \dots, \bar{e}_n^*\} = \{\bar{e}_j^*\}$$
$$\text{tq. } \forall E \quad con \quad B_E = \{\bar{e}_1, \dots, \bar{e}_n\} = \{\bar{e}_j\}$$
$$\bar{e}_j^*(\bar{e}_i) \;= \delta_{i,j} = \begin{cases} 1 & \text{si} j = i \\ 0 & \text{si} j \neq i \end{cases}$$

Nota. $\delta_{i,j}$ se denomina "delta de Kronecker"

Para calcular la base dual $B_{E^*} = \{\bar{e}_1^*, \dots, \bar{e}_n^*\}$ asociada a una base $B_E = \{\bar{e}_1, \dots, \bar{e}_n\}$ se fuerza la obtención de la delta de Kronecker en la combinatoria de productos de vectores duales por vectores no duales $(\bar{e}_j^*(\bar{e}_i) = \delta_{i,j})$. Por cada vector $\bar{e}_j^*$ de la base dual se plantea un sistema de n ecuaciones con n incógnitas (cada una de los n $\delta_{j,:}$) que permite despejar los coeficientes de $\bar{e}_j^*$.

Ejemplo
Calcular la base dual $B_{E^*} = \{\bar{e}_1^*, \bar{e}_2^*, \bar{e}_3^*\}$ de la base $B_E = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\} = \{(1, 0, 0), (0, 1, 3), (1, 0, 2)\}$.

$$\begin{matrix} \bar{e}_1^* = (e_{1,1}^*, e_{1,2}^*, e_{1,3}^*) \\ \bar{e}_2^* = (e_{2,1}^*, e_{2,2}^*, e_{2,3}^*) \\ \bar{e}_3^* = (e_{3,1}^*, e_{3,2}^*, e_{3,3}^*) \end{matrix} \quad \longrightarrow \quad \forall \bar{x} = (x, y, z) \in E \quad \begin{matrix} \bar{e}_1^*(\bar{x}) = (e_{1,1}^*x + e_{1,2}^*y + e_{1,3}^*z) \\ \bar{e}_2^*(\bar{x}) = (e_{2,1}^*x + e_{2,2}^*y + e_{2,3}^*z) \\ \bar{e}_3^*(\bar{x}) = (e_{3,1}^*x + e_{3,2}^*y + e_{3,3}^*z) \end{matrix}$$

$$\begin{cases} 1 = \bar{e}_1^*(\bar{e}_1) \\ 0 = \bar{e}_1^*(\bar{e}_2) \\ 0 = \bar{e}_1^*(\bar{e}_3) \end{cases} \rightarrow \begin{cases} 1 = (1 \cdot e_{1,1}^* + 0 \cdot e_{1,2}^* + 0 \cdot e_{1,3}^*) \\ 0 = (0 \cdot e_{1,1}^* + 1 \cdot e_{1,2}^* + 3 \cdot e_{1,3}^*) \\ 0 = (1 \cdot e_{1,1}^* + 0 \cdot e_{1,2}^* + 2 \cdot e_{1,3}^*) \end{cases} \rightarrow \begin{cases} 1 = e_{1,1}^* \\ 0 = e_{1,2}^* + 3e_{1,3}^* \\ 0 = e_{1,1}^* + 2e_{1,3}^* \end{cases} \quad \bar{e}_1^* = (1, \tfrac{3}{2}, -\tfrac{1}{2})$$

$$\begin{cases} 0 = \bar{e}_2^*(\bar{e}_1) \\ 1 = \bar{e}_2^*(\bar{e}_2) \\ 0 = \bar{e}_2^*(\bar{e}_3) \end{cases} \rightarrow \begin{cases} 0 = (1 \cdot e_{2,1}^* + 0 \cdot e_{2,2}^* + 0 \cdot e_{2,3}^*) \\ 1 = (0 \cdot e_{2,1}^* + 1 \cdot e_{2,2}^* + 3 \cdot e_{2,3}^*) \\ 0 = (1 \cdot e_{2,1}^* + 0 \cdot e_{2,2}^* + 2 \cdot e_{2,3}^*) \end{cases} \rightarrow \begin{cases} 0 = e_{2,1}^* \\ 1 = e_{2,2}^* + 3e_{2,3}^* \\ 0 = e_{2,1}^* + 2e_{2,3}^* \end{cases} \quad \bar{e}_2^* = (0, 1, 0)$$

$$\begin{cases} 0 = \bar{e}_3^*(\bar{e}_1) \\ 0 = \bar{e}_3^*(\bar{e}_2) \\ 1 = \bar{e}_3^*(\bar{e}_3) \end{cases} \rightarrow \begin{cases} 0 = (1 \cdot e_{3,1}^* + 0 \cdot e_{3,2}^* + 0 \cdot e_{3,3}^*) \\ 0 = (0 \cdot e_{3,1}^* + 1 \cdot e_{3,2}^* + 3 \cdot e_{3,3}^*) \\ 1 = (1 \cdot e_{3,1}^* + 0 \cdot e_{3,2}^* + 2 \cdot e_{3,3}^*) \end{cases} \rightarrow \begin{cases} 0 = e_{3,1}^* \\ 0 = e_{3,2}^* + 3e_{3,3}^* \\ 1 = e_{3,1}^* + 2e_{3,3}^* \end{cases} \quad \bar{e}_3^* = (0, -\tfrac{3}{2}, \tfrac{1}{2})$$

$$B_{E^*} = \{(1, \tfrac{3}{2}, -\tfrac{1}{2}), (0, 1, 0), (0, -\tfrac{3}{2}, \tfrac{1}{2})\}$$

Veamos que, debido al ya mencionado cumplimiento de la delta de Kronecker, siempre se cumple que:

$$\forall \bar{x} = (\alpha_1\bar{e}_1 + \dots + \alpha_j\bar{e}_j + \dots + \alpha_n\bar{e}_n) \in E \qquad \forall \bar{e}^* = (\lambda_1\bar{e}_1^* + \dots + \lambda_j\bar{e}_j^* + \dots + \lambda_n\bar{e}_n^*) \in E^*$$

- $\bar{e}_j^*(\bar{x}) = \bar{e}_j^*(\alpha_1\bar{e}_1 + \dots + \alpha_j\bar{e}_j + \dots + \alpha_n\bar{e}_n) = (\alpha_1\cancelto{0}{\bar{e}_j^*\bar{e}_1} + \dots + \alpha_j\cancelto{1}{\bar{e}_j^*\bar{e}_j} + \dots + \alpha_n\cancelto{0}{\bar{e}_j^*\bar{e}_n}) = \alpha_j$
- $\bar{e}^*(\bar{x}) = \bar{e}^*(\alpha_1\bar{e}_1) + \dots + \bar{e}^*(\alpha_j\bar{e}_j) + \dots + \bar{e}^*(\alpha_n\bar{e}_n) = \alpha_1(\bar{e}^*\bar{e}_1) + \dots + \alpha_j(\bar{e}^*\bar{e}_j) + \dots + \alpha_n(\bar{e}^*\bar{e}_n) =$
$= \alpha_1(\lambda_1\bar{e}_1^*\bar{e}_1) + \dots + \alpha_j(\lambda_j\bar{e}_j^*\bar{e}_j) + \dots + \alpha_n(\lambda_n\bar{e}_n^*\bar{e}_n) = \alpha_1(\lambda 1) + \dots + \alpha_j(\lambda_j) + \dots + \alpha_n(\lambda_n)$

También se puede observar que, si contamos con un espacio vectorial E expresado en su base canónica, la base de E^* es exactamente la misma que la de E.

Ejemplo
Calcular la base dual $Bc(\mathbb{R}^{3*}) = \{\bar{e}_1^*, \bar{e}_2^*, \bar{e}_3^*\}$ de la base $Bc(\mathbb{R}^3) = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

$$\begin{cases} 1 = \bar{e}_1^*(\bar{e}_1) \\ 0 = \bar{e}_1^*(\bar{e}_2) \\ 0 = \bar{e}_1^*(\bar{e}_3) \end{cases} \rightarrow \begin{cases} 1 = (1 \cdot e_{1,1}^* + 0 \cdot e_{1,2}^* + 0 \cdot e_{1,3}^*) \\ 0 = (0 \cdot e_{1,1}^* + 1 \cdot e_{1,2}^* + 0 \cdot e_{1,3}^*) \\ 0 = (0 \cdot e_{1,1}^* + 0 \cdot e_{1,2}^* + 1 \cdot e_{1,3}^*) \end{cases} \rightarrow \begin{cases} 1 = e_{1,1}^* \\ 0 = e_{1,2}^* \\ 0 = e_{1,3}^* \end{cases} \quad \bar{e}_1^* = (1, 0, 0)$$

$$\begin{cases} 0 = \bar{e}_2^*(\bar{e}_1) \\ 1 = \bar{e}_2^*(\bar{e}_2) \\ 0 = \bar{e}_2^*(\bar{e}_3) \end{cases} \rightarrow \begin{cases} 0 = (1 \cdot e_{2,1}^* + 0 \cdot e_{2,2}^* + 0 \cdot e_{2,3}^*) \\ 1 = (0 \cdot e_{2,1}^* + 1 \cdot e_{2,2}^* + 0 \cdot e_{2,3}^*) \\ 0 = (0 \cdot e_{2,1}^* + 0 \cdot e_{2,2}^* + 1 \cdot e_{2,3}^*) \end{cases} \rightarrow \begin{cases} 0 = e_{2,1}^* \\ 1 = e_{2,2}^* \\ 0 = e_{2,3}^* \end{cases} \quad \bar{e}_2^* = (0, 1, 0)$$

$$\begin{cases} 0 = \bar{e}_3^*(\bar{e}_1) \\ 0 = \bar{e}_3^*(\bar{e}_2) \\ 1 = \bar{e}_3^*(\bar{e}_3) \end{cases} \rightarrow \begin{cases} 0 = (1 \cdot e_{3,1}^* + 0 \cdot e_{3,2}^* + 0 \cdot e_{3,3}^*) \\ 0 = (0 \cdot e_{3,1}^* + 1 \cdot e_{3,2}^* + 0 \cdot e_{3,3}^*) \\ 1 = (0 \cdot e_{3,1}^* + 0 \cdot e_{3,2}^* + 1 \cdot e_{3,3}^*) \end{cases} \rightarrow \begin{cases} 0 = e_{3,1}^* \\ 0 = e_{3,2}^* \\ 1 = e_{3,3}^* \end{cases} \quad \bar{e}_3^* = (0, 0, 1)$$

$$Bc(\mathbb{R}^{3*}) = \{(1, 0, 0, (0, 1, 0), (0, 0, 1)\} = Bc(\mathbb{R}^3)$$

2.7.3. Aplicación traspuesta

Sean dos espacios vectoriales E y F, y f una aplicación lineal de E en F de matriz asociada $M(f)$, entonces, se define como aplicación traspuesta de f la aplicación lineal f^t de F^* en E^* y de matriz asociada $M(f)^T$.

$$\forall \alpha \in F^*$$
$$f^t[\alpha] = \alpha \circ f : E \longmapsto \mathbb{R}$$
$$f^t[\alpha](\bar{e}) = (\alpha \circ f)(\bar{e}) = \alpha(f(\bar{e})) = \alpha(\bar{v}) \qquad \forall \bar{e} \in E \text{ tq. } f(\bar{e}) = \bar{v} \in F$$

DIAGRAMA

$$\begin{tikzcd} \bar{e} \in E \arrow[r, dash] \arrow[d, "f"'] & f^t[\alpha] \in E^* \arrow[r] & \mathbb{R} \\ \bar{v} \in F \arrow[r, dash] & \alpha \in F^* \arrow[u, "f^t"'] \arrow[ur, bend right] & \end{tikzcd}$$

Nota. Para poder emplear la aplicación traspuesta es necesario que los dos espacios vectoriales que relaciona la aplicación tengan la misma dimensión. En caso contrario la aplicación traspuesta resultante no será biyectiva.

DEMOSTRACIÓN

Sean $f : E \mapsto F$ y $f^t : F^* \mapsto E^*$ una aplicación y su traspuesta de matrices asociadas A y A^T respectivamente, tomando $B_E = \{\bar{e}_1, \dots, \bar{e}_n\}$, $B_F = \{\bar{v}_1, \dots, \bar{v}_m\}$, $B_{E^*} = \{\bar{e}_1^*, \dots, \bar{e}_n^*\}$ y $B_{F^*} = \{\bar{v}_1^*, \dots, \bar{v}_m^*\}$, se satisface que:

$$f(\bar{e}_j) = f_j = (a_{1,j}, \dots, a_{m,j})_{B_F} = \sum_{i=1}^{m} a_{i,j}\bar{v}_i \qquad i = \{1, \dots, m\} \qquad A = \underset{(\;f_1 \;|\; \cdots \;|\; f_n\;)}{\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix}}$$

Esto significa que la imagen de un vector $\bar{e}_j \in B_E$ respecto de $f : E \mapsto F$ es una combinación lineal de los vectores de la base de F, B_F, es decir, que $f(\bar{e}_j) \in F$.

Para probar que A^T es la matriz de $f^t : F^* \mapsto E^*$, debemos probar que $f^t(\bar{v}_i^*) = A^T \cdot \bar{v}_i^*$, es decir, que se cumple la igualdad:

$$f^t(\bar{v}_i^*) = \sum_{j=1}^{n} a_{i,j}\bar{e}_j^* \qquad j = \{1, \dots, n\}$$

Como $f^t[\bar{v}_i] = ev[f^t(\bar{v}_i)] \in E^*$ y $\sum_{j=1}^{n} a_{i,j}\bar{e}_j^* = ev[\sum_{j=1}^{n} a_{i,j}\bar{e}_j^*] \in E^*$ son aplicaciciones lineales de E en $\mathbb{R}$, para ver que son iguales, basta comprobar que coinciden sobre una base de E (comprobar que evaluando el mismo vector en la misma base en las dos aplicaciones, ambas arrojan el mismo escalar de $\mathbb{R}$):

$$\text{Sea } B' \text{ una base de } E: \qquad \bar{e}_h \in B' \qquad h = \{1, \dots, n\}$$

Evaluamos el vector e_h en $f^t[\bar{v}_i^*]$:

$$f^t[\bar{v}_i^*](\bar{e}_h) =^{\text{def. } f^t} (\bar{v}_i^* \circ f)(\bar{e}_h) = \bar{v}_i^*(f(\bar{e}_h)) = \bar{v}_i^*(a_{1,h}\bar{v}_1 + \dots + a_{m,h}\bar{v}_m) =$$
$$= a_{1,h}\bar{v}_i^*\bar{v}_1 + \dots + a_{m,h}\bar{v}_i^*\bar{v}_m =^* a_{i,h}$$

Nota.* $\quad \bar{v}_i^*\bar{v}_j = \begin{matrix} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{matrix} \qquad a_{i,h}\bar{v}_i^*\bar{v}_j = \begin{matrix} a_{i,h}, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{matrix}$

Evaluamos el vector e_h en $\sum_{j=1}^{n} a_{i,j}\bar{e}_j^*$:

$$\left(\sum_{j=1}^{n} a_{i,j}\bar{e}_j^*\right)(\bar{e}_h) = a_{i,1}\bar{e}_1^*(\bar{e}_h) + \dots + a_{i,n}\bar{e}_n^*(\bar{e}_h) =^* a_{i,h}$$

Nota.* $\quad \bar{e}_j^*\bar{e}_h = \begin{matrix} 1, & \text{si } j = h \\ 0, & \text{si } j \neq h \end{matrix} \qquad a_{i,j}\bar{e}_j^*\bar{e}_h = \begin{matrix} a_{i,h}, & \text{si } j = h \\ 0, & \text{si } j \neq h \end{matrix}$

Como el escalar $a_{i,h}$ coincide, podemos concluir que la matriz de la aplicación f^t es A^T, es decir, $M(f)^T$.

$\square$

2.7.4. Calcular la base del espacio dual utilizando la aplicación traspuesta

Un cambio de base puede entenderse como una aplicación lineal. De esta forma, conociendo la aplicación de cambio de una base B' a una base B (siendo B y B' bases de un espacio vectorial E) podemos emplear la aplicación traspuesta para cambiar el espacio dual $E*$ de la base B^* a la base B'^*.

Dado que siempre conocemos la base canónica dual Bc^* de un espacio vectorial E expresado en su base canónica Bc (ya que Bc^* y Bc son iguales), la aplicación traspuesta nos ofrece un método alternativo al de los sistemas de ecuaciones para calcular la base dual B^* de una base B cualquiera de E.

Sabemos que la matriz $M_{Bc \leftarrow B}$ de cambio de base de B a Bc se construye colocando los vectores de la base de B en columnas. Por tanto, su aplicación traspuesta tiene como matriz asociada a $(M_{Bc \leftarrow B})^T$. Para hallar los vectores de la base dual B^* basta con muliplicar cada uno de los vectores de Bc^* por $M_{B^* \leftarrow Bc^*} = (M_{Bc \leftarrow B})^T$.

Ejemplo
Calcular la base B^* del espacio vectorial $\mathbb{R}^3$ en base B.

$B = \{(2, 2, 1), (3, 4, 2), (1, 2, 1)\}$

1. Obtenemos la matriz de cambio de base de $Bc(\mathbb{R}^{3*})$ a B^*:

$$M_{B^* \leftarrow Bc^*} = (M_{Bc \leftarrow B})^T = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}^T = \begin{pmatrix} 2 & 2 & 1 \\ 3 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

2. Evaluamos cada uno de los vectores de la base canónica dual $Bc(\mathbb{R}^{3*})$ en $M_{B^* \leftarrow Bc^*}$:

$$Bc(\mathbb{R}^{3*}) = Bc(\mathbb{R}^3) = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

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

3. Obtenemos la base B^*:

$$B^* = \{(2, 3, 1), (2, 4, 2), (1, 2, 1)\}$$

2.8. Espacio doble dual

El espacio doble dual (o dual-dual) asociado a un espacio vectorial (E) es el espacio dual de su espacio dual E^*. Es decir, el espacio de los homomorfismos que toman vectores del espacio dual E^* y los transforman en escalares de $\mathbb{R}$. Se denota por E^{**}.

$$E^{**} = Hom(E^*, \mathbb{R})$$

2.8.1. Noción de evaluador del espacio doble dual

Sea E un espacio vectorial, al igual que disponemos de una noción de "evaluador" para el espacio dual E^*, también existe una noción de "evaluador" asociada al espacio doble dual E^{**}. En este caso cada vector $\bar{e} \in E$ define un una aplicación de E^* en $\mathbb{R}$, que denotamos por $\phi[\bar{e}]$, y que transforma un vector cualquiera $\bar{v}^* \in E^*$ en $\lambda = \bar{v}^* \cdot \bar{e}$.

$$\text{Sean:} \quad \bar{v}^* \text{ (genérico) } \in E^* \quad y \quad \bar{e} \text{ (fijo) } \in E$$

$$\phi[\bar{e}] : E^* \longmapsto \mathbb{R}$$

$$\phi[\bar{e}](\bar{v}^*) = \bar{v}^* \cdot \bar{e} = \lambda$$

DIAGRAMA

$$\begin{tikzcd}
\bar{e} \in E \arrow[rr, "ev[\bar{e}^*](\bar{e})"] & & \mathbb{R} \\
& \bar{e}^* \in E^* \arrow[ur, bend right, "\phi[\bar{e}](\bar{e}^*)"] & \\
& & \bar{e}^{**} \in E^{**}
\end{tikzcd}$$

Nota. $\quad ev[\bar{e}^*] \sim e^* \in E^* \qquad \phi[\bar{e}] \sim e^{**} \in E^{**}$

2.8.2. Base del espacio doble dual

Sea $\{e_1, \dots, e_n\}$ una base de un espacio vectorial E y sea $\{e_1^*, \dots, e_n^*\}$ la base de su espacio dual E^*. Como E^{**} es el espacio dual de E^* su base $\{e_1^{**}, \dots, e_n^{**}\}$ viene determinada por las siguientes condiciones:

$$\bar{e}_j^{**}(\bar{e}_i^*) \; = \delta_{j,i} = \begin{cases} 1 & \text{si} j = i \\ 0 & \text{si} j \neq i \end{cases}$$

La base de E^{**} depende de la base de E^*, que a su vez depende de la base de E. Recurrir al "evaluador" $\phi[\bar{e}_j] \in E^{**}$ nos permite fijar un vector $\bar{e}_j$ de la base de E y evaluar en su correspondiente vector $\phi[\bar{e}_j] = \bar{e}_j^{**} \in E^{**}$ un vector $\bar{e}^*$ cualquiera de E^* (en el siguiente desarrollo, evaluaremos en $\phi[\bar{e}_j]$ los vectores $\bar{e}_i^*$ de la base dual). Tomamos $\bar{e}_j^{**} = \phi[\bar{e}_j]$ y sustituimos en la expresión anterior, obteniendo:

$$\delta_{j,i} = \bar{e}_j^{**}(\bar{e}_i^*) \Rightarrow \phi[\bar{e}_j](\bar{e}_i^*) = \bar{e}_i^* \cdot \bar{e}_j = \bar{e}_j \cdot \bar{e}_i^* = \bar{e}_j(\bar{e}_i^*) \;\; = \delta_{j,i} = \begin{cases} 1 & \text{si} j = i \\ 0 & \text{si} j \neq i \end{cases}$$

$$\bar{e}_j^{**}\cancel{(\bar{e}_i^*)} = \bar{e}_j\cancel{(\bar{e}_i^*)} \quad \longrightarrow \quad \bar{e}_j^{**} = \bar{e}_j \quad \longrightarrow \quad \{\bar{e}_j^{**}\}_{j=1}^n = \{\bar{e}_j\}_{j=1}^n \quad \longrightarrow \quad B_E = B_{E^{**}}$$

Por lo que podemos concluir que los vectores de las bases de un espacio vectorial E y de su espacio doble dual E^{**} son iguales.

Analíticamente, podemos asegurarnos de que esta relación entre las bases de E y E^{**} tiene sentido comprobando que se cumple la igualdad $\left[\phi[\bar{e}](\bar{e}^*) = \bar{e}^* \cdot \bar{e} = \lambda = \bar{e} \cdot \bar{e}^*\right]$:

$$\text{Sean:} \quad \phi[\bar{e}](\bar{e}^*) \in E^{**} \qquad \bar{e}^* = (e_1^*, \dots, e_n^*) \in E^* \qquad \bar{e} = (e_1, \dots, e_n) \in E$$

$$\phi[\bar{e}](\bar{e}^*) = \bar{e}^* \cdot \bar{e} = (e_1^*, \dots, e_n^*) \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = (e_1^* \cdot e_1 + \dots + e_n^* \cdot e_n) = \lambda = (e_1 \cdot e_1^* + \dots + e_n \cdot e_n^*) = (e_1, \dots, e_n) \begin{pmatrix} e_1^* \\ \vdots \\ e_n^* \end{pmatrix} = \bar{e} \cdot \bar{v}^*$$

2.8.3. Teorema de reflexividad

Si analizamos la construcción del evaluador $\phi[\bar{e}] \in E^{**}$ podemos ver que mantiene una relación directa con el vector fijado $\bar{e} \in E$. De esta forma, podemos definir la aplicación ϕ que toma vectores $\bar{e} \in E$ y los lleva a su evaluador $\phi[\bar{e}] \in E^{**}$:

$$\text{Sean:} \quad \bar{e} \in E \quad y \quad \phi[\bar{e}] \in E^{**}$$

$$\phi : E \longmapsto E^{**}$$

$$\phi(\bar{e}) = \phi[\bar{e}]$$

DIAGRAMA

$$\begin{tikzcd}
\bar{e} \in E \arrow[rr] \arrow[dr, bend right, "\phi"'] & ev[\bar{e}^*] \in E^* \arrow[r, bend right] & \mathbb{R} \\
& \phi[\bar{e}] \in E^{**} &
\end{tikzcd}$$

TEOREMA DE REFLEXIVIDAD
La aplicación $\phi : E \longmapsto E^{**}$ que toma vectores $\bar{e} \in E$ y los lleva a su evaluador $\phi[\bar{e}] \in E^{**}$ es un isomorfismo canónico.

Una isomorfismo canónico es un isomorfismo (aplicación lineal biyectiva) que no depende de la base del espacio preimagen. Es decir, que un cambio en la base del dominio (espacio vectorial preimagen) no requiere un cambio en la base del isomorfismo para que los vectores de la imagen se sigan correspondiendo con los que se correspondían de la preimagen antes del cambio de base.

DEMOSTRACIÓN DE QUE ϕ ES ISOMORFISMO
Un isomorfismo es una aplicación lineal biyectiva (inyectiva + sobreyectiva):

1. Probar que la aplicación ϕ es lineal:

   Veamos que dados dos vectores $\bar{u}, \bar{v} \in E$, para cualquiera que sea $\bar{w}^* \in E^*$ tenemos:

   $$\phi[\lambda\bar{u} + \mu\bar{v}] = \bar{w}^* \cdot (\lambda\bar{u} + \mu\bar{v}) = \lambda(\bar{w}^* \cdot \bar{u}) + \mu(\bar{w}^* \cdot \bar{v}) = \lambda\phi[\bar{u}](\bar{w}^*) + \mu\phi[\bar{v}](\bar{w}^*) = (\lambda\phi[\bar{u}] + \mu\phi[\bar{v}])(\bar{w}^*)$$

   Lo cual prueba que las aplicaciones $\phi(\lambda\bar{u} + \mu\bar{v})$ y $\lambda\phi(\bar{u}) + \mu\phi(\bar{v})$ son iguales, por lo que ϕ cumple el teorema de caracterización, y, por tanto, es lineal.

2. Probar que la aplicación ϕ es inyectiva:
   $\forall \bar{e} \in E - \{\bar{0}\}, \;\; \exists \bar{w}^* \in E^*$ tal que $\bar{w}^* \cdot \bar{e} = \phi[e](\bar{w}^*) \neq 0$, por lo que podemos decir que:

   $$\forall \bar{e} \in E \qquad \phi(\bar{e}) = \phi[e] \neq \bar{0}$$

3. Probar que la aplicación ϕ es sobreyectiva:
   Si E y E^{**} tienen la misma dimensión y a cada vector $\bar{e} \in E$ le corresponde un vector $\phi[\bar{e}] \in E^{**}$, a cada vector $\phi[\bar{e}] \in E^{**}$ le debe corresponder un vector $\bar{e} \in E$.

DEMOSTRACIÓN DE QUE ϕ ES CANÓNICO
Esta característica del teorema es algo más sutil. Que un isomorfismo f sea canónico, significa que un cambio en la base del espacio vectorial del dominio (espacio vectorial preimagen) no requiere un cambio en la base del isomorfismo para que las imágenes de los vectores (del dominio) se sigan correspondiendo con los vectores que deben corresponderse.

Esta es una cualidad bastante poco frecuente, ya que un cambio de la base de la preimagen siempre implica un cambio en los vectores generados por la aplicación (si la base de la aplicación o la base del subespacio imagen no varían también). Sin embargo, en el caso de la aplicación $\phi : E \mapsto E^{**}$ el concepto de "correspondencia correcta" está desligado del concepto de "correspondencia con el mismo vector que antes del cambio de base".

En el caso del espacio del espacio doble dual E^{**}, su base depende directamente de la base de E (en concreto es la misma), ya que un vector $\phi[\bar{e}_i] \in E^{**}$ debe anular al anulador $ev[\bar{e}_i^*] \in E^*$ que anula a $\bar{e}_i \in E$, y la única forma de que esto suceda es que los coeficientes de $\bar{e}_i^{**}$ coincidan exactamente con los de $\bar{e}_i$.

De esta forma, un cambio en la base de E trae implícito un cambio en la base de E^{**}, por lo que la base de ambos espacios es siempre exactamente la misma, independientemente de la base en la que se exprese E, y por tanto, siempre se mantiene la "correspondencia correcta" entre vectores. A efectos prácticos, esto significa que el espacio E^{**} es una copia exacta de E, aunque formalmente no sea el mismo.

2.9. Anulador

El anulador de un subespacio vectorial $V \in E$, que se denota por $V^{\perp}$, es el conjunto de formas lineales $\bar{a}$ que transforman cualquier vector de dicho subespacio vectorial en 0. Véase que, como $0 \in \mathbb{R}$, el anulador de V está contenido en el espacio dual E^*.

$$\bar{a} \in V^{\perp} : V \longmapsto 0 \qquad V^{\perp} \in E^*$$

2.9.1. Relación entre anulador y ecuadiones implícitas

Una evaluación de un vector genérico $\bar{x} = (x, \ldots, z)$ de V en un vector $\bar{a} = (a_1, \ldots, a_n)$ de su anulador $V^{\perp}$ siempre toma la siguiente forma:

$$(a_1, \ldots, a_n) \begin{pmatrix} x \\ \vdots \\ z \end{pmatrix} = (a_1 \cdot x + \cdots + a_n \cdot z) = 0$$

La expresión $(a_1 \cdot x + \cdots + a_n \cdot z = 0)$, nos indica una combinación lineal entre coordenadas que se cumple para todos los vectores contenidos en V y cuyo resultado es 0. De esta forma, la expresión anterior, coincide con la estructura una ecuación implícita del subespacio V. Es por esto que decimos, que existe una relación de equivalencia entre los coeficientes de las ecuaciones implícitas de V y las coordenadas de todos los vectores $\bar{a}_i$ de una base cualquiera de su anulador.

$$\text{Base Anulador } (V^{\perp}) \quad \longleftrightarrow \quad \text{Ecuaciones implícitas } V$$

De manera equivalente, existe una relación de equivalencia entre las ecuaciones implícitas del anulador de V y las coordenadas de los vectores de la base de V:

$$\text{Base Subespacio } (V) \quad \longleftrightarrow \quad \text{Ecuaciones implícitas } V^{\perp}$$

Como el subespacio anulador de V está contenido en E^*, el subespacio anulador del anulador de V está contenido en E^{**}, por lo que su base se corresponde con la del espacio V (Por el teorema de reflexividad). Al coincidir los vectores de la base del anulador con los coeficientes de unas implícitas V, haciendo el anulador de dichos vectores (el anulador del anulador) podemos obtener los vectores de una base de V (ya que una base del anulador del anulador de V coincide con una base de V). De esta forma, podemos pasar de ecuaciones paramétricas a implícitas de un subespacio vectorial V y viceversa simplemente calculando los anuladores.

DIAGRAMA

$$\begin{tikzcd}
V \arrow[d, bend right, leftrightarrow, "An."'] \arrow[dr, leftrightarrow] & Imp.V \arrow[d, bend left, leftrightarrow, "An."] \arrow[dl, leftrightarrow] \\
V^{\perp} & Imp.V^{\perp}
\end{tikzcd}$$

2.9.2. Relación entre anulador y kernel

Veamos que, la forma de la evaluación de un vector genérico $\bar{x} \in V$ en un vector $\bar{a}$ de su anulador $V^{\perp}$, además de coincidir con la expresión de una ecuación implícita de V, también nos da una ecuación que describe aquellos vectores $\bar{x}$ que al ser evaluados en la forma lineal $\bar{a}$ resultan en 0, es decir, una ecuación implícita del kernel de $\bar{a}$.

$$(a_1, \ldots, a_n) \begin{pmatrix} x \\ \vdots \\ z \end{pmatrix} = (a_1 \cdot x + \cdots + a_n \cdot z) = 0$$

De la misma forma, si tomanos todos los vectores $\bar{a}_i$ de una base de $V^{\perp}$ y formamos con ellos una matriz (que podemos entender como la matriz asociada a una aplicación lineal f), la expresión resultante de multiplicar la matriz por un vector genérico $\bar{x}$ representa las ecuaciones implícitas del kernel de f:

$$\begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ & \vdots & \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix} \begin{pmatrix} x \\ \vdots \\ z \end{pmatrix} = \begin{pmatrix} a_{1,1} \cdot x & \ldots & a_{1,n} \cdot z \\ & \vdots & \\ a_{m,1} \cdot x & \ldots & a_{m,n} \cdot z \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Es por esto que calcular el anulador de un subespacio vectorial V es equivalente a calcular el kernel de la aplicación cuya matriz asociada es el resultado de apilar horizontalmente los vectores de una base cualquiera de dicho espacio vectorial.

Nota. Esta relación entre anulador y kernel es particularmente útil cuando se cuenta con un ordenador ya que muchos programas matemáticos disponen de funciones que permiten calcular un anulador automáticamente.

Ejemplo
Calcular el anulador del subespacio vectorial dado V usando el kernel de la aplicación correspondiente:

$$V = L\{(1, 0, 2), (3, 4, 1)\}$$

1. Calculando el kernel de la matriz correspondiente obtenemos las ecuaciones implícitas del subespacio anulador $V^{\perp}$:

$$ker \begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \longrightarrow \left\{ \begin{array}{lcl} x_1 + 2x_3 & = & 0 \\ 3x_1 + 4x_2 + x_3 & = & 0 \end{array} \right.$$

2. A partir de las ecuaciones implícitas obtenemos una base generadora del subespacio anulador:

$$\left\{ \begin{array}{lcl} x_1 + 2x_3 & = & 0 \\ 3x_1 + 4x_2 + x_3 & = & 0 \end{array} \right. \longrightarrow V^{\perp} = L\{(-2, {}^5/_4, 1)\}$$

Ejemplo
Dados una aplicación lineal f y un subespacio V, hallar una base del subespacio $f^{-1}(V)$:

$$M(f) = \begin{pmatrix} 4 & 1 & 0 \\ -4 & 0 & 2 \\ -2 & -1 & 0 \end{pmatrix} \qquad V = L\{(1, 1, 1), (1, 2, 3)\}$$

DESARROLLO EMPLEADO PARA LA RESOLUCIÓN DEL PROBLEMA

Vamos a aplicar el siguiente desarrollo (empleando teoría de anuladores y aplicación traspuesta):

$$\begin{tikzcd}
V \arrow[r, dash] & V^{\perp} \in V^* \arrow[r] \arrow[d, "f^t"] & 0 \\
f^{-1}(V) \arrow[u, "f"] \arrow[r, dash] & \left(f^{-1}(V)\right)^{\perp} \arrow[ur, bend right] &
\end{tikzcd} \qquad f^{-1}(V) = \left(\left(f^{-1}(V)\right)^{\perp}\right)^{\perp} = \left(f^t V^{\perp}\right)^{\perp}$$

Primero obtenemos $\left(f^{-1}(V)\right)^{\perp}$ simplemente evaluando el anulador de V en la aplicación traspuesta de f. Y después, aprovechándonos de la equivalencia entre ecuaciones implícitas de $f^{-1}(V)$ y generadores de $\left(f^{-1}(V)\right)^{\perp}$, hallamos una base de $f^{-1}(V)$ mediante el anulador de $\left(f^{-1}(V)\right)^{\perp}$.

1. Calculando el kernel correspondiente obtenemos el subespacio anulador $V^{\perp}$:

$$ker \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \longrightarrow \left\{ \begin{array}{lcl} x_1 + x_2 + x_3 & = & 0 \\ x_1 + 2x_2 + 3x_3 & = & 0 \end{array} \right. \longrightarrow V^{\perp} = L\{(1, -2, 1)\}$$

2. Evaluando el subespacio $V^{\perp}$ en la aplicación traspuesta obtenemos un generador de $\left(f^{-1}(V)\right)^{\perp}$:

$$\begin{pmatrix} 4 & -4 & -2 \\ 1 & 0 & -1 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 14 \\ 0 \\ -4 \end{pmatrix} \longrightarrow \left(f^{-1}(V)\right)^{\perp} = L\{(14, 0, -4)\}$$

3. Calculando el anulador de $\left(f^{-1}(V)\right)^{\perp}$ obtenemos los coeficientes de las implícitas de V:

$$ker \begin{pmatrix} 14 & 0 & -4 \end{pmatrix} \longrightarrow \left\{ \begin{array}{lcl} 14x_1 - 4x_3 & = & 0 \end{array} \right. \longrightarrow \left(f^{-1}(V)\right)^{\perp} = \left\{ \begin{array}{lcl} x_2 & = & 0 \\ \frac{2}{7}x_1 + x_3 & = & 0 \end{array} \right.$$

4. El generador de $f^{-1}(V)$ coincide con los coeficientes de las implíctas de $\left(f^{-1}(V)\right)^{\perp}$:

$$\left(f^{-1}(V)\right)^{\perp} = \left\{ \begin{array}{lcl} x_2 & = & 0 \\ \frac{2}{7}x_1 + x_3 & = & 0 \end{array} \right. \longrightarrow V^{\perp} = L\{(0, 1, 0), ({}^2/_7, 0, 1)\}$$

2.9.3. Propiedades del anulador

1. Para todo sistema $S \in E$, $S^{\perp}$ es un subespacio vectorial de E^*
2. $S^T = L(S)^T$
3. Sea E un espacio vectorial de dimensión máxima, tenemos $\bar{0}^{\perp} = E$ y $E^{\perp} = \bar{0} \in E^*$
4. Si E tiene dimensión finita y V es un subespacio vectorial:

$$dim(V^{\perp}) = dim(E) - dim(V)$$

5. Si E tiene dimensión finita, identificando E con E^{**}, tenemos que $(V^{\perp})^{\perp} = V$
6. Si $V, W \in E$ son subespacios vectoriales, entonces:

$$(V + W)^{\perp} = V^T \cap W^T \qquad (V \cap W)^{\perp} = V^T + W^T$$

7. Si $f : E \mapsto F$ es lineal, entonces $kerf^t = Imf^t$, e $Imf^t = (kerf^t)$. En particular, $rg(f^t) = rg(f)$

Capítulo 3

Diagonalización

3.1. Definición de matrices semejantes

Dos matrices A y B son semejantes si existe una matriz P regular tal que:

$$A = P^{-1}BP$$

A P se la denomina **matriz de paso**.

Nota. Véase que si: $A = P^{-1}BP$ entonces: $B = PAP^{-1}$.

Ejemplo

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

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

Nota. Todas las matrices asociadas a un endomorfismo $f : E \longmapsto E$ en distintas bases son semejantes.

3.2. Propiedades de matrices semejantes

1. Si A y B son dos matrices semjantes, entonces $|A| = |B|$.

   DEMOSTRACIÓN

   $$\text{Nota.} \quad |M^{-1}| = \frac{1}{|M|}$$

   $$B = P^{-1}AP \quad \rightarrow \quad |B| = |P^{-1}||A||P| \quad \rightarrow \quad |B| = \cancelto{1}{|P^{-1}||P|}\,|A| \quad \rightarrow \quad |B| = |A|$$

   $\square$

   ---

2. Si A y B son matrices semejantes, entonces A^n y B^n también son semejantes con la misma matriz de cambio de paso.

   DEMOSTRACIÓN

   $$B^n = (P^{-1}AP)^n \quad \rightarrow \quad B \cdot B \cdots B \;=\; (P^{-1}AP) \cdots (P^{-1}AP) \;=\; P^{-1}A\,\cancelto{I}{[PP^{-1}]}\,A \cdots AP$$

   $$\rightarrow \quad B \cdot B \cdots B = P^{-1}AIAI \cdots IAIAP \quad \rightarrow \quad B^n = P^{-1}A^nP$$

   $\square$

   ---

3. Si A es semejante a A' y B es semejante a B' con la misma matriz de paso P, entonces $\lambda A + \mu B$ es semejante a $\lambda A' + \mu B'$ con esa misma matriz de paso.

   DEMOSTRACIÓN

   $$\lambda A + \mu B \;=\; \lambda P^{-1}AP + \mu P^{-1}BP \;=\; P^{-1}\lambda AP + P^{-1}\mu BP \;=\; P^{-1}(\lambda A + \mu B)P$$

   $\square$

3.3. Forma canónica de Jordan

(No se profundiza en lógica detrás de la construcción de las matrices de paso.)

La forma canónica de Jordan es una expresión de matrices semejantes en donde la matriz de paso (P) está construida a partir de los autovectores de la matriz original (A) colocados en columnas, y donde la matriz semejante (J) es una matriz diagonal (En el caso estándar de que (A) sea estrictamente diagonalizable) compuesta por los autovalores de la matriz original.

$$A = PJP^{-1} \quad \longrightarrow \quad J = P^{-1}AP$$

3.3.1. Autovectores, autovalores

Un **autovector** es un vector que al ser multiplicado por la matriz original A da como resultado un vector múltiplo de sí mismo.

$$A\bar{u} = \lambda\bar{u}$$

Podemos deducir que todo autovector $\bar{u}_i$ está contenido en el kernel de $(A - \lambda_i I)$, denominado "primer autoespacio asociado a λ_i", y que se denota por (N_{1,λ_i}).

$$A\bar{u}_i = \lambda_i\bar{u}_i \quad \to \quad A\bar{u}_i - \lambda_i\bar{u}_i = \bar{0} \quad \to \quad (A-\lambda_i I)\bar{u}_i = \bar{0} \quad \Longleftrightarrow \quad \bar{u}_i \subseteq Ker(A-\lambda_i I)$$

Al (λ_i) asociado al autovector $\bar{u}_i$ se lo denomina **autovalor** asociado al autovector. Todos los autovectores de una matriz linealmente independientes entre sí tienen exactamente un autovalor asociado, por tanto, una matriz puede tener hasta tantos autovalores como su dimensión.

Como $\bar{u}_i \subseteq Ker(A-\lambda_i I)$, para hallar los autovalores de la matriz A basta con resolver la ecuación:

$$|(A-\lambda_i I)| = 0$$

Ya que:

$$(A-\lambda_i I)\bar{u}_i \sim f(\bar{u}_i) \qquad \text{Y si: } kerf \neq \{\bar{0}\} \text{ (Porque } \bar{u}_i \in kerf) \;\to\; rango(M(f)) = 0 \;\to\; |M(f)| = 0$$

Nota. Multiplicar el vector $(\bar{u}_i)$ por una matriz $(A-\lambda I)$ es equivalente a evaluar en una aplicación lineal f.

Al resolver la ecuación, se obtienen los autovalores de la matriz. Si alguno de ellos está repetido n-veces se dice que dicho autovalor λ_i tiene multiplicidad n.

Nota. Véase que si 0 es autovalor de A, $\exists\bar{u}_0$ tal que $\bar{u}_0 \in ker(A-0\cdot I) \;\to\; (A-0\cdot I)\bar{u}_0 = A\bar{u}_0 = \bar{0}$, es decir, que $rango(A) \neq$ máx y $|A| = 0$.

Ejemplo

Hallar los autovalores de la matriz $A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix}$

Resolvemos la ecuación $|A - \lambda I| = 0$.

$$\begin{vmatrix} 2-\lambda & 1 & 1 \\ 2 & 3-\lambda & 2 \\ 3 & 3 & 4-\lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & 1 & 1 \\ -1+\lambda & 3-\lambda & 2 \\ 0 & 3 & 4-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 1 & 1 & 1 \\ -1 & 3-\lambda & 2 \\ 0 & 3 & 4-\lambda \end{vmatrix} =$$

$$= (1-\lambda)\begin{vmatrix} 1 & 1 & 1 \\ 0 & 4-\lambda & 3 \\ 0 & 3 & 4-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 4-\lambda & 3 \\ 3 & 4-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 7-\lambda & 7-\lambda \\ 3 & 4-\lambda \end{vmatrix} =$$

$$= (1-\lambda)(7-\lambda)\begin{vmatrix} 1 & 1 \\ 3 & 4-\lambda \end{vmatrix} = \frac{1}{3}(1-\lambda)(7-\lambda)\begin{vmatrix} 3 & 3 \\ 3 & 4-\lambda \end{vmatrix} = \frac{1}{3}(1-\lambda)(7-\lambda)\begin{vmatrix} 3 & 3 \\ 0 & 1-\lambda \end{vmatrix} =$$

$$= (1-\lambda)^2(7-\lambda) = 0 \qquad \text{autovalores: } \begin{cases} \lambda = 1 \text{ con multiplicidad 2} \\ \lambda = 7 \text{ con multiplicidad 1} \end{cases}$$

3.3.2. Autoespacio

Cuando un autovalor λ_i de una matriz A tiene un único autovector asociado (Es decir, existe un único autovector $\bar{u}_i$ tal que $A\bar{u}_i = \lambda_i\bar{u}_i$ por lo que la multiplicidad de $\lambda_i = 1$), entonces, dicho autovector $\bar{u}_i$ siempre está contenido en el **primer autoespacio asociado** $(ker(A-\lambda_i) = N_{1,\lambda_i})$.

$$\bar{u}_i \subseteq Ker(A-\lambda_i I)$$

Sin embargo, cuando la multiplicidad de un autovalor λ_i es mayor que 1, este tiene más de un único vector $\bar{u}_i$ asociado. Cuando esto sucede se dan dos casos posibles:

1. Todos los autovectores asociados a λ_i están contenidos en el primer autoespacio asociado (N_{1,λ_i}).

   (Si la multiplicidad del autovalor coincide con la dimensión del primer autoespacio)

   En este caso la matriz A es **estrictamente diagonalizable**, es decir, que J es una matriz diagonal. El autoespacio asociado es un espacio vectorial compuesto por uno o varios autovectores $\bar{u}_i$ de la matriz. Para hallarlos podemos extraer unas ecuaciones implícitas del autoespacio $(Ker(A-\lambda_i) = N_{1,\lambda_i})$.

   Determinar las ecuaciones implícitas del primer autoespacio:

   $$\text{Sean:} \quad \bar{u}_i = (x,\ldots,z) \quad \text{y} \quad (A-\lambda_i I) = \begin{pmatrix} a_{1,1}-\lambda_i & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n}-\lambda_i \end{pmatrix}$$

   $$(A-\lambda_i I)\begin{pmatrix} x \\ \vdots \\ z \end{pmatrix} = \bar{0} \quad \longrightarrow \quad \begin{cases} (a_{1,1}-\lambda_i)x + \cdots + (a_{1,n})z = 0 \\ \vdots \\ (a_{n,1})x + \cdots + (a_{n,n}-\lambda_i)z = 0 \end{cases}$$

   Por el **teorema de las dimensiones** sabemos que la dimensión del primer autoespacio asociado es igual a la diferencia entre la dimensión de la matriz A (equivalente al espacio vectorial de dimensión máxima) y el número de ecuaciones implícitas resultantes (que coincide con el rango de $(A-\lambda_i I)$).

   $$dim(N_{1,\lambda_i}) = dim(A) - rango(A-\lambda_i I)$$

   Además, sabemos que, para estar en el caso 1 (Que todos los autovectores estén contenidos en el primer autoespacio), la dimensión del primer autoespacio N_{1,λ_i} (Que es lo que nos indica cuantos autovectores linealmente independientes podemos encontrar en el primer autoespacio asociado a λ_i) debe coincidir con la multiplicidad del autovalor λ_i (Que es lo que nos indica cuantos vectores linealmente independientes asociados a λ_i tiene la matriz).

   $$dim(N_{1,\lambda_i}) = dim(ker(A-\lambda_i)) = multiplicidad\,\lambda_i$$

2. No todos los vectores asociados a λ_i están contenidos en el primer autoespacio asociado.

   (Si la multiplicidad del autovalor NO coincide con la dimensión del primer autoespacio)

   En este caso la matriz A **no es estrictamente diagonalizable**, es decir, que J no es una matriz completamente diagonal. Además, es necesario profundizar hasta el k-ésimo autoespacio asociado $N_{k,\lambda_i} con\ k \in \mathbb{N} > 1$ (También llamado **autoespacio generalizado**) para encontrar todos los vectores asociados a λ_i.

   Sabemos que la dimensión del último autoespacio N_{k,λ_i} (Que en este caso no coincide con el primer autoespacio N_{1,λ_i}) debe coincidir con la multiplicidad del autovalor λ_i (que es lo que nos indica cuantos autovectores linealmente independientes asociados a λ_i tiene la matriz).

   $$dim(N_{k,\lambda_i}) = dim(ker((A-\lambda_i)^k)) = multiplicidad\,\lambda_i$$

   El conjunto de los k autoespacios asociados a un autovalor λ_i, denotado por $\{N_{j,\lambda_i}\}_{j=1}^k$, es el conjunto de los kernels de las potencias de $(A-\lambda_i I)$:

   $$\{Ker((A-\lambda_i I)^j)\}_{j=1}^k \quad \forall j \in \{1,\ldots,k\}$$

   Donde el autoespacio generalizado $Ker(A-\lambda_i I)^K$ es el primer autoespacio de dimensión igual a la multiplicidad de λ_i.

   Nota. Al hacer la potencia de una matriz M con $|M| = 0$ esta siempre reduce su dimensión. Como la matriz $(A-\lambda_i I)^j$ representa los coeficientes de las ecuaciones implícitas del j-ésimo autoespacio asociado a λ_i, N_{j,λ_i}, por el **teorema de dimensiones** [dim(N_{j,λ_i}) = dim(A) - nº ecs. imp. de N_{j,λ_i}], al hacer cualquier potencia de $(A-\lambda_i I)$, la dimensión del autoespacio asociado N_{j,λ_i} cuyas implícitas represente dicha potencia siempre será mayor a la dimensión del primer autoespacio asociado N_{1,λ_i}. De esta forma, eventualmente, algún autoespacio asociado N_{k,λ_i} alcanzará la dimensión necesaria para contener a todos los vectores asociados a λ_i.

   Nótese que los vectores asociados a λ_i contenidos en $ker((A-\lambda_i I)^j)$ con $j \in \mathbb{N} > 1$ realmente no son autovectores, ya que la consecuencia de ser autovector $(A\bar{u} = \lambda\bar{u})$ es estar contenido en $ker(A-\lambda I)$.

3.3.3. Forma canónica de Jordan de una matriz estrictamente diagonalizable

Ejemplo

Hallar la forma canónica de Jordan de la matriz $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

1. Buscar los autovalores de A:

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

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

   Los autovalores son: $\begin{cases} \lambda = 0 \text{ (mult. = 1)} \\ \lambda = 2 \text{ (mult. = 2)} \end{cases}$

2. Buscar los autovectores:

   - $N_{1,2}$

     $$\begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longrightarrow \{x - y = 0 \longrightarrow N_{1,2} = L\{(1,1,0),(0,0,1)\}$$

   - $N_{1,0}$

     $$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longrightarrow \begin{cases} x + y = 0 \\ z = 0 \end{cases} \longrightarrow N_{1,2} = L\{(1,-1,0)\}$$

3. Colocamos los autovectores y lo autovalores como corresponde para construir las matrices J y P de la forma de Jordan:

   Nota. El orden en el que se coloquen los autovalores en la matriz J es irrelevante siempre y cuando sus correspondientes autovectores en la misma posición en la matriz P. Estas son algunas de las combinaciones válidas para las matrices J y P:

   $$J = \left(\begin{array}{c|c|c} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{array}\right) \qquad P = \left(\begin{array}{c|c|c} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)$$

   $$J = \left(\begin{array}{c|c|c} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{array}\right) \qquad P = \left(\begin{array}{c|c|c} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{array}\right)$$

   $$J = \left(\begin{array}{c|c|c} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{array}\right) \qquad P = \left(\begin{array}{c|c|c} 0 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 0 \end{array}\right)$$

3.3.4. Forma canónica de Jordan de una matriz no estrictamente diagonalizable

En el caso de las matrices no estrictamente diagonalizables no contamos con suficientes autovectores contenidos en N_{1,λ_i} como para completar la matriz P. Lo más parecido que podemos tomar a varios autovectores contenidos en $(A-\lambda I)$ es una serie de vectores contenidos en los autoespacios $\{(A-\lambda_i I)^j\}_{j=1}^k$. Para estos espacios k en ningún caso puede superar la multiplicidad de λ_i (por la propiedad de que la potencia de una matriz M con $|M| = 0$ siempre reduce su rango).

El problema de buscar los vectores contenidos en los espacios $\{N_{j,\lambda_i}\}_{j=1}^k$ se encuentra en que el autoespacio N_{j+1,λ_i} siempre contiene a todos los vectores linealmente independientes contenidos en N_{j,λ_i} más, al menos, un vector más. De esta forma, cada vez que se profundiza un nuevo autoespacio N_{j+1,λ_i}, para obtener el o los nuevos vectores, hay que hacer la comprobación de independencia lineal con todos los vectores ya conocidos contenidos en el autoespacio N_{j,λ_i}.

Para evitar este problema nos aprovechamos precisamente de la propiedad de que siempre existe al menos un vector contenido en N_{j+1,λ_i} pero no en N_{j,λ_i}:

Notación: $\bar{u}^{\text{autovalor asociado}}_{\text{posición en la cadena desde el primer vector } \bar{u}_0}$

$$\exists\bar{u}_0^{\lambda_i} \in N_{k,\lambda_i} \backslash\backslash \in N_{k-1,\lambda_i} \quad \text{tq.} \quad (A-\lambda_i I)^k\bar{u}_0^{\lambda_i} = 0 \quad \text{pero} \quad (A-\lambda_i I)^j\bar{u}_0^{\lambda_i} \neq 0 \quad \forall j < k$$

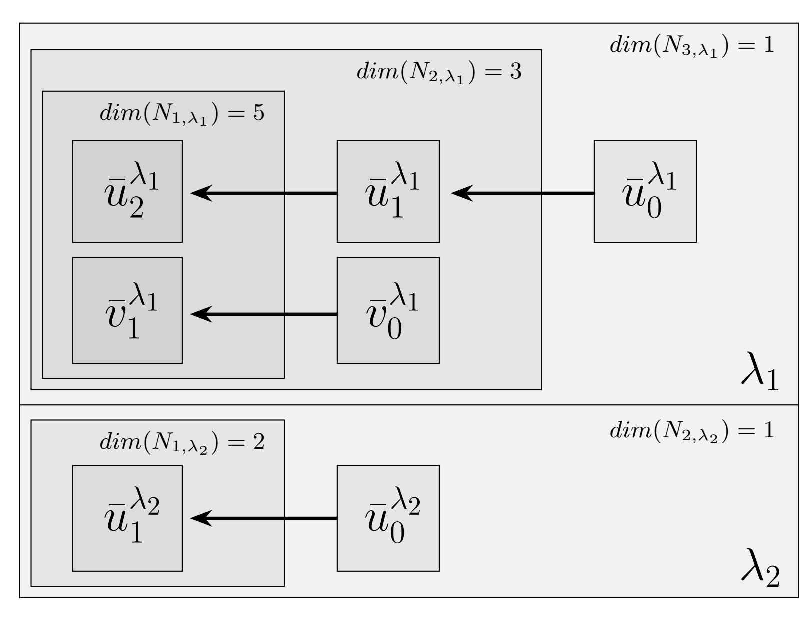
$$\exists\bar{u}_{k-j}^{\lambda_i} \in N_{j,\lambda_i} \backslash\backslash \in N_{j-1,\lambda_i} \quad \text{tq.} \quad (A-\lambda_i I)^j\bar{u}_{k-j}^{\lambda_i} = 0 \quad \text{pero} \quad (A-\lambda_i I)^t\bar{u}_{k-j}^{\lambda_i} \neq 0 \quad \forall t < j$$

La ventaja de este método es que una vez que hemos encontrado el $\bar{u}_0^{\lambda_i} \in N_{k,\lambda_i} \backslash\backslash \in N_{k-1,\lambda_i}$ podemos hallar cualquier $\bar{u}_{k-j}^{\lambda_i} \in N_{j,\lambda_i} \backslash\backslash \in N_{j-1,\lambda_i}$ simplemente multiplicando por la matriz $(A-\lambda_i I)$:

$$\bar{0} = (A-\lambda_i I)^k\bar{u}_0^{\lambda_i} \quad \to \quad \bar{0} = (A-\lambda_i I)^{k-1}(A-\lambda_i I)\bar{u}_0^{\lambda_i} = (A-\lambda_i I)^{k-1}\bar{u}_1^{\lambda_i} \quad \to \quad \cdots \quad \to \quad \bar{0} = (A-\lambda_i I)\bar{u}_{k-1}^{\lambda_i}$$

De esta forma, tal y como se muestra en el esquema, tomando $\bar{u}_0$ y sus sucesivas imágenes podemos obtener en cadena el conjunto de vectores que necesitamos para completar la matriz J.

Simplemente debemos tener en cuenta que es posible que, en algún momento, siguiendo la cadena se llegue a un autoespacio N_{j,λ_i} que pase de incrementar dimensiones de 1 en 1 a incrementarlas de n en n. En este caso habría que buscar más vectores $\{\bar{v}_0^{\lambda_i}, \bar{w}_0^{\lambda_i}, ...\}$ linealmente independientes a $\bar{v}_j^{\lambda_i}$ que inicien nuevas cadenas.



A la hora de construir la matriz J debemos colocar los autovalores pertenecientes a vectores de una misma cadena en bloque, con una diagonal inferior $(\bar{u}_0^{\lambda_i} \to \bar{u}_{k-1}^{\lambda_i})$ o superior $(\bar{u}_{k-1}^{\lambda_i} \to \bar{u}_0^{\lambda_i})$ de 1(s). Los distintos bloques se pueden mezclar indiferentemente entre sí, tal y como sucede con la matriz J de las matrices estrictamente diagonalizables.

Ejemplo

Hallar la forma canónica de Jordan de la matriz $A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 3 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

1. Buscar los autovalores de A:

   $$\begin{vmatrix} 1-\lambda & 2 & 0 \\ -1 & 3-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & 2 & 0 \\ 0 & 3-\lambda & 1 \\ 1-\lambda & 1 & 1-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 1 & 2 & 0 \\ 0 & 3-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} =$$

   $$= (1-\lambda)\begin{vmatrix} 1 & 2 & 0 \\ 0 & 3-\lambda & 1 \\ 0 & -1 & 1-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 3-\lambda & 1 \\ -1 & 1-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 2-\lambda & 2-\lambda \\ -1 & 1-\lambda \end{vmatrix} =$$

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

   Los autovalores son: $\begin{cases} \lambda = 1 \text{ (mult. = 1)} \\ \lambda = 2 \text{ (mult. = 2)} \end{cases}$

2. Buscar los autovectores:

   - $N_{1,1}$

     $$\begin{pmatrix} 0 & 2 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \bar{0} \longrightarrow \begin{cases} x - 2y - z = 0 \\ y = 0 \end{cases} \longrightarrow N_{1,2} = L\{(1,0,1)\}$$

   - $N_{1,2}$

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

     $$1 = \dim N_{1,2} \neq \text{mult.}(\lambda = 2) = 2$$

     Debemos profundizar a $N_{2,2}$ ya que $N_{1,2}$ no tiene dimensión suficiente.

   - $N_{2,2}$

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

     $$\bar{u}_1^2 \quad \to \quad \text{vector contenido en } N_{2,2} \text{ pero no en } N_{1,2}$$

     Por ejemplo: $u_1^2 = (2,0,1)$.

     $$\bar{u}_2^2 \quad \to \quad \text{vector resultante del producto } (A-2I)\,\bar{u}_1^2$$

     $$u_2^2 = \begin{pmatrix} -1 & 2 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ -1 \end{pmatrix}$$

3. Colocamos los autovectores y lo autovalores como corresponde para construir las matrices J y P de la forma de Jordan:

   Nota. Si se coloca en la matriz P el autovector $u_{\lambda_1}^k$ (el del último autoespacio N_{k,λ_i}) a la izquierda y $u_{\lambda_i}^1$ (el del primer autoespacio N_{1,λ_i}) a la derecha, entonces, los 1(s) de la matriz J se colocan en la diagonal inferior.

   $$J = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \end{array}\right) \qquad P = \left(\begin{array}{c|cc} 1 & 2 & -2 \\ 0 & 0 & -1 \\ 1 & 1 & -1 \end{array}\right)$$

   Nota. Si cambiamos el orden de los autovectores de λ_i, entonces, los 1(s) de la matriz J se colocan en la diagonal superior.

   $$J = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{array}\right) \qquad P = \left(\begin{array}{c|cc} 1 & -2 & 2 \\ 0 & -1 & 0 \\ 1 & -1 & 1 \end{array}\right)$$

3.4. Propiedades de autovalores y autovectores

1. A y A^T tienen los mismos autovalores.

2. Si $\bar{u}$ es autovector de A, λA coon $\lambda \in \mathbb{R}$ tiene por autovector $\lambda \bar{u}$.

3. Si λ es autovalor de A, $(A - kI)$ con $k \in \mathbb{R}$ tiene por autovalor $\lambda - k$.

4. Si λ es autovalor de A y A es regular, $\frac{1}{\lambda}$ es autovalor de A^{-1}.

5. Si A y B son matrices semejantes, tienen el mismo polinomio característico y los mismos autovalores.

6. Si λ es autovalor de A, λ^n es autovalor de A^n.

7. $\lambda = 0$ es autovalor de A si y solo si $ker(A) \neq \{\bar{0}\}$, verificándose que $N_{1,0} = ker(A)$.

3.5. Cadenas de Markov

El método de cadenas de Markov sirve para calcular la probabilidad final de un agente de encontrarse en cada uno de los nodos de un grafo con probabilidades fijas de transición de cualquiera de sus nodos hacia cada uno de los demás (con independencia del nodo inicial del agente) tras un número n de transiciones entre nodos con n tendiendo a infinito.

Para ello se utilizan los siguientes elementos:

- La matriz P: indica las probabilidades de pasar de cada uno de los nodos del grafo a otro.

$$P = \begin{pmatrix} P(A_{n+1}|A_n) & \cdots & P(A_{n+1}|Z_n) \\ \vdots & \ddots & \vdots \\ P(Z_{n+1}|A_n) & \cdots & P(Z_{n+1}|Z_n) \end{pmatrix}$$

  Donde $P(X_{n+1}|Y_n)$ es la probabilidad del agente de encontrarse en el nodo X en el paso $n+1$ tras haberse encontrado en el nodo Y en el paso n.

  Nota. La matriz P es una matriz estocaústica (matriz en la que todas las columnas suman 1) ya que es una matriz que representa probabilidades.

- El vector $\bar{p}_0$: indica las probabilidades iniciales del agente de encontrarse en cada uno de los nodos

- El vector $\bar{p}_n$: indica las probabilidades del agente de encontrarse en cada uno de los nodos en el paso n.

De esta forma, cada vez que multiplicamos $\bar{p}_n$ por P obtenemos las prababilidades del agente de encontrarse en cada uno de los nodos en el paso $n+1$.

$$\bar{p}_{n+1} = P \cdot \bar{p}_n \qquad \longrightarrow \qquad \bar{p}_n = P^n \cdot \bar{p}_0$$

$$\begin{pmatrix} P(A_{n+1}|A_n) & \cdots & P(A_{n+1}|Z_n) \\ \vdots & \ddots & \vdots \\ P(Z_{n+1}|A_n) & \cdots & P(Z_{n+1}|Z_n) \end{pmatrix} \begin{pmatrix} P(A_n) \\ \vdots \\ P(Z_n) \end{pmatrix} = \begin{pmatrix} P(A_{n+1}|A_n)P(A_n) + \cdots + P(A_{n+1}|Z_n)P(Z_n) \\ \vdots \\ P(Z_{n+1}|A_n)P(A_n) + \cdots + P(Z_{n+1}|Z_n)P(Z_n) \end{pmatrix} = \begin{pmatrix} P(A_{n+1}) \\ \vdots \\ P(Z_{n+1}) \end{pmatrix}$$

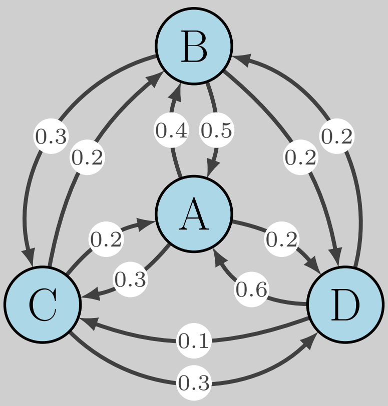
Al tratarse de un grafo de probabilidades invariables, cuando n tiende a infinito, el vector de probabilidades del agente converge. Llamemos $\bar{Q}$ al vector de probabilidades en la convergencia.

$$\bar{Q} = \lim_{n\to\infty} P^n \cdot \bar{p}_0 \qquad \longrightarrow \qquad \bar{Q} = P \cdot \bar{Q}$$

Veamos entonces que el vector $\bar{Q}$ es un autovector asociado al autovalor $\lambda = 1$ de la matriz P (Esto tiene sentido ya que las matrices estocaústicas siempre tienen el autovalor $\lambda = 1$). Por lo que:

$$\bar{Q} \in ker(P - I)$$

Ejemplo
Calcular las probabilidades de que un agente acabe en cada uno de los nodos del siguiente grafo tras un número elevado de iteraciones.



1. Construir la matriz P:

$$P = \begin{pmatrix} P(A_{n+1}|A_n) & P(A_{n+1}|B_n) & P(A_{n+1}|C_n) & P(A_{n+1}|D_n) \\ P(B_{n+1}|A_n) & P(B_{n+1}|B_n) & P(B_{n+1}|C_n) & P(B_{n+1}|D_n) \\ P(C_{n+1}|A_n) & P(C_{n+1}|B_n) & P(C_{n+1}|C_n) & P(C_{n+1}|D_n) \\ P(D_{n+1}|A_n) & P(D_{n+1}|B_n) & P(D_{n+1}|C_n) & P(D_{n+1}|D_n) \end{pmatrix} = \begin{pmatrix} 0{,}1 & 0{,}5 & 0{,}2 & 0{,}6 \\ 0{,}4 & 0{,}0 & 0{,}2 & 0{,}2 \\ 0{,}3 & 0{,}3 & 0{,}3 & 0{,}1 \\ 0{,}2 & 0{,}2 & 0{,}3 & 0{,}1 \end{pmatrix}$$

2. Calcular el $ker(P - I)$:

$$ker(P - I) = ker \begin{pmatrix} -0{,}9 & 0{,}5 & 0{,}2 & 0{,}6 \\ 0{,}4 & -1{,}0 & 0{,}2 & 0{,}2 \\ 0{,}3 & 0{,}3 & -0{,}7 & 0{,}1 \\ 0{,}2 & 0{,}2 & 0{,}3 & -0{,}9 \end{pmatrix} = (-0{,}6235, -0{,}4324, -0{,}5103, -0{,}4048)$$

3. Normalizamos el vector de probabilidades (ya que la suma de probabilidades tiene que ser 1):

$$\bar{p}_f = \frac{\bar{Q}}{sum(\bar{Q})} = (0{,}3163, 0{,}2194, 0{,}2589, 0{,}2054)$$

Capítulo 4

Aplicaciones Multilineales y Álgebra Tensorial

4.1. Aplicación Multilineal

Una aplicación es una transformación que toma elementos de uno o varios conjuntos (C_i) dando lugar a elementos de otro u otros conjuntos (G_i).

$$a : C_1 \times \cdots \times C_n \longmapsto G_1 \times \cdots \times G_n$$

Una aplicación multilineal es una aplicación que toma vectores de varios espacios vectoriales E_i distintos y los transforma en un vector de otro espacio vectorial F cumpliendo una serie de características de linealidad en todos ellos.

$$g : E_1 \times E_2 \times \cdots \times E_n \longmapsto F$$

Definición: Si el espacio imagen es $\mathbb{R}$, entonces decimos que la aplicación se trata de una **forma multilineal**.

$$g : E_1 \times E_2 \times \cdots \times E_n \longmapsto \mathbb{R}$$

Nota.	$g : E \mapsto \mathbb{R}$	$\longrightarrow$	Forma lineal
	$g : E_1 \times E_2 \mapsto \mathbb{R}$	$\longrightarrow$	Forma bilineal

Definición: Al número de espacios vectoriales que toma la aplicación multilineal se lo denomina orden o grado de la aplicación, y se lo denota por "d".

Definición: Si $d = 1$ decimos que tenemos una aplicación lineal.

Definición: Si $d = 2$ decimos que tenemos una aplicación bilineal.

Definición: Si $d = n$ decimos que tenemos una aplicación n-lineal.

4.2. Teorema de caracterización de aplicaciones multilineales

Para que una aplicación $g : E_1 \times \cdots \times E_n \longmapsto F$ sea multilineal debe cumplir las siguientes 2 propiedades en todas sus variables:

$$g(\bar{x}_1, \ldots, \bar{x}_j + \bar{y}_j, \ldots, \bar{x}_n) = g(\bar{x}_1, \ldots, \bar{x}_j, \ldots, \bar{x}_n) + g(\bar{x}_1, \ldots, \bar{y}_j, \ldots, \bar{x}_n)$$
$$g(\bar{x}_1, \ldots, \lambda\bar{x}_j, \ldots, \bar{x}_n) = \lambda g(\bar{x}_1, \ldots, \bar{x}_j, \ldots, \bar{x}_n)$$

Lo que es equivalente a:

$$\forall \bar{x}_j, \bar{y}_j \in E_j \quad \text{y } \forall \lambda, \mu \in \mathbb{R}$$
$$g(\bar{x}_j, \ldots, \lambda\bar{x}_j + \mu\bar{y}_j, \ldots, \bar{x}_n) = \lambda g(\bar{x}_1, \ldots, \bar{x}_j, \ldots, \bar{x}_n) + \mu g(\bar{x}_1, \ldots, \bar{y}_j, \ldots, \bar{x}_n)$$

Nota. Si una aplicación es multilineal toma el $\bar{0}$ como uno de sus vectores de entrada el resultado debe ser $\bar{0}$.

$$g(\bar{x}_1, \ldots, \bar{0}, \ldots, \bar{x}_d) = \bar{0}$$

Ejemplo

Demostrar que la aplicación $g(\bar{x}, \bar{y}) = ((x_1, x_2)y_2, (x_1, x_2)y_1)$ es multilineal.

1. Veamos las propiedades de vectores que vamos a utilizar para la demostración:
   - 1 - $(a + a', b + b') = (a, b) + (a', b')$
   - 2 - $(\lambda a, \lambda b) = \lambda(a, b)$
2. Demostrar que la aplicación es lineal en la primera variable:

$$\begin{gathered}
g(\lambda\bar{x} + \mu\bar{x}', \bar{y}) = \\
= g(\lambda(x_1, x_2) + \mu(x_1', x_2'), (y_1, y_2)) = g((\lambda x_1 + \mu x_1', \lambda x_2 + \mu x_2'), (y_1, y_2)) = \\
= [(\lambda x_1 + \mu x_1', \lambda x_2 + \mu x_2')y_2, (\lambda x_1 + \mu x_1', \lambda x_2 + \mu x_2')y_1] = \\
= [(\lambda x_1 + \mu x_1')y_2, (\lambda x_2 + \mu x_2')y_2], [(\lambda x_1 + \mu x_1')y_1, (\lambda x_2 + \mu x_2')y_1] =^1 \\
= [(\lambda x_1 + \lambda x_2)y_2 + (\mu x_1' + \mu x_2')y_2], [(\lambda x_1 + \lambda x_2)y_1 + (\mu x_1' + \mu x_2')y_1] =^2 \\
= [\lambda(x_1 + x_2)y_2 + \mu(x_1' + x_2')y_2], [\lambda(x_1 + x_2)y_1 + \mu(x_1' + x_2')y_1] =^1 \\
= [\lambda(x_1 + x_2)y_2 + \lambda(x_1 + x_2)y_1], [\mu(x_1' + x_2')y_2 + \mu(x_1' + x_2')y_1] =^2 \\
= \lambda[(x_1 + x_2)y_2 + (x_1 + x_2)y_1], \mu[(x_1' + x_2')y_2 + (x_1' + x_2')y_1] = \\
= \lambda g((x_1, x_2), (y_2, y_1)) + \mu g((x_1', x_2'), (y_2, y_1)) = \\
= \lambda g(\bar{x}, \bar{y}) + \mu g(\bar{x}', \bar{y})
\end{gathered}$$

3. Demostrar que la aplicación es lineal en la segunda variable:

$$\begin{gathered}
g(\bar{x}, \lambda\bar{y} + \mu\bar{y}') = \\
= g((x_1, x_2), \lambda(y_1, y_2) + \mu(y_1', y_2')) = g((x_1, x_2), (\lambda y_1 + \mu y_1', \lambda y_2 + \mu y_2')) = \\
= [(x_1, x_2)(\lambda y_2 + \mu y_2'), (x_1, x_2)(\lambda y_1 + \mu y_1')] = \\
= [(x_1(\lambda y_2 + \mu y_2'), x_2(\lambda y_2 + \mu y_2')), (x_1(\lambda y_1 + \mu y_1'), x_2(\lambda y_1 + \mu y_1'))] =^1 \\
= [(\lambda x_1 y_2, \lambda x_2 y_2) + (\mu x_1 y_2', \mu x_2 y_2'), (\lambda x_1 y_1, \lambda x_2 y_1) + (\mu x_1 y_1', \mu x_2 y_1')] =^2 \\
= [\lambda(x_1 y_2, x_2 y_2) + \mu(x_1 y_2', x_2 y_2'), \lambda(x_1 y_1, x_2 y_1) + \mu(x_1 y_1', x_2 y_1')] =^1 \\
= [(\lambda(x_1 y_2, x_2 y_2), \lambda(x_1 y_1, x_2 y_1)) + (\mu(x_1 y_2', x_2 y_2'), \mu(x_1 y_1', x_2 y_1'))] =^2 \\
= [\lambda((x_1 y_2, x_2 y_2), (x_1 y_1, x_2 y_1)) + \mu((x_1 y_2', x_2 y_2'), (x_1 y_1', x_2 y_1'))] = \\
= [\lambda((x_1, x_2)y_2, (x_1, x_2)y_1) + \mu((x_1, x_2)y_2', (x_1, x_2)y_1')] = \\
= [\lambda((x_1, x_2)(y_2, y_1)) + \mu((x_1, x_2)(y_2', y_1'))] = \lambda g(\bar{x}, \bar{y}) + \mu g(\bar{x}, \bar{y}')
\end{gathered}$$

Como g cumple el teorema de caracterización en ambas variables, podemos afirmar que sí es multilineal.

4.3. Teorema de unicidad de aplicaciones multilineales

El teorema de unicidad nos dice que una aplicación multilineal $f : E_1 \times \cdots \times E_n \mapsto F$ queda perfectamente definida por las imágenes de la combinatoria de los vectores canónicos de sus dominios (Imágenes de las combinatoria de los vectores de los espacios origen E_j), es decir, que queda perfectamente definida por su matriz (o tensor) asociada (Ya que esta se forma con las imágenes de las combinaciones de los elementos de las bases de los espacios E_j).

$$g : E_1 \times \cdots \times E_n \mapsto F \qquad B(E_j) = \{\bar{e}_{j,1}, \ldots \bar{e}_{j,mj}\} \qquad \begin{array}{l} E_j \rightarrow \text{Espacio vectorial } j \rightarrow \{E_1, \ldots, E_d\} \\ m_j \rightarrow \text{dimensión del espacio vectorial } E_j \end{array}$$

$$\text{Sea: } \bar{x}_j = \sum_{i=1}^{mj} x_{j,i}\bar{e}_{j,i} = (x_{j,1}\bar{e}_{j,1} + \cdots + x_{j,m_j}\bar{e}_{j,m_j}) = (x_{j,1}, \ldots, x_{j,m_j})_{B(E_j)}$$

Entonces tenemos:

$$g(\bar{x}_1, \ldots, \bar{x}_d) = g\left(\sum_{i_1=1}^{m_1} x_{1,i_1}\bar{e}_{1,i_1}, \ldots, \sum_{i_d=1}^{m_d} x_{d,i_d}\bar{e}_{d,i_d}\right) \Rightarrow \left\{ \begin{array}{c} \text{Propiedades de aplicaciones multilineales} \\ g(\ldots, v_i + v_j, \ldots) = g(\ldots, v_i, \ldots) + g(\ldots, v_j, \ldots) \\ g(\ldots, \lambda \cdot v_i, \ldots) = \lambda \cdot g(\ldots, v_i, \ldots) \end{array} \right\} \Rightarrow$$

$$\Rightarrow \sum_{i_1=1}^{m_1}\left(\ldots\left(\sum_{i_d=1}^{m_d}(x_{1,i_1}\cdots x_{d,i_d})g(\bar{e}_{1,i1}, \ldots, \bar{e}_{d,id})\right)\right) = \underbrace{\sum_{i_1=1,\ldots,i_j=1,\ldots,i_d=1}^{m_1,\ldots,m_j,\ldots,m_d}}_{*}\left(\prod_{j=1}^{d} x_{j,i_j}\right) g(\bar{e}_{1,i1}, \ldots, \bar{e}_{d,id})$$

Nota.* El sumatorio representa la suma de todos los elementos con la combinatoria de valores en $\{i_1, \ldots, i_d\}$

$$\left[\left(\prod_{j=1}^{d} x_{j,i_j}\right) g(\bar{e}_{1,i1}, \ldots, \bar{e}_{d,id})\right] \text{ representa a cada uno de los elementos de la matriz de la aplicación } g.$$

Utilizando una matriz solamente podemos representar aplicaciones lineales o formas bilineales. Para representar cualquier otro tipo de aplicación multilineal necesitamos matrices de más de 2 dimensiones geométricas. La forma que tenemos de representar matrices n-dimensionales es a través de la utilización de los tensores.

$$\text{Si } d > 2: \quad M(g) \;\; \text{Matriz de } g \quad \rightarrow \quad G \;\; \text{Tensor de orden } d \text{ asociado a } g$$

4.4. Aplicaciones lineales simétricas

Decimos que una aplicación multilineal es simétrica respecto a 2 variables (Espacios vectoriales) cuando podemos cambiar el orden en el que toma los vectores de dichas variables sin alterar el resultado final.

$$g(\bar{x}_1, \ldots, \bar{x}_i, \ldots, \bar{x}_j, \ldots, \bar{x}_d) = g(\bar{x}_1, \ldots, \bar{x}_j, \ldots, \bar{x}_i, \ldots, \bar{x}_d)$$

Para demostrar la linealidad de una aplicación multilineal respecto a 2 variables simétricas entre sí, basta con demostrar la linealidad de una de ellas.

Ejemplo
Demostrar que el producto escalar de $\mathbb{R}_2 : \bar{x} \circ \bar{y} = (x_1 + y_1, x_2 + y_2)$ es una aplicación multilineal.

1. Demostrar que es simétrico respecto a sus dos variables:

$$\bar{x} \circ \bar{y} = (x_1, x_2) \circ (y_1, y_2) = (x_1 \cdot y_1 + x_2 \cdot y_2) = (y_1 \cdot x_1 + y_2 \cdot x_2) = (y_1, y_2) \circ (x_1, x_2) = \bar{y} \circ \bar{x}$$

2. Demostrar que es lineal en alguna de sus variables:

$$\begin{gathered}(\lambda \bar{x} + \mu \bar{x}') \circ (\bar{y}) = \\ = (\lambda(x_1, x_2) + \mu(x_1', x_2')) \circ (y_1, y_2) = ((\lambda x_1 + \mu x_1', \lambda x_2 + \mu x_2')) \circ (y_1, y_2) = \\ = (\lambda x_1 + \mu x_1')y_1 + (\lambda x_2 + \mu x_2')y_2 = (\lambda x_1 y_1 + \lambda x_2 y_2) + (\mu x_1' y_1 + \mu x_2' y_2 = \\ = \lambda(x_1 y_1 + x_2 y_2) + \mu(x_1' y_1 + x_2' y_2) = \lambda((x_1, x_2) \circ (y_1, y_2)) + \mu((x_1', x_2') \circ (y_1, y_2)) = \\ = \lambda(\bar{x} \circ \bar{y}) + \mu(\bar{x}' \circ \bar{y})\end{gathered}$$

3. Al ser simétrico y lineal en una de sus variables se demuestra que el producto escalar es lineal en sus dos variables y por tanto una aplicación multilineal.

4.4.1. Aplicaciones lineales antisimétricas

La antisimetría de aplicaciones multilineales es una propiedad equivalente a la simetría, solo que en este caso alterar el orden de los vectores de las 2 variables produce una alteración en el signo del resultado. Es decir, la imagen resultante es el producto de la imagen original por (-1).

$$g(\bar{x}_1, \ldots, \bar{x}_i, \ldots, \bar{x}_j, \ldots, \bar{x}_d) = -g(\bar{x}_1, \ldots, \bar{x}_j, \ldots, \bar{x}_i, \ldots, \bar{x}_d)$$

4.5. Formas bilinales

Las formas bilineales son aplicaciones mulitilineales de la forma:

$$g : E_1 \times E_2 \longmapsto \mathbb{R}$$

De manera que se cumple que:

$$\text{Sean:} \quad B(E_1) = \{\bar{e}_{1,i}\}_{i=1}^{n} = \{\bar{e}_{1,1}, \dots, \bar{e}_{1,n}\} \qquad B(E_2) = \{\bar{e}_{2,j}\}_{j=1}^{m} = \{\bar{e}_{2,1}, \dots, \bar{e}_{2,m}\}$$

$$g(\bar{x}, \bar{y}) = g((x_1, \dots, x_n), (y_1, \dots, y_m)) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j g_{i,j} \qquad \text{Donde: } g_{i,j} = g(\bar{e}_{1,i}, \bar{e}_{2,j})$$

4.5.1. Matriz de una forma bilineal

La matriz de una forma bilineal debe de construirse de tal forma que al multiplicarse por los vectores $\bar{x} \in E_1$ e $\bar{y} \in E_2$ el resultado sea $\left[\sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j g_{i,j}\right]$. Por lo que se tienen que cumplir las siguientes condiciones:

- Cada coeficiente x_i debe multiplicar los correspondientes $(g_{i,:})$
- Cada coeficiente y_j debe multiplicar los correspondientes $(g_{:,j})$

Ambas se cumplen al llevar a cabo el siguiente producto:

$$g((x_1, \dots, x_n), (y_1, \dots, y_m)) = \begin{pmatrix} x_1, & \cdots & , x_n \end{pmatrix} \begin{pmatrix} g_{1,1} & \cdots & g_{1,m} \\ \vdots & \ddots & \vdots \\ g_{n,1} & \cdots & g_{n,m} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

Por tanto, tomamos como matriz de una forma bilineal la matriz resultante de colocar ordenadamente las imágenes de todas las combinaciones posibles de los vectores canónicos de los espacios del dominio de (g):

$$M(g)_{B_1 \otimes B_2} = \begin{pmatrix} g_{1,1} & \cdots & g_{1,m} \\ \vdots & \ddots & \vdots \\ g_{n,1} & \cdots & g_{n,m} \end{pmatrix}$$

$$\begin{array}{ll} g_{1,1} = g((1,0,\dots,0),(1,0,\dots,0)) & g_{1,m} = g((1,0,\dots,0),(0,\dots,0,1)) \\ g_{n,1} = g((0,\dots,0,1),(1,0,\dots,0)) & g_{n,m} = g((0,\dots,0,1),(0,\dots,0,1)) \end{array}$$

Veamos que, al multiplicar por un vector a cada lado de la matriz, las dimensiones encajan dando lugar a un escalar como resultado, tal y como sabemos que debe suceder para una forma bilineal:

$$\left[\ {\scriptstyle 1\times n}\ \right] \left(\left[\ {\scriptstyle m\times n}\ \right] \underset{m\times 1}{\left[\ \right]} \right) = \left[\ {\scriptstyle 1\times n}\ \right] \underset{m\times 1}{\left[\ \right]} = \underset{1\times 1}{\left[\ \right]} \in \mathbb{R}$$

Cuando las dos variables de una forma bilineal toman vectores del mismo espacio E en la misma base B, su matriz se expresa con la siguiente notación:

$$M(g)_{B\otimes B} \quad \longrightarrow \quad M(g)_B$$

4.5.2. Formas bilineales simétricas y antisimétricas

Si tenemos una aplicación multilineal con 2 variables simétricas, dicha aplicación cumple que:

$$g(\bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_j, \dots, \bar{x}_d) = g(\bar{x}_1, \dots, \bar{x}_j, \dots, \bar{x}_i, \dots, \bar{x}_d)$$

Si la aplicación es en concreto una forma bilineal, entonces tenemos que:

$$g(\bar{x}, \bar{y}) = g(\bar{y}, \bar{x})$$

Por lo que, si hacemos la descomposición de la aplicación, obtenemos lo siguiente:

$$\cancel{x_{1,i} y_{2,j}}\, g(\bar{e}_{1,i}, \bar{e}_{2,j}) = \cancel{x_{1,i} y_{2,j}}\, g(\bar{e}_{2,j}, \bar{e}_{1,i}) \quad \longrightarrow \quad g_{i,j} = g_{j,i}$$

De esta forma, podemos concluir que si tenemos una forma bilineal simétrica, su matriz asociada será una matriz **simétrica** S_n.

Análogamente, la matriz asociada a una forma bilineal antisimétrica será una matriz **antisimétrica** A_n.

4.5.3. Cambios de base en formas bilineales

Una forma bilineal toma vectores de dos espacios vectoriales distintos, por tanto, cuando aplicamos un cambio de base sobre la matriz asociada a una aplicación bilineal, realmente, solo estamos aplicando un cambio de base sobre una de sus dos variables.

CAMBIO DE BASE EN LA PRIMERA VARIABLE

Cuando tenenemos la matriz de una forma bilineal g expresada en las bases $B_1 \otimes B_2$ y el vector $\bar{x}$ expresado en una base B_1', necesitamos cambiar a base $B_1' \otimes B_2$ la matriz para que esta pueda tomar el vector $\bar{x}_{B_1'}$. Para obtener la expresión de la nueva matriz tengamos en cuenta que, un cambio de base de la matriz $M(g)$ de $B_1 \otimes B_2$ a $B_1' \otimes B_2$ es equivalente a cambiar de base el vector $\bar{x}_{B_1'}$ de B_1' A B_1 para luego aplicarlo a la matriz $M(g)$ en base $B_1 \otimes B_2$:

Sea: $g : E_1 \times E_2 \longmapsto \mathbb{R}$
tq.: $[\forall \bar{x}_{B_1} \in E_1 \text{ y } \forall \bar{y}_{B_2} \in E_2] \quad g(\bar{x}_{B_1}, \bar{y}_{B_2}) = \bar{x}_{B_1}^T (M(g)_{B1\otimes B2}) \bar{y}_{B_2}$

Expresamos el vector $\bar{x}_{B_1}$ en función del vector $\bar{x}_{B_1'}$:
Nota. $(AB)^T = A^T B^T$
$\bar{x}_{B_1}^T = (M_{B_1 \leftarrow B_1'} \bar{x}_{B_1'})^T = (\bar{x}_{B_1'}^T M_{B_1 \leftarrow B_1'}^T)$

Y sustituimos la expresión anterior en la ecuación de la forma bilineal:

$$\bar{x}_{B_1}^T (M(g)_{B1\otimes B2}) \bar{y}_{B_2} = (\bar{x}_{B_1'}^T M_{B_1 \leftarrow B_1'}^T)(M(g)_{B1\otimes B2}) \bar{y}_{B_2} = \bar{x}_{B_1'}^T \left(M_{B_1 \leftarrow B_1'}^T M(g)_{B1\otimes B2} \right) \bar{y}_{B_2} = \bar{x}_{B_1'}^T (M(g)_{B1'\otimes B2}) \bar{y}_{B_2}$$

CAMBIO DE BASE EN LA SEGUNDA VARIABLE

Análogamente, un cambio de base de la matriz $M(g)$ de $B_1 \otimes B_2$ a $B_1 \otimes B_2'$ (en la variable del vector $\bar{y}$) es equivalente a cambiar de base el vector $\bar{y}_{B_2'}$ de B_2' A B_2 para luego aplicarlo a la matriz en base $B_1 \otimes B_2$:

$\bar{y}_{B_2} = (M_{B_2 \leftarrow B_2'} \bar{y}_{B_2'})$
$\bar{x}_{B_1}^T (M(g)_{B1\otimes B2}) \bar{y}_{B_2} = \bar{x}_{B_1}^T (M(g)_{B1\otimes B2})(M_{B_2 \leftarrow B_2'} \bar{y}_{B_2'}) = \bar{x}_{B_1}^T \left(M(g)_{B1\otimes B2} M_{B_2 \leftarrow B_2'} \right) \bar{y}_{B_2'} = \bar{x}_{B_1}^T (M(g)_{B1\otimes B2'}) \bar{y}_{B_2'}$

CAMBIO DE BASE EN AMBAS VARIABLES AL MISMO TIEMPO

Si queremos realizar ambos cambios de base de forma simultánea multiplicamos por una matriz a cada lado:
$\bar{x}_{B_1}^T (M(g)_{B1\otimes B2}) \bar{y}_{B_2} = (\bar{x}_{B_1'}^T M_{B_1 \leftarrow B_1'}^T)(M(g)_{B1\otimes B2})(M_{B_2 \leftarrow B_2'} \bar{y}_{B_2'}) = \bar{x}_{B_1'}^T \left(M_{B_1 \leftarrow B_1'}^T M(g)_{B1\otimes B2} M_{B_2 \leftarrow B_2'} \right) \bar{y}_{B_2'} =$
$= \bar{x}_{B_1'}^T (M(g)_{B1'\otimes B2'}) \bar{y}_{B_2'}$

Ejemplo
Dada la forma bilineal : $g((x_1, x_2, x_3), (y_1, y_2, y_3)) = (3x_1y_2 + x_2y_2 + 4x_3y_3)$ expresada en la base canónica de $\mathbb{R}^3$, calcular su matriz asociada en base $B = \{(1,1,-1), (2,0,1), (1,0,1)\}$.

1. Obtener la matriz asociada a g en base B_c:

$$M(g)_{B_c} = \begin{pmatrix} 0 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

2. Multiplicar por las matriz de cambio de base B a base B_c a ambos lados de la matriz de g:

$$M(g)_B = (M_{Bc \leftarrow B})^T M(g)_{B_c} M_{Bc \leftarrow B}$$

$$M(g)_B = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}^T \begin{pmatrix} 0 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \\ -1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 8 & -4 & -4 \\ 2 & 4 & 4 \\ -1 & 4 & 4 \end{pmatrix}$$

Ejemplo
Sea $g : \mathbb{R}^3 \times \mathbb{R}^3 \longmapsto \mathbb{R}$ una forma bilineal simétrica que cumple las siguientes propiedades:

1. La matriz de g en la base canónica de $\mathbb{R}^3$ tiene rango 1
2. $g((1,-2,2)(1,-2,2)) = 3$
3. $g((2,-4,4)(\alpha, \beta-\alpha, -\beta)) = 0 \ \ \forall \alpha, \beta \in \mathbb{R}$

Hallar la matriz de g en la base $B = \{(1,0,0), (1,1,0), (1,1,1)\}$

1. Debido a la primera característica, que nos dice que la forma bilineal tiene rango = 1, sabemos que la forma de su matriz debe ser:

$$M_{B_c \otimes B_c} = \begin{pmatrix} m_1 & m_2 & m_3 \\ m_4 & m_5 & m_6 \\ m_7 & m_8 & m_9 \end{pmatrix} \Longrightarrow \begin{pmatrix} \lambda a_1 & \lambda a_2 & \lambda a_3 \\ \mu a_1 & \mu a_2 & \mu a_3 \\ \delta a_1 & \delta a_2 & \delta a_3 \end{pmatrix}$$

2. Utilizando las características segunda y tercera y aplicando las propiedades de aplicaciones multilineales podemos extraer ecuaciones que relacionan todos los elementos de una misma fila:

$$\begin{array}{rcl} 2 \cdot g((1,2,2),(1,-2,2)) = 2 \cdot 3 & \longrightarrow & g((2,-4,4),(1,-2,2)) = 6 \\ g((2,-4,4),(\alpha,\beta-\alpha,-\beta)) = 0 & \xrightarrow{\alpha = 0, \beta = 2} & g((2,-4,4),(0,2,-2)) = 0 \\ g((2,-4,4),(\alpha,\beta-\alpha,-\beta)) = 0 & \xrightarrow{\alpha = (-1), \beta = 2} & g((2,-4,4),(-1,3,-2)) = 0 \\ g((2,-4,4),(\alpha,\beta-\alpha,-\beta)) = 0 & \xrightarrow{\alpha = (-1), \beta = 1} & g((2,-4,4),(-1,2,-1)) = 0 \end{array}$$

$$g((2,-4,4),(1,-2,2)) + g((2,-4,4),(0,2,-2)) = g((2,-4,4),(1,0,0)) = 6$$
$$2 \cdot g((1,0,0),(1,0,0)) + (-4) \cdot g((0,1,0),(1,0,0)) + 4 \cdot g((0,0,1),(1,0,0)) = 2m_1 - 4m_2 + 4m_3 = 6$$

$$g((2,-4,4),(1,-2,2)) + g((2,-4,4),(-1,3,-2)) = g((2,-4,4),(0,1,0)) = 6$$
$$2 \cdot g((1,0,0),(1,0,0)) + (-4) \cdot g((0,1,0),(1,0,0)) + 4 \cdot g((0,0,1),(1,0,0)) = 2m_4 - 4m_5 + 4m_6 = 6$$

$$g((2,-4,4),(1,-2,2)) + g((2,-4,4),(-1,2,-1)) = g((2,-4,4),(0,0,1)) = 6$$
$$2 \cdot g((1,0,0),(1,0,0)) + (-4) \cdot g((0,1,0),(1,0,0)) + 4 \cdot g((0,0,1),(1,0,0)) = 2m_7 - 4m_8 + 4m_9 = 6$$

Las tres ecuaciones anteriores se pueden simplificar:

$$2m_i - 4m_j + 4m_k = 6 \quad \longrightarrow \quad m_i - 2m_j + 2m_k = 3$$

3. Aplicamos la multiplicidad que sabemos que existe entre las filas de la matriz para hallar el valor de las constantes $\{\lambda, \mu, \delta\}$:

$$\begin{array}{c} m_1 - 2m_2 + 2m_3 = (\lambda a_1) - 2(\lambda a_2) + 2(\lambda a_3) = \lambda(a_1 - 2a_2 + 2a_3) = 3 \\ m_4 - 2m_5 + 2m_6 = (\mu a_1) - 2(\mu a_2) + 2(\mu a_3) = \mu(a_1 - 2a_2 + 2a_3) = 3 \\ m_7 - 2m_8 + 2m_9 = (\delta a_1) - 2(\delta a_2) + 2(\delta a_3) = \delta(a_1 - 2a_2 + 2a_3) = 3 \end{array}$$

$$\lambda(a_1 - 2a_2 + 2a_3) = \mu(a_1 - 2a_2 + 2a_3) = \delta(a_1 - 2a_2 + 2a_3) \iff \lambda = \mu = \delta$$

$$M_{B_c \otimes B_c} = \begin{pmatrix} \lambda a_1 & \lambda a_2 & \lambda a_3 \\ \mu a_1 & \mu a_2 & \mu a_3 \\ \delta a_1 & \delta a_2 & \delta a_3 \end{pmatrix} \Longrightarrow \begin{pmatrix} \lambda a_1 & \lambda a_2 & \lambda a_3 \\ \lambda a_1 & \lambda a_2 & \lambda a_3 \\ \lambda a_1 & \lambda a_2 & \lambda a_3 \end{pmatrix}$$

4. Debido a la primera propiedad, que nos dice que la forma bilineal es simétrica, sabemos que los elementos de su matriz deben cumplir:

$$\lambda a_1 = \lambda a_2 = \lambda a_3$$

$$M_{B_c \otimes B_c} = \begin{pmatrix} \lambda a_1 & \lambda a_2 & \lambda a_3 \\ \lambda a_1 & \lambda a_2 & \lambda a_3 \\ \lambda a_1 & \lambda a_2 & \lambda a_3 \end{pmatrix} \Longrightarrow \begin{pmatrix} \lambda a_1 & \lambda a_1 & \lambda a_1 \\ \lambda a_1 & \lambda a_1 & \lambda a_1 \\ \lambda a_1 & \lambda a_1 & \lambda a_1 \end{pmatrix}$$

5. Recurriendo a la ecuación de antes podemos hallar el valor de λa_1 y obtener la matriz $M_{B\otimes B}$:

$$m_i - 2m_j + 2m_k = 3 \quad \longrightarrow \quad (\lambda a_1) - 2(\lambda a_1) + 2(\lambda a_1) = 3 \quad \longrightarrow \quad (\lambda a_1) = 3$$

$$M_{B_c \otimes B_c} = \begin{pmatrix} \lambda a_1 & \lambda a_1 & \lambda a_1 \\ \lambda a_1 & \lambda a_1 & \lambda a_1 \\ \lambda a_1 & \lambda a_1 & \lambda a_1 \end{pmatrix} \Longrightarrow \begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix}$$

6. Finalmente obtenemos la matriz de la forma bilineal expresada en la base B:

$$M_{B\otimes B} = M_{B_c \leftarrow B}^T M_{B_c \otimes B_c} M_{B_c \leftarrow B}$$

$$M_{B\otimes B} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 9 \\ 6 & 12 & 18 \\ 9 & 18 & 27 \end{pmatrix}$$

4.6. Formas cuadráticas

Una forma cuadrática q o forma bilineal simétrica es una aplicación que asigna a cada vector $\bar{x}$ de un espacio vectorial E el escalar resultante de evaluar el vector $\bar{x}$ con sí mismo en una forma bilineal simétrica g:

$$q : E \longmapsto \mathbb{R}$$
$$q(\bar{x}) = g(\bar{x}, \bar{x}) \in \mathbb{R}$$

4.6.1. Matriz de una forma cuadrática

La matriz de una forma cuadrática q coincide exactamente con la matriz de su forma bilineal simétrica g asociada. Si contamos con una forma cuadrática q dada de forma analítica y queremos obtener su matriz debemos tener en cuenta que, debido a la condición de simetría, cada elemento $g_{i,j}$ de la matriz debe coincidir con su complementario $g_{j,i}$. Si uno de los dos no aparece en la expresión analítica de q deberemos obtenerlo dividiendo el otro por la mitad.

El cambio de base de una forma cuadrática q se lleva a cabo de la misma forma que el de una forma bilineal, multiplicando por la matriz de cambio de base a un lado y por su traspuesta al otro.

Ejemplo
Dada la forma cuadrática q en su forma analítica, hallar su matriz asociada expresada en la base B, y obtener la imagen del vector $\bar{v} = (1, 0, 1)$ dado en la base B:

$$q(x, y, z) = x^2 + 4xy + 2yz + 3z^2$$
$$B = \{(7, 3, 5), (1, 1, 1), (6, 0, 4)\}$$

1. Obtener la matriz de la forma cuadrática q en base canónica:

$$4xy \longrightarrow (2xy + 2yx) \qquad\qquad 2yz \longrightarrow (1yz + 1zy)$$

$$M(q)_{B_c} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

2. Multiplicar por las matriz de cambio de base B a base B_c a ambos lados de la matriz de q:

$$M(q)_B = (M_{Bc \leftarrow B})^T M(q)_{B_c} M_{Bc \leftarrow B}$$

$$M(q)_B = \begin{pmatrix} 7 & 3 & 5 \\ 1 & 1 & 1 \\ 6 & 0 & 4 \end{pmatrix} \begin{pmatrix} 0 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 238 & 50 & 150 \\ 50 & 10 & 34 \\ 150 & 34 & 84 \end{pmatrix}$$

3. Obtener la imagen del vector $\bar{v}$ multiplicando por la matriz de q en base B:

$$\begin{pmatrix} 238 & 50 & 150 \\ 50 & 10 & 34 \\ 150 & 34 & 84 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 388 \\ 84 \\ 234 \end{pmatrix}$$

4.6.2. Carácter de una forma cuadrática

Una forma cuadrática puede tener carácter positivo, negativo o indefinido. El carácter de una forma cuadrática se determina de la misma forma que el carácter de un producto escalar. (Ver producto escalar en el tema 5).

4.7. Producto Tensorial

Para entender correctamente el concepto de **tensor** primero se debe entender bien la operación con la que se construye, el **producto tensorial**. El producto tensorial es una operación diseñada con la idea de hacer posible el "combinar" / "relacionar" vectores (pertenecientes a distintos espacios vectoriales) o espacios vectoriales completamente distintos entre sí respetando una serie de propiedades que resultan útiles a la hora de operar con ellos. Se denota por $\otimes$.

4.7.1. Propiedades del producto tensorial

Las dos propieades que cumple el producto tensorial son las siguientes:

1. Asociativo en la suma: $\longrightarrow$ $(\bar{e}_1 + \bar{e}_2) \otimes \bar{e}_3 = \bar{e}_1 \otimes \bar{e}_3 + \bar{e}_2 \otimes \bar{e}_3$

2. Asociativo en el producto por escalar: $\longrightarrow$ $(\lambda\bar{e}_1) \otimes \bar{e}_2 = \bar{e}_1 \otimes (\lambda\bar{e}_2) = \lambda(\bar{e}_1 \otimes \bar{e}_2)$

4.7.2. Idea intuitiva del producto tensorial

Al hablar de producto tensorial se debe diferenciar entre dos tipos:

1. El producto tensorial de espacios vectoriales da lugar a una combinacion tensorial de dichos espacios que podemos entender como un **"espacio tensorial" de los espacios vectoriales originales**.

   Nota. Existen una serie de equivalencias isomórficas que nos permiten describir cualquier aplicación multilineal como un producto tensorial de espacios vectoriales

2. El producto tensorial de vectores contenidos en distintos espacios vectoriales es un objeto contenido en la combinación de espacios vectoriales anterior (1). Este objeto recibe el nombre de **tensor**, y según como queramos trabajar con él nos será conveniente respresentarlo de una u otra forma. Las más frecuentes son, como un **vector** (representación canónica) contenido en el "espacio tensorial" anterior (1), como una **descomposición en matrices simples** de una matriz de d dimensiones geométricas, o en función del propio **producto tensorial** de los d vectores contenidos en sus d respectivos espacios vectoriales.

4.8. Tensor

Un tensor es un objeto algebraico que representa la "matriz d-dimensional"* resultante del producto tensorial de d vectores pertenecientes a d distintos espacios vectoriales. Como la representación geométrica de una matriz d-dimensional generalmente es imposible ($d > 3$) o poco matemáticamente operativa ($d = 3$), se utlilizan otros métodos para representarlos. Los más frecuentes son, mediante un vector (representación canónica), mediante un conjunto de matrices, o en función de un conjunto de productos tensoriales.

Nota.* En este caso "matriz d-dimensional", no quiere decir "matriz $(\bar{f} \otimes \bar{c})$ de dimensión $m \times n = d$", sino, "matriz $(\bar{v}_1 \otimes \bar{v}_2 \otimes \cdots \otimes \bar{v}_d)$ de dimensión $n_1 \otimes n_2 \times \cdots \times n_d$" siendo (n_i) la dimensión de cada vector $\bar{v}_i$. Es decir, que "dimensión", significa "dimensión geométrica" y no "número de elementos de la matriz".

Al número de vectores que forman un tensor se lo define como "orden del tensor".

4.8.1. Representaciones del tensor

REPRESENTACIÓN EN FUNCIÓN DE UN PRODUCTO TENSORIAL (SIN DESCOMPONER):

De esta forma, se representa el tensor en función del producto tensorial de los vectores de los distintos espacios. La ventaja de representar así el tensor es que podemos distinguir claramente a que espacio pertenece cada vector, y por tanto, trabajar con cada uno de ellos por separado.

Ejemplo

$$\underset{1}{(1,3)} \otimes \underset{2}{(1,2)} \otimes \underset{3}{(0,1)} = T$$

$$\bar{v}^1 \in E_1 \qquad \bar{v}^2 \in E_2 \qquad \bar{v}^3 \in E_3$$

REPRESENTACIÓN CANÓNICA:

Esta es la representación más compacta del tensor, ya que toma la forma de un vector. El problema en este caso es determinar el espacio vectorial en el que está contenido este vector. Para ello, surge la noción de producto tensorial de espacios vectoriales, cuyo resultado se puede considerar un nuevo espacio vectorial (que se denota por el propio producto tensorial). En este nuevo espacio vectorial está contenido el vector.

Ejemplo

$$\underset{1}{(0,1)} \otimes \underset{2}{(1,2)} \otimes \underset{3}{(1,3)} = (0,1,0,2)^T \otimes (1,3) = (0,1,0,2,0,3,0,6)^T = T$$

$$\bar{v}^1 \in E_1 \quad \bar{v}^2 \in E_2 \quad \bar{v}^3 \in E_3 \qquad\qquad T \in E_1 \otimes E_2 \otimes E_3$$

La representación canónica tiene la peculiaridad de que el producto tensorial de vectores contenidos en espacios vectoriales duales expanden el tensor de forma horizontal en lugar de vertical:

Ejemplo

$$\underset{1}{(0,-1)} \otimes \underset{2}{(3,2)} \otimes \underset{3}{(2,1)} = \begin{pmatrix} 0 & -3 \\ 0 & -2 \end{pmatrix} \otimes (2,1) = \begin{pmatrix} 0 & -6 & 0 & -3 \\ 0 & -4 & 0 & -2 \end{pmatrix} = T$$

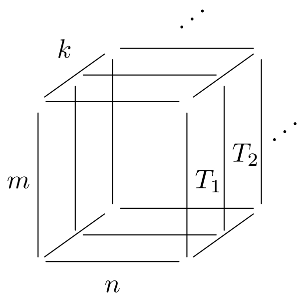
$$\bar{v}^1 \in E_1^* \quad \bar{v}^2 \in E_2 \quad \bar{v}^3 \in E_3^* \qquad\qquad T \in E_1^* \otimes E_2 \otimes E_3^*$$

Nota. Para obtener la representación canónica se empieza a multiplicar los vectores desde delante.

Nota. Una matriz de una aplicación lineal $f : E_1 \longmapsto E_2$ se puede considerar la representación canónica de un tensor contenico en $(E_2 \otimes E_1^*)$

REPRESENTACIÓN MATRICIAL:

Esta representación del tensor ofrece una descripción simplicicada de la matriz (de d dimensiones geométricas) asociada. La simplificación se lleva a cabo a través de una descomposición laminar de la matriz en matrices de 2 dimensiones geométricas, tal y como se puede observar en el siguiente ejemplo:



Esta representación del tensor es útil por ofrecer la visualización más cómoda posible del mismo. Además, puede ayudarnos a la hora de realizar a mano algunas operaciones como la contracción o el trenzado.

Ejemplo

$$(1,3) \otimes (1,2) \otimes (0,1) = (1,3) \begin{pmatrix} 1 \\ 2 \end{pmatrix} (0,1) = (1,3) \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} =$$

$$= T = \left[T_{1,:,:} = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} ; \quad T_{2,:,:} = \begin{pmatrix} 0 & 3 \\ 0 & 6 \end{pmatrix} \right]$$

Nota. Para obtener la representación canónica se empieza a multiplicar los vectores desde atrás.

REPRESENTACIÓN EN FUNCIÓN DE UN PRODUCTO TENSORIAL (DESCOMPUESTO):

De esta forma, se representa el tensor en función del producto tensorial de los vectores de los distintos espacios. La ventaja de representar así el tensor es que podemos distinguir claramente a que espacio pertenece cada vector, y por tanto, trabajar con cada uno de ellos por separado.

Ejemplo

$$(1,3) \otimes (1,2) \otimes (0,1) = (1,3) \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} = (1,3) \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} (0,1) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,2) \right] =$$

$$= (1,0) \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} (0,1) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,2) \right] + (0,3) \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} (0,1) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,2) \right] =$$

$$= (1,0) \otimes \big[(1,0) \otimes (0,1) + (0,1) \otimes (0,2) \big] + (0,3) \otimes \big[(1,0) \otimes (0,1) + (0,1) \otimes (0,2) \big] = T$$

4.9. Contracción de dos espacios vectoriales de un tensor

En álgebra multilineal, una contracción tensorial es una **operación** sobre un tensor que consiste en la **evaluación** de un vector v_i de un espacio vectorial de dimensión finita en un vector v_j de su espacio dual, y se denota por $ev_{i,j}(T)$. Dado que la evaluación de un vector en otro de su espacio dual resulta en un escalar, efectuar una contracción consiste en sutituir dos productos tensoriales por un producto por escalar, lo que reduce el orden del tensor original T en 2 unidades.

Para que la operación de contracción de dos vectores v_i y v_j (evaluación de un vector en el otro) tenga sentido es necesario que uno de ellos esté contenido en el espacio dual del espacio vectorial del otro.

$$Sea \quad T = v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_j \otimes \cdots \otimes v_d \quad tq. \quad T \in E_1 \otimes \cdots \otimes E_i \otimes \cdots \otimes E_j \otimes \cdots \otimes E_d$$
$$\exists \ ev_{i,j}(T) \quad \Longleftrightarrow \quad E_j = E_i^* \text{ ò } E_i = E_j^*$$

Ejemplo
Dado el tensor T en función de varios productos tensoriales, efectuar la contracción de tantos vectores del mismo como sea posible.

$T = (1,2,3) \otimes (2,-1) \otimes (1,-3,0) \in \mathbb{R}^3 \otimes \mathbb{R}^2 \otimes \mathbb{R}^{3*}$

Como tenemos un vector de $\mathbb{R}^3$ y otro de su dual $\mathbb{R}^{3*}$, podemos contraerlos.

$$ev_{1,3}\left((1,2,3)\otimes(2,-1)\otimes(1,-3,0)\right) \ \Rightarrow \ \left\{(1,2,3)\begin{pmatrix}1\\-3\\0\end{pmatrix} = (-5)\right\} \ \Rightarrow \ (-5)\cdot(2,-1) = (-10,5)$$

Ejemplo
Dado el tensor T en su representación matricial, efectuar la contracción de tantos vectores del mismo como sea posible.

$T = \left[T_{1,:,:} = \begin{pmatrix}1 & 0 & 2\\ 2 & 3 & 0\end{pmatrix};\ T_{2,:,:} = \begin{pmatrix}3 & 1 & 1\\ 0 & 1 & 0\end{pmatrix}\right] \in \mathbb{R}^{2*} \otimes \mathbb{R}^2 \otimes \mathbb{R}^3$

Como tenemos un vector de $\mathbb{R}^2$ y otro de su dual $\mathbb{R}^{2*}$, podemos contraerlos:

Nota. El teorema de reflexividad nos dice que un espacio vectorial E es canónicamente isomorfo a su doble dual E^{**}, y que por tanto podemos entenderlo como tal. Veamos entonces que evaluar un vector de E en uno de E^* es equivalente a evaluar uno de E^* en uno de E^{**}, por lo que el orden en el que multipliquemos ambos vectores es irrelevante.

$$ev_{1,2}\left(\left[T_{1,:,:} = \begin{pmatrix}1 & 0 & 2\\ 2 & 3 & 0\end{pmatrix};\ T_{2,:,:} = \begin{pmatrix}3 & 1 & 1\\ 0 & 1 & 0\end{pmatrix}\right]\right) =$$
$$= ev_{1,2}\left(\left[(1,0)\otimes\begin{pmatrix}1 & 0 & 2\\ 2 & 3 & 0\end{pmatrix} + (0,1)\otimes\begin{pmatrix}3 & 1 & 1\\ 0 & 1 & 0\end{pmatrix}\right]\right) =$$
$$= ev_{1,2}\left(\left[(1,0)\otimes\left[(1,0)\otimes(1,0,2)+(0,1)\otimes(2,3,0)\right] + (0,1)\otimes\left[(1,0)\otimes(3,1,1)+(0,1)\otimes(0,1,0)\right]\right]\right) =$$
$$\Longrightarrow \quad \left\{(1,0)\begin{pmatrix}1\\0\end{pmatrix} = 1; \quad (1,0)\begin{pmatrix}0\\1\end{pmatrix} = 0; \quad (0,1)\begin{pmatrix}1\\0\end{pmatrix} = 0; \quad (0,1)\begin{pmatrix}0\\1\end{pmatrix} = 1\right\} \quad \Longrightarrow$$
$$= \left[1\cdot(1,0,2) + 0\cdot(2,3,0) + 0\cdot(3,1,1) + 1\cdot(0,1,0)\right] = (1,1,2)$$

4.9.1. Relación de la contracción tensorial con la traza

Cuando tenemos un producto tensorial de dos vectores expresado como una matriz, realizar la contracción de dichos vectores es equivalente a calcular la traza de dicha matriz.

$$Sea \ \ A = \begin{pmatrix}a_{1,1} & \dots & a_{1,n}\\ \vdots & \ddots & \vdots\\ a_{n,1} & \dots & a_{n,n}\end{pmatrix}$$

$$ev_{1,2}(A) = ev_{1,2}\big((1,0,\dots,0)\otimes(a_{1,1}\dots a_{1,n}) + \cdots + (0,\dots,0,1)\otimes(a_{n,1}\dots a_{n,n})\big) =$$
$$= a_{1,1} + \cdots + a_{i,i} + \cdots + a_{n,n} = tr(A)$$

Ejemplo
Dado el tensor T en su representación matricial, efectuar la contracción de tantos vectores del mismo como sea posible.

$T = \left[T_{1,:,:} = \begin{pmatrix}1 & 1\\ 2 & 1\end{pmatrix};\ T_{2,:,:} = \begin{pmatrix}3 & 1\\ 2 & 1\end{pmatrix};\ T_{3,:,:} = \begin{pmatrix}3 & 0\\ 0 & 5\end{pmatrix}\right] \in \mathbb{R}^3 \otimes \mathbb{R}^2 \otimes \mathbb{R}^{2*}$

Como tenemos un vector de $\mathbb{R}^2$ y otro de su dual $\mathbb{R}^{2*}$, podemos contraerlos:

$$ev_{2,3}\left(\left[T_{1,:,:} = \begin{pmatrix}1 & 1\\ 2 & 1\end{pmatrix};\ T_{2,:,:} = \begin{pmatrix}3 & 1\\ 2 & 1\end{pmatrix};\ T_{3,:,:} = \begin{pmatrix}3 & 0\\ 0 & 5\end{pmatrix}\right]\right) =$$
$$= ev_{2,3}\left(\left[(1,0,0)\otimes\begin{pmatrix}1 & 1\\ 2 & 1\end{pmatrix};+(0,1,0)\otimes\begin{pmatrix}3 & 1\\ 2 & 1\end{pmatrix};+(0,0,1)\otimes\begin{pmatrix}3 & 0\\ 0 & 5\end{pmatrix}\right]\right) =$$
$$= \left[(1,0,0)\otimes tr\begin{pmatrix}1 & 1\\ 2 & 1\end{pmatrix};+(0,1,0)\otimes tr\begin{pmatrix}3 & 1\\ 2 & 1\end{pmatrix};+(0,0,1)\otimes tr\begin{pmatrix}3 & 0\\ 0 & 5\end{pmatrix}\right] =$$
$$= \left[2\cdot(1,0,0) + 4\cdot(0,1,0) + 8\cdot(0,0,1)\right] = (2,4,8)$$

4.10. Trenzado de espacios vectoriales de un tensor

El trenzado es una operación que consiste en la alteración del orden de los espacios vectoriales que se multiplican tensorialmente dando lugar a un tensor. Consecuentemente se produce la misma alteración en los vectores pertenecientes a dichos espacios vectoriales. Es vectores (multiplicados tensorialmente entre sí) que ocupan las posiciones i y j de producto tensorial. La denotaremos por $\tau_{i,j}$.

Ejemplo

$$\tau_{2,3}\Big(\underset{1}{(2,5)}\otimes\underset{2}{(1,-1,0)}\otimes\underset{3}{(6,3)}\otimes\underset{4}{(4,2,1)}\Big) = \underset{1}{(2,5)}\otimes\underset{3}{(6,3)}\otimes\underset{2}{(1,-1,0)}\otimes\underset{4}{(4,2,1)}$$

Puede pasar que no contemos con el tensor en función de productos tensoriales sino en su expresión matricial. En estos casos no podemos cambiar de posición los vectores directamente. En su lugar debemos cambiar los valores del tensor índice por índice.

Ejemplo

$$\tau_{2,3}\begin{bmatrix} \begin{pmatrix} \underset{1,1,1,1}{2} & \underset{1,1,1,2}{5} \\ \underset{1,1,2,1}{0} & \underset{1,1,2,2}{2} \end{pmatrix} & \begin{pmatrix} \underset{1,2,1,1}{1} & \underset{1,2,1,2}{1} \\ \underset{1,2,2,1}{3} & \underset{1,2,2,2}{2} \end{pmatrix} \\ \begin{pmatrix} \underset{2,1,1,1}{3} & \underset{2,1,1,2}{5} \\ \underset{2,1,2,1}{0} & \underset{2,1,2,2}{1} \end{pmatrix} & \begin{pmatrix} \underset{2,2,1,1}{3} & \underset{2,2,1,2}{4} \\ \underset{2,2,2,1}{1} & \underset{2,2,2,2}{7} \end{pmatrix} \end{bmatrix} = \begin{bmatrix} \begin{pmatrix} \underset{1,1,1,1}{2} & \underset{1,1,1,2}{5} \\ \underset{1,1,2,1}{1} & \underset{1,1,2,2}{1} \end{pmatrix} & \begin{pmatrix} \underset{1,2,1,1}{0} & \underset{1,2,1,2}{2} \\ \underset{1,2,2,1}{3} & \underset{1,2,2,2}{2} \end{pmatrix} \\ \begin{pmatrix} \underset{2,1,1,1}{3} & \underset{2,1,1,2}{5} \\ \underset{2,1,2,1}{3} & \underset{2,1,2,2}{4} \end{pmatrix} & \begin{pmatrix} \underset{2,2,1,1}{0} & \underset{2,2,1,2}{1} \\ \underset{2,2,2,1}{1} & \underset{2,2,2,2}{7} \end{pmatrix} \end{bmatrix}$$

Cuando queremos cambiar simultáneamente de orden más que únicamente las posiciones de dos índices se cambia la denotación de τ indicando directamente la disposición final de los índices (en lugar de los dos índices que se desean intercambiar). Los índices que no se van a cambiar de posición se indican con (:).

Ejemplo

$$\tau_{3,:,4,1}\Big(\underset{1}{(2,5)}\otimes\underset{2}{(1,-1,0)}\otimes\underset{3}{(6,3)}\otimes\underset{4}{(4,2,1)}\Big) = \underset{3}{(6,3)}\otimes\underset{2}{(1,-1,0)}\otimes\underset{4}{(4,2,1)}\otimes\underset{1}{(2,5)}$$

4.11. Equivalencias isomórficas del producto tensorial

Existen una serie de equivalencias isomórficas entre aplicaciones multilineales y espacios vectoriales relacionados por productos tensoriales (equivalencias que afirman que cualquier aplicación multilineal es un isomorfismo ¿canónico (creo que sí)? de un conjunto de espacios vectoriales relacionados por productos tensoriales) que nos permiten llevar a cabo la tensorización de las aplicaciones multilineales.

1. $E \otimes \mathbb{R} \cong E$
2. $(E_1 \otimes \cdots \otimes E_n)^* \cong E_1^* \otimes \cdots \otimes E_n^*$
3. $Hom(E_1 \times \cdots \times E_d, \mathbb{R}) \cong E_1^* \otimes \cdots \otimes E_n^*$
4. $Hom(E, F) \cong F \otimes E^*$
5. $Hom(E_1 \times \cdots \times E_n, F) \cong E_1^* \otimes \cdots \otimes E_n^* \otimes F$

Esto nos va a ser útil para saber en que "espacio tensorial" vive una aplicación multilineal.

Ejemplo

Aplicando la definición y propiedades del producto tensorial de espacios vectoriales, describir los siguientes espacios como tensores en términos de los espacios vectoriales indicados y sus duales.

a) Espacio de las aplicaciones bilineales que toman un polinomio de grado a lo sumo 2 y una matriz 3×3 y devuelven un vector de $\mathbb{R}^7$.

$Hom(\mathbb{R}_{2[x]} \otimes M_3, \mathbb{R}^7) \cong \mathbb{R}_{2[x]}^* \otimes M_3^* \otimes \mathbb{R}^7$

b) Espacio de aplicaciones trilineales de $\mathbb{R}^3$ a $\mathbb{R}^5$.

$Hom(\mathbb{R}^3 \otimes \mathbb{R}^3 \otimes \mathbb{R}^3, \mathbb{R}^5) \cong \mathbb{R}^{3*} \otimes \mathbb{R}^{3*} \otimes \mathbb{R}^{3*} \otimes \mathbb{R}^5 \cong (\mathbb{R}^{3*})^{\otimes 3} \otimes \mathbb{R}^5$

c) Espacio de las aplicaciones multilineales que toman 2 vectores de $\mathbb{R}^5$ y devuelven una forma lineal de $\mathbb{R}^2$.

$Hom(\mathbb{R}^5 \otimes \mathbb{R}^5, Hom(\mathbb{R}_{2[x]}, \mathbb{R})) \cong (\mathbb{R}^5 \otimes \mathbb{R}^5)^* \otimes Hom(\mathbb{R}_{2[x]}, \mathbb{R})) \cong \mathbb{R}^{5*} \otimes \mathbb{R}^{5*} \otimes \mathbb{R}_{2[x]}^*$

d) Espacio de endomorfismos de $End(S_3)$ (es decir, de $Hom(S_3, S_3)$.

$End(End(S_3)) = End(S_3) \otimes End(S_3)^* = S_3 \otimes S_3^* \otimes S_3 \otimes S_3^*$

e) Espacio de aplicaciones bilineales que toman una forma bilineal de $\mathbb{R}^2$ y una forma bilineal de $\mathbb{R}^4$ y devuelven un endomorfismo de $\mathbb{R}^3$.

$Hom(Hom(\mathbb{R}^2 \otimes \mathbb{R}^2, \mathbb{R}) \otimes Hom(\mathbb{R}^4 \otimes \mathbb{R}^4, \mathbb{R}), End(\mathbb{R}^3)) = Hom((\mathbb{R}^{2*})^{\otimes 2} \otimes (\mathbb{R}^{4*})^{\otimes 2}, \mathbb{R}^3 \otimes \mathbb{R}^{3*}) =$
$= (\mathbb{R}^2)^{\otimes 2} \otimes (\mathbb{R}^4)^{\otimes 2} \otimes \mathbb{R}^3 \otimes \mathbb{R}^{3*}$

f) Espacio dual de $Hom(E, F)$.

$Hom(Hom(E, F), \mathbb{R}) = (Hom(E, F))^* = (F \otimes E^*)^* = F^* \otimes E$

g) Espacio de aplicaciones multilineales que toman un endomorfismo de E, una forma lineal de E y una forma bilineal de E y devuelven un homomorfismo de E a $E \otimes E \otimes E$.

$Hom(End(E) \otimes E^* \otimes Hom(E \otimes E, \mathbb{R}), Hom(E, E \otimes E \otimes E)) =$
$= Hom(E \otimes (E^*)^{\otimes 4}, (E)^{\otimes 3} \otimes E^*) = E^* \otimes (E)^{\otimes 7} \otimes E^*$

4.12. Aplicación de los tensores al álgebra multilineal

Cuando trabajamos con una aplicación multilineal nos encontramos con el problema de ser incapaces de representarla de forma matricial cuando toma más de 1 vector del dominio (o cuando toma 2 siendo una forma bilineal).

Para resolver este problema, el álgebra tensorial es una herramienta muy útil, ya que nos permite descomponer una aplicación multilineal de orden d en d aplicaciones lineales. El proceso se divide en dos pasos:

1. Transformación tanto de la aplicación multilineal como del conjunto de vectores (que va a tomar la aplicación multilineal) en tensores (tensorización).
2. Aplicación o conjunto de aplicaciones lineales que toman ambos tensores y los llevan al resultado de evaluar los vectores originales en la aplicación multilineal.

DIAGRAMA

$$\begin{tikzcd}
E_1 \times \cdots \times E_n \arrow[rr, "g"] \arrow[d] & & F \\
T = E_1 \otimes \cdots \otimes E_n & G = E_1^* \otimes \cdots \otimes E_n^* \otimes F \arrow[ur, bend right] &
\end{tikzcd}$$

tensorización

Toda aplicación multilineal se puede describir como un tensor contenido en: el producto tensorial de, **AL MENOS, los espacios duales de los espacios vectoriales en los que viven los vectores que toma la aplicación**. Es decir, una aplicación multilineal g que toma vectores de los espacios $E_1, \dots, E_n$ cuple que:

Sean $\bar{v}_1 \in E_1, \dots, \bar{v}_n \in E_n \quad g(\bar{v}_1, \dots, \bar{v}_n) \in E_1^* \otimes \cdots \otimes E_n^* \otimes (\text{Posible(s) espacio(s) vectorial(es) no dual(es)})$

$$g(\bar{v}_1, \dots, \bar{v}_n) : E_1 \times \cdots \times E_n \longmapsto (\text{Posible(s) espacio(s) vectorial(es) no dual(es)})$$

Al conjunto de los m posibles espacios vectoriales no duales que forman parte de la definición del "espacio tensorial" en el que esta contenido el tensor asociado a g (que pueden ser, desde ningún espacio vectorial hasta infinitos de ellos) lo denotaremos por F. Estos espacios vectoriales también estan relacionados por productos tensoriales.

$$g(v_1, \dots, v_n) \in E_1^* \otimes \cdots \otimes E_n^* \otimes F = E_1^* \otimes \cdots \otimes E_*n^* \otimes V_1 \otimes \cdots \otimes V_m \qquad \text{Con } m \in \{0, \mathbb{N}\}$$

4.12.1. Aplicación multilineal mediante contracciones tensoriales

El cálculo de una aplicación multilineal mediante contracción tensorial se lleva a cabo de la siguiente manera:

DESARROLLO

Dada una aplicación multilineal $g(\bar{v}_1, \dots, \bar{v}_n)$:

$$g(\bar{v}_1, \dots, \bar{v}_n) : E_1 \times \cdots \times E_n \longmapsto F \quad \in \quad E_1^* \otimes \cdots \otimes E_n^* \otimes F$$

Nota. F puede ser $\mathbb{R}$, un espacio vectorial, o varios espacios vectoriales (multiplicados entre sí tensorialmente) (y posiblemente intercalados con los espacios duales). Por ejemplo, en el caso de una aplicación lineal $Hom(E_1, E_2)$ (que hemos visto que se pueden entender como un tensor $E_2 \otimes E_1^*$), F es el espacio E_2.

1. Se obtine un tensor T asociado a los vectores $v_1, \dots, v_n$ que va a tomar la aplicación multilineal g:
$$T = \bar{v}_1 \otimes \cdots \otimes \bar{v}_n$$
2. Se obtine un tensor G asociado a la forma multilineal. Este tensor internamente está formado por vectores $\bar{v}_i^*$ contenidos en los espacios duales de los vectores de T y por los posibles vectores no duales $\bar{u}_i$ contenidos en F:
$$G = \bar{v}_1^* \otimes \cdots \otimes \bar{v}_n^* \otimes T_F$$
Nota. Como el número de vectores $\bar{u}_i$ es indefinido, denotaremos a $\bar{u}_1 \otimes \cdots \otimes \bar{u}_n$ como el tensor T_F
$$T_F = \bar{u}_1 \otimes \cdots \otimes \bar{u}_n$$
Nota. Nótese que T_F puede no estar formado por ningún vector y ser solo un escalar (incluyendo el 1).
3. Se juntan ambos tensores en uno solo, el tensor $T \otimes G$:
$$T \otimes G = \bar{v}_1 \otimes \cdots \otimes \bar{v}_n \otimes \bar{v}_1^* \otimes \cdots \otimes \bar{v}_n^* \otimes T_F$$
4. Se realizan todas las contracciones de cada vector $\bar{v}_i$ con el correspondiente vector $\bar{v}_i^*$ de su dual:
$$g(v_1, \dots, v_n) = ev_{i_n, j_n}\big(\dots ev_{i_1, j_1}\big(\bar{v}_1 \otimes \cdots \otimes \bar{v}_n \otimes \bar{v}_1^* \otimes \cdots \otimes \bar{v}_n^* \otimes T_F\big)\big)$$

Véase que, de esta forma, hemos sustituido una aplicación multilineal g por, primero una transformación en tensores de los vectores de entrada y de la propia aplicación (tensorización) y después un conjunto de aplicaciones lineales (evaluaciones) utilizando dichos tensores.

Ejemplo
Calcula la imagen de los vectores $\bar{v}_1 = (1,1)$ y $\bar{v}_2 = (1,2,1)$ respecto de la forma bilineal g de matriz asociada: $\text{M} = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & -1 \end{pmatrix}$

- RESOLUCIÓN UTILIZANDO LA MATRIZ DE LA FORMA BILINEAL:
$$(1,1)\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = (1,1)\begin{pmatrix} 3 \\ 2 \end{pmatrix} = 5$$
- RESOLUCIÓN UTILIZANDO CONTRACCIONES TENSORIALES:

1 - Obtenemos el tensor de los vectores de entrada:

$T = \underset{1}{(1,1)} \otimes \underset{2}{(1,2,1)} \in \underset{1}{\mathbb{R}^2} \otimes \underset{2}{\mathbb{R}^3}$

2 - Obtenemos el tensor de la aplicación multilineal:

$g : \mathbb{R}^2 \times \mathbb{R}^3 \longmapsto \mathbb{R}$

$g \in Hom(\mathbb{R}^2 \times \mathbb{R}^3, \mathbb{R}) \cong (\mathbb{R}^2 \otimes \mathbb{R}^3)^* \cong \underset{3}{\mathbb{R}^{2*}} \otimes \underset{4}{\mathbb{R}^{3*}}$

$G = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}(1,0,2) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(1,1,-1)\right] = \underset{3}{(1,0)} \otimes \underset{4}{(1,0,2)} + \underset{3}{(0,1)} \otimes \underset{4}{(1,1,-1)}$

Juntamos los 2 tensores anteriores en uno solo, y contraemos cada vector con su dual:

$T \otimes G = \underset{1}{(1,1)} \otimes \underset{2}{(1,2,1)} \otimes \big[\underset{3}{(1,0)} \otimes \underset{4}{(1,0,2)} + \underset{3}{(0,1)} \otimes \underset{4}{(1,1,-1)}\big] \in \underset{1}{\mathbb{R}^2} \otimes \underset{2}{\mathbb{R}^3} \otimes \underset{3}{\mathbb{R}^{2*}} \otimes \underset{4}{\mathbb{R}^{3*}}$

$ev_{1,3}\Big(\underset{1}{(1,1)} \otimes \underset{2}{(1,2,1)} \otimes \big[\underset{3}{(1,0)} \otimes \underset{4}{(1,0,2)} + \underset{3}{(0,1)} \otimes \underset{4}{(1,1,-1)}\big]\Big) \Rightarrow \left\{ \begin{array}{l} (1,0)\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \\ (0,1)\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \end{array} \right\} \Rightarrow$

$\Longrightarrow \underset{1}{(1,2,1)} \otimes \big[1 \cdot \underset{2}{(1,0,2)} + 1 \cdot \underset{2}{(1,1,-1)}\big] \in \underset{1}{\mathbb{R}^3} \otimes \underset{2}{\mathbb{R}^{3*}}$

$ev_{1,2}\Big(\underset{1}{(1,2,1)} \otimes \big[1 \cdot \underset{2}{(1,0,2)} + 1 \cdot \underset{2}{(1,1,-1)}\big]\Big) \Rightarrow \big(1 \cdot 3 + 1 \cdot 2\big) = 5$

Ejemplo
Calcula la imagen de los vectores $\bar{v}_1 = (1,0,0)$ y $\bar{v}_2 = (1,-1)$ respecto de la aplicación multilineal descrita por el tensor G:

$$G = \left[G_{1,1,:,:}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \quad G_{1,2,:,:}\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \quad G_{1,3,:,:}\begin{pmatrix} 3 & 2 \\ 1 & 3 \end{pmatrix}\right] \in \mathbb{R}^{3*} \otimes \mathbb{R}^{2*} \otimes \mathbb{R}^2$$

- RESOLUCIÓN UTILIZANDO CONTRACCIONES TENSORIALES:

1 - Obtenemos el tensor de los vectores de entrada:

$T = \underset{1}{(1,1)} \otimes \underset{2}{(1,2,1)} \in \underset{1}{\mathbb{R}^2} \otimes \underset{2}{\mathbb{R}^3}$

2 - Obtenemos el tensor de la aplicación multilineal:

$g : \mathbb{R}^3 \times \mathbb{R}^2 \longmapsto \mathbb{R}^2$

$g \in Hom(\mathbb{R}^3 \times \mathbb{R}^2, \mathbb{R}^2) \cong (\mathbb{R}^3 \otimes \mathbb{R}^2)^* \otimes \mathbb{R}^2 \cong \underset{3}{\mathbb{R}^{3*}} \otimes \underset{4}{\mathbb{R}^{2*}} \otimes \underset{5}{\mathbb{R}^2}$

$G = \Big[(1,0,0) \otimes \Big[\begin{pmatrix} 1 \\ 0 \end{pmatrix}(1,2) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(2,-1)\Big] + (0,1,0) \otimes \Big[\begin{pmatrix} 1 \\ 0 \end{pmatrix}(2,0) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(1,1)\Big] +$

$+ (0,0,1) \otimes \Big[\begin{pmatrix} 1 \\ 0 \end{pmatrix}(3,2) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(1,3)\Big]\Big] =$

$= \underset{3}{(1,0,0)} \otimes \big[\underset{4}{(1,0)} \otimes \underset{5}{(1,2)} + \underset{4}{(0,1)} \otimes \underset{5}{(2,-1)}\big] + \underset{3}{(0,1,0)} \otimes \big[\underset{4}{(1,0)} \otimes \underset{5}{(2,0)} + \underset{4}{(0,1)} \otimes \underset{5}{(1,1)}\big] +$

$+ \underset{3}{(0,0,1)} \otimes \big[\underset{4}{(1,0)} \otimes \underset{5}{(3,2)} + \underset{4}{(0,1)} \otimes \underset{5}{(1,3)}\big]$

3 - Juntamos los 2 tensores anteriores en uno solo, y contraemos cada vector con su dual:

$T \otimes G \in \underset{1}{\mathbb{R}^3} \otimes \underset{2}{\mathbb{R}^2} \otimes \underset{3}{\mathbb{R}^{3*}} \otimes \underset{4}{\mathbb{R}^{2*}} \otimes \underset{5}{\mathbb{R}^2}$

$T \otimes G = (1,0,0) \otimes (1,-1) \otimes \Big[\underset{3}{(1,0,0)} \otimes \big[\underset{4}{(1,0)} \otimes \underset{5}{(1,2)} + \underset{4}{(0,1)} \otimes \underset{5}{(2,-1)}\big] + \underset{3}{(0,1,0)} \otimes$

$\otimes \big[\underset{4}{(1,0)} \otimes \underset{5}{(2,0)} + \underset{4}{(0,1)} \otimes \underset{5}{(1,1)}\big] + \underset{3}{(0,0,1)} \otimes \big[\underset{4}{(1,0)} \otimes \underset{5}{(3,2)} + \underset{4}{(0,1)} \otimes \underset{5}{(1,3)}\big]\Big]$

$ev_{1,3}\big(T \otimes G\big) \Longrightarrow 1 \cdot \underset{1}{(1,-1)} \otimes \big[\underset{2}{(1,0)} \otimes \underset{3}{(1,2)} + \underset{2}{(0,1)} \otimes \underset{3}{(2,-1)}\big] \in \underset{1}{\mathbb{R}^2} \otimes \underset{2}{\mathbb{R}^{2*}} \otimes \underset{3}{\mathbb{R}^2}$

$ev_{1,2}\Big(\underset{1}{(1,-1)} \otimes \big[\underset{2}{(1,0)} \otimes \underset{3}{(1,2)} + \underset{2}{(0,1)} \otimes \underset{3}{(2,-1)}\big]\Big) \Rightarrow \big(1 \cdot (1,2) + (-1) \cdot (2,-1)\big) = (-1,3)$

Ejemplo (Caso particular de una aplicación lineal)
Calcula la imagen del vector $\bar{v}_1 = (2,1,3)$ respecto de la aplicación lineal f de matriz asociada M:

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

- RESOLUCIÓN UTILIZANDO LA MATRIZ DE LA APLICACIÓN LINEAL:
$$\begin{pmatrix} 1 & 3 & 2 \\ 1 & 3 & 0 \\ 4 & 1 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 5 \\ 15 \end{pmatrix}$$
- RESOLUCIÓN UTILIZANDO CONTRACCIONES TENSORIALES:

1 - Obtenemos el tensor del vector de entrada:

$T = \underset{1}{(2,1,3)} \in \underset{1}{\mathbb{R}^3}$

2 - Obtenemos el tensor de la aplicación multilineal:

$g : \mathbb{R}^3 \longmapsto \mathbb{R}^3$

$g \in End(\mathbb{R}^3) = Hom(\mathbb{R}^3, \mathbb{R}^3) \cong \underset{2}{\mathbb{R}^3} \otimes \underset{3}{\mathbb{R}^{3*}}$

$G = \left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}(1,3,2) + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}(1,3,0) + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}(4,1,2)\right]$

$G = \underset{2}{(1,0,0)} \otimes \underset{3}{(1,3,2)} + \underset{2}{(0,1,0)} \otimes \underset{3}{(1,3,0)} + \underset{2}{(0,0,1)} \otimes \underset{3}{(4,1,2)}$

Juntamos los 2 tensores anteriores en uno solo, y contraemos cada vector con su dual:

$T \otimes G = \underset{1}{(2,1,3)} \otimes \big[\underset{2}{(1,0,0)} \otimes \underset{3}{(1,3,2)} + \underset{2}{(0,1,0)} \otimes \underset{3}{(1,3,0)} + \underset{2}{(0,0,1)} \otimes \underset{3}{(4,1,2)}\big] \in \underset{1}{\mathbb{R}^3} \otimes \underset{2}{\mathbb{R}^3} \otimes \underset{3}{\mathbb{R}^{3*}}$

$ev_{1,3}\Big(\underset{1}{(2,1,3)} \otimes \big[\underset{2}{(1,0,0)} \otimes \underset{3}{(1,3,2)} + \underset{2}{(0,1,0)} \otimes \underset{3}{(1,3,0)} + \underset{2}{(0,0,1)} \otimes \underset{3}{(4,1,2)}\big]\Big) \Longrightarrow$

$\Longrightarrow \; 11 \cdot (1,0,0) + 5 \cdot (0,1,0) + 15 \cdot (0,0,1) = (11,5,15)$

4.12.2. Aplicación multilineal mediante tensores en la forma canónica

El cálculo de una aplicación multilineal mediante los tensores en su forma canónica se lleva a cabo de la siguiente manera:

DESARROLLO

Dada una aplicación multilineal $g(\bar{v}_1, \dots, \bar{v}_n)$:

$$g(\bar{v}_1, \dots, \bar{v}_n) : E_1 \times \cdots \times E_n \longmapsto F \quad \in \quad E_1^* \otimes \cdots \otimes E_n^* \otimes F$$

Nota. F puede ser $\mathbb{R}$, un espacio vectorial, o un producto tensorial de espacios vectoriales

1. Se obtine un tensor T de los vectores $v_1, \dots, v_n$ que va a tomar la aplicación multilineal g. En este caso, representamos el vector en su forma canónica:
$$T = \bar{v}_1 \otimes \cdots \otimes \bar{v}_n = \big(v_{1,1} \cdot (\dots (\bar{v}_n)) \;\big|\; \dots \;\big|\; v_{1,n} \cdot (\dots (\bar{v}_n))\big)$$
2. Se obtine un tensor G de la forma multilineal. Este tensor internamente está formado por vectores $\bar{v}_i^*$ contenidos en los espacios duales de los vectores de T y por los posibles vectores $\bar{u}_i$ contenidos en F.

   Para representar un tensor G de una **aplicación multilineal g en su forma canónica** se deben tener en cuenta una serie de peculiaridades.

   - Los vectores $\bar{v}_i^*$ pertenecientes a los espacios duales E_i^* amplían la matriz de la forma canónica del tensor de manera horizontal.
   - Los vectores $\bar{u}_i$ pertenecientes a los posibles espacios vectoriales F amplían la matriz de la forma canónica del tensor de manera vertical.

$$\text{Sea} \quad G = \bar{v}_1^* \otimes \cdots \otimes \bar{v}_n^* \otimes T_F \qquad \text{Donde } T_F \text{ es el tensor de todos los vectores de } F$$
$$\text{Sea} \quad T_F = \bar{u}_1 \otimes \cdots \otimes \bar{u}_n = (u_{1,1}, \dots, u_{1,m_1}) \otimes \cdots \otimes (u_{n,1}, \dots, u_{n,m_n})$$

$$G = \begin{bmatrix} \begin{bmatrix} u_{1,1} \cdot u_{n,1} \cdot \big(v_{1,1}^* \cdot (\dots (\bar{v}_n^*)) \;\big|\; \dots \;\big|\; v_{1,n} \cdot (\dots (\bar{v}_n))\big) \\ \vdots \\ u_{1,m_1} \cdot u_{n,1} \cdot \big(v_{1,1}^* \cdot (\dots (\bar{v}_n^*)) \;\big|\; \dots \;\big|\; v_{1,n}^* \cdot (\dots (\bar{v}_n^*))\big) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} u_{1,1} \cdot u_{n,m_n} \cdot \big(v_{1,1}^* \cdot (\dots (\bar{v}_n^*)) \;\big|\; \dots \;\big|\; v_{1,n}^* \cdot (\dots (\bar{v}_n^*))\big) \\ \vdots \\ u_{1,m_1} \cdot u_{n,m_n} \cdot \big(v_{1,1}^* \cdot (\dots (\bar{v}_n^*)) \;\big|\; \dots \;\big|\; v_{1,n}^* \cdot (\dots (\bar{v}_n^*))\big) \end{bmatrix} \end{bmatrix}$$

Ejemplo
Dada una aplicación bilineal g de tensor asociado G, representar el tensor de la aplicación de forma canónica:

$$G = (-1,1) \otimes (1,2) \otimes (2,3) \otimes (1,1) \in \underset{1}{\mathbb{R}^{2*}} \otimes \underset{2}{\mathbb{R}^2} \otimes \underset{3}{\mathbb{R}^2} \otimes \underset{4}{\mathbb{R}^{2*}}$$

- El vector $(-1,1)$ está contenido en un espacio dual (se coloca en horizontal)
$$\big(-1, 1\big)$$
- El vector $(1,2)$ está contenido en un espacio no dual (amplía la matriz en vertical)
$$\begin{pmatrix} 1 \cdot (-1,1) \\ 2 \cdot (-2,3) \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix}$$
- El vector $(2,3)$ está contenido en un espacio no dual (amplía la matriz en vertical)
$$\begin{pmatrix} 2 \cdot \begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} \\ 3 \cdot \begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ -4 & 6 \\ -3 & 3 \\ -6 & 9 \end{pmatrix}$$
- El vector $(1,1)$ está contenido en un espacio dual (amplía la matriz en horizontal)
$$\left(1 \cdot \begin{pmatrix} -2 & 2 \\ -4 & 6 \\ -3 & 3 \\ -6 & 9 \end{pmatrix} \;\middle|\; 1 \cdot \begin{pmatrix} -2 & 2 \\ -4 & 6 \\ -3 & 3 \\ -6 & 9 \end{pmatrix} \right) = \begin{pmatrix} -2 & 2 & -2 & 2 \\ -4 & 6 & -4 & 6 \\ -3 & 3 & -3 & 3 \\ -6 & 9 & -6 & 9 \end{pmatrix}$$

Nota. Para que el tercer paso (la multiplicación matricial) tenga sentido los vectores duales del tensor de G deben multiplicarse (tensorialmente) en el mismo orden que se multiplican sus correspondientes vectores no duales en el tensor de T. En caso de que el orden no coincida deberemos modificar uno de los 2. A la acción de cambiar el orden de multiplicación tensorial de los vectores de un tensor se le denomina trenzado.

Ejemplo
Dada una aplicación bilineal $g(\bar{v}_1, \bar{v}_2)$ (de tensor asociado G) y el tensor T para evaluar en ella, hacer las reordenaciones pertinentes para poder hacer la evaluación usando los tensores en su forma canónica:

$$G = (3,1,6) \otimes (-2,-5) \otimes (3,-1,-3) \otimes (5,4) \in \mathbb{R}^{3*} \otimes \mathbb{R}^2 \otimes \mathbb{R}^3 \otimes \mathbb{R}^{2*}$$
$$T = (1,4) \otimes (0,1,1) \in \mathbb{R}^2 \otimes \mathbb{R}^3$$

- REORDENANDO EL TENSOR DE LA APLICACIÓN LINEAL

$\tau_{1,4}\Big((3,1,6) \otimes (-2,-5) \otimes (3,-1,-3) \otimes (5,4)\Big) = (5,4) \otimes (-2,-5) \otimes (3,-1,-3) \otimes (3,1,6)$

$G' = (5,4) \otimes (-2,-5) \otimes (3,-1,-3) \otimes (3,1,6) \in \mathbb{R}^{2*} \otimes \mathbb{R}^2 \otimes \mathbb{R}^3 \otimes \mathbb{R}^{3*}$

$T = (1,4) \otimes (0,1,1) \in \mathbb{R}^2 \otimes \mathbb{R}^3$

- REORDENANDO EL TENSOR DE LOS VECTORES DE ENTRADA

$\tau_{1,2}\Big((1,4) \otimes (0,1,1)\Big) = (0,1,1) \otimes (1,4)$

$G = (3,1,6) \otimes (-2,-5) \otimes (3,-1,-3) \otimes (5,4) \in \mathbb{R}^{3*} \otimes \mathbb{R}^2 \otimes \mathbb{R}^3 \otimes \mathbb{R}^{2*}$

$T' = (0,1,1) \otimes (1,4) \in \mathbb{R}^3 \otimes \mathbb{R}^2$

Nota. El orden de los vectores no duales de G influirá en el orden de los vectores en el tensor solución F, pero es irrelevante en que la multiplicación matricial tenga sentido o no.

3. Multiplicamos (ambos en su representación canónica) el tensor T de los vectores de entrada por el tensor G de la aplicación multilineal:
$$G \cdot T$$
Como los vectores $\bar{v}_i$ de un espacio E tienen la misma dimensión que los vectores $\bar{v}_i^*$ de su espacio dual E^* la dimensión horizontal de la matriz de la representación canónica de G siempre coincide con la dimensión del vector de la representación canónica de T. Como las dimeniones coinciden, siempre se puede multiplicar uno por el otro.
$$\underset{dim(F) \times dim(\bar{v}_1 \otimes \cdots \otimes \bar{v}_n)}{\begin{bmatrix} & & \\ & G & \\ & & \end{bmatrix}} \cdot \underset{dim(\bar{v}_1 \otimes \cdots \otimes \bar{v}_n) \times 1}{\begin{bmatrix} \\ T \\ \\ \end{bmatrix}} = \underset{dim(F) \times 1}{\begin{bmatrix} \\ F \\ \\ \end{bmatrix}}$$

Véase que, de esta forma, hemos sustituido una aplicación multilineal g por, primero una transformación en tensores de los vectores de entrada y de la propia aplicación (tensorización) y después una aplicación lineal utilizando dichos tensores en su representación canónica.

Ejemplo
Calcula la imagen de los vectores $v_1 = (1,1)$ y $v_2 = (1,2,1)$ respecto de la forma bilineal g de matriz asociada: $\text{M} = \left(\begin{smallmatrix} 1 & 0 & 2 \\ 1 & 1 & -1 \end{smallmatrix}\right)$

1. RESOLUCIÓN UTILIZANDO LA MATRIZ DE LA FORMA BILINEAL:
$$(1,1)\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = (1,1)\begin{pmatrix} 3 \\ 2 \end{pmatrix} = 5$$
2. RESOLUCIÓN MEDIANTE LA FORMA CANÓNICA DEL TENSOR:

$T = (1,1) \otimes (1,2,1) = (1,2,1,1,2,1)^T \in \mathbb{R}^2 \otimes \mathbb{R}^3$

$g : \mathbb{R}^2 \times \mathbb{R}^3 \longmapsto \mathbb{R}$

$g \in Hom(\mathbb{R}^2 \times \mathbb{R}^3, \mathbb{R}) \cong (\mathbb{R}^2 \otimes \mathbb{R}^3)^* \cong \mathbb{R}^{2*} \otimes \mathbb{R}^{3*}$

$G = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}(1,0,2) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(1,1,-1)\right] = (1,0) \otimes (1,0,2) + (0,1) \otimes (1,1,-1) = (1,0,2,1,1,-1)$

$$g(v_1, v_2) = G \cdot T = \big(1,0,2,1,1,-1\big)\begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \\ 2 \\ 2 \end{pmatrix} = 5$$

Ejemplo
Calcula la imagen de los vectores $\bar{v}_1 = (1,-1)$ y $\bar{v}_2 = (1,0,0)$ y respecto de la aplicación multilineal descrita por el tensor G:

$$G = \left[G_{1,1,:,:}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \quad G_{1,2,:,:}\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \quad G_{1,3,:,:}\begin{pmatrix} 3 & 2 \\ 1 & 3 \end{pmatrix}\right] \in \mathbb{R}^{3*} \otimes \mathbb{R}^{2*} \otimes \mathbb{R}^2$$

- RESOLUCIÓN UTILIZANDO TENSORES EN SU REPRESENTACIÓN CANÓNICA:

1 - Obtenemos el tensor de los vectores de entrada
(haciendo corresponder su orden con el de sus duales en G):

$T = (1,-1) \otimes (1,0,0) \xrightarrow{\text{cambiamos el orden}} \tau_{1,2}\Big((1,-1) \otimes (1,0,0)\Big) = (1,0,0) \otimes (1,-1)$

$T = (1,0,0) \otimes (1,-1) = (1,0,0,-1,0,0)^T \in \mathbb{R}^3 \otimes \mathbb{R}^2$

2 - Obtenemos el tensor de la aplicación multilineal:

$g : \mathbb{R}^3 \times \mathbb{R}^2 \longmapsto \mathbb{R}^2$

$g \in Hom(\mathbb{R}^3 \times \mathbb{R}^2, \mathbb{R}^2) \cong (\mathbb{R}^3 \otimes \mathbb{R}^2)^* \otimes \mathbb{R}^2 \cong \underset{1}{\mathbb{R}^{3*}} \otimes \underset{2}{\mathbb{R}^{2*}} \otimes \underset{3}{\mathbb{R}^2}$

$G = \Big[(1,0,0) \otimes \Big[\begin{pmatrix} 1 \\ 0 \end{pmatrix}(1,2) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(2,-1)\Big] + (0,1,0) \otimes \Big[\begin{pmatrix} 1 \\ 0 \end{pmatrix}(2,0) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(1,1)\Big] +$

$+ (0,0,1) \otimes \Big[\begin{pmatrix} 1 \\ 0 \end{pmatrix}(3,2) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(1,3)\Big]\Big] =$

$= \underset{1}{(1,0,0)} \otimes \big[\underset{2}{(1,0)} \otimes \underset{3}{(1,2)} + \underset{2}{(0,1)} \otimes \underset{3}{(2,-1)}\big] + \underset{1}{(0,1,0)} \otimes \big[\underset{2}{(1,0)} \otimes \underset{3}{(2,0)} + \underset{2}{(0,1)} \otimes \underset{3}{(1,1)}\big] +$

$+ \underset{1}{(0,0,1)} \otimes \big[\underset{2}{(1,0)} \otimes \underset{3}{(3,2)} + \underset{2}{(0,1)} \otimes \underset{3}{(1,3)}\big] \rightarrow \left\{ \begin{array}{l} \text{vector dual (expande en horizontal)} \\ \text{vector no dual (expande en vertical)} \end{array} \right.$

$G = \left[\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{pmatrix}\right] + \left[\begin{pmatrix} 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}\right.$

$\left. + \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}\right] + \left[\begin{pmatrix} 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}\right] =$

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

3 - Evaluamos el vector del tensor T en la aplicación de matriz asociada la representación canónica de G:

$$g(v_1, v_2) = G \cdot T = \begin{pmatrix} 1 & 2 & 3 & 2 & 1 & 1 \\ 2 & 0 & 2 & -1 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \in \mathbb{R}^2$$

4.13. Cambio de base de tensores

Al igual que cuando tratamos con espacios vectoriales trabajamos con los vectores en una base concreta, cuando tenemos un tensor todos los vectores que lo forman también se encuentran en una base. Puede darse una situación en la que nos surja la necesidad de cambiar estos vectores de base.

Para cambiar un tensor de base, lo más cómodo es trabajar con la representación del tensor en función de un producto tensorial, ya que de esta forma podemos cambiar de base los vectores uno por uno.

Ejemplo
Dado el tensor T en las bases $B_1 \otimes B_c(M_2)$, hallar este mismo tensor expresado en las bases $B_c(\mathbb{R}^2) \otimes B_2$:

$$T = (3, -1) \otimes (1, 2, 3, 4) \in \mathbb{R}^2 \times M_2$$

$$B_1 = \{(0, 1), (-2, 1)\}$$
$$B_2 = \{(0, 1, 1, 0), (1, 1, 1, 0), (2, 1, 2, 0), (1, 3, 2, 1)\}$$

1. Hallamos las dos matrices de cambio de base necesarias:

$$M_{B_c \leftarrow B_1} = \begin{pmatrix} 0 & -2 \\ 1 & 1 \end{pmatrix} \qquad M_{B_2 \leftarrow B_c} = \begin{pmatrix} -1 & 0 & 1 & -1 \\ 1 & 2 & -2 & -3 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

2. Cambiamos los vectores de base uno por uno empleando las matrices anteriores:

$$\begin{pmatrix} 0 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \qquad \begin{pmatrix} -1 & 0 & 1 & -1 \\ 1 & 2 & -2 & -3 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 12 \\ 18 \\ 17 \\ 4 \end{pmatrix}$$

3. Obtenemos el Tensor expresado en las nuevas bases $B_c(\mathbb{R}^2) \times B_2$:

$$T = (2, 2) \otimes (12, 18, 17, 4)$$

Si solo contamos con la representación canónica del tensor no podemos cambiar los vectores de base uno a uno. En su lugar, deberemos construir una aplicación multilineal de tensor asociado G que aplique el cambio de base sobre todos los vectores a la vez.

La aplicación multilineal G se obtiene a partir del producto tensorial ordenado de todas las matrices de cambio de base asociadas a cada uno de los vectores (no duales) que componen el tensor T. Si alguno de estos tensores no se quiere cambiar de base su matriz de cambio de base asociada será la identidad.

Nota. Una matriz de aplicación lineal (y por tanto también una de cambio de base) se puede considerar como un tensor $(E \otimes E'^*)$ expresado en su forma canónica. Por tanto, al ser multiplicada tensorialmente para dar lugar a un tensor mayor en forma canónica, expande en una dimensión horizontalmente y en otra verticalmente.

Ejemplo
Dado el tensor T en las bases $B_c(\mathbb{R}^2) \otimes B$, hallar este mismo tensor expresado en las bases $B_c(\mathbb{R}^2) \otimes B_c(\mathbb{R}^2)$:

$$T = (2, 3, 0, 1) \in \mathbb{R}^2 \otimes M_2$$

$$B = \{(-1, 1), (2, 3)\}$$

1. Hallamos las matrices de cambio de base necesarias (en este caso, una es la identidad):

$$M_{B_c \leftarrow B} = \begin{pmatrix} -1 & 2 \\ 1 & 3 \end{pmatrix} \qquad M_{B_c \leftarrow B_c} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

2. Construimos el tensor asociado a la aplicación multilineal de G :

$$G = \begin{pmatrix} -1 & 2 \\ 1 & 3 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 2 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

3. Multiplicamos G por T y obtenemos el tensor expresado en las nuevas bases $B_c(\mathbb{R}^2) \otimes B_c(\mathbb{R}^2)$:

$$\begin{pmatrix} -1 & 2 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 11 \\ 2 \\ 3 \end{pmatrix} \qquad T = (4, 11, 2, 3)^T$$

Capítulo 5

Espacio Euclídeo

5.1. Espacio euclídeo

Un espacio euclídeo es un espacio vectorial E sobre el que se ha definido un producto escalar ($\circ$). El producto escalar dota a los vectores del espacio vectorial E de las nociones geométricas de dirección y módulo (longitud).

5.1.1. Producto escalar

Sea E un espacio vectorial, se llama producto escalar sobre E a cualquier forma bilineal $\circ : E \times E \mapsto \mathbb{R}$ simétrica y definida positiva.

Sea E un espacio vectorial. Se dice que una forma bilineal $\circ : E \times E \mapsto \mathbb{R}$ es definida positiva si para todo $\bar{x} \in E$ con $\bar{x} \neq \bar{0}$:

$$g(\bar{x}, \bar{x}) > 0$$

Ejemplo
Producto escalar estándar en $\mathbb{R}^3$

$$\bar{u} \circ \bar{v} = (u_1, u_2, u_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = (u_1 \cdot v_1 + u_2 \cdot v_2 + u_3 \cdot v_3)$$

Otro producto escalar posible de $\mathbb{R}^3$

$$\bar{u} \circ \bar{v} = (u_1, u_2, u_3) \begin{pmatrix} 4 & 2 & 0 \\ 2 & 4 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \big(\bar{u}_1(4v_1 + 2v_2) + \bar{u}_2(2v_1 + 4v_2 + v_3) + \bar{u}_3(v_2 + v_3)\big)$$

Nota. La matriz del producto escalar estándar es la identidad.

5.1.2. Criterio de Sylvester

Es posible que solo contemos con la matriz de la forma bilineal en su forma matricial. En estos casos, para saber si es definida positiva, basta con aplicar el criterio de Sylvester.

El criterio de Sylvester nos dice que, sea G la matriz de una aplicación bilineal $g : E \times E \mapsto \mathbb{R}$ en una base B cualquiera, y sean $G_1, \ldots, G_n$ los menores principales de G, entonces g es definida positiva si y solo si:

$$|G_1| > 0 \; ; \; |G_2| > 0 \; ; \; |G_3| > 0 \; ; \; \ldots \; ; \; G_i > 0$$

Recíprocamente, decimos que g es definida negativa si $(-g)$ es definida positiva. Es decir, que una forma bilineal h es definida negativa si su matriz asociada H cumple:

$$|H_1| < 0 \; ; \; |H_2| > 0 \; ; \; |H_3| < 0 \; ; \; \ldots \; ; \; \begin{matrix} |H_i| > 0 \text{ si } i \text{ es par} \\ |H_i| < 0 \text{ si } i \text{ es impar} \end{matrix}$$

Nota. Ya que para una matriz $A_{n\times n}$ cualquiera: $\begin{matrix} |A| = |-A| \text{ si } n \text{ es par} \\ |A| = -|A| \text{ si } n \text{ es impar} \end{matrix}$

Si una matriz de una forma bilineal no es ni definida positiva ni definida negativa decimos que es una forma bilineal indefinida.

5.1.3. Norma (o módulo) de un vector

Dado un vector $\bar{u}$ de un espacio vectorial euclídeo, se define como norma o módulo de dicho vector $\bar{u}$ a la raiz cuadrada del producto escalar del vector por sí mismo, se denota por $|\bar{u}|$.

$$|\bar{u}| = \sqrt{|\bar{u}|^2} = \sqrt{\bar{u} \circ \bar{u}}$$

Al trabajar con espacios euclídeos de 2 o 3 dimensiones generalmente la norma se asocia al concepto de distancia entre los 2 puntos que delimitan el vector. Es por esto que decimos que el producto escalar dota a los vectores de un espacio vectorial de la noción de "distancia".

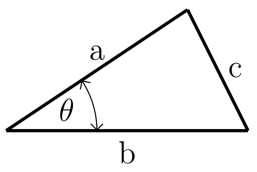
5.1.4. Ángulo entre dos vectores

Dados dos vectores $\bar{u}$ y $\bar{v}$ de un espacio vectorial euclídeo, la noción de "ángulo" entre dichos vectores nos indica la magnitud de la divergencia entre sus "direcciones". Su valor se obtiene a partir de la siguiente expresión:

$$\theta = \arccos\left(\frac{\bar{u} \circ \bar{v}}{|\bar{v}||\bar{u}|}\right)$$

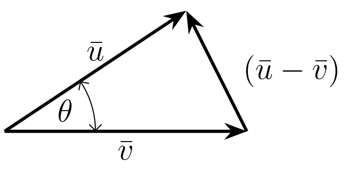
DEMOSTRACIÓN

El **teorema del coseno** nos permite conocer la longitud de un lado de un triángulo a partir de la longitud de los otros dos lados y la magnitud del ángulo opuesto.



$$c^2 = a^2 + b^2 - 2ab\,cos(\theta)$$

Podemos identificar cada uno de los lados del triángulo con un vector, y, dado que estamos trabajando en un espacio euclídeo, la longitud de los lados coincidirá con el módulo de los vectores.



$$|\bar{v} - \bar{u}|^2 - |\bar{v}|^2 - |\bar{u}|^2 = -2|\bar{v}||\bar{u}|cos(\theta)$$

Desarrollando el producto escacar de la diferencia y simplificando en la ecuación anterior se obtiene:

$$|\bar{v} - \bar{u}|^2 - |\bar{v}|^2 - |\bar{u}|^2 = (-2)|\bar{v}||\bar{u}|cos(\theta) \quad \rightarrow$$

$$\rightarrow \quad (\bar{v} - \bar{u}) \circ (\bar{v} - \bar{u}) - |\bar{v}|^2 - |\bar{u}|^2 = (-2)|\bar{v}||\bar{u}|cos(\theta) \quad \rightarrow$$

$$\rightarrow \quad \cancel{|\bar{v}|^2} - \bar{v} \circ \bar{u} - \bar{u} \circ \bar{v} + \cancel{|\bar{u}|^2} - \cancel{|\bar{v}|^2} - \cancel{|\bar{u}|^2} = (-2)|\bar{v}||\bar{u}|cos(\theta) \quad \rightarrow$$

$$\rightarrow \quad \cancel{(-2)}(\bar{u} \circ \bar{v}) = \cancel{(-2)}|\bar{v}||\bar{u}|cos(\theta)$$

Y finalmente, despejando el ángulo θ llegamos a la expresión del ángulo entre dos vectores:

$$\cancel{-2}(\bar{u} \circ \bar{v}) = \cancel{-2}|\bar{v}||\bar{u}|cos(\theta) \quad \rightarrow$$

$$\rightarrow \quad \frac{(\bar{u} \circ \bar{v})}{|\bar{v}||\bar{u}|} = cos(\theta) \quad \rightarrow$$

$$\rightarrow \quad \theta = \arccos\left(\frac{\bar{u} \circ \bar{v}}{|\bar{v}||\bar{u}|}\right)$$

Como se puede observar en el desarrollo, para obtener el ángulo es necesario contar primero con la noción de módulo del vector. Es por esto que, indirectamente, también es el producto escalar del espacio el que dota al vector de "dirección".

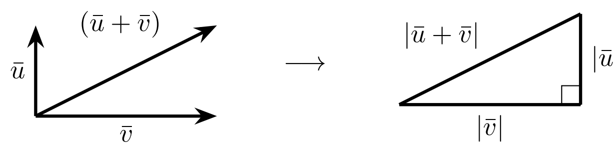
5.2. Perpendicularidad en un espacio euclídeo

Cuando trabajamos en un **espacio euclídeo** decimos que dos vectores $\bar{u}$ y $\bar{v}$ son perpendiculares entre sí si verifican que:

$$\bar{u} \circ \bar{v} = 0$$

DEMOSTRACIÓN

Sean $\bar{u}$ y $\bar{v}$ dos vectores distintos de $\bar{0}$, según el teorema de Pitágoras se cumple que:



$$\bar{u} \perp \bar{v} \quad \Leftrightarrow \quad |\bar{u}+\bar{v}|^2 = |\bar{u}|^2 + |\bar{v}|^2$$

Manipulando la ecuación anterior podemos obtener lo siguiente:

$$\begin{aligned} |\bar{u}+\bar{v}|^2 = |\bar{u}|^2 + |\bar{v}|^2 \quad &\longrightarrow \quad |\bar{u}+\bar{v}|^2 - |\bar{u}|^2 - |\bar{v}|^2 = 0 \\ \longrightarrow \quad &(\bar{u}+\bar{v}) \circ (\bar{u}+\bar{v}) - \bar{u} \circ \bar{u} - \bar{v} \circ \bar{v} \;\; = 0 \\ \longrightarrow \quad &\bar{u} \circ \bar{u} + \bar{u} \circ \bar{v} + \bar{v} \circ \bar{v} + \bar{v} \circ \bar{u} - \bar{u} \circ \bar{u} - \bar{v} \circ \bar{v} = 0 \\ 2(\bar{u} \circ \bar{v}) = 0 \quad &\Longleftrightarrow \quad \bar{u} \circ \bar{v} = 0 \end{aligned}$$

Por tanto, dos vectores son perpendiculares entre sí solo si su producto escalar es igual a 0.

$\square$

Como la noción geométrica de perpendicularidad implica independencia lineal, podemos decir que si en un conjunto de vectores $\{\bar{x}_1, \ldots, \bar{x}_n\}$ son todos ellos perpendiculares entre sí, entonces también forman una base.

DEMOSTRACIÓN

Dados un conjunto de vectores $\{\bar{x}_1, \ldots, \bar{x}_n\}$ distintos de $\bar{0}$ que verifican que:

$$\bar{x}_i \circ \bar{x}_j = 0 \quad \forall i, j \;\; tq. \;\; i \neq j$$

Si los multiplicamos por un conjunto de escalares λ_i cumpliéndose que:

$$\lambda_1 \bar{x}_1 + \cdots + \lambda_i \bar{x}_i + \cdots + \lambda_n \bar{x}_n = \bar{0}$$

Multiplicando por $\bar{x}_i$ toda la expresión obtenemos:

$$\begin{aligned} \bar{0} \;\; &= (\lambda_1 \bar{x}_1 + \cdots + \lambda_i \bar{x}_i + \cdots + \lambda_n \bar{x}_n) \circ \bar{x}_i = \\ &= \lambda_1 (\bar{x}_1 \circ \bar{x}_i) + \cdots + \lambda_i (\bar{x}_i \circ \bar{x}_i) + \cdots + \lambda_n (\bar{x}_n \circ \bar{x}_i) = \\ &= 0 + \cdots + 0 + \lambda_i (\bar{x}_i)^2 + 0 + \cdots + 0 = \bar{0} \quad \Longleftrightarrow \quad \lambda_i = 0 \quad \longrightarrow \quad \text{Def. de sistema libre} \end{aligned}$$

Por tanto podemos concluir que si se verifica $\bar{x}_i \circ \bar{x}_j = \bar{0}$ las condiciones iniciales el sistema forma una base.

$\square$

5.3. Relación entre ortogonalidad y anuladores

Existe una estrecha relación entre los vectores ortogonales a un vector dado y el anulador de dicho vector. Para ayudamos en la demostración de esta relación antes vamos a definir la aplicación **bemol**.

DEFINICIÓN DE BEMOL

Tenemos que un producto escalar es una forma bilineal de $E \times E$ en $\mathbb{R}$:

$$\begin{aligned} \circ : E \times E &\longmapsto \mathbb{R} \\ \circ(u, v) &\longmapsto \bar{u} \circ \bar{v} \end{aligned}$$

En un espacio euclídeo de **producto escalar ESTÁNDAR** se define como bemol en $\bar{v}$ una aplicación lineal que toma vectores $\bar{x}$ y devuelve su producto escalar con un vector fijo $\bar{v}$. Lo denotaremos por $b[\bar{v}]$:

$$\begin{aligned} \flat[\bar{v}] : E &\longmapsto \mathbb{R} \\ \flat[\bar{v}](\bar{x}) &= \bar{x} \circ \bar{v} \end{aligned}$$

Veamos que, dado que estamos trabajando con el producto escalar estándar, la aplicación bemol en $\bar{v}$ es también un vector contenido en el espacio dual del espacio euclídeo.

Nota. Se define como **bemol** $\flat$ al isomorfismo que lleva un vector $\bar{v}$ a su bemol correspondiente $\flat[\bar{v}]$.

$$\begin{aligned} \flat : E &\longmapsto E^* \\ \flat(\bar{v}) &= \flat[\bar{v}] \end{aligned}$$

DEMOSTRACIÓN

Suponiendo que $\circ : E \times E \longmapsto \mathbb{R}$ sea un producto escalar estándar y que los vectores $\bar{x}$ y $\bar{y}$:

$$\bar{x} \circ \bar{y} = (x_1, \ldots, x_n)_B \circ (y_1, \ldots, y_n)_B = 0$$

Entonces, empleando el bemol en $\bar{x}$ tenemos que:

$$b[\bar{x}](\bar{y}) = \big(x_1, \ldots, x_n\big)_{B^*} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_B = 0 \iff b[\bar{x}] \in \bar{y}^{\perp}$$

Por tanto, en un espacio euclídeo de producto escalar estándar, si el producto de dos vectores es 0, entonces ambos están mutuamente contenidos en el anulador del otro (y por tanto en el espacio dual).

$\square$

Ejemplo
Dado el vector de $\mathbb{R}^3$ $v = (1, 2, 5)$, su bemol correspondiente será:

$$b[\bar{v}](x, y, z) = \bar{v} \circ (x, y, z) = (1, 2, 5) \circ (x, y, z) = x + 2 \cdot y + 5 \cdot z$$
$$b[v] = (1, 2, 5)_{B_{\mathbb{R}^{3*}}} \in \mathbb{R}^{3*}$$

5.4. Bases ortogonales y ortonormales de un espacio euclídeo

Toda base de un espacio vectorial no euclídeo (Es decir, de un espacio en el que no existe la noción de ángulo entre vectores ni de módulo) sin importar su conformación puede adquirir ciertas cualidades de ortogonalidad y longitud al definirse para el espacio un producto escalar concreto. Es decir, cualquier base puede ser ortogonal u ortonormal si se define el producto escalar adecuado.

5.4.1. Base ortogonal de un espacio euclídeo

Una base de un espacio vectorial euclídeo es ortogonal (BO) si todos sus vectores son ortogonales entre sí.

$$BO = \{\bar{v}_k\}_{k=1}^{n} \quad tq. \quad \bar{v}_i \circ \bar{v}_j = \begin{cases} \lambda_i \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases} \quad \text{Siendo } \lambda_i \text{ un escalar cualquiera}$$

DETERMINACIÓN DE UN PRODUCTO ESCALAR PARA OBTENER BASE ORTOGONAL

Para convertir una base en ortogonal debemos definir el producto escalar forzando las condiciones anteriores. Es decir, un producto escalar que, al multiplicar dos vectores distíntos dé la base de como resultado 0, y que al multiplicar un vector de la base por sí mismo dé un escalar λ_i.

1. Para imponer las condiciones a toda la combinatoria de vectores de la base al mismo tiempo podemos definir la matriz $B = \left(\bar{v}_1 \,\middle|\, \dots \,\middle|\, \bar{v}_n \right)$ y multiplicarla (escalarmente) por ella misma traspuesta.

$$B^T \circ B \to B^T M(\circ) B = \begin{pmatrix} v_{1,1} & \dots & v_{1,n} \\ \vdots & \ddots & \vdots \\ v_{n,1} & \dots & v_{n,n} \end{pmatrix} \begin{pmatrix} & M(\circ) & \end{pmatrix} \begin{pmatrix} v_{1,1} & \dots & v_{n,1} \\ \vdots & \ddots & \vdots \\ v_{1,n} & \dots & v_{n,n} \end{pmatrix} = \begin{pmatrix} \bar{v}_1 \circ \bar{v}_1 & \dots & \bar{v}_1 \circ \bar{v}_n \\ \vdots & \ddots & \vdots \\ \bar{v}_n \circ \bar{v}_1 & \dots & \bar{v}_n \circ \bar{v}_n \end{pmatrix}$$

2. A continuación forzamos que el resultado sea el que nos garantiza la ortogonalidad de la base.

$$\begin{pmatrix} \bar{v}_1 \circ \bar{v}_1 & \dots & \bar{v}_1 \circ \bar{v}_n \\ \vdots & \ddots & \vdots \\ \bar{v}_n \circ \bar{v}_1 & \dots & \bar{v}_n \circ \bar{v}_n \end{pmatrix} \xrightarrow{\bar{v}_i \circ \bar{v}_j = \begin{cases} \lambda_i \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix} = A$$

3. Por último, despejamos en la ecuación la matriz del producto escalar que hace ortogonal a la base.

$$B^T M(\circ) B = A \quad \longrightarrow \quad M(\circ) = (B^T)^{-1} A (B)^{-1}$$

5.4.2. Base ortonormal de un espacio euclídeo

Una base de un espacio vectorial euclídeo es ortonormal (BON) si es ortogonal y además todos sus vectores son unitarios.

$$BON = \{\bar{v}_k\}_{k=1}^{n} \quad tq. \quad \bar{v}_i \circ \bar{v}_j = \delta_{i,j} \quad \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

Nota. La base canónica de cualquier espacio vectorial es ortonormal con el producto estándar.

DETERMINACIÓN DE UN PRODUCTO ESCALAR PARA OBTENER BASE ORTONORMAL

De manera análoga al caso de la conversión de una base en ortogonal, para convertir una base en ortonormal debemos definir un producto escalar que fuerce que los vectores de la base cumplan las características necesarias. Es decir, un producto escalar que, al multiplicar dos vectores distíntos de la base de 0 y que al multiplicar un vector de la base por sí mismo de 1.

1. Veamos que para que ambas condiciones se satisfagan se debe cumplir la siguiente ecuación.

$$B^T M(\circ) B = Id \quad \longrightarrow \quad \begin{pmatrix} v_{1,1} & \dots & v_{1,n} \\ \vdots & \ddots & \vdots \\ v_{n,1} & \dots & v_{n,n} \end{pmatrix} \begin{pmatrix} & M(\circ) & \end{pmatrix} \begin{pmatrix} v_{1,1} & \dots & v_{n,1} \\ \vdots & \ddots & \vdots \\ v_{1,n} & \dots & v_{n,n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

2. Despejamos en la ecuación la matriz $M(\circ)$ del producto escalar que hace ortonormal a la base.

$$B^T M(\circ) B = Id \quad \longrightarrow \quad M(\circ) = (B^T)^{-1} (B)^{-1}$$

5.4.3. Algoritmo de Gram Schmidt

El algoritmo de Gram Schmidt nos permite, dados una base de un espacio o subespacio vectorial y un producto escalar, obtener la base ortogonal u ortonormal más similar a la base. La base resultante siempre genera el mismo espacio o subespacio que la base original.

La lógica seguida para la obtención de la base ortogonal consiste en ir rotando los vectores uno a uno hasta posiciones ortogonales a todos los vectores ya rotados previamente. Finalmente, si queremos una base ortonormal simplemente normalizamos los vectores.

Partiendo de la base original B del espacio, deberemos llevar a cabo los siguientes pasos:

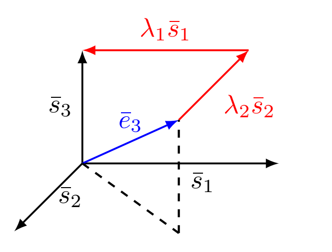
$$B = \{\bar{e}_1, \dots, \bar{e}_n\} \xrightarrow{\text{ortogonalización}} BO = \{\bar{s}_1, \dots, \bar{s}_n\} \xrightarrow{\text{normalización}} BON = \left\{ \frac{\bar{s}_1}{|\bar{s}_1|}, \dots, \frac{\bar{s}_n}{|\bar{s}_n|} \right\}$$

PROCEDIMIENTO DEL ALGORITMO

El primer vector ($\bar{e}_1$) no tenemos por que rotarlo, ya que, al no tener todavía ninguno colocado en su posición final, no tenemos ninguno cuya ortogonalidad respecto a $\bar{e}_1$ tengamos que garantizar:

$$\bar{e}_1 = \bar{s}_1$$

Para cualquier otro vector ($\bar{e}_i$) tenemos que garantizar la ortogonalidad con todos los vectores ya colocados. Para hacer esto veamos que $\bar{s}_i$ (perpendicular a todos los $\bar{s}_{i-k}$ $[\forall k \leq i]$ anteriores) se puede expresar en función del vector $\bar{e}_i$ menos una combinación de los $\{\bar{s}_j\}_{j=1}^{i-1}$ anteriores:



$$\bar{s}_i = \bar{e}_i - \lambda_1 \cdot \bar{s}_1 - \dots - \lambda_j \bar{s}_j - \dots - \lambda_{i-1} \bar{s}_{i-1}$$

Para obtener el valor de los escalares λ_j debemos aplicar el siguiente desarrollo, construido a partir de la búsqueda de perpendicularidad entre $\bar{s}_i$ y los vectores $\{\bar{s}_j\}_{j=1}^{i-1}$ ya rotados:

$$\bar{s}_j \circ \bar{s}_i = 0 \quad \to \quad \bar{s}_j \circ (\bar{e}_i - \lambda_j \bar{s}_j) = 0 \quad \to \quad \bar{s}_j \circ \bar{e}_i - \lambda_j (\bar{s}_j \circ \bar{s}_j) = 0$$
$$\longrightarrow \quad \bar{s}_j \circ \bar{e}_i = \lambda_j (\bar{s}_j \circ \bar{s}_j) \quad \longrightarrow \quad \lambda_j = \frac{\bar{e}_i \circ \bar{s}_j}{\bar{s}_j \circ \bar{s}_j}$$

Finalmente, aplicando la equivalencia anterior, obtenemos la expresión de $\bar{s}_i$ en función de los vectores $\{\bar{s}_j\}_{j=1}^{i-1}$:

$$\bar{s}_i = \bar{e}_i - \frac{\bar{e}_i \circ \bar{s}_1}{\bar{s}_1 \circ \bar{s}_1} \bar{s}_1 - \dots - \frac{\bar{e}_i \circ \bar{s}_j}{\bar{s}_j \circ \bar{s}_j} \bar{s}_j - \dots - \frac{\bar{e}_i \circ \bar{s}_{i-1}}{\bar{s}_{i-1} \circ \bar{s}_{i-1}} \bar{s}_{i-1}$$

Una vez que hemos garantizado la ortogonalidad de todos los vectores (Y por tanto obtenido una base ortogonal), podemos obtener una base ortonormal dividiendo cada vector por su norma:

$$BO = \{\bar{s}_1, \dots, \bar{s}_n\} \xrightarrow{\text{normalización}} BON = \left\{ \frac{\bar{s}_1}{|\bar{s}_1|}, \dots, \frac{\bar{s}_n}{|\bar{s}_n|} \right\}$$

Ejemplo
Sea E un espacio euclídeo de producto vectorial estándar y $B = \{\bar{e}_1 = (0,2,0), \bar{e}_2 = (4,-1,-3), \bar{e}_3 = (3,2,4)\}$ una base de E, obtener la base ortonormal más próxima a B:

1. Tomamos $\bar{s}_1$ como $\bar{e}_1$ ya que, al ser el primer vector, no hay que garantizar ninguna ortogonalidad:

$$\bar{s}_1 = \bar{e}_1 = (0,2,0)$$

2. Calculamos $\bar{s}_2$ garantizando la ortogonalidad con $\bar{s}_1$:

$$\bar{s}_2 = \bar{e}_2 - \frac{\bar{e}_2 \circ \bar{s}_1}{\bar{s}_1 \circ \bar{s}_1} \cdot \bar{s}_1 = (4,-1,-3) - \frac{(-2)}{4}(0,2,0) = (4,-1,-3) + (0,1,0) = (4,0,-3)$$

3. Calculamos $\bar{s}_3$ garantizando la ortogonalidad con $\bar{s}_1$ y con $\bar{s}_2$:

$$\bar{s}_2 = \bar{e}_3 - \frac{\bar{e}_3 \circ \bar{s}_1}{\bar{s}_1 \circ \bar{s}_1} \cdot \bar{s}_1 - \frac{\bar{e}_3 \circ \bar{s}_2}{\bar{s}_2 \circ \bar{s}_2} \cdot \bar{s}_2 =$$
$$= (3,2,4) - \frac{4}{4}(0,2,0) - \frac{0}{25}(0,2,0) = (3,2,4) - (0,2,0) - (0,0,0) = (3,3,4)$$

4. Dividimos por la norma los vectores $\{\bar{s}_1, \bar{s}_2, \bar{s}_3\}$ para obtener la base ortonormal:

$$\bar{s}_1 \circ \bar{s}_1 = 4 \qquad \bar{s}_2 \circ \bar{s}_2 = 25 \qquad \bar{s}_3 \circ \bar{s}_3 = 34$$

$$\text{BON} = \left\{ \frac{\bar{s}_1}{|\bar{s}_1|} = \frac{(0,2,0)}{2}, \ \frac{\bar{s}_2}{|\bar{s}_2|} = \frac{(4,0,-3)}{5}, \ \frac{\bar{s}_3}{|\bar{s}_3|} = \frac{(3,3,4)}{\sqrt{34}} \right\}$$

Ejemplo
Sea E un espacio euclídeo de producto vectorial NO estándar con matriz asociada $M(\circ)$ y $B = \{\bar{e}_1 = (1,2), \bar{e}_2 = (-1,1)\}$ una base de E, obtener la base ortonormal más próxima a B:

$$M(\circ) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

1. Tomamos $\bar{s}_1$ como $\bar{e}_1$ ya que, al ser el primer vector, no hay que garantizar ninguna ortogonalidad:

$$\bar{s}_1 = \bar{e}_1 = (1,2)$$

2. Calculamos $\bar{s}_2$ garantizando la ortogonalidad con $\bar{s}_1$:

$$\bar{s}_1 \circ \bar{e}_2 = (1,2) \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 3 \qquad \bar{s}_1 \circ \bar{s}_1 = (1,2) \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 12$$

$$\bar{s}_2 = \bar{e}_2 - \frac{\bar{e}_2 \circ \bar{s}_1}{\bar{s}_1 \circ \bar{s}_1} \cdot \bar{s}_1 = (0,1) - \frac{3}{12}(1,2) = (-1,1) - ({}^{1}/_{4}, {}^{1}/_{2}) = (-{}^{5}/_{4}, {}^{1}/_{2})$$

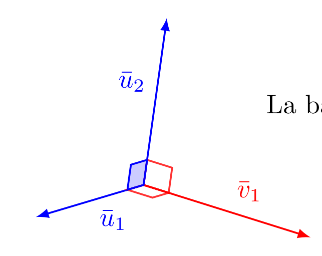
3. Dividimos por la norma los vectores $\{\bar{s}_1, \bar{s}_2\}$ para obtener la base ortonormal:

$$\bar{s}_1 \circ \bar{s}_1 = (1,2) \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 12 \qquad \bar{s}_2 \circ \bar{s}_2 = ({}^{-5}/_{4}, {}^{1}/_{2}) \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} {}^{-5}/_{4} \\ {}^{1}/_{2} \end{pmatrix} = 3$$

$$\text{BON} = \left\{ \frac{\bar{s}_1}{|\bar{s}_1|} = \frac{(1,2)}{\sqrt{12}}, \ \frac{\bar{s}_2}{|\bar{s}_2|} = \frac{({}^{-5}/_{4}, {}^{1}/_{2})}{\sqrt{3}} \right\}$$

5.4.4. Base ortogonal a un subespacio vectorial

Se define como base ortogonal a un subespacio vectorial V a una base cuyos vectores son perpendiculares a todos los vectores de la base de V y perpendiculares entre sí mismos. Se denota por $V^{\perp\circ}$.



La base $\{\bar{u}_1, \bar{u}_2\}$ es ortogonal a la base $\{\bar{v}_1\}$

5.4.5. Obtener base ortogonal a un subespacio (producto escalar estándar)

Sea E un espacio euclídeo de producto escalar $\circ : E \times E \longmapsto \mathbb{R}$ **estándar**. Si $V \in E$, entonces se cumple que:

$$\flat[V^{\perp\circ}] \cong V^{\perp}$$

Nota. $V^{\perp}$ denota el anulador de V.

Ya que $\flat[V^{\perp\circ}]$ puede ser un vector $\flat[\bar{w}]$ con $\bar{w} = V^{\perp\circ}$ o un conjunto de vectores $\flat[\bar{w}_i]$ con $\bar{w}_i \in V^{\perp\circ}$ $\forall i$.

De esta forma, si el producto escalar de E es estándar, para calcular la base ortonormal del subespacio V basta con calcular el anulador de V.

Ejemplo
Sea V un subespacio vectorial contenido en un espacio euclídeo de producto vectorial estándar, obtener el subespacio ortogonal a V:

$$V = \begin{cases} x_1 + x_2 - x_3 + x_4 = 0 \\ 2x_1 + x_2 - x_3 + 3x_4 = 0 \end{cases}$$

1. Obtenemos el anulador de V. Como existe una relación entre los coeficientes de las implícitas y el anulador (Ver espacios duales) lo tomamos directamente :

$$V^{\perp} = L\{(1,1,-1,1), (2,1,-1,3)\}$$

2. Como el producto es estándar el anulador de V coincide con el subespacio ortogonal a V:

$$V^{\perp\circ} = L\{(1,1,-1,1), (2,1,-1,3)\}$$

5.4.6. Obtener base ortogonal a un subespacio (producto escalar NO estándar)

Sea E un espacio euclídeo cuyo producto escalar $\circ : E \times E \longmapsto \mathbb{R}$ no es estándar, entonces no se cumple la equivalencia $\flat[V^{\perp\circ}] \cong V^{\perp}$. Para calcular el subespacio ortogonal $V^{\perp\circ}$ de un subespacio $V = L\{\bar{v}_i\}_{i=1}^{n}$ necesitamos encontrar el conjunto de vectores $\bar{w}_i$ que cumplan:

$$\bar{w}_i^T \cdot M(\circ) \cdot \bar{v}_j = \left(w_{i,1}, \dots, w_{i,n}\right) \begin{pmatrix} & M(\circ) & \end{pmatrix} \begin{pmatrix} v_{i,1} \\ \vdots \\ v_{i,n} \end{pmatrix} = 0 \qquad \forall \bar{v}_j \in V$$

Dado que conocemos el subespacio V, tenemos los vectores $\bar{v}_j$. Con cada vector $\bar{v}_j$ podemos obtener una serie de ecuaciones implícitas que luego se simplifiquen entre sí dándonos la base ortogonal que buscamos.

Ejemplo
Sea V un subespacio vectorial contenido en un espacio euclídeo de producto vectorial no estándar con matriz asociada $M(\circ)$, obtener el subespacio ortogonal a V:

$$V = L\{(1,2,0), (2,2,3)\} \qquad M(\circ) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

1. Obtenemos las ecuaciones implícitas de todo vector $\bar{w}$ perperndicular a $(1,2,0)$ y $(2,2,3)$ forzando que cumpla $\bar{w} \cdot M(\circ) \cdot \bar{v}_i = 0$:

$$(x_1, x_2, x_3) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \left(\begin{array}{c|c} 1 & 2 \\ 2 & 2 \\ 0 & 3 \end{array} \right) = \bar{0} \quad \longrightarrow \quad \begin{pmatrix} x_1 + 4x_2 + 3x_3 \\ 5x_1 + 7x_2 + 10x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\bar{w} = \begin{cases} x_1 + 4x_2 + 3x_3 = 0 \\ 5x_1 + 7x_2 + 10x_3 = 0 \end{cases}$$

2. Obtenemos el subespacio vectorial que describen las ecuaciones implícitas anteriores:

$$V^{\perp\circ} = L\{({}^{-19}/_{13}, {}^{-5}/_{13}, 1)\}$$

5.5. Proyecciones

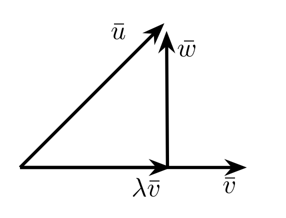
5.5.1. Proyección de un vector sobre otro

La ecuación de la proyección de un vector sobre otro puede demostrarse de más de una única forma. A continuación se muestran dos de ellas. La primera con un enfoque más vectorial y la otra con un enfoque más geométrico.

Dados dos vectores $\bar{u}$ y $\bar{v}$, se pide hallar la proyección del vector $\bar{u}$ sobre el vector $\bar{v}$:

1. DESARROLLO VECTORIAL

   Veamos que podemos reescribir $\bar{u}$ en función de la suma de un múltiplo de $\bar{v}$ más un vector $\bar{w}$ perpendicular a $\bar{v}$.



$$\bar{u} = P_{L\{\bar{v}\}}(\bar{u}) + \bar{w} = \lambda \cdot \bar{v} + \bar{w}$$

   Ahora multipliquemos por el vector $\bar{v}$ a ambos lados de la igualdad:

$$\bar{u} \circ \bar{v} = (\lambda \cdot \bar{v} + \bar{w}) \circ \bar{v} = (\lambda \cdot \bar{v}) \circ \bar{v} + \cancelto{0}{\bar{w} \circ \bar{v}} = \lambda \cdot (\bar{v} \circ \bar{v})$$

   Despejando lambda obtenemos la siguiente igualdad:

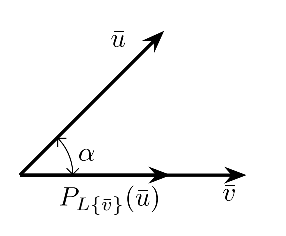
$$\lambda = \frac{\bar{u} \circ \bar{v}}{\bar{v} \circ \bar{v}}$$

   Por tanto, la proyección de $\bar{u}$ sobre $\bar{v}$ es:

$$P_{L\{\bar{v}\}}(\bar{u}) = \lambda \cdot \bar{v} = \frac{\bar{u} \circ \bar{v}}{\bar{v} \circ \bar{v}} \cdot \bar{v}$$

2. DESARROLLO GEOMÉTRICO

   Para este desarrollo debemos expresar la proyección en función del coseno del ángulo que forman los vectores:



$$P_{L\{\bar{v}\}}(\bar{u}) = \lambda \cdot \bar{v} = (\cos\alpha \cdot |\bar{u}|) \cdot \vec{v} = (\cos\alpha \cdot |\bar{u}|) \cdot \frac{\bar{v}}{|\bar{v}|}$$

   nota: donde $\vec{v}$ denota vector unitario en la dirección de $\bar{v}$

   Y sustituimos empleando la ecuación del ángulo entre dos vectores:

$$\cos\alpha = \frac{\bar{u} \circ \bar{v}}{|\bar{u}| \cdot |\bar{v}|}$$

$$P_{L\{\bar{v}\}}(\bar{u}) = (\cos\alpha \cdot |\bar{u}|) \cdot \frac{\bar{v}}{|\bar{v}|} \longrightarrow \left(\frac{\bar{u} \circ \bar{v}}{\cancel{|\bar{u}|} \cdot |\bar{v}|} \cdot \cancel{|\bar{u}|}\right) \cdot \frac{\bar{v}}{|\bar{v}|} = \frac{\bar{u} \circ \bar{v}}{\bar{v} \circ \bar{v}} \cdot \bar{v}$$

   nota: ya que $\bar{v} \circ \bar{v} = |\bar{v}|^2$

Ejemplo
Dados dos vectores $\bar{u}$ y $\bar{v}$ de un espacio vectorial euclídeo de producto estándar, hallar la proyección de $\bar{u}$ sobre $\bar{v}$:

$$\bar{u} = (-2, 5, 2) \qquad \bar{v} = (1, 0, 3)$$

1. Obtenemos el resultado de los productos escalares que necesitamos para obtener la proyección:

$$\bar{u} \circ \bar{v} = (-2, 5, 2) \circ (1, 0, 3) = 4 \qquad \bar{u} \circ \bar{v} = (1, 0, 3) \circ (1, 0, 3) = 10$$

2. Aplicamos la ecuación de proyección de un vector sobre otro:

$$P_{L\{\bar{v}\}}(\bar{u}) = \frac{\bar{u} \circ \bar{v}}{\bar{v} \circ \bar{v}} \cdot \bar{u} = \frac{4}{10} \cdot (1, 0, 3) = \left({}^{4}/_{10}, 0, {}^{6}/_{5}\right)$$

5.5.2. Proyección de un vector sobre un subespacio vectorial

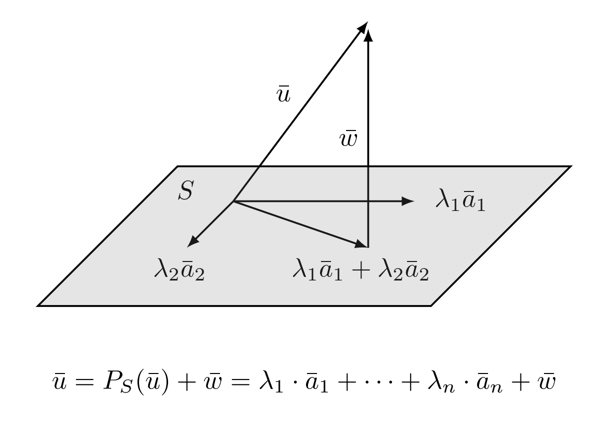
TEOREMA
Sea $\bar{u}$ un vector de un espacio vectorial euclídeo y S un subespacio vectorial de dicho espacio vectorial, la proyección de $\bar{u}$ sobre S es el vector de S más cercano a $\bar{u}$.

Nota. Hablamos de subespacio y no de plano porque se pueden proyectar vectores de dimensión mayor que 3.

El método algebráico para la obtención de la proyección de un vector $\bar{u}$ sobre un subespacio vectorial $S = L\{\bar{a}_1, \dots, \bar{a}_n\}$, que denotamos por $P_S(\bar{u})$, es una generalización del procedimiento empleado para la obtención de la proyección de un vector sobre otro vector.

DESARROLLO

Veamos que podemos expresar $\bar{u}$ en función de una combinación lineal de los vectores de la base de S más un vector $\bar{w}$ perpendicular a todos los vectores de la base de S.



$$\bar{u} = P_S(\bar{u}) + \bar{w} = \lambda_1 \cdot \bar{a}_1 + \dots + \lambda_n \cdot \bar{a}_n + \bar{w}$$

Ahora formemos un sistema de ecuaciones multiplicando a ambos lados de la igualdad por cada uno de los vectores de la base del subespacio S:

$$\left\{\begin{array}{c} \bar{u} \circ \bar{a}_1 = (\lambda_1 \cdot \bar{a}_1 + \dots + \lambda_n \cdot \bar{a}_n + \bar{w}) \circ \bar{a}_1 \\ \vdots \\ \bar{u} \circ \bar{a}_n = (\lambda_1 \cdot \bar{a}_1 + \dots + \lambda_n \cdot \bar{a}_n + \bar{w}) \circ \bar{a}_n \end{array}\right. \longrightarrow \left\{\begin{array}{c} \bar{u} \circ \bar{a}_1 = \lambda_1 \cdot (\bar{a}_1 \circ \bar{a}_1) + \dots + \lambda_n \cdot (\bar{a}_n \circ \bar{a}_1) + \cancelto{0}{\bar{w} \circ \bar{a}_1} \\ \vdots \\ \bar{u} \circ \bar{a}_n = \lambda_1 \cdot (\bar{a}_1 \circ \bar{a}_n) + \dots + \lambda_n \cdot (\bar{a}_n \circ \bar{a}_n) + \cancelto{0}{\bar{w} \circ \bar{a}_n} \end{array}\right.$$

Reescirbimos el sistema de ecuaciones como una ecuación matricial y despejamos el vector de coeficientes:

$$A^T M(\circ)\bar{u} = (A^T M(\circ) A)\bar{X}$$

$$\begin{pmatrix} \bar{u} \circ \bar{a}_1 \\ \vdots \\ \bar{u} \circ \bar{a}_n \end{pmatrix} = \begin{pmatrix} \bar{a}_1 \circ \bar{a}_1 & \dots & \bar{a}_n \circ \bar{a}_1 \\ \vdots & \ddots & \vdots \\ \bar{a}_1 \circ \bar{a}_n & \dots & \bar{a}_n \circ \bar{a}_n \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \longrightarrow \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \bar{a}_1 \circ \bar{a}_1 & \dots & \bar{a}_n \circ \bar{a}_1 \\ \vdots & \ddots & \vdots \\ \bar{a}_1 \circ \bar{a}_n & \dots & \bar{a}_n \circ \bar{a}_n \end{pmatrix}^{-1} \begin{pmatrix} \bar{u} \circ \bar{a}_1 \\ \vdots \\ \bar{u} \circ \bar{a}_n \end{pmatrix}$$

$$\bar{X} = (A^T M(\circ) A)^{-1} A^T M(\circ)\bar{u}$$

Una vez que tenemos los coeficientes, obtenemos la proyección de $\bar{u}$ sustituyendo en la expresión:

$$P_S(\bar{u}) = \lambda_1 \cdot \bar{a}_1 + \dots + \lambda_n \cdot \bar{a}_n$$

También podemos obtener directamente la proyección del vector multiplicando escalarmente todo por A:

$$P_S(\bar{u}) = AM(\circ)\bar{X} = A(A^T M(\circ) A)^{-1} A^T M(\circ)\, \bar{u}$$

Ejemplo
Dado un subespacio vectorial V contenido en un espacio euclídeo E con producto estándar y un vector $\bar{u} \in E$, hallar la proyección de $\bar{v}$ sobre V:

$$V = L\{a_1, a_2\} = L\{(2, 0, 1), (0, -3, 2)\} \qquad \bar{u} = (1, 2, 2)$$

1. Construimos la ecuación matricial resultante de multiplicar la ecuación $\bar{u} = P_V + \bar{w}$ por cada uno de los vectores de V:

$$\left\{\begin{array}{l} \bar{u} \circ \bar{a}_1 = \lambda_1 \cdot \bar{a}_1 \circ \bar{a}_1 + \lambda_2 \cdot \bar{a}_2 \circ \bar{a}_1 \\ \bar{u} \circ \bar{a}_2 = \lambda_1 \cdot \bar{a}_1 \circ \bar{a}_2 + \lambda_2 \cdot \bar{a}_2 \circ \bar{a}_2 \end{array}\right. \longrightarrow \begin{pmatrix} \bar{u} \circ \bar{a}_1 \\ \bar{u} \circ \bar{a}_2 \end{pmatrix} = \begin{pmatrix} \bar{a}_1 \circ \bar{a}_1 & \bar{a}_2 \circ \bar{a}_1 \\ \bar{a}_1 \circ \bar{a}_2 & \bar{a}_2 \circ \bar{a}_2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

2. Despejamos el vector de coeficientes de la ecuación anterior y sustituimos los productos escalares conocidos:

$$\bar{u} \circ \bar{a}_1 = 4 \qquad \bar{u} \circ \bar{a}_2 = -2 \qquad \bar{a}_1 \circ \bar{a}_1 = 5 \qquad \bar{a}_2 \circ \bar{a}_2 = 13 \qquad \bar{a}_1 \circ \bar{a}_2 = 9$$

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \bar{a}_1 \circ \bar{a}_1 & \bar{a}_2 \circ \bar{a}_1 \\ \bar{a}_1 \circ \bar{a}_2 & \bar{a}_2 \circ \bar{a}_2 \end{pmatrix}^{-1} \begin{pmatrix} \bar{u} \circ \bar{a}_1 \\ \bar{u} \circ \bar{a}_2 \end{pmatrix} \longrightarrow \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 5 & 9 \\ 9 & 13 \end{pmatrix}^{-1} \begin{pmatrix} 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 10 \end{pmatrix}$$

3. Una vez obtenidos los coeficientes los sustituimos en la ecuación de la proyección:

$$P_V(\bar{u}) = \lambda_1 \cdot a_1 + \lambda_2 \cdot a_2 = 2 \cdot (2, 0, 1) + 10 \cdot (0, -3, 2) = (4, -30, 22)$$

Ejemplo
Sea $E = C^0([0, 2\pi])$ el espacio vectorial de las funciones continuas definidas en el intervalo $[0, 2\pi]$. $\forall f, g \in E$ se define el producto escalar $f \circ g = \int_0^{2\pi} f(x)g(x)dx$. Dado un sistema S de funciones de E , se pide:

a) Demostrar que S es un sistema ortogonal de vectores no nulos

b) Hallar la proyección de $f(x) = x$ sobre $L\{S\}$

$$S = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\} = \{1, cos(x), sen(x)\}$$

a) Para que S sea un sistema ortogonal de vectores no nulos, la matriz resultante de la multiplicación SS^T debe ser diagonal y distinta de $\bar{0}_{M_3}$.

1 Obtenemos los productos escalares resultantes de la combinatoria de los elementos de S:

$$1 \circ 1 = \int_0^{2\pi} 1dx = 2\pi \qquad 1 \circ cos(x) = \int_0^{2\pi} cos(x)dx = sen(x)\Big|_0^{2\pi} = 0$$

$$1 \circ sen(x) = \int_0^{2\pi} sen(x)dx = -cos(x)\Big|_0^{2\pi} = 0$$

$$cos(x) \circ cos(x) = \int_0^{2\pi} cos^2(x)dx = \frac{1}{2}(x - \frac{cos(2x)}{2})\Big|_0^{2\pi} = \pi$$

$$cos(x) \circ sen(x) = \int_0^{2\pi} sen(x) \cdot cos(x)dx = \frac{sen^2(x)}{2}\Big|_0^{2\pi} = 0$$

$$sen(x) \circ sen(x) = \int_0^{2\pi} sen^2(x)dx = 2\pi - \int_0^{2\pi} cos^2(x)dx = 2\pi - \pi = \pi$$

2 Construimos la matriz SS^T utilizando los productos escalares calculados previamente:

$$S \cdot S^T = \begin{pmatrix} \bar{e}_1 \circ \bar{e}_1 & \bar{e}_1 \circ \bar{e}_2 & \bar{e}_1 \circ \bar{e}_3 \\ \bar{e}_2 \circ \bar{e}_1 & \bar{e}_2 \circ \bar{e}_2 & \bar{e}_2 \circ \bar{e}_3 \\ \bar{e}_3 \circ \bar{e}_1 & \bar{e}_3 \circ \bar{e}_2 & \bar{e}_3 \circ \bar{e}_3 \end{pmatrix} = \begin{pmatrix} 1 \circ 1 & 1 \circ sen(x) & 1 \circ cos(x) \\ sen(x) \circ 1 & sen(x) \circ cos(x) & sen(x) \circ cos(x) \\ cos(x) \circ 1 & cos(x) \circ sen(x) & cos(x) \circ cos(x) \end{pmatrix}$$

$$S \cdot S^T = \begin{pmatrix} 2\pi & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & \pi \end{pmatrix}$$

3 Dado que la matriz es diagonal podemos concluir que el sistema S sí es ortogonal

b) Para hallar la proyección de $f(x) = x$ sobre $L\{S\}$ aprovechamos la igualdad $x = P_S(x) + g(x)$. Multiplicando a ambos lados de la ecuación por cada uno de los vectores de S podemos formar un sistema de ecuaciones del que obtener la proyección:

$$x = P_S(x) + g(x) = \lambda_1 \cdot 1 + \lambda_2 \cdot sen(x) + \lambda_3 \cdot cos(x) + g(x)$$

1 Construimos el sistema de ecuaciones y despejamos el vector de coeficientes $(\lambda_1, \lambda_2, \lambda_3)$:

$$\left\{\begin{array}{ccc} x \circ 1 & = & \lambda_1 \cdot (1 \circ 1) + \lambda_2 \cdot (1 \circ sen(x)) + \lambda_3 \cdot (1 \circ cos(x)) \\ x \circ cos(x) & = & \lambda_1 \cdot (sen(x) \circ 1) + \lambda_2 \cdot (sen(x) \circ cos(x)) + \lambda_3 \cdot (sen(x) \circ cos(x)) \\ x \circ sen(x) & = & \lambda_1 \cdot (cos(x) \circ 1) + \lambda_n \cdot (cos(x) \circ sen(x)) + \lambda_3 \cdot (cos(x) \circ cos(x)) \end{array}\right.$$

$$\begin{pmatrix} x \circ 1 \\ x \circ cos(x) \\ x \circ sen(x) \end{pmatrix} = \begin{pmatrix} 1 \circ 1 & 1 \circ sen(x) & 1 \circ cos(x) \\ sen(x) \circ 1 & sen(x) \circ cos(x) & sen(x) \circ cos(x) \\ cos(x) \circ 1 & cos(x) \circ sen(x) & cos(x) \circ cos(x) \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}$$

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 1 \circ 1 & 1 \circ sen(x) & 1 \circ cos(x) \\ sen(x) \circ 1 & sen(x) \circ cos(x) & sen(x) \circ cos(x) \\ cos(x) \circ 1 & cos(x) \circ sen(x) & cos(x) \circ cos(x) \end{pmatrix}^{-1} \begin{pmatrix} x \circ 1 \\ x \circ cos(x) \\ x \circ sen(x) \end{pmatrix}$$

2 Sustituyendo por los productos escalares que ya hemos calculado o que podemos calcular y haciendo el producto matricial obtenermos los coeficientes $\{\lambda_1, \lambda_2, \lambda_3\}$:

$$x \circ 1 = \int_0^{2\pi} xdx = \frac{x^2}{2}\Big|_0^{2\pi} = 2\pi^2$$

$$x \circ cos(x) = \int_0^{2\pi} x \cdot cos(x)dx = x \cdot sen(x) + cos(x)\Big|_0^{2\pi} = 0$$

$$x \circ sen(x) = \int_0^{2\pi} x \cdot sen(x)dx = -x \cdot cos(x) - sen(x)\Big|_0^{2\pi} = -2\pi$$

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 2\pi & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & \pi \end{pmatrix}^{-1} \begin{pmatrix} 2\pi^2 \\ 0 \\ -2\pi \end{pmatrix} = \begin{pmatrix} \pi \\ 0 \\ -2 \end{pmatrix}$$

3 Multiplicando cada coeficiente por su correspondiente vector de S obtenemos la proyección:

$$P_{L\{S\}}(x) = \lambda_1 \cdot 1 + \lambda_2 \cdot sen(x) + \lambda_3 \cdot cos(x) = \pi - 2cos(x)$$

5.5.3. Matriz de proyección ortogonal

La matriz de proyección ortogonal sobre un subespacio vectorial S, que denotamos por P_S, es una matriz cuyo producto por cualquier vector $\bar{u}$ del espacio vectorial E (con $S \in E$) resulta en la proyección de $\bar{u}$ sobre S

Tomando la ecuación matricial que nos da la proyección de un vector u sobre un subespacio S (de matriz A):

$$P_S(\bar{u}) = A(A^T M(\circ) A)^{-1} A^T M(\circ)\, \bar{u}$$

A la matriz resultante del producto $A(A^T M(\circ) A)^{-1} A^T M(\circ)$ se la define como matriz de proyección ortogonal sobre el subespacio $S = L\{\bar{a}_1, \dots, \bar{a}_n\}$. Esta matriz verifica que al multiplicarse por cualquier vector $\bar{u}$ (No necesariamente contenido en S) se obtiene la proyección de dicho vector sobre S, $P_S(\bar{u})$:

$$P_S \cdot \bar{u} = P_S(\bar{u})$$

Nota. Si la base del espacio euclídeo en el que trabajamos es ortonormal (BON) se cumple $M(\circ) = Id$, por lo que:

$$A(A^T M(\circ) A)^{-1} A^T M(\circ) \longrightarrow A(A^T Id\, A)^{-1} A^T Id \longrightarrow A(A^T A)^{-1} A^T$$

Ejemplo
Dado el subespacio vectorial S, contenido en un espacio vectorial euclídeo con producto estándar, hallar la matriz de proyección ortogonal asociada a S:

$$S = L\{(1, 2, 0), (1, 0, -1)\}$$

Aplicamos la ecuación de la matriz de proyección ortogonal (Dado que el producto escalar es estándar podemos omitir las matrices del producto escalar):

$$P_S = (A^T A)^{-1} A^T = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 0 & -1 \end{pmatrix} \left(\begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 0 & -1 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & -1 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 5 & 2 & -4 \\ 2 & 8 & 2 \\ -4 & 2 & 5 \end{pmatrix}$$

5.5.4. Propiedades de la matriz de proyección

Sea P una matriz $n \times n$, se verifica que P es una matriz es de proyección si y solo si es simétrica e idempotente:

$$P \in M_{n \times n} \text{ es de proyección} \iff \left\{\begin{array}{l} P = P^T \\ P = P^2 \end{array}\right.$$

5.5.5. Matriz de proyección sobre un subespacio de base ortonormal

Tomemos la expresión de una matriz de proyección sobre un subespacio vectorial $S = L\{\bar{a}_1, \dots, \bar{a}_n\}$ en un espacio euclídeo de dimensión m y base ortonormal:

$$P_S = A(A^T A)^{-1} A^T$$

Si además se cumple que los vectores de la base del subespacio (no del espacio euclídeo) son también ortonormales, es decir, que las columnas de la matriz $A = \left(\bar{a}_1 | \dots | \bar{a}_n\right) \in M_{m \times n}$ (Matriz no cuadrada ya que un subespacio no tiene dimesión máxima) son ortonormales:

Podemos aprovecharnos de la siguiente equivalencia, derivada de las propiedades de una matriz ortogonal *, para simplificar la expresión:

$$A^T A = Id \in M_{m \times m}$$

Nota*. Una matriz ortogonal $M_{n \times n}$ cumple que $M^T M = Id \in M_{n \times n}$ (Ver propiedades de matrices ortogonales), análogamente si tenemos una "semimatriz ortogonal" $M_{m \times n}$ se cumple que $M^T M = Id \in M_{m \times m}$.

Aplicando la equivalencia anterior la expresión de la matriz de proyección se puede simplificar de la siguiente manera:

$$P_S = A(A^T A)A^T \longrightarrow P_S = A \cdot A^T$$

Nota. Al ser A una "semimatriz ortogonal" $M_{m \times n}$ y no una matriz ortogonal completa $M_{m \times m}$, se cumple que $A^T A = Id$ pero no que $AA^T = Id$ (Como sí pasa con una matriz ortogonal).

5.6. Ajuste de regresión mediante una proyección

EXPLICACIÓN DEL MSE

Una regresión polinomial es un modelo de predicción en el que la relación entre la variable independiente x y la variable dependiente $\hat{y}$ viene dada por un polinomio de n-ésimo grado:

$$\hat{y} = \lambda_0 + \lambda_1 \cdot x + \lambda_2 \cdot x^2 + \cdots + \lambda_n \cdot x^n$$
$$\bar{X}^T = [\ \lambda_1 \, , \, \ldots \, , \, \lambda_n \ ]$$

Para "ajustar" (obtener los coeficientes $\{\lambda_j\}_{j=0}^{n}$) la regresión es necesario contar con el set de datos que se desea modelar, que debe estar compuesto por un conjunto de datos de la variable no dependiente $\{z_i\}_{i=1}^{m}$ y sus correspondientes datos de la variable dependiente $\{y_i\}_{i=1}^{m}$:

$$\bar{Z}^T = [\ z_1 \, , \, \ldots \, , \, z_m \ ]$$
$$\bar{Y}^T = [\ y_1 \, , \, \ldots \, , \, y_m \ ]$$

El principal método empleado para el ajuste de este tipo de regresiones es el de optimización del error cuadrático medio, o "mean square error (MSE)" en inglés, que comete el modelo :

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad \text{Donde } y_i \text{ pertenece al set de datos e } \hat{y}_i \text{ es una predicción del modelo}$$

Nota. $\frac{1}{n}$ es constante. Véase por tanto que optimizar $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ es equivalente a optimizar el MSE

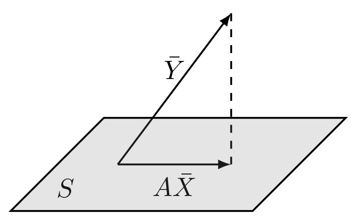
Si tenemos un sistema de ecuaciones con mayor número de ecuaciones que incógnitas (sistema sobredeterminado) es normalmente incompatible. Sobre un modelo de regresión polinomial esto se traduce en que si el número de datos z_i del set es mayor que el número de coeficientes λ_j del modelo, entonces será poco probable que el modelo pueda predecir con total exactitud el conjunto de valores dependientes a partir de los valores independientes:

$$A \cdot \bar{X} = \bar{Y} \quad \longrightarrow \quad \begin{pmatrix} 1 & z_1 & \ldots & z_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_m & \ldots & z_m^n \end{pmatrix} \begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\text{Si } m > n \quad \longrightarrow \quad \text{Sistema sobredeterminado} \quad \longrightarrow \quad A \cdot \bar{X} \neq \bar{Y} \quad \forall \bar{X}$$

Ante la inexistencia de un ajuste del modelo que describa a la perfección el conjunto de los datos del set, deberemos buscar aquel ajuste del modelo que cometa el menor error posible en sus predicciones (La técnica de búsqueda más frecuente es la optimización del MSE).

Para llevar a cabo el ajuste algebráicamente debemos trasladar los vectores anteriores a un espacio euclídeo (canónico y de producto estándar) donde se cumple que las columnas de la matriz A representan un subespacio vectorial y el vector $A \cdot \bar{X}$ es la proyección de $\bar{Y}$ sobre dicho subespacio.



Intuitivamente, el subespacio vectorial S descrito por las columnas de A representa los potenciales resultados que podrá producir el modelo de regresión (A partir de los valores independientes del set de datos) al ser ajustados los coeficientes. El vector $A{\cdot}\bar{X}$ representa los valores producidos por la regresión lo más óptimamente posible ajustada. Y el vector $\bar{Y}$ representa los valores dependientes del set.

Veamos que existe una equivalencia entre la expresión de la distancia de $\bar{Y}$ a cualquier otro vector $\bar{v}$ de S, $|\bar{Y} - \bar{v}|$ y la ecuación del MSE. Por lo tanto, dado que $A \cdot \bar{X}$ es el vector de S más cercano a $\bar{Y}$ (ya que es su proyección) también es, de acuerdo con el MSE, el ajuste óptimo del modelo de regresión.

$$|\bar{Y} - \bar{v}| = \sqrt{|\bar{Y} - \bar{v}|^2} \quad \xrightarrow{\ *\ } \quad |\bar{Y} - \bar{v}|^2 = (\bar{Y} - \bar{v}) \circ (\bar{Y} - \bar{v}) = \sum_{i=1}^{n}(y_i - v_i)^2$$
$$|\bar{Y} - \bar{v}|^2 = \sum_{i=1}^{n}(y_i - v_i)^2 \ \sim \ \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = MSE$$

Nota.* Optimizar $\sqrt{|\bar{Y} - \bar{v}|^2}$ es equivalente a optimizar $|\bar{Y} - \bar{v}|^2$

En consecuencia, para ajustar la regresión aplicando algebráicamente la optimización del MSE solo tenemos que obtener los coeficientes del vector $\bar{X}$ a partir de la proyección de $\bar{Y}$ sobre S.

$$A \cdot \bar{X} = P_S(\bar{Y}) \quad \longrightarrow \quad A \cdot \bar{X} = A(A^T A)A^T \bar{Y} \quad \longrightarrow \quad \bar{X} = (A^T A)A^T \bar{Y}$$

Una vez que tenemos los coeficientes los sutituimos en la ecuación de la regresión polinomial de n-grado:

$$\hat{y} = \lambda_0 + \lambda_1 \cdot x + \lambda_2 \cdot x^2 + \cdots + \lambda_n \cdot x^n$$

Ejemplo
En una exposición de automóviles, un visitante decidió realizar un conjunto de observaciones relacionando el precio de los vehículos b_i con sus pesos a_i. Dados los datos de dichas observaciones se pide realizar una predicción del precio de un coche de 1,6 toneladas empleando un modelo de **regresión lineal**:

a_i(peso en toneladas)	0,8	1	1,2	1,3
b_i(precios en millones)	1	2	3	5

1. Construimos el sistema sobredeterminado correspondiente a una regresión lineal del set dado:

$$A \cdot \bar{X} = \bar{Y} \quad \longrightarrow \quad \begin{pmatrix} 1 & 0{,}8 \\ 1 & 1 \\ 1 & 1{,}2 \\ 1 & 1{,}3 \end{pmatrix} \begin{pmatrix} \lambda_0 \\ \lambda_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 5 \end{pmatrix}$$

2. Obtenemos los coeficientes $\{\lambda_0, \lambda_1\}$ de la proyección de $\bar{Y}$ sobre el subespacio vectorial S:

$$\bar{X} = (A^T A)^{-1} A^T \bar{Y} = \left(\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0{,}8 & 1 & 1{,}2 & 1{,}3 \end{pmatrix} \begin{pmatrix} 1 & 0{,}8 \\ 1 & 1 \\ 1 & 1{,}2 \\ 1 & 1{,}3 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0{,}8 & 1 & 1{,}2 & 1{,}3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} -5{,}08 \\ 7{,}29 \end{pmatrix}$$

3. Construimos el modelo utilizando los coeficientes obtenidos y hacemos la predicción:

$$\hat{y} = \lambda_0 + \lambda_1 \cdot x = (-5{,}08) + 7{,}29 \cdot x$$
$$\text{(Predicción de 1,6 T)} = (-5{,}08) + 7{,}29 \cdot 1{,}6 = 6{,}58M$$

Ejemplo
En una exposición de automóviles, un visitante decidió realizar un conjunto de observaciones relacionando el precio de los vehículos b_i con sus pesos a_i. Dados los datos de dichas observaciones, se pide realizar una predicción del precio de un coche de 1,6 toneladas empleando un modelo de **regresión polinomial de** 2 **grado**:

a_i(peso en toneladas)	0,8	1	1,2	1,3
b_i(precios en millones)	1	2	3	5

1. Construimos el sistema sobredeterminado correspondiente a una regresión polinomial de 2 grado:

$$A \cdot \bar{X} = \bar{Y} \quad \longrightarrow \quad \begin{pmatrix} 1 & 0{,}8 & 0{,}8^2 \\ 1 & 1 & 1^2 \\ 1 & 1{,}2 & 1{,}2^2 \\ 1 & 1{,}3 & 1{,}3^2 \end{pmatrix} \begin{pmatrix} \lambda_0 \\ \lambda_1 \end{pmatrix} = \begin{pmatrix} 1 & 0{,}8 & 0{,}64 \\ 1 & 1 & 1 \\ 1 & 1{,}2 & 1{,}44 \\ 1 & 1{,}3 & 1{,}69 \end{pmatrix} \begin{pmatrix} \lambda_0 \\ \lambda_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 5 \end{pmatrix}$$

2. Obtenemos los coeficientes $\{\lambda_0, \lambda_1, \lambda_2\}$ de la proyección de $\bar{Y}$ sobre el subespacio vectorial S:

$$\bar{X} = (A^T A)^{-1} A^T \bar{Y} = \begin{pmatrix} 11{,}72 & -13{,}07 & -6{,}39 & 8{,}74 \\ -20{,}57 & 26{,}88 & 12{,}17 & -18{,}49 \\ 8{,}92 & -13{,}07 & -5{,}4 & 9{,}55 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 10{,}12 \\ -22{,}74 \\ 14{,}32 \end{pmatrix}$$

3. Construimos el modelo utilizando los coeficientes obtenidos y hacemos la predicción:

$$\hat{y} = \lambda_0 + \lambda_1 \cdot x + \lambda_2 \cdot x^2 = 10{,}11 - 22{,}73 \cdot x + 14{,}32 \cdot x^2$$
$$\text{(Predicción de 1,6 T)} = 10{,}11 - 22{,}73 \cdot 1{,}6 + 14{,}32 \cdot 1{,}6^2 = 10{,}40M$$

5.7. Matriz ortogonal

Se define como **matriz ortogonal** a una matriz cuadrada M si verifica $M^T = M^{-1}$, es decir, si se cumple que:

$$M \cdot M^T = M \cdot M^{-1} = Id$$

Nota. Al ser una matriz cuadrada, también se verifica:

$$M \cdot M^T = M^T \cdot M = Id$$

Toda matriz ortogonal $M \in M_{n\times n}$ presenta la estructura $M = \left(\bar{v}_1 | \ldots | \bar{v}_n\right)$, donde $\{\bar{v}_i\}_{i=1}^n$ constituye una base canónicamente ortonormal (ortonormal con el producto escalar estándar).

DEMOSTRACIÓN

Dada una matriz M de columnas $\{\bar{c}_1, \ldots, \bar{c}_n\}$, si se multiplica por ella misma traspuesta se tiene que:

$$M^T M = \left(\begin{array}{c|c|c} \bar{c}_1 & \cdots & \bar{c}_n \end{array}\right)^T \left(\begin{array}{c|c|c} \bar{c}_1 & \ldots & \bar{c}_n \end{array}\right) = \begin{pmatrix} \bar{c}_1 \circ \bar{c}_1 & \ldots & \bar{c}_1 \circ \bar{c}_n \\ \vdots & \ddots & \vdots \\ \bar{c}_n \circ \bar{c}_1 & \ldots & \bar{c}_n \circ \bar{c}_n \end{pmatrix}$$

($\circ$) denota el producto **estándar**, ya que sino $\bar{c}_i \circ \bar{c}_j \neq \bar{c}_i^T \cdot \bar{c}_j$

Y dado que $M^T M = Id$, podemos hacer la siguiente deducción:

$$\begin{pmatrix} \bar{c}_1 \circ \bar{c}_1 & \ldots & \bar{c}_1 \circ \bar{c}_n \\ \vdots & \ddots & \vdots \\ \bar{c}_n \circ \bar{c}_1 & \ldots & \bar{c}_n \circ \bar{c}_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{pmatrix} \iff \bar{c}_i \circ \bar{c}_j = \delta_{i,j} \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases} \iff \{\bar{c}_1, \ldots, \bar{c}_n\} = BON$$

Por tanto, las columnas de toda matriz ortogonal forman una base ortonormal con el producto estándar.

□

5.7.1. Propiedades de una matriz ortogonal

Si una matriz M es ortogonal, se verifican las siguientes propiedades:

1. El determinante de M valdrá ± 1, $|M| = \pm 1$
   - $|M| = +1 \quad \longrightarrow \quad$ Matriz directa
   - $|M| = -1 \quad \longrightarrow \quad$ Matriz inversa
2. A^{-1} y A^T también son matrices ortogonales
3. Si A y B son ortogonales, AB también es ortogonal

5.8. Factorización QR de una matriz

Toda matriz M cuadrada puede reescribirse como un producto $M = QR$ donde:

- Q es una matriz ortogonal
- R es una matriz triangular superior

PROCEDIMIENTO DE LA FACTORIZACIÓN QR

Para obtener la factorización QR de una matriz M

1. Descomponemos la matriz M por columnas, obteniendo los vectores $\{\bar{c}_i\}_{i=1}^{n}$:

$$M = \left(\begin{array}{c|c|c} \bar{c}_1 & \dots & \bar{c}_n \end{array} \right) \quad \longrightarrow \quad \{\bar{c}_1, \dots, \bar{c}_n\}$$

2. Aplicando Gram Schmidt sobre los vectores $\{\bar{c}_i\}_{i=1}^{n}$ se obtiene la base ortonormal $\{\bar{v}_i\}_{i=1}^{n}$:

$$\{\bar{c}_1, \dots, \bar{c}_n\} \quad \xrightarrow{\text{Gram Schmidt}} \quad \{\bar{v}_1, \dots, \bar{v}_n\}$$

3. Tomamos como matriz Q y matriz R:

$$Q = \left(\begin{array}{c|c|c} \bar{v}_1 & \dots & \bar{v}_n \end{array} \right) \qquad R = \begin{pmatrix} \bar{v}_1 \circ \bar{c}_1 & \dots & \bar{v}_1 \circ \bar{c}_n \\ 0 & \ddots & \vdots \\ \ddots & 0 & \bar{v}_n \circ \bar{c}_n \end{pmatrix}$$

5.9. Transformaciones ortogonales

Sea $(E, \circ)$ un espacio vectorial euclídeo. Una aplicación $f : E \mapsto$ es una transformación ortogonal si es lineal y conserva el producto escalar, es decir:

$$\forall \bar{x}, \bar{y} \in E \qquad \bar{x} \cdot \bar{y} = f(\bar{x}) \circ f(\bar{y})$$

$f : E \mapsto E$ es una matriz ortogonal si y solo si la matriz asociada a f en cualquier base ortonormal es ortogonal.

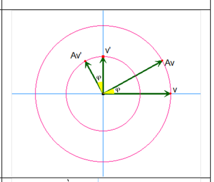
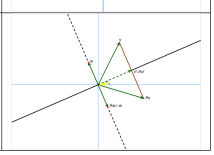
5.9.1. Matrices de transformaciones

COMILLAS
UNIVERSIDAD PONTIFICIA
ICAI
DEPARTAMENTO DE MATEMÁTICA APLICADA

TITULACIÓN
Grado en Ingeniería Matemática e Inteligencia Artificial
ÁLGEBRA Y GEOMETRÍA

Transformaciones ortogonales en $\mathbb{R}^2$ y $\mathbb{R}^3$
ESPACIO VECTORIAL EUCLÍDEO

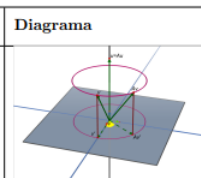
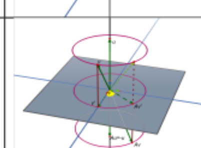
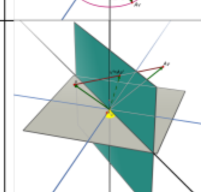
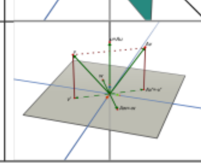
Clasificación de transformaciones ortogonales en $\mathbb{R}^2$

Tipo	**Matriz**	**Determinante/Traza**	**Autovalores**	**Elementos Geométricos**	**Diagrama**
Rotación	$A = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$	$\det(A) = 1$ $\mathrm{tr}(A) = 2\cos(\varphi)$	$\lambda_1 = e^{\varphi i}$ $\lambda_2 = e^{-\varphi i}$	Ángulo: φ	
Simetría respecto a una recta	$A = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\det(A) = -1$ $\mathrm{tr}(A) = 0$	$\lambda_1 = 1$ $\lambda_2 = -1$	Eje de simetría: $N_{1,1}$	

Casos particulares

- Rotación con $\varphi = 0 \iff$ Identidad ($\det(A) = 1$, $\mathrm{tr}(A) = 2$)
- Rotación con $\varphi = \pi \iff$ Simetría central ($\det(A) = 1$, $\mathrm{tr}(A) = -2$)

Clasificación de transformaciones ortogonales en $\mathbb{R}^3$

Tipo	**Matriz**	**Determinante/Traza**	**Autovalores**	**Elementos Geométricos**	**Diagrama**
Rotación	$A \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix}$	$\det(A) = 1$ $\mathrm{tr}(A) = 1 + 2\cos(\varphi)$	$\lambda_1 = 1$ $\lambda_2 = e^{\varphi i}$ $\lambda_3 = e^{-\varphi i}$	Eje de giro: $N_{1,1}$ Ángulo: φ	
Composición de rotación y simetría	$A \sim \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix}$	$\det(A) = -1$ $\mathrm{tr}(A) = -1 + 2\cos(\varphi)$	$\lambda_1 = -1$ $\lambda_2 = e^{\varphi i}$ $\lambda_3 = e^{-\varphi i}$	Eje de giro: $N_{1,-1}$ Ángulo: φ	
Simetría respecto a un plano	$A \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$	$\det(A) = -1$ $\mathrm{tr}(A) = 1$	$\lambda_1 = \lambda_2 = 1$ $\lambda_3 = -1$	Plano de simetría: $N_{1,1}$	
Simetría respecto a una recta	$A \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$	$\det(A) = 1$ $\mathrm{tr}(A) = -1$	$\lambda_1 = 1$ $\lambda_2 = \lambda_3 = -1$	Eje de simetría: $N_{1,1}$	

Casos particulares

- Rotación con $\varphi = 0 \iff$ Identidad ($\det(A) = 1$, $\mathrm{tr}(A) = 3$)
- Rotación con $\varphi = \pi \iff$ Simetría respecto a una recta ($\det(A) = 1$, $\mathrm{tr}(A) = -1$)
- Composición de rotación y simetría con $\varphi = 0 \iff$ Simetría respecto a un plano ($\det(A) = -1$, $\mathrm{tr}(A) = 1$)
- Composición de rotación y simetría con $\varphi = \pi \iff$ Simetría central ($\det(A) = -1$, $\mathrm{tr}(A) = -3$)

Capítulo 6

Espacio Afín

6.1. Espacio afín

Dado un conjunto no vacío A a cuyos elementos definimos como "puntos", decimos que dicho conjunto A es un espacio afín asociado a un espacio vectorial E si se tienen las siguientes aplicaciones:

$$\begin{aligned} + : A \times E &\longmapsto A \\ (P + \bar{v}) &\longmapsto Q \in A \end{aligned} \qquad\qquad \begin{aligned} \phi : A \times A &\longmapsto E \\ \phi(P, Q) &\longmapsto P - Q = \overline{PQ} = \bar{v} \in E \end{aligned}$$

Para las cuales, si fijamos un punto P, obtenemos una aplicación **biyectiva**:

$$\begin{aligned} +_p : E &\longmapsto A \\ +_p(\bar{v}) = (P + \bar{v}) &\longmapsto Q \in A \end{aligned} \qquad\qquad \begin{aligned} \phi_p : A &\longmapsto E \\ \phi_p(Q) &\longmapsto P - Q = \overline{PQ} = \bar{v} \in E \end{aligned}$$

$$\begin{aligned} (+_p)^{-1} : A &\longmapsto E \\ (+_p)^{-1}(Q) = (Q - P) &\longmapsto \bar{v} \in E \end{aligned} \qquad\qquad \begin{aligned} (\phi_p)^{-1} : E &\longmapsto A \\ (\phi_p)^{-1}(\overline{PQ}) &\longmapsto \overline{PQ} - P = Q \in A \end{aligned}$$

De esta forma, veamos que todo conjunto no nulo A, para ser espacio afín, debe cumplir lo siguiente:

1. $\forall\, P, Q \in A \;\longrightarrow\; \exists!\, \bar{v} \in E$ tal que $Q = P + \bar{v}$, ya que ϕ_p es inyectiva ($\phi_p(Q) = \bar{v}$)
2. $\forall\, \bar{v}, \bar{u} \in E \;\&\; P \in A \;\longrightarrow\; (P + \bar{v}) + \bar{u} = P + (\bar{v} + \bar{u})$

6.1.1. Propiedades de un espacio afín

1. $P + \bar{v} = P \iff \bar{v} = 0$
2. $\forall P, Q \in A \implies \overline{PQ} = -\overline{QP}$
3. Relación de Chasles
$$\forall P, Q \in A \implies \overline{PQ} + \overline{QR} = \overline{PR}$$
4. $\forall P, Q, P', Q' \in A$
$$\text{Si } \overline{PQ} = \overline{P'Q'} \implies \overline{PP'} = \overline{QQ'}$$

6.2. Referencia afín

Sea A un espacio afín asociado a un espacio vectorial E. Una referencia afín de A es un conjunto formado por un punto $O \in A$, llamado origen de la referencia y una base $\{\bar{e}_1, \dots, \bar{e}_n\}$ de E.

$$R = \{O;\ \bar{e}_1, \dots, \bar{e}_n\}$$

Nota. Tomamos como dimensión de un subespacio afín A la dimensión de su espacio vectorial asociado E

Establecer una referencia sobre un espacio afín A nos permite utilizar un sitema de coordenadas para describir los puntos contenidos en él. Las coordenadas afines de un punto $P \in A$ en la referencia R coinciden exactamente con las coordenadas del vector $\overline{OP}$ en la base $\{\bar{e}_1, \dots, \bar{e}_n\}$ de E.

$$\overline{OP} = \alpha_1 \cdot \bar{e}_1 + \cdots + \alpha_n \cdot \bar{e}_n$$

$$P = (\alpha_1, \dots, \alpha_n)_R = O + \alpha_1 \cdot \bar{e}_1 + \cdots + \alpha_n \cdot \bar{e}_n$$

Generalizando, podemos decir que todo punto P de un espacio (o subespacio) afín A de espacio vectorial asociado E se puede describir a través de la siguiente ecuación:

$$\text{Punto de } A\ (P) = (\text{Solución particular } (O)) + (\text{Solución homogénea } (M_E \bar{X}^T))$$

- Solución particular (O): Punto cualquiera del espacio afín A. Lo definimos como **particular** porque se puede obtener del sistema particular (No homogéneo) de ecuaciones implícitas que describe el espacio afín.

$$O = (O_1, \dots, O_n)$$

- Solución homogénea $(M_E \bar{X}^T)$: Vector contenido en el espacio vectorial E asociado al espacio afín A. Lo definimos como **solución homogénea** porque se puede obtener resolviendo sistema homogéneo de ecuaciones implícitas (igualado al vector $\bar{0}$) que describe el espacio asociado.

$$E = \{\bar{e}_1, \dots, \bar{e}_k\} \qquad M_E = \left(\bar{e}_1 | \dots | \bar{e}_k\right)^T$$

$$\bar{X} = (\alpha_1, \dots, \alpha_k)$$

Nota. k es menor (Subespacio vectorial) o igual (Espacio vectorial) que n

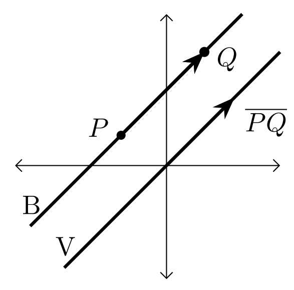
6.2.1. Referencia canónica

El sistema de referencia canónico de $\mathbb{R}^n$ es el sistema que cuenta con el vector $\bar{0}$ como punto de origen y que tiene por base asociada la base canónica del espacio vectorial $\mathbb{R}^n$.

$$R_c = \{\bar{0};\ (1, 0, \dots, 0), \dots, (0, \dots, 0, 1)\}$$

6.3. Subespacio afín

Sea A un espacio afín asociado al espacio vectorial E. Un subespacio afín de A es cualquier subcojunto $B \in A$ asociado a un subespacio vectorial $V \in E$.



6.3.1. Teorema de caracterización de subespacios afines

Un subconjunto B es un subespacio afín de A asociado al subespacio V de E si y solo si:

1. $\forall P, Q \in B \quad \overline{PQ} \in V$
2. $\forall P \in B \ \& \ \forall \bar{v} \in V \quad P + \bar{v} \in B$

Ejemplo
Demostrar que el conjunto de soluciones del sistema de ecuaciones S es un subespacio afín de $\mathbb{A}^n$ cuando el sistema es compatible.

$$S = \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= c_1 \\ \vdots \\ a_{1,m}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_n &= c_m \end{cases}$$

1. Para que S describa un subespacio afín B el sistema dado debe representar unas ecuaciones implícitas del mismo. De esta forma, se tiene que $O = (c_1, \ldots, c_n)$, y que para cualquier punto $X = (x_1, \ldots, x_n)$ de B se cumple:
$$AX = O$$
Además, los vectores resultantes de la diferencia entre cualquiera dos puntos de B verificarán el sistema homogéneo:
$$A\overline{X_1X_2} = \bar{0}$$

2. Ahora, para demostrar que es subespacio afín, debemos probar el teorema de caracterización:

a) $\forall P, Q \in B \quad \overline{PQ} \in V$ (Por reducción al absurdo)
$$P - Q \neq \overline{QP} \longrightarrow A(P-Q) \neq A\overline{QP} \longrightarrow AP - AQ \neq A\overline{QP} \longrightarrow O - O \neq \bar{0} \longrightarrow \bar{0} = \bar{0} \ !!$$

b) $\forall P \in B \ \& \ \forall \bar{v} \in V \quad P + \bar{v} \in B$ (Por reducción al absurdo)
$$A(P + \overline{PQ}) \neq O \longrightarrow AP + A\overline{PQ} \neq O \longrightarrow AP + \bar{0} \neq O \longrightarrow AP \neq O!!$$

Ejemplo
Sea $AX = C$ un sistema de ecuaciones lineales con $A \in M_{4\times3}$ y $rg(A) = rg(A|C) = 2$ se pide:

a) Sabiendo que $\bar{X}_1, \bar{X}_2 y \bar{X}_3$ son 3 soluciones de dicho sistema tales que $\bar{X}_1 - \bar{X}_2 = (1,2,1)$ y $-2\bar{X}_1 + 2\bar{X}_2 + \bar{X}_3 = (1,3,-1)$, determinar el conjunto B de soluciones del mismo.

b) Hallar el hperplano que pasa por $(1,-1,2)$ y es perpendicular a B.

c) Determinar un posible sistema de ecuaciones lineales $AX = C$ con $A \in M_{4\times3}$ cuyo conjunto de soluciones sea B.

a) Para obtener el subespacio afín B necesitamos un punto $P \in B$ y un vector $\bar{v}$ contenido en su subespacio asociado V.

1 En primer lugar veamos que $dim(B) = dim(\mathbb{R}^3) - rg(A|C) = 3 - 2 = 1$, ya que el sistema $AX = C$ representa las ecuaciones implícitas de B.

2 Como $dim(B) = 1$ el subespacio vectorial V asociado a B está generado por un único vector $\bar{v}$, que coincidirá con la diferencia de cualesquiera dos puntos contenidos en B:
$$\bar{v} = P - Q \quad \forall P, Q \in B \quad \longrightarrow \quad \bar{v} = \bar{X}_1 - \bar{X}_2 = (1,2,1)$$

3 Para obtener un punto de B podemos combinar las dos expresiones con las que contamos:
$$\bar{X}_1 - \bar{X}_2 = (1,2,1) \ \rightarrow \ (-2)(\bar{X}_1 - \bar{X}_2) = (-2)(1,2,1) \ \rightarrow \ -2\bar{X}_1 - 2\bar{X}_2 = (-2,-4,-2)$$
$$-2\bar{X}_1 + 2\bar{X}_2 + \bar{X}_3 = (1,3,-1) \ \rightarrow \ (-2,-4,-2) + \bar{X}_3 = (1,3,-1) \ \rightarrow \ \bar{X}_3 = (3,-1,1)$$

4 Por tanto, podemos describir el subespacio afín B como el siguiente:
$$B = (3,-1,1) + L\{(1,2,1)\}$$

b) Para obtener el hiperplano H que pasa por el punto $(1,-1,2)$ (Ver casos particulares de subespacios afines) podemos simplemente emplear la ecuación dada por el vector normal (que en este caso es el vector generador de V).
$$(x-1) + 2(y+1) + (z-2) = 0 \ \longrightarrow \ x + 2y + z - 1 = 0$$

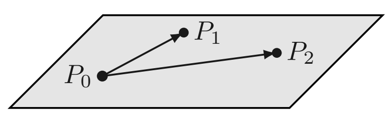
c) Como B tiene dimensión 1, para obtener un sistema de 4 ecuaciones que describa a B deberemos obtener primero las 2 ecuaciones implícitas de B y luego añadir otras dos ecuaciones redundantes.

1 Obtenemos las 2 ecuaciones implícitas de B:
$$B = (3,7,1) + L\{(1,2,1)\} \rightarrow \begin{cases} x_1 = \alpha + 3 \\ x_2 = 2\alpha + 7 \\ x_3 = \alpha + 1 \end{cases} \rightarrow \begin{cases} x_1 - x_3 = \cancel{\alpha} + 2 \\ x_2 - 2x_3 = \cancel{2\alpha} + 5 \\ x_3 = \alpha + 1 \end{cases} \rightarrow \begin{cases} x_1 - x_3 = 2 \\ x_2 - 2x_3 = 5 \end{cases}$$

2 Construimos un posible sistema:
$$A\bar{X} = C \quad \Longrightarrow \quad \begin{cases} x_1 - x_3 = 2 \\ x_2 - 2x_3 = 5 \\ \lambda ec_1 + \mu ec_2 \\ \lambda ec_1 - \mu ec_2 \end{cases}$$

6.3.2. Puntos afínmente independientes

Decimos que un conjunto de puntos $\{P_0, \ldots, P_n\}$ de un espacio afín A son afínmente independientes (a.i) o que están en posición general si y sólo si los vectores $\{\overline{P_0P_1}, \ldots, \overline{P_0P_n}\}$ son linealmente independientes.



Se denomina subespacio afín generado por los puntos $\{P_0, P_1, \ldots, P_n\}$ y se representa por $B\{P_0, P_1, \ldots, P_n\}$ al subespacio afín:
$$B\{P_0, P_1, \ldots, P_n\} = L\{\overline{P_1P_0}, \ldots, \overline{P_nP_0}\}$$

6.3.3. Casos particulares de subespacios afines

Sea A un espacio afín de dimensión n, se definen los siguientes subespacios afines:

- **Recta**: subespacio afín de dimensión 1
Una recta r se describe mediante una suma de un punto $P \in r$ y un múltiplo de un vector director $\bar{v}$
$$r = P + L\{\bar{v}\} \qquad X = P + \lambda\bar{v}$$
Nota, X denota un punto general de la recta

- **Plano**: subespacio afín de dimensión 2
Un plano π se describe mediante una suma de un punto $P \in \pi$ más un subespacio vectorial $L\{\bar{u}, \bar{v}\}$
$$\pi = P + L\{\bar{u}, \bar{v}\} = \{X = P + \lambda\bar{u} + \mu\bar{v} \ | \ \lambda, \mu \in \mathbb{R}\}$$

- **Hiperplano**: subespacio afín de dimensión $n-1$
Un hiperplano H se puede describir de 2 formas distintas:

1. Mediante la suma de un punto $P \in H$ más un subespacio vectorial $L\{\bar{u}_1, \ldots, \bar{u}_n\}$
$$H = P + L\{\bar{u}_1, \ldots, \bar{u}_{n-1}\}$$

2. Mediante un punto $P \in H$ y un vector normal o característico $\bar{c} = (c_1, \ldots, c_n)$ (vector perpendicular al hiperplano)
$$c_1(x_1 - p_1) + c_2(x_2 - p_2) + \cdots + c_n(x_n - p_n) = 0$$
$$c_1x_1 + c_2x_2 + \cdots + c_nn + D = 0$$
Nota. Se representa con una ecuación implícita porque le falta 1 dimensión para tener la máxima

6.4. Ecuaciones paramétricas e implícitas de un espacio afín

6.4.1. Ecuaciones paramétricas de un espacio afín

Las ecuaciones paramétricas de un subespacio (o espacio) afín B (de subespacio o espacio vectorial asociado $V = L\{\bar{e}_1, \ldots, \bar{e}_k\}$) describen, en función de unos ciertos parámetros variables $\{\alpha_1, \ldots, \alpha_k\}$, las posibles coordenadas que puede poseer un punto $P = (p_1, \ldots, p_n)$ contenido en B.

$$\forall P \in B \qquad \text{tq.} \qquad P = (p_1, \ldots, p_n)$$

$$P \;=\; O \;+\; M_V \bar{X}$$

$$\begin{cases} p_1 &= & O_1 &+ & v_{1,1}\alpha_1 + \cdots + v_{k,1}\alpha_k \\ \vdots &= & \vdots &+ & \vdots \\ p_n &= & O_n &+ & v_{1,n}\alpha_1 + \cdots + v_{k,n}\alpha_k \end{cases}$$

Nota. k es menor (Subespacio vectorial) o igual (Espacio vectorial) que n

Nota. Las ecuaciones paramétricas no solo son útiles para describir los puntos de un subespacio afín, sino que también nos pueden servir para llevar a cabo cambios de referencia sin necesidad de emplear matrices.

6.4.2. Ecuaciones implícitas de un espacio afín

Las ecuaciones implícitas de un subespacio afín $B = L\{\bar{e}_1, \ldots, \bar{e}_k\}$ establecen las relaciones que se tienen que cumplir entre las coordenadas de un punto $P = (p_1, \ldots, p_n)$ para que dicho punto esté contenido en B. Por tanto, si queremos comprobar si un punto P está en B, basta con ver si sus coordenadas satisfacen las ecuaciones implícitas de B.

Las ecuaciones de un subespacio afín deben formar un sistema de ecuaciones **no homogéneo**, es decir, que los términos independientes de todas las ecuaciones sean distintos de 0. Esto se debe a que precisamente lo que diferencia a un subespacio afín B de su subespacio vectorial asociado V es que el espacio afín no puede contener el punto $\bar{0}$.

Nota. A diferencia de su uso en espacios vectoriales, donde no se emplean para describir espacios de dimensión máxima (ya que sabemos que en este caso siempre van a describir el punto $\bar{0}$), cuando trabajamos en un espacio afín podemos usar un sistema de n ecuaciones y n incógnitas para definir cualquier punto distinto del $\bar{0}$.

$$\forall P \in B \qquad \text{tq.} \qquad P = (p_1, \ldots, p_n) \qquad B = \begin{cases} \lambda_{1,1}p_1 + \cdots + \lambda_{1,n}p_n &= & \lambda_{1,0} \\ \vdots & & \vdots \\ \lambda_{m,1}p_1 + \cdots + \lambda_{m,n}p_n &= & \lambda_{m,0} \end{cases}$$

$$\text{Donde:} \qquad \forall \lambda_{i,j} \;\; \lambda_{i,j} \in \mathbb{R} \qquad \text{y} \qquad m = n - k$$

No obstante, veamos que la diferencia de cualesquiera dos puntos P y Q pertenecientes a un mismo subespacio afín B, al tratarse de un vector contenido en V, sí que debe satisfacer un sistema de ecuaciones implícitas homogéneo. En concreto, el sistema homogéneo de ecuaciones implícitas que describe al subespacio asociado V es el mismo que el sistema que describe a B pero sustituyendo los coeficientes $\lambda_{i,0}$ por 0.

$$\forall P, Q \in B \qquad \text{tq.} \qquad P - Q = \overline{PQ} = \bar{v} = (v_1, \ldots, v_n) \qquad V = \begin{cases} \lambda_{1,1}v_1 + \cdots + \lambda_{1,n}v_n &= & 0 \\ \vdots & & \vdots \\ \lambda_{m,1}v_1 + \cdots + \lambda_{m,n}v_n &= & 0 \end{cases}$$

$$\text{Donde:} \qquad \forall \lambda_{i,j} \;\; \lambda_{i,j} \in \mathbb{R} \qquad \text{y} \qquad m = n - k$$

6.5. Cambio de referencia afín

Cuando hablamos de cambios de referencia nos referimos al proceso que consiste en pasar de expresar un punto o subespacio afín en una referencia R_1 a expresarlo en otra referencia R_2. (Ambas referencias de espacios afines de dimensión máxima).

Para llevar a cabo el cambio de referencia, tanto de un punto como de un subespacio afín, generalmente usamos las matrices de cambio de referencia, aunque también podemos hacerlo mediante un sistema de ecuaciones.

6.5.1. Mediante un sistema de ecuaciones paramétricas

Para llevar a cabo un cambio de referencia de un punto expresado en una referencia R_1 sin recurrir a una matriz de cambio de referencia debemos primero obtener el punto en la referencia canónica y posteriormente recurrir a un sistema de ecuaciones paramétricas para expresarlo en la referencia R_2.

$$R_1 \longrightarrow R_c \longrightarrow R_2$$

DESARROLLO

Sea $P_{R_1} = (p'_1, \dots, p'_n)_{R_1}$ un punto expresado en la referencia $R_1 = \{O_1;\ \bar{e}'_1, \dots, \bar{e}'_n\}$, para obtener este mismo punto expresado en la base $R_2 = \{O_2;\ \bar{e}''_1, \dots, \bar{e}''_n\}$ debemos seguir el siguiente desarrollo:

1. Obtenemos el punto P_{R_1}, expresado en la referencia R_1, en la referencia canónica $R_c = \{\bar{0};\ \bar{e}_1, \dots, \bar{e}_n\}$ multiplicando cada coordenada por su correspondiente vector de la referencia R_1 y sumando el origen:

$$P_{R_1} = (p'_1, \dots, p'_n)_{B_1}$$

$$P_{R_c} = (O_1 + p'_1 \cdot \bar{e}'_1 + \dots + p'_n \cdot \bar{e}'_n) = (p'_1, \dots, p'_n)_{R_c}$$

2. Utilizamos un sistema de ecuaciones paramétricas para obtener el punto P_{B_c}, expresado en la referencia canónica, en la referencia $R_2 = \{O_2;\ \bar{e}''_1, \dots, \bar{e}''_n\}$:

$$P_{B_c} = (p_1, \dots, p_n)_{R_c}$$

$$\begin{cases} p_1 &= O_{2,1} + e''_{1,1} \cdot p''_1 \cdot + \dots + e''_{n,1} \cdot p''_n \\ \vdots & \qquad \vdots \\ p_n &= O_{2,n} + e''_{1,n} \cdot p''_1 + \dots + e''_{n,n} \cdot p''_n \end{cases}$$

Al tratarse de un sistema compatible y determinado, ya que tiene mismo número de ecuaciones que de incógnitas, tiene exactamente una solución (una combinación de $\{p''_1, \dots, p''_n\}$ que satisface el sistema).

$$P_{R_2} = (p''_1, \dots, p''_n)_{R_2}$$

Ejemplo

Hallar el $P = (4, -1)$, expresado en la referencia R_1, en la referencia R_2.

$B_1 = \{(3,0);\ (-1,2), (2,3)\}$
$B_2 = \{(5,-2);\ (1,1), (-4,1)\}$

1. Obtenemos el punto P_{R_1} en referencia canónica:

$$\bar{v}_{B_c} = ((3,0) + 4 \cdot (-1,2) - 1 \cdot (2,3)) = (-3,5)_{B_c}$$

2. Utilizamos un sistema de ecuaciones paramétricas para obtener el punto P_{R_c} en la referencia R_2:

$$\begin{cases} -3 = 5 + 1 \cdot p''_1 - 4 \cdot p''_2 \\ 5 = -2 + 1 \cdot p''_1 + 1 \cdot p''_2 \end{cases} \longrightarrow \begin{array}{l} p''_1 = 4 \\ p''_2 = 3 \end{array} \longrightarrow P_{R_2} = (4,3)_{R_2}$$

6.5.2. Mediante una matriz de cambio de referencia

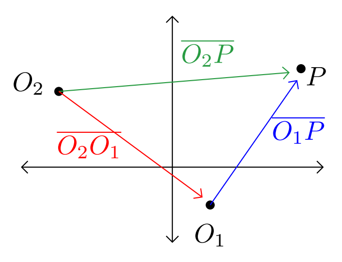
La matriz de cambio de la referencia R_1 a la base R_2 se denota de la siguiente forma ($M_{B_2 \leftarrow B_1}$), y al ser multiplicada por un punto en una referencia R_1 produce ese mismo punto en una referencia R_2:

$$\begin{pmatrix} p''_1 \\ p''_2 \\ p''_3 \\ \vdots \end{pmatrix}_{R_2} = M_{R_2 \leftarrow R_1} \begin{pmatrix} p'_1 \\ p'_2 \\ p'_3 \\ \vdots \end{pmatrix}_{R_1}$$

CONSTRUCCIÓN DE LA MATRIZ DE CAMBIO DE REFERENCIA

Dado un punto P en un sistema de referencia $R_1 = \{O_1;\ \bar{e}'_1, \dots, \bar{e}'_n\}$, para hallar las coordenadas de ese mismo punto en otro sistema de referencia $R_2 = \{O_2;\ \bar{e}''_1, \dots, \bar{e}''_n\}$ debemos obtener las coordenadas de el vector $\overline{O_2P}$ expresadas en la base $B_2 = \{\bar{e}''_1, \dots, \bar{e}''_n\}$.

Para obtener el vector $\overline{O_2P}$ en base B_2 nos aprovechamos de la relación de Chasles, que nos da la igualdad:



$$P_{R_2} = (\overline{O_2P})_{B_2} = (\overline{O_2O_1})_{B_2} + (\overline{O_1P})_{B_2}$$

Por tanto, para obtener el punto P en la referencia R_2 basta con obtener los vectores $\overline{O_2O_1}$ y $\overline{O_1P}$ en la base B_2 y posteriormente sumarlos.

1. Para obtener el vector $\overline{O_1P}$ en base B_2 simplemente usamos una matriz de cambio de base:

$$(\overline{O_1P})_{B_2} = M_{B_2 \leftarrow B_1} \cdot (\overline{O_1P})_{B_1} = M_{B_2 \leftarrow B_1} \cdot P_{R_1}$$

2. Para obtener el vector $\overline{O_2O_1}$ en base B_2 podemos calcular el vector $\overline{O_1O_2}$ en base B_1, pasarlo a base B_2 (también con la matriz de cambio de base), y después multiplicarlo por (-1). No obstante, veamos que para calcular el vector $\overline{O_1O_2}$ en base B_1 primero debemos obtener el punto O_2 en la referencia R_1:

$$(\overline{O_2O_1})_{B_2} = -(\overline{O_1O_2})_{B_2} = M_{B_2 \leftarrow B_1} \cdot (-\overline{O_1O_2})_{B_1} = (-1) \cdot M_{B_2 \leftarrow B_1} \cdot (O_2)_{R_1}$$

Nota. Para simplificar la notación denotemos al vector $(\overline{O_2O_1})_{B_2}$ por $\bar{b} = (b_1, \dots, b_n)$.

Explicado el procedimiento, veamos como trasladarlo a una matriz que permita hacerlo todo al mismo tiempo:

$$P_{R_2} = \bar{b} + M_{B_2 \leftarrow B_1} \cdot P_{R_1} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} + \begin{pmatrix} & M_{B_2 \leftarrow B_1} & \end{pmatrix} \begin{pmatrix} p'_1 \\ \vdots \\ p'_n \end{pmatrix}_{R_1}$$

Podemos eliminar la suma de los dos vectores después del producto de $M_{B_2 \leftarrow B_1}$ por P_{R_1} haciendo lo siguiente:

$$\begin{pmatrix} p''_1 \\ \vdots \\ p''_n \end{pmatrix}_{R_2} = 1 \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} + \begin{pmatrix} & M_{B_2 \leftarrow B_1} & \end{pmatrix} \begin{pmatrix} p'_1 \\ \vdots \\ p'_n \end{pmatrix}_{R_1} = \left(\begin{array}{c|c} b_1 & \\ \vdots & M_{B_2 \leftarrow B_1} \\ b_n & \end{array}\right) \begin{pmatrix} 1 \\ \hline p'_1 \\ \vdots \\ p'_n \end{pmatrix}_{R_1}$$

El problema que surge al aplicar la equivalencia anterior es que el número de coordenadas de el punto P_{R_1} es mayor que el número de coordenadas del punto P_{R_2} (debido al 1 añadido al inicio de las coordenadas de P_{R_1}). Esta inconsistencia nos impide concatenar matrices de cambio de referencia una tras otra para cambiar pasando por una referencia intermedia como hacíamos con las matrices de cambio de base.

$$M_{R_3 \leftarrow R_1} = M_{R_3 \leftarrow R_2} \cdot M_{R_2 \leftarrow R_1}$$

Para arreglar el problema de la inconsistencia de coordenadas añadimos una fila de $(1, 0, \dots, 0)$ sobre la matriz, de forma que, tras la multiplicación por el punto P_{R_1}, el punto P_{R_2} se obtenga con el mismo formato:

$$\left(\begin{array}{c|c} b_1 & \\ \vdots & M_{B_2 \leftarrow B_1} \\ b_n & \end{array}\right) \begin{pmatrix} 1 \\ \hline p'_1 \\ \vdots \\ p'_n \end{pmatrix}_{R_1} \longrightarrow \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline b_1 & \\ \vdots & M_{B_2 \leftarrow B_1} \\ b_n & \end{array}\right) \begin{pmatrix} 1 \\ \hline p'_1 \\ \vdots \\ p'_n \end{pmatrix}_{R_1}$$

Por tanto, toda matriz de cambio de entre dos referencias R_1 y R_2 verifica la siguiente ecuación:

$$\begin{pmatrix} 1 \\ \hline p''_1 \\ \vdots \\ p''_n \end{pmatrix}_{R_2} = \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline b_1 & \\ \vdots & M_{B_2 \leftarrow B_1} \\ b_n & \end{array}\right) \begin{pmatrix} 1 \\ \hline p'_1 \\ \vdots \\ p'_n \end{pmatrix}_{R_1}$$

Nota. Del resultado del producto por la matriz las coordenadas del punto en la nueva referencia son únicamente $(p''_1, \dots, p''_n)$ (omitiendo el 1, ya que este solo es una adición para construir la matriz de cambio de referencia).

Ejemplo

Obtener la matriz de cambio de la referencia $R_1 = \{(1,1,1);\ (2,3,-1), (2,2,-4), (0,3,2)\}$ canónica a $R_2 = \{(1,2,4);\ (1,1,0), (1,0,2), (-1,-4,2)\}$.

1. Obtener el punto O_2 en la referencia R_1:

$$\begin{cases} 1 + 2\alpha + 2\beta &= 1 \\ 1 + 3\alpha + 2\beta + 3\gamma &= 2 \\ 1 - \alpha - 4\beta + 2\gamma &= 4 \end{cases} \longrightarrow \begin{cases} 2\alpha + 2\beta &= 0 \\ 3\alpha + 2\beta + 3\gamma &= 1 \\ \alpha - 4\beta + 2\gamma &= 3 \end{cases} \longrightarrow \begin{cases} \alpha = 1 \\ \beta = -1 \\ \gamma = 0 \end{cases}$$

2. Obtener la matriz de cambio de B_1 a B_2. Siendo $R_1 = \{O_1; B_1\}$ y $R_2 = \{O_2; B_2\}$:

$$M_{B_2 \leftarrow B_1} = M_{B_2 \leftarrow B_c} \cdot M_{B_c \leftarrow B_1} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & -4 \\ 0 & 2 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 2 & 0 \\ 3 & 2 & 3 \\ -1 & -4 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 6 & -5 \\ -1/4 & -3 & 3 \\ -1/4 & 1 & -2 \end{pmatrix}$$

3. Obtener el vector $\overline{O_2O_1}$ en base B_2:

$$(\overline{O_2O_1})_{B_2} = (-1) \cdot M_{B_2 \leftarrow B_1} \cdot (O_2)_{R_1} = (-1) \begin{pmatrix} 2 & 6 & -5 \\ -1/4 & -3 & 3 \\ -1/4 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ -11/4 \\ 5/4 \end{pmatrix}$$

4. Construir la matriz de cambio de referencia R_1 a R_2:

$$M_{R_1 \leftarrow R_2} = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 4 & 2 & 6 & -5 \\ -11/4 & -1/4 & -3 & 3 \\ 5/4 & -1/4 & 1 & -2 \end{array}\right)$$

6.5.3. Caso específico de cambio de referencia R a R_c

Para evitar las dificultades que supone la obtención del vector $\bar{b}$, generalmente se obtienen primero las matrices de cambio de referencia a la referencia canónica R_c (que debido a la lógica seguida para la construcción de la matriz son especialmente fáciles de obtener) y luego se trabaja con las propiedeades de las matrices de aplicaciones afines* para obtener otros cambios de referencia.

Nota.* Las propiedades de matrices de aplicaciones afines son equivalentes a las propiedades de matrices de aplicaciones lineales:

$$M_{R_i \leftarrow R_j} = (M_{R_j \leftarrow R_i})^{-1} \qquad M_{R_i \leftarrow R_k} = M_{R_i \leftarrow R_j} \cdot M_{R_j \leftarrow R_k}$$

Una matriz de cambio de una referencia cualquiera $R = \{O;\ \bar{e}_1, \dots, \bar{e}_n\}$ a la referencia canónica $R_c = \{O_c;\ (1,0,\dots,0), \dots, (0,\dots,0,1)\}$ es especialmente fácil de obtener debido a las siguientes siguientes dos características:

1. La matriz de cambio de base B (base asociada a la referencia R) a B_c (sea cuál sea la base B) se construye colocando los vectores de la base B en columnas (Ver tema 1).

$$M_{B_c \leftarrow B} = (\bar{e}_1 | \dots | \bar{e}_n)$$

2. El vector $\bar{b} = \overline{O_cO}$ coincide con O, ya que en la referencia canónica R_c el punto origen O_c es el $\bar{0}$.

$$\overline{O_cO} = \overline{(\bar{0})O} = O - \bar{0} = O$$

Teniendo ambas en cuenta podemos concluir que la matriz de cambio de referencia $M_{R_c \leftarrow R}$ tiene la forma:

$$M_{R_c \leftarrow R} = \left(\begin{array}{c|c|c|c} 1 & 0 & \cdots & 0 \\ \hline O_1 & & & \\ \vdots & \bar{e}_1 & \cdots & \bar{e}_n \\ O_n & & & \end{array}\right)$$

Ejemplo

Obtener la matriz de cambio de la referencia $R_1 = \{(1,1,1);\ (2,3,-1), (2,2,-4), (0,3,2)\}$ canónica a $R_2 = \{(1,2,4);\ (1,1,0), (1,0,2), (-1,-4,2)\}$ por medio de la referencia canónica.

1. Obtener las matrices de R_1 y de R_2 a la referencia canónica:

$$M_{R_c \leftarrow R_1} = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 2 & 2 & 0 \\ 1 & 3 & 2 & 3 \\ 1 & -1 & -4 & 2 \end{array}\right) \qquad M_{R_c \leftarrow R_2} = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & -1 \\ 2 & 1 & 0 & -4 \\ 4 & 0 & 2 & 2 \end{array}\right)$$

2. Obtener la matriz de cambio de referencia R_1 a referencia R_2 ($M_{R_2 \leftarrow R_1} = M_{R_2 \leftarrow R_c} M_{R_c \leftarrow R_1}$):

$$M_{R_2 \leftarrow R_1} = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 2 & 2 & 0 \\ 1 & 3 & 2 & 3 \\ 1 & -1 & -4 & 2 \end{array}\right)^{-1} \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & -1 \\ 2 & 1 & 0 & -4 \\ 4 & 0 & 2 & 2 \end{array}\right) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 4 & 2 & 6 & -5 \\ -11/4 & -1/4 & -3 & 3 \\ 5/4 & -1/4 & 1 & -2 \end{array}\right)$$

6.6. Aplicaciones afines

Sean A_1 y A_2 dos espacios afines de espacios vectoriales asociados E_1 y E_2, una aplicación afín f de A_1 en A_2 es una aplicación compuesta por una transformación de vectores (aplicación lineal) y un desplazamiento del origen. En concreto, la aplicación $f : A_1 \mapsto A_2$ es afín siempre y cuando exista la aplicación asociada $\bar{f} : E_1 \mapsto E_2$.

$$f(Q) - f(P) = \overline{f(P)f(Q)} = \bar{f}(\overline{PQ}) \in E_2 \qquad \forall P, Q \in A_1$$

Nota. Ya que $\overline{PQ} \in E_1$

A la aplicación lineal $\bar{f} : E_1 \longmapsto E_2$ se la denomina como aplicación lineal asociada a la aplicación afín f.

6.6.1. Teorema de caracterización de aplicaciones afines

Una aplicación $f : A_1 \longmapsto A_2$ es una aplicación afín si existe un punto $O_1 \in A_1$ y una aplicación lineal $\bar{f} : E_1 \longmapsto E_2$ tal que para todo punto $P \in A_1$ se cumpla (independientemente de la referencia):

$$f(P) = f(O_1) + \bar{f}(\overline{O_1 P})$$

6.6.2. Matriz de una aplicación afín

La forma de una matriz de aplicación afín es la misma que la de una matriz de cambio de referencia, ya que de hecho, los cambios de referencia son un caso particular de aplicación afín. Sea A un espacio afín de referencia $R = \{O; \bar{e}_1, \ldots, \bar{e}_n\}$, su matriz cualquier matriz de aplicación afín f de A en A' sigue la estructura:

$$M(f) = \left(\begin{array}{c|c|c|c} 1 & 0 & \cdots & 0 \\ \hline f(O) & \bar{f}(\bar{e}_1) & \cdots & \bar{f}(\bar{e}_n) \end{array}\right) = \left(\begin{array}{c|c} 1 & 0 \ \cdots \ 0 \\ \hline f(O) & M(\bar{f}) \end{array}\right)$$

6.6.3. Propiedades de las matrices de aplicaciones afines

Cualquiera dos matrices de aplicación afín del conjunto de aplicaciones $\{f\}_{x=1}^{n}$ cumple las siguientes dos propiedades análogas a las propiedades de las matrices de aplicación lineal:

1. Relación inversa entre cambios de referencia opuestos:

$$M(f_i) = (M(f_i))^{-1}$$

2. Concatenación de matrices de cambio de referencia:

$$M(f_i \circ f_j) = M(f_i) \cdot M(f_j)$$

Ejemplo
Dada la aplicación $f : \mathbb{R}^2 \longmapsto M_2$ (ambos espacios vectoriales en base canónica), calcular la matriz de entre las referencias R y R'.

$$f(p) = \begin{pmatrix} p(0) & 1 \\ p'(0) & \int_0^1 p(x)dx + 3 \end{pmatrix}$$

$$R = \{1;\ 1 + 3x, 1 + x^2, 3 - 2x + x^2\}$$

$$R' = \left\{ I;\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right\}$$

1. Obtener las matrices de cambio de referencia de R y de R' a la referencia canónica:

$$M_{R_c \leftarrow R} = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 3 \\ 0 & 3 & 0 & -2 \\ 0 & 0 & 1 & 1 \end{array}\right) \qquad M_{R_c \leftarrow R'} = \left(\begin{array}{c|cccc} 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 \\ 1 & 1 & -1 & 0 & 0 \end{array}\right)$$

2. Obtener la aplicación asociada $\bar{f}$ a la aplicación afín f:

$$f(p) = f(O) + \bar{f}(p) = \begin{pmatrix} p(0) & 1 \\ p'(0) & \int_0^1 p(x)dx + 3 \end{pmatrix}$$

$$f(O) \quad \longrightarrow \quad O = \bar{0} \quad \longrightarrow \quad f(O) = f(\bar{0}) = \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix}$$

Nota. Un punto O expresado en una referencia R = { O ; B} es igual a $\bar{0}$

$$\bar{f}(p) = f(p) - f(O) = f(p) - f(\bar{0}) = \begin{pmatrix} p(0) & 0 \\ p'(0) & \int_0^1 p(x)dx \end{pmatrix}$$

3. Obtener la matriz de la aplicación asociada $\bar{f}$ para después construir la matriz de f en R_c:

$$\left\{ \bar{f}(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ \ \bar{f}(x) = \begin{pmatrix} 0 & 0 \\ 1 & {}^1/_2 \end{pmatrix},\ \ \bar{f}(x^2) = \begin{pmatrix} 0 & 0 \\ 0 & {}^1/_3 \end{pmatrix} \right\} \to M(\bar{f}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & {}^1/_2 & {}^1/_3 \end{pmatrix}$$

$$M(f) = \left(\begin{array}{c|c} 1 & 0 \ \cdots \ 0 \\ \hline f(O) & M(\bar{f}) \end{array}\right) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 3 & 1 & {}^1/_2 & {}^1/_3 \end{array}\right)$$

4. Obtener la matriz de la aplicación afín f en referencias $R' \leftarrow R$:

$$M(f)_{R' \leftarrow R} = M_{R' \leftarrow R_c} M(f)_{R_c} M_{R_c \leftarrow R}$$

$$\left(\begin{array}{c|cccc} 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 \\ 1 & 1 & -1 & 0 & 0 \end{array}\right)^{-1} \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 3 & 1 & {}^1/_2 & {}^1/_3 \end{array}\right) \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 3 \\ 0 & 3 & 0 & -2 \\ 0 & 0 & 1 & 1 \end{array}\right) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline {}^3/_2 & {}^7/_4 & {}^7/_6 & {}^8/_3 \\ {}^{-3}/_2 & {}^{-3}/_4 & {}^{-1}/_6 & {}^1/_3 \\ 0 & -3 & 0 & 2 \\ 1 & -3 & 0 & 2 \end{array}\right)$$

6.7. Transformaciones afines

Se dice que una aplicación afín $f : A \longmapsto A$ es una transformacióon afín si el endomorfismo asociado $\bar{f}$ es biyectivo (Ya que esto es condición necesaria para considerar iguales a los espacios imagen y preimagen).

$$f : A \mapsto A \text{ es transformación afín} \quad \longrightarrow \quad \bar{f} : E \mapsto E \text{ es biyectivo}$$

6.7.1. Propiedades de una transformación afín

Si una aplicación f es una transformación afín cumple las siguientes propiedades:

1. La aplicación f es biyectiva
2. La composición de dos transformaciones afines es otra transformación afín
3. La aplicación inversa de una transformación afín es otra transformación afín
4. Una transformación afín tiene 0, 1 o infinitos puntos fijos

6.7.2. Estudio de los puntos fijos

Sea $f : A \longmapsto A$ una transformación afín. Un punto I de A es un punto fijo o invariante por f si:

$$f(P) = P \qquad \forall P \in A$$

Conocer el número de puntos fijos de una transformación afín puede aportarnos información muy útil, especialmente para el estudio y la clasificación de movimientos.

ESTUDIO DEL NÚMERO DE PUNTOS FIJOS DE UNA APLICACIÓN AFÍN

Sea $f : A \mapsto A$ una transformación afín de matriz G, y f su aplicación asociada $f : E \mapsto E$ de matriz M:

$$G = \left(\begin{array}{c|c} 1 & 0 \ \cdots \ 0 \\ \hline f(O) & M(\bar{f}) \end{array}\right)$$

Nota. Para simplificar la notación denotaremos a $M(\bar{f})$ por M.

Si tomamos un punto cualquiera P del espacio A, para ser punto fijo debe satisfacer la siguiente sistema:

$$P \text{ es punto fijo} \iff GP = P \implies f(O) + MP = P \longrightarrow MP - P = -f(O) \longrightarrow (M - I)P = -f(O)$$

$$\begin{pmatrix} & & \\ & M - I & \\ & & \end{pmatrix} \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} = - \begin{pmatrix} f(O_1) \\ \vdots \\ f(O_n) \end{pmatrix} \longrightarrow \left\{ \begin{array}{ccccc} (m_{1,1} - 1) \cdot p_1 & \cdots & (m_{1,n}) \cdot p_n & = & -f(O_1) \\ & \vdots & = & & \vdots \\ (m_{n,1}) \cdot p_1 & \cdots & (m_{n,n} - 1) \cdot p_n & = & -f(O_1) \end{array} \right.$$

Atendiendo al número de posibles soluciones para P en el sistema anterior podemos distinguir tres casos:

- 1 punto fijo $\big(|M - I| \neq 0 \longrightarrow$ Sistema Compatible Determinado $\big)$
  Si el sistema es SCD tiene una sola solución, y por tanto existe un único punto fijo.
- 0 puntos fijos $\big(|M - I| = 0$ y $rg(Amp.) \neq rg(M - I) \longrightarrow$ Sistema Incompatible $\big)$
  Si el sistema es SI tiene ecuaciones incompatibles, por tanto, al no haber ningún punto P que las satisfaga todas, no existe ningún punto fijo.
- Infinitos puntos fijos $\big(|M - I| = 0$ y $rg(Amp.) = rg(M - I) \longrightarrow$ Sistema Compatible Indeterminado $\big)$
  Si el sistema es SCI tiene ecuaciones redundantes, por tanto, al no haber suficiente concreción, existen infinitas soluciones al sistema, y consecuentemente infinitos puntos fijos.

Ejemplo

a) Calcular el transformado del punto $(1, 2)$ por el centro de giro de centro $(1, 1)$ y el ángulo $\frac{\pi}{4}$

b) Calcular el centro del giro de ángulo $\frac{\pi}{4}$ que transforma el punto $(0, 4)$ en el punto $(2, 2)$ que transforma el punto (0, 4) en el punto (2, 2).

a) Para calcular el transformado del punto $(1, 2)$ debemos construir primero la matriz $M(f)$ del movimiento, para lo que necesitamos hallar $f(O)$ y $M(\bar{f})$.

1 Lógicamente el centro de giro C tiene que ser un punto fijo (ya que es el centro), por tanto, se ha de verificar:

$$C = f(O) + M(\bar{f})C \longrightarrow (1, 1)^T = f(O) + M(\bar{f})(1, 1)^T$$

2 Aplicando lo que sabemos de matrices de transformaciones ortogonales (Ver tema 5) y el ángulo que nos dan $\frac{\pi}{4}$ podemos hallar $M(\bar{f})$:

$$M(\bar{f}) = \begin{pmatrix} cos(\theta) & -sen(\theta) \\ sen(\theta) & cos(\theta) \end{pmatrix} \xrightarrow{\pi = ^{\pi} / _4} \begin{pmatrix} \sqrt{2}/_2 & -\sqrt{2}/_2 \\ \sqrt{2}/_2 & \sqrt{2}/_2 \end{pmatrix}$$

3 Despejamos $f(O)$ de la ecuación:

$$C = f(O) + M(\bar{f})C \longrightarrow -f(O) = \big(M(\bar{f}) - I\big)C$$

$$f(O) = \begin{pmatrix} \sqrt{2}/_2 - 1 & -\sqrt{2}/_2 \\ \sqrt{2}/_2 & \sqrt{2}/_2 - 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = - \begin{pmatrix} -1 \\ \sqrt{2} - 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 - \sqrt{2} \end{pmatrix}$$

4 Construimos la matriz del movimiento y evaluamos el vector $(1, 2)$:

$$M(f) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 1 & \sqrt{2}/_2 - 1 & -\sqrt{2}/_2 \\ 1 - \sqrt{2} & \sqrt{2}/_2 & \sqrt{2}/_2 - 1 \end{array}\right)$$

$$\left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 1 & \sqrt{2}/_2 - 1 & -\sqrt{2}/_2 \\ 1 - \sqrt{2} & \sqrt{2}/_2 & \sqrt{2}/_2 - 1 \end{array}\right) \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 - ^{\sqrt{2}} /2 \\ 1 + ^{\sqrt{2}} /2 \end{pmatrix} \rightarrow f(1, 2) = \left(1 - \frac{\sqrt{2}}{2}, 1 + \frac{\sqrt{2}}{2}\right)$$

b) Para hallar el centro C del giro de $^{\pi}/4$ que transforma el punto $(0, 4)$ en el $(2, 2)$ debemos construir primero la matriz $M(f)$ del movimiento, y después despejar C en la ecuación:

$$f(0, 4) = (2, 2)$$

1 Obtenemos la imagen del origen $f(O)$ en función del centro de giro $C = (c_1, c_2)$:

$$f(O) = -\big(M(\bar{f}) - I\big)C = \begin{pmatrix} \sqrt{2}/_2 - 1 & -\sqrt{2}/_2 \\ \sqrt{2}/_2 & \sqrt{2}/_2 - 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \left(1 - \frac{\sqrt{2}}{2}\right) c_1 + \left(\frac{\sqrt{2}}{2}\right) c_2 \\ \left(-\frac{\sqrt{2}}{2}\right) c_1 + \left(1 - \frac{\sqrt{2}}{2}\right) c_2 \end{pmatrix}$$

2 Obtenemos la matriz $M(f)$ del movimiento en función del centro de giro $C = (c_1, c_2)$:

$$M(f) = \left(\begin{array}{c|c} 1 & 0 \ \cdots \ 0 \\ \hline f(O) & M(\bar{f}) \end{array}\right) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline \left(1 - \frac{\sqrt{2}}{2}\right) c_1 + \left(\frac{\sqrt{2}}{2}\right) c_2 & \sqrt{2}/_2 & -\sqrt{2}/_2 \\ \left(-\frac{\sqrt{2}}{2}\right) c_1 + \left(1 - \frac{\sqrt{2}}{2}\right) c_2 & \sqrt{2}/_2 & \sqrt{2}/_2 \end{array}\right)$$

3 Reescribimos la ecuación $f(0, 4) = (2, 2)$ en función de C y despejamos el centro de giro C:

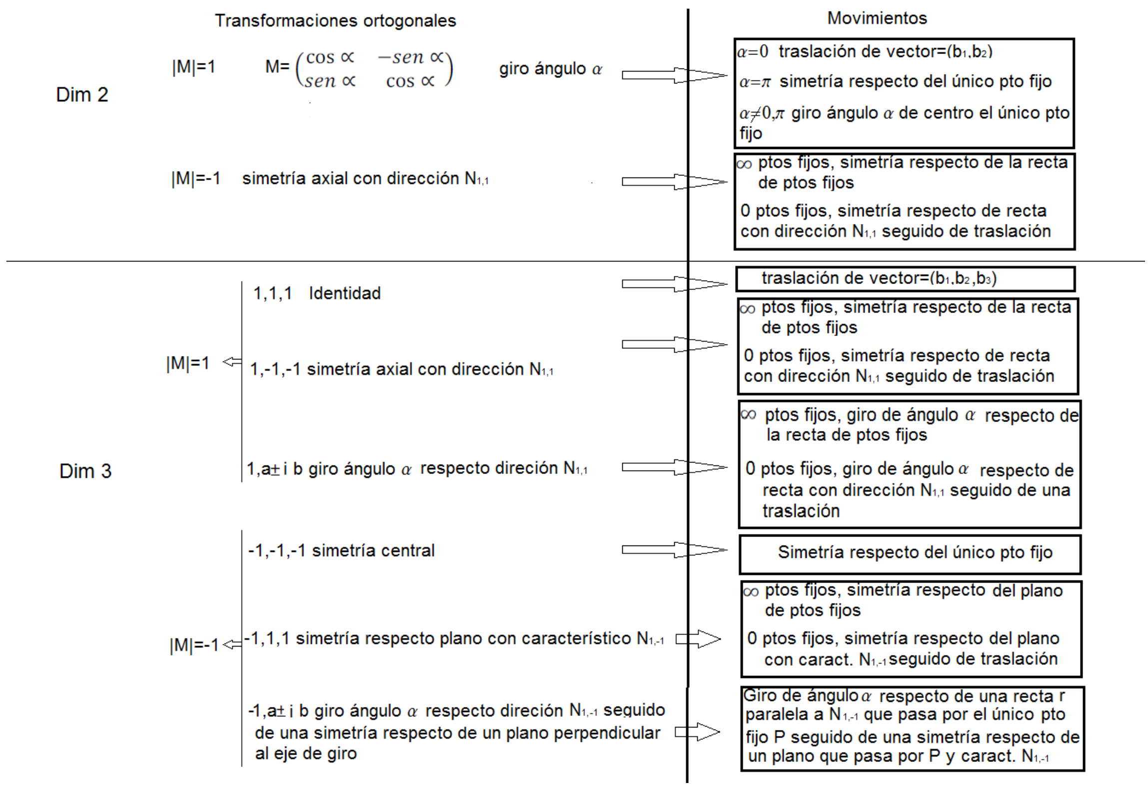
$$f(0, 4) = (2, 2) \longrightarrow M(f)(1, 0, 4)^T = (1, 2, 2)^T$$

$$\left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline \left(1 - \frac{\sqrt{2}}{2}\right) c_1 + \left(\frac{\sqrt{2}}{2}\right) c_2 & \sqrt{2}/_2 & -\sqrt{2}/_2 \\ \left(-\frac{\sqrt{2}}{2}\right) c_1 + \left(1 - \frac{\sqrt{2}}{2}\right) c_2 & \sqrt{2}/_2 & \sqrt{2}/_2 \end{array}\right) \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \iff C = (2 + \sqrt{2}, 4 + \sqrt{2})$$

6.8. Movimientos

Un movimiento es una transformación afín f tal que el endomorfismo asociado es una transformación ortogonal (una transformación cuya matriz asociada es ortogonal), es decir, que no varía la longitud de los vectores.

6.8.1. Matrices de movimientos



. .