Álgebra

Apuntes de clase

Ingeniería Biomédica

Guillermo Vera de Salas

Curso 2024-2025

Índice general

TEMA 1

Sistemas de ecuaciones, matrices y determinantes

1. Sistemas de ecuaciones lineales

Llamamos **ecuación lineal con n incógnitas x_i** a la expresión

$$a_1x_1 + a_2x_2 + \cdots a_nx_n = b.$$

Así, llamaremos **sistema de m ecuaciones lineales y n incógnitas** a la expresión

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

donde $a_{ij} \in \mathbb{R}$ son los coeficientes, $b_j \in \mathbb{R}$ los términos independientes y $x_i \in \mathbb{R}$ las incógnitas. Si todos los términos independientes son cero, se dirá que el sistema de ecuaciones es **homogéneo**. Llamaremos **solución de un sistema de ecuaciones lineales** a $s_1, ..., s_n \in \mathbb{R}$ si, al reemplazarlo por las incógnitas $x_1, ..., x_n$ respectivamente, se verifican las igualdades. Veremos en la siguiente sección de matrices cómo resolver un sistema de ecuaciones lineales.

Por ejemplo, se puede comprobar que los sistemas de dos ecuaciones y tres incógnitas

$$\left\{\begin{array}{rcl} x+y-2z & = & 0 \\ x-y+z & = & 0 \end{array}\right. \quad \text{y} \quad \left\{\begin{array}{rcl} 3x+y-3z & = & 0 \\ 5x+y-4z & = & 0 \end{array}\right.$$

tienen como solución $x = \lambda$, $y = 3\lambda$ y $z = 2\lambda$, con $\lambda \in \mathbb{R}$. En estos casos, cuando dos sistemas distintos tengan el mismo conjunto de soluciones diremos que son dos **sistemas de ecuaciones equivalentes**.

Resolver un sistema de ecuaciones lineales es una tarea mecánica y sencilla. La simpleza reside en que sabemos resolver ecuaciones lineales de una única incógnita, como por ejemplo $3x + 7 = 2$, donde basta con "despejar" la incógnita x, obteniendo que $x = -\frac{5}{3}$. Entonces, si tenemos un sistema con la siguiente estructura:

$$\left\{\begin{array}{rcl} 2x+y-z & = & 1 \\ y+z & = & 3 \\ 2z & = & 1 \end{array}\right.$$

será igual de simple que resolver tres ecuaciones lineales con una única incógnita. De nuevo, "despejando" en la última ecuación sabemos que $z = \frac{1}{2}$, así, el sistema lo podemos reescribir como:

$$\left\{\begin{array}{rcl} 2x+y-\frac{1}{2} & = & 1 \\ y+\frac{1}{2} & = & 3 \\ z & = & \frac{1}{2} \end{array}\right.$$

que, repitiendo el proceso ahora en la penúltima ecuación tenemos que $y = \frac{5}{2}$. Finalmente, el sistema puede reescribirse como:

$$\left\{\begin{array}{rcl} 2x + \frac{5}{2} - \frac{1}{2} & = & 1 \\ y & = & \frac{5}{2} \\ z & = & \frac{1}{2} \end{array}\right.$$

donde ya podemos deducir que $x = -\frac{1}{2}$. Es decir,

$$\left\{\begin{array}{rcr} x & = & -\frac{1}{2} \\ y & = & \frac{5}{2} \\ z & = & \frac{1}{2} \end{array}\right.$$

es la solución al sistema de ecuaciones. Cuando un sistema tenga la estructura anterior, diremos que es un **sistema de ecuaciones escalonado**, y se ha resuelto por el **método de remonte** o **método de sustitución regresiva**.

Puede ocurrir que el número de ecuaciones en un sistema de ecuaciones escalonado sea menor que el número de incógnitas, en ese caso hablaremos de **incógnitas independientes y dependientes**. Por ejemplo, en el sistema

$$\left\{\begin{array}{rcl} x - y + z & = & 1 \\ y + 2z & = & 2 \end{array}\right.$$

consideraremos x e y como incógnitas independientes y z como incógnita dependiente. Para evitar confusiones, identificaremos las **incógnitas dependientes por parámetros**, $z = \lambda \in \mathbb{R}$. Así, podemos reescribir el sistema como:

$$\left\{\begin{array}{rcl} x - y + \lambda & = & 1 \\ y + 2\lambda & = & 2 \end{array}\right.$$

que, resolviendo como el caso anterior, tenemos que

$$\left\{\begin{array}{rcr} x & = & 3 - 3\lambda \\ y & = & 2 - 2\lambda \\ z & = & \lambda \end{array}\right., \text{ con } \lambda \in \mathbb{R}.$$

Con estos ejemplos queda claro cómo resolver un sistema de ecuaciones escalonado. De modo que para resolver un sistema de ecuaciones cualquiera bastará con encontrar un sistema de ecuaciones escalonado equivalente. Esto se trabajará con el método de Gauss en la siguiente sección de matrices.

Los sistemas de ecuaciones lineales se pueden clasificar según el número de soluciones que tienen. Se dirá **compatible** si tiene solución, y más concretamente, **compatible determinado** si la solución es única, y **compatible indeterminado** si existen infinitas soluciones (dependientes de un/os parámetro/s). Si el sistema no tiene solución se dice **incompatible**.

Por otro lado, hay veces que resulta más conveniente trabajar con la expresión de sistema de ecuaciones que con la solución del mismo. Esto puede ocurrir cuando queremos hallar, por ejemplo, la intersección de dos planos, o más generalmente, entre dos subespacios vectoriales.

Ejemplo 1. Vamos a construir un sistema de ecuaciones lineales tal que

$$x = 2\lambda + \mu, \quad y = \lambda - 2\mu, \quad z = \lambda + \mu,$$

con $\lambda, \mu \in \mathbb{R}$, sea solución. Una forma podría ser la siguiente: Despejamos el parámetro μ de la primera identidad y la sustituimos en la segunda y tercera.

$$y = \lambda - 2(x - 2\lambda) = -2x + 5\lambda, \quad z = \lambda + x - 2\lambda = x - \lambda.$$

Despejando finalmente el parámetro λ de la identidad con z y sustituyendo en la primera.

$$y = -2x + 5(x - z) = 3x - 5z \text{ o, equivalentemente, } 3x - y - 5z = 0.$$

2. Matrices

Una matriz puede entenderse como una tabla de m filas y n columnas compuesta por escalares (números) que usualmente denotaremos con letra mayúscula A, a la entrada de la fila i y columna j con letra minúscula a_{ij}. Diremos que la matriz A es de tamaño $m \times n$ y la representaremos como

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

En función del tamaño, y sus coeficientes, las matrices pueden catalogarse como:

1. A es **cuadrada** si $m = n$.
2. A es **triangular superior** si es cuadrada y $a_{ij} = 0$ para todo $j < i$.
3. A es **triangular inferior** si es cuadrada y $a_{ij} = 0$ para todo $i < j$.
4. A es **diagonal** si es cuadrada y $a_{ij} = 0$ para todo $i \neq j$.

Hablaremos de **submatriz** de otra matriz a cualquier matriz generada a partir de restringir los índices de filas y columnas. Por ejemplo, para la matriz $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ las matrices

$$\begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{pmatrix}, \quad \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 \\ 4 & 6 \\ 7 & 9 \end{pmatrix}, \quad \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix}$$

son submatrices de A, pero

$$\begin{pmatrix} 1 & 3 & 2 \\ 4 & 6 & 5 \end{pmatrix}, \quad \begin{pmatrix} 4 & 5 \\ 1 & 2 \end{pmatrix}$$

no son submatrices de A, pues se ha permutado alguna fila y/o columna.

Llamaremos **matriz identidad de tamaño** n, I_n, a la matriz cuadrada de tamaño $n \times n$ diagonal con 1 en su diagonal, es decir,

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

y **matriz nula de tamaño** $m \times n$, $O_{m\times n}$, a la matriz de tamaño $m \times n$ donde todos sus coeficientes son cero. Si $m = n$, la denotaremos únicamente por O_m. Por ejemplo,

$$O_{1\times 2} = \begin{pmatrix} 0 & 0 \end{pmatrix}, \quad O_{2\times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad O_{3\times 2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad O_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Se pueden definir las siguientes operaciones matriciales:

1. **Producto por un escalar**. Sea $\alpha \in \mathbb{K}$ y $A = (a_{ij})$ una matriz de tamaño $m \times n$. Se define la matriz αA como la matriz de tamaño $m \times n$ con entradas αa_{ij}.
2. **Suma de dos matrices**. Sean $A = (a_{ij})$ y $B = (b_{ij})$ dos matrices del mismo tamaño $m \times n$. Se define la matriz suma $C = A + B$ como la matriz de tamaño $m \times n$ con entradas $c_{ij} := a_{ij} + b_{ij}$.
3. **Producto de dos matrices**. Sea $A = (a_{ij})$ una matriz de tamaño $m \times n$ y $B = (b_{ij})$ una matriz de tamaño $n \times p$. Se define la matriz producto $C = AB$ como la matriz de tamaño $m \times p$ con entradas $c_{ij} := \sum_{k=1}^{m} a_{ik} b_{kj}$. **Observación importante:** El producto de matrices no es conmutativo. Hay veces que ni siquiera está bien definido.

Ejemplo 2. Sean $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 3 \\ 7 \end{pmatrix}$, $C = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \end{pmatrix}$ e $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Entonces

- $A + B$ no está definida, y $A + C = \begin{pmatrix} 2 & -2 \\ 2 & 1 \end{pmatrix}$.
- $AB = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \cdot 3 + (-1) \cdot 7 \\ 1 \cdot 3 + 0 \cdot 7 \end{pmatrix} = 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 7 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -7 \\ 3 \end{pmatrix}$.
- $CX = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 2 \\ 1 \end{pmatrix} + y \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2x - y \\ x + y \end{pmatrix}$.
- El producto matricial no es conmutativo
$$AC = \begin{pmatrix} -1 & -1 \\ 2 & -1 \end{pmatrix}, \quad CA = \begin{pmatrix} -1 & -2 \\ 1 & -1 \end{pmatrix}.$$
- $AI_2 = I_2 A = A$. Por otro lado, $I_2 B = B$, pero BI_2 no está definida.

Otra manera útil de construir o denotar matrices es por bloques. Para ello utilizaremos barras horizontales y verticales para separar los distintos bloques. Por ejemplo, si $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y $B = \begin{pmatrix} 8 \\ 9 \end{pmatrix}$ entonces

$$\left(\begin{array}{c|c} 1 & O_{1\times 2} \\ \hline B & A \end{array}\right) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 8 & 1 & 2 \\ 9 & 3 & 4 \end{array}\right)$$

es una matriz cuadrada de tamaño 3×3. Otros ejemplos recurrentes pueden ser,

$$\left(A \mid B \right) = \left(\begin{array}{cc|c} 1 & 2 & 8 \\ 3 & 4 & 9 \end{array}\right), \quad \left(A \mid I_2 \right) = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right), \quad \left(\begin{array}{c} A \\ \hline I_2 \end{array}\right) = \left(\begin{array}{cc} 1 & 2 \\ 3 & 4 \\ \hline 1 & 0 \\ 0 & 1 \end{array}\right)$$

matrices de tamaño 2×3, 2×4 y 4×2, respectivamente.

Será importante entender el producto matricial en términos de la notación por bloques.

Ejemplo 3. Consideremos una matriz de tamaño $m \times n$ $A = \left(\, C_1 \mid C_2 \mid \cdots \mid C_n \, \right)$ y otra matriz cualquiera P de tamaño $m \times m$. Entonces

$$PA = \left(\, PC_1 \mid PC_2 \mid \cdots \mid PC_n \, \right).$$

Este es hecho importante que se utilizará de manera habitual. También, si en lugar de preocuparnos por las columnas, nos preocupamos por las filas, es decir, $A = \begin{pmatrix} F_1 \\ \hline \vdots \\ \hline F_m \end{pmatrix}$ y sea Q una matriz cuadrada cualquiera de tamaño $n \times n$, entonces

$$AQ = \begin{pmatrix} F_1 Q \\ \hline \vdots \\ \hline F_m Q \end{pmatrix}.$$

Definición. Sea $A = (a_{ij})$ una matriz de tamaño $m \times n$. Definimos la **matriz traspuesta de** A, denotada por $A^t = (a^t_{ij})$, como aquella de tamaño $n \times m$ tal que $a^t_{ij} := a_{ji}$.

La trasposición es lo que comúnmente nos referimos a intercambiar las filas por las columnas. Por ejemplo,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad A^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Propiedades. *Sean A y B dos matrices de tamaño $m \times n$:*

1. $(A^t)^t = A$,
2. $(A+B)^t = A^t + B^t$,
3. $(\alpha A)^t = \alpha A^t$,
4. $(AB)^t = B^t A^t$,
5. $\left(\, A \mid B \, \right)^t = \begin{pmatrix} A^t \\ \hline B^t \end{pmatrix}$.

Definición. Sea A una matriz cuadrada de tamaño $n \times n$. Diremos que A es **simétrica** si $A^t = A$. Por otra parte, diremos que A es **antisimétrica** si $A^t = -A$.

Ejemplo 4. Sea A una matriz de tamaño $m \times n$. Entonces las matrices AA^t y A^tA son simétricas. Esto es fácil de comprobar con las propiedades 1 y 4 enunciadas previamente:

$$(AA^t)^t = (A^t)^t A^t = AA^t$$

teniendo así que AA^t es simétrica.

También, cualquier matriz cuadrada B de tamaño n, puede ser expresada como una suma de matriz simétrica y antisimétrica:

$$B = \frac{B + B^t}{2} + \frac{B - B^t}{2},$$

donde $B + B^t$ es simétrica y $B - B^t$ es antisimétrica.

Definición. Sea A una matriz cuadrada de tamaño $n \times n$. Se define la **traza de** A a la suma de los elementos de su diagonal, es decir, $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}$.

Al igual que el determinante, la traza de una matriz es un número que se le asocia y tendrá su interés más adelante.

Ejemplo 5. La traza de la matriz $A = \begin{pmatrix} 2 & 1 & -4 \\ 2 & 4 & 3 \\ -1 & -2 & -7 \end{pmatrix}$ es $\operatorname{tr}(A) = 2 + 4 - 7 = -1$.

Definición. Sea A una matriz cuadrada de tamaño $n \times n$. Diremos que A es **invertible** si existe otra matriz B cuadrada de tamaño $n \times n$ tal que $AB = I_n = BA$. A la matriz B la denotaremos por A^{-1} y la llamaremos **matriz inversa de** A.

Propiedades. *Si A y B son cuadradas e invertibles, se tiene que*

1. $(A^{-1})^{-1} = A$,
2. $(AB)^{-1} = B^{-1}A^{-1}$,
3. $(A^t)^{-1} = (A^{-1})^t$.

2.1. Método de Gauss. Para describir el método de Gauss por filas, o la eliminación gaussiana por filas, debemos introducir primero las **transformaciones elementales** por filas que pueden aplicarse a una matriz. Hay tres tipos:

I. Permutar dos filas. Denotada por $F_i \leftrightarrow F_j$.
II. Multiplicar una fila por un escalar distinto de cero. Denotada por $F_i^N = \alpha F_i$, con $\alpha \neq 0$.
III. Sumar a una fila otra multiplicada por un escalar. Denotada por $F_i^N = F_i + \alpha F_j$.

Por ejemplo,

$$\begin{pmatrix} 1 & 2 & 8 \\ 3 & 4 & 9 \end{pmatrix} \xrightarrow{F_2^N = F_2 - 3F_1} \begin{pmatrix} 1 & 2 & 8 \\ 0 & -2 & -15 \end{pmatrix} \xrightarrow{F_1^N = F_1 + F_2} \begin{pmatrix} 1 & 0 & -7 \\ 0 & -2 & -15 \end{pmatrix} \xrightarrow{F_2^N = -\frac{1}{2}F_2} \begin{pmatrix} 1 & 0 & -7 \\ 0 & 1 & \frac{15}{2} \end{pmatrix}$$

El método de Gauss consiste en aplicar transformaciones elementales por filas a una matriz hasta obtener, lo que llamaremos, una forma escalonada. El algoritmo aplicado a una matriz A consiste en:

1. Escoger un elemento distinto de cero de la primera columna ignorando las columnas con pivotes y, mediante permutaciones de filas sobre la matriz A, colocarlo en la primera fila ignorando las filas con pivotes anteriores. A este elemento lo llamaremos **pivote**.
2. Mediante transformaciones elementales por filas del tipo III sobre la matriz A, hacer ceros debajo del pivote.
3. Si tras ignorar las filas y columnas con pivotes, hay elementos distintos de cero, volvemos al paso 1., en caso contrario, hemos terminado.

La matriz final que se obtiene de realizar el método de Gauss a una matriz A la llamaremos **matriz escalonada de** A. Observar que esta matriz escalonada no es única.

Por ejemplo, para la matriz del ejemplo anterior $\begin{pmatrix} 1 & 2 & 8 \\ 3 & 4 & 9 \end{pmatrix}$, tres matrices escalonadas asociadas son

$$\begin{pmatrix} \boxed{1} & 2 & 8 \\ 0 & \boxed{-2} & -15 \end{pmatrix} \quad , \quad \begin{pmatrix} \boxed{1} & 0 & -7 \\ 0 & \boxed{-2} & -15 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} \boxed{1} & 0 & -7 \\ 0 & \boxed{1} & \frac{15}{2} \end{pmatrix}.$$

Donde hemos recuadrado los pivotes para hacerlos más evidentes. Aunque una matriz tenga muchas formas escalonadas asociadas, el número de pivotes es único. Ese número de pivotes juega un papel fundamental en la asignatura, y es lo que llamaremos **rango de** A, que denotaremos por $\mathrm{rg}(A)$. Una hecho importante del cual se hará un uso implícito será $\mathrm{rg}(A) = \mathrm{rg}(A^t)$.

Ejemplo 6. Hallemos el rango de la matriz $A = \begin{pmatrix} 0 & -4 & 3 \\ 3 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix}$ aplicando el método de Gauss por filas.

Etapa 1.1. Seleccionamos un pivote, que debe ser un elemento distinto de 0 de la primera columna y realizamos permutaciones filas para colocarlo en la primera fila:

$$\begin{pmatrix} 0 & -4 & 3 \\ 3 & -1 & 0 \\ \boxed{1} & 1 & -1 \end{pmatrix} \xrightarrow{F_1 \leftrightarrow F_3} \begin{pmatrix} \boxed{1} & 1 & -1 \\ 3 & -1 & 0 \\ 0 & -4 & 3 \end{pmatrix}$$

Etapa 1.2. Realizamos transformaciones elementales de tipo III para hacer ceros debajo del pivote:

$$\begin{pmatrix} \boxed{1} & 1 & -1 \\ 3 & -1 & 0 \\ 0 & -4 & 3 \end{pmatrix} \xrightarrow{F_2^N = F_2 - 3F_1} \begin{pmatrix} \boxed{1} & 1 & -1 \\ 0 & -4 & 3 \\ 0 & -4 & 3 \end{pmatrix}$$

Etapa 1.3. Ignorando la primera fila y columna, sigue habiendo elementos distintos de cero, por lo tanto volvemos al paso 1.

Etapa 2.1. Seleccionamos un pivote, que debe ser un elemento distinto de 0 ignorando la columna con pivote:

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

Etapa 2.2 Realizamos transformaciones elementales de tipo III para hacer ceros debajo del pivote:

$$\begin{pmatrix} \boxed{1} & 1 & -1 \\ 0 & \boxed{-4} & 3 \\ 0 & -4 & 3 \end{pmatrix} \xrightarrow{F_3^N = F_3 - F_2} \begin{pmatrix} \boxed{1} & 1 & -1 \\ 0 & \boxed{-4} & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

Etapa 2.3 Ignorando las primeras y segundas filas y columnas sólo tenemos un cero y por tanto hemos terminado.

Entonces $\begin{pmatrix} \boxed{1} & 1 & -1 \\ 0 & \boxed{-4} & 3 \\ 0 & 0 & 0 \end{pmatrix}$ es una matriz escalonada de A teniendo así que $\mathrm{rg}(A) = 2$.

Las transformaciones elementales *por filas* pueden describirse en términos de productos matriciales *por la izquierda*. Y se pueden ver como transformaciones elementales aplicadas a la matriz identidad. Por ejemplo, para describir las transformaciones elementales por filas en una matriz con tres filas:

1. Permutar la primera y segunda fila:
$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{pmatrix}.$$
2. Multiplicar la primera fila por un escalar distinto de cero:
$$\begin{pmatrix} \alpha & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} \alpha & 2\alpha & 3\alpha \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$
3. Sumar a la tercera fila la primera multiplicada por -7:
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -7 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & -6 & -12 \end{pmatrix}.$$

A estas matrices se les llaman **matrices elementales**. *Todas las matrices elementales son invertibles.*

Ejemplo 7. Recuperando las tres transformaciones elementales por filas realizadas durante el método de Gauss en el ejemplo anterior sobre la matriz $A = \begin{pmatrix} 0 & -4 & 3 \\ 3 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix}$, tenemos que
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -4 & 3 \\ 3 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & -4 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$
Es importante tener en cuenta el orden, ya que el producto matricial no es conmutativo.

Definición. Sean A y B dos matrices de tamaño $m \times n$. Diremos que A y B son **equivalentes por filas** (resp. por columnas) si existe una matriz invertible P de tamaño $m \times m$ (resp. de tamaño $n \times n$) tal que $PA = B$ (resp. $AP = B$).

2.2. Método de Gauss-Jordan. Una vez obtenida una matriz escalonada tras aplicar el método de Gauss, podemos utilizar los pivotes de manera regresiva para, mediante transformaciones elementales por filas, obtener ceros por encima del pivote correspondiente. Por ejemplo:
$$\begin{pmatrix} 0 & -4 & 3 \\ 3 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix} \xrightarrow{Gauss} \begin{pmatrix} \boxed{1} & 1 & -1 \\ 0 & \boxed{-4} & 3 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{F_1^N = F_1 + \frac{1}{4}F_2} \begin{pmatrix} \boxed{1} & 0 & -\frac{1}{4} \\ 0 & \boxed{-4} & 3 \\ 0 & 0 & 0 \end{pmatrix}$$
Además, si hacemos que los pivotes sean igual a 1, mediante transformaciones elementales de tipo 2, obtendremos la que llamaremos **matriz escalonada reducida**. En el caso anterior, la

matriz escalonada reducida de $\begin{pmatrix} 0 & -4 & 3 \\ 3 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix}$ es $\begin{pmatrix} \boxed{1} & 0 & -\frac{1}{4} \\ 0 & \boxed{1} & -\frac{3}{4} \\ 0 & 0 & 0 \end{pmatrix}$. Una observación importante es que la *matriz escalonada reducida es única*, cosa que no es cierto para una matriz escalonada cualquiera, como ya se vió anteriormente.

2.3. Cálculo de la inversa de una matriz. Del hecho que las transformaciones lineales por filas se identifican con el producto matricial de matrices elementales por la izquierda, se puede hallar la inversa de una matriz. Por ejemplo:

$$\begin{pmatrix} \boxed{1} & 2 \\ 3 & 4 \end{pmatrix} \xrightarrow{F_2^N = F_2 - 3F_1} \begin{pmatrix} \boxed{1} & 2 \\ 0 & \boxed{\text{-2}} \end{pmatrix} \xrightarrow{F_1^N = F_1 + F_2} \begin{pmatrix} \boxed{1} & 0 \\ 0 & \boxed{\text{-2}} \end{pmatrix} \xrightarrow{F_2^N = -\frac{1}{2} F_2} \begin{pmatrix} \boxed{1} & 0 \\ 0 & \boxed{1} \end{pmatrix}$$

que visto en términos de matrices elementales, el cálculo anterior se identifica con:

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

es decir,

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

es aquella matriz que al multiplicarla por $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ se obtiene la matriz identidad, en otras palabras, $\begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$ es la inversa de A, y escribiremos

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

Sin embargo, la manera habitual de presentar el cálculo de la inversa de una matriz es aprovechando las transformaciones elementales. Para ello, ya que el producto matricial por bloques se comporta tal y como se describió en el Ejemplo 3, se trabaja sobre la matriz definida por bloque $\left(A \mid I_n \right)$. En este caso:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} \boxed{1} & 2 & 1 & 0 \\ 0 & \boxed{\text{-2}} & -3 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} \boxed{1} & 0 & -2 & 1 \\ 0 & \boxed{\text{-2}} & -3 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} \boxed{1} & 0 & -2 & 1 \\ 0 & \boxed{1} & \frac{3}{2} & -\frac{1}{2} \end{array}\right).$$

A modo resumen, si se quiere hallar la inversa de una matriz invertible A, entonces se halla la matriz escalonada reducida de la matriz por bloques $\left(A \mid I_n \right)$ obteniendo que la matriz escalonada reducida será de la forma $\left(I_n \mid A^{-1} \right)$. Por la unicidad de la matriz escalonada reducida, si ésta no coincide con la identidad, entonces la matriz no será invertible.

Más generalmente, éste razonamiento se puede utilizar para una matriz A cualquiera y deducir que, partiendo de la matriz por bloques $\left(A \mid I_n \right)$ $\left(\text{resp. } \left(\frac{A}{I_n}\right)\right)$ y mediante transformaciones elementales por filas (resp. por columnas) se llega a la matriz $\left(B \mid P \right)$ $\left(\text{resp. } \left(\frac{B}{Q}\right)\right)$ entonces se verifica que $PA = B$ (resp. $AQ = B$).

2.4. Matrices asociadas a sistemas de ecuaciones. Del sistema de ecuaciones de m ecuaciones y n incógnitas

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

podemos construir las siguientes matrices:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

que llamaremos **matriz de coeficientes**, **matriz de incógnitas** y **matriz de términos independientes** del sistema de ecuaciones, respectivamente. Ocurre que, siguiendo la definición del producto matricial, el sistema se puede expresar matricialmente como

$$AX = B.$$

Además, llamaremos **matriz ampliada** del sistema a la matriz por bloques $\left(\begin{array}{c|c} A & B \end{array}\right)$.

Si la matriz A es invertible, entonces de $AX = B$ se puede despejar y obtener que $X = A^{-1}B$ es la única solución del sistema. Dicho de otra manera, si la matriz A es invertible, el sistema de ecuaciones es compatible determinado.

Se ha visto en la primera sección que resolver un sistema de ecuaciones escalonado es sencillo utilizando el método de remonte. Por tanto, para resolver cualquier sistema de ecuaciones bastará con realizar el método de Gauss a la matriz ampliada del sistema, obteniendo así una matriz escalonada equivalente, que corresponderá con un sistema de ecuaciones escalonado equivalente al sistema inicial.

EJEMPLO 8. Consideramos el sistema de 2 ecuaciones y 2 incógnitas

$$\left\{\begin{array}{cccc} x & +2y & = & 8 \\ 3x & +4y & = & 9 \end{array}\right.$$

tiene asociada las matrices

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \end{pmatrix}, \quad B = \begin{pmatrix} 8 \\ 9 \end{pmatrix}, \quad \left(\begin{array}{c|c} A & B \end{array}\right) = \left(\begin{array}{cc|c} 1 & 2 & 8 \\ 3 & 4 & 9 \end{array}\right),$$

aplicando la eliminación gaussiana a la matriz ampliada:

$$\left(\begin{array}{cc|c} \boxed{1} & 2 & 8 \\ 3 & 4 & 9 \end{array}\right) \overset{F_2^N = F_2 - 3F_1}{\longrightarrow} \left(\begin{array}{cc|c} \boxed{1} & 2 & 8 \\ 0 & \boxed{-2} & -15 \end{array}\right)$$

que corresponde con el sistema de ecuaciones

$$\left\{\begin{array}{cccc} x & +2y & = & 8 \\ & -2y & = & -15 \end{array}\right.$$

teniendo como solución, tras aplicar el método de remonte, $y = \frac{15}{2}$ y $x = -7$. Podríamos haber hallado la matriz escalonada reducida y obtener la solución de manera inmediata.

Ejemplo 9. Consideramos el sistema de ecuaciones homogéneo de 3 ecuaciones y 3 incógnitas

$$\left\{\begin{array}{rrrcl} & -4y & +3z & = & 0 \\ 3x & -y & & = & 0 \\ x & +y & -z & = & 0 \end{array}\right.$$

tiene asociada las matrices

$$A = \begin{pmatrix} 0 & -4 & 3 \\ 3 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \left(A \mid B \right) = \left(\begin{array}{ccc|c} 0 & -4 & 3 & 0 \\ 3 & -1 & 0 & 0 \\ 1 & 1 & -1 & 0 \end{array}\right),$$

y podemos hallar la matriz escalonada reducida de la matriz ampliada:

$$\left(\begin{array}{ccc|c} 0 & -4 & 3 & 0 \\ 3 & -1 & 0 & 0 \\ 1 & 1 & -1 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} \boxed{1} & 0 & -\frac{1}{4} & 0 \\ 0 & \boxed{1} & -\frac{3}{4} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

que corresponde con el sistema de ecuaciones homogéneo de 2 ecuaciones y 3 incógnitas

$$\left\{\begin{array}{rrrcl} x & & -\frac{1}{4}z & = & 0 \\ & y & -\frac{3}{4}z & = & 0 \end{array}\right. .$$

Asignaremos como incógnitas independientes aquellas donde se ha pivotado, y dependientes las demás. En nuestro caso, x e y son las incógnitas independientes y z la incógnita dependiente, y por tanto la consideraremos un parámetro $z = \lambda \in \mathbb{R}$. Así, resolviendo por el método de remonte

$$\left\{\begin{array}{rcl} x & = & \frac{1}{4}\lambda \\ y & = & \frac{3}{4}\lambda \\ z & = & \lambda \end{array}\right. , \text{ con } \lambda \in \mathbb{R}.$$

En los sistemas de ecuaciones homogéneos, la matriz ampliada no aporta ninguna información relevante, ya que en el adjetivo homogéneo ya da a conocer que su última columna es nula y, además, en los cálculos que se realizan mediante transformaciones elementales por filas, esta columna permanecerá siempre nula. Por ello, para trabajar en la resolución de sistemas de ecuaciones homogéneos, únicamente utilizaremos su matriz de coeficientes asociada.

Ésta relación entre matrices y soluciones de sistemas de ecuaciones se puede formalizar en el siguiente teorema, donde el rango juega un papel crucial.

Teorema 1.1 (de Rouché-Frobenius). *Sea $AX = B$ la expresión matricial de un sistema de ecuaciones de m ecuaciones y n incógnitas.*

1. *Si* $\operatorname{rg}(A) = \operatorname{rg}(A|B)$ *entonces el sistema es compatible. Además,*
   a) *si* $\operatorname{rg}(A) = n$ *entonces es determinado y,*
   b) *si* $\operatorname{rg}(A) < n$ *entonces es indeterminado y tiene* $n - \operatorname{rg}(A)$ *incógnitas dependientes, que llamaremos grados de libertad.*
2. *Si* $\operatorname{rg}(A) \neq \operatorname{rg}(A|B)$ *entonces el sistema es incompatible.*

3. Determinantes

Vamos a introducir una notación únicamente para la definición de determinante. Sea A una matriz cuadrada de tamaño n. Podemos considerar una nueva matriz cuadrada de tamaño $n-1$ resultante de eliminar la fila i y columna j de la matriz A, que denotaremos por A_{ij}. Por ejemplo,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad A_{13} = \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix}, \quad A_{21} = \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix}, \quad A_{31} = \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix}$$

Definición. Sea A una matriz cuadrada de tamaño n. Se define, de manera recursiva, el **determinante de** A como

$$\det(A) = |A| := a_{11}\alpha_{11} + a_{21}\alpha_{21} + \cdots + a_{n1}\alpha_{n1}$$

donde $\alpha_{ij} := (-1)^{i+j}\det(A_{ij})$ es llamado el **cofactor** i, j **de** A**.** Y con caso base $\det(\lambda) = \lambda$ donde $\lambda \in \mathbb{R}$.

En la definición se ha expresado el determinante *desarrollando la primera columna*. Se puede calcular desarrollando cualquier fila o columna.

Ejemplo 10. Hallemos el determinante de una matriz triangular superior $A = \begin{pmatrix} a & \star & \star \\ 0 & b & \star \\ 0 & 0 & c \end{pmatrix}$.

Aplicando la definición

$$\det(A) = a_{11}\alpha_{11} + a_{21}\alpha_{21} + a_{31}\alpha_{31} = a\det\begin{pmatrix} b & \star \\ 0 & c \end{pmatrix} = ab\det(c) = abc,$$

es decir, el determinante de una matriz triangular es el producto de la diagonal.

El ejemplo nos indica una técnica de cálculo de determinante, triangularizar la matriz. Para ello nos apoyaremos en el método de Gauss y en las siguientes propiedades.

Propiedades. *Sean A y C dos matrices cuadradas de tamaño n. Sean E_I, E_{II} y E_{III} matrices elementales de tipo I, II y III respectivamente.*

1. $\det(E_I) = -1$, $\det(E_{II}) = \alpha$ *y* $\det(E_{III}) = 1$,
2. $\det(A) = \det(A^t)$,
3. $\det(AC) = \det(A)\det(C)$,
4. *A es invertible si y sólo si* $\det(A) \neq 0$,
5. *El rango de una matriz coincide con el tamaño de la mayor submatriz cuadrada con determinante no nulo.*
6. *Regla de Cramer. Sea $AX = B$ un sistema de ecuaciones y $A = (A_1|\cdots|A_n)$. Entonces*
$$x_i = \frac{\det(A_1|\cdots|B|\cdots|A_n)}{\det(A)}.$$

Ejemplo 11. Utilizando el valor del determinante para las matrices elementales y la propiedad del determinante del producto, podemos hallar fácilmente el determinante de una matriz aplicando el método de Gauss:

$$\begin{pmatrix} 0 & 5 & 1 \\ \boxed{1} & 4 & 5 \\ -3 & 2 & 7 \end{pmatrix} \xrightarrow{F_1 \leftrightarrow F_2} \begin{pmatrix} \boxed{1} & 4 & 5 \\ 0 & 5 & 1 \\ -3 & 2 & 7 \end{pmatrix} \xrightarrow{F_3^N = F_3 + 3F_1} \begin{pmatrix} \boxed{1} & 4 & 5 \\ 0 & \boxed{5} & 1 \\ 0 & 14 & 22 \end{pmatrix} \xrightarrow{F_3^N = 5F_3 - 14F_2} \begin{pmatrix} \boxed{1} & 4 & 5 \\ 0 & \boxed{5} & 1 \\ 0 & 0 & \boxed{96} \end{pmatrix}.$$

Que, traducido a matrices elementales sería:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -14 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 5 & 1 \\ 1 & 4 & 5 \\ -3 & 2 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 5 \\ 0 & 5 & 1 \\ 0 & 0 & 96 \end{pmatrix},$$

y utilizando las propiedades del producto de determinantes:

$$\det \begin{pmatrix} 0 & 5 & 1 \\ 1 & 4 & 5 \\ -3 & 2 & 7 \end{pmatrix} = \frac{\det \begin{pmatrix} 1 & 4 & 5 \\ 0 & 5 & 1 \\ 0 & 0 & 96 \end{pmatrix}}{\det \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -14 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right)} = \frac{1 \cdot 5 \cdot 96}{1 \cdot 5 \cdot 1 \cdot (-1)} = -96.$$

Es importante remarcar que en el último paso del método de Gauss se han realizado dos transformaciones elementales al mismo tiempo, una de tipo II y otra de tipo III, pero combinadas. Esto puede ocurrir siempre que las matrices elementales conmuten entre sí, si no, no es posible.

Para evitar tener que hacer el razonamiento anterior cada vez que se quiera calcular un determinante, es más cómodo expresarlo como una cadena de igualdades de la siguiente manera:

$$\begin{vmatrix} 0 & 5 & 1 \\ 1 & 4 & 5 \\ -3 & 2 & 7 \end{vmatrix} \overset{F_1 \leftrightarrow F_2}{=\!=\!=} - \begin{vmatrix} 1 & 4 & 5 \\ 0 & 5 & 1 \\ -3 & 2 & 7 \end{vmatrix} \overset{F_3^N = F_3 + 3F_1}{=\!=\!=\!=\!=} - \begin{vmatrix} 1 & 4 & 5 \\ 0 & 5 & 1 \\ 0 & 14 & 22 \end{vmatrix} \overset{F_3^N = 5F_3 - 14F_2}{=\!=\!=\!=\!=\!=} -\frac{1}{5} \begin{vmatrix} 1 & 4 & 5 \\ 0 & 5 & 1 \\ 0 & 0 & 96 \end{vmatrix} = -96.$$

No obstante, esta técnica no siempre resultará útil.

Ejemplo 12. Un ejercicio recurrente será el de hallar las raíces del polinomio resultante de determinantes como el siguiente:

$$\begin{vmatrix} 1-\lambda & \frac{1}{2} & -1 \\ \frac{1}{2} & 1-\lambda & -1 \\ -1 & -1 & \frac{5}{2}-\lambda \end{vmatrix} \overset{C_2^N = C_2 - C_1}{=\!=\!=\!=} \begin{vmatrix} 1-\lambda & \lambda-\frac{1}{2} & -1 \\ \frac{1}{2} & \frac{1}{2}-\lambda & -1 \\ -1 & 0 & \frac{5}{2}-\lambda \end{vmatrix} \overset{F_2^N = F_2 + F_1}{=\!=\!=\!=} \begin{vmatrix} 1-\lambda & \lambda-\frac{1}{2} & -1 \\ \frac{3}{2}-\lambda & 0 & -2 \\ -1 & 0 & \frac{5}{2}-\lambda \end{vmatrix}$$

$$= -\left(\lambda - \frac{1}{2}\right) \begin{vmatrix} \frac{3}{2}-\lambda & -2 \\ -1 & \frac{5}{2}-\lambda \end{vmatrix} = \left(\frac{1}{2} - \lambda\right)\left(\lambda^2 - 4\lambda + \frac{7}{4}\right) = \left(\frac{1}{2} - \lambda\right)^2 \left(\frac{7}{2} - \lambda\right)$$

teniendo así que las raíces son $\frac{1}{2}$ doble y $\frac{7}{2}$ simple.

TEMA 2

Espacios vectoriales

1. Definiciones básicas

DEFINICIÓN. Sea V un conjunto cualquiera y $\mathbb{K}$ el cuerpo real o complejo. Diremos que la terna $(V, +, \cdot_{\mathbb{K}})$ es un $\mathbb{K}$**-espacio vectorial** si, para cualesquiera $\vec{u}, \vec{v}, \vec{w} \in V$ y $\lambda, \mu \in \mathbb{K}$, se verifica:

1. La suma es asociativa. $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$.
2. La suma es conmutativa. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$.
3. La suma tiene neutro. $\vec{0} + \vec{u} = \vec{u}$.
4. La suma tiene opuesto. Para todo $\vec{u} \in V$ existe $\vec{v} \in V$ tal que $\vec{u} + \vec{v} = \vec{0}$.
5. El producto es asociativo. $\lambda \cdot_{\mathbb{K}} (\mu \cdot_{\mathbb{K}} \vec{u}) = (\lambda \cdot_{\mathbb{K}} \mu) \cdot_{\mathbb{K}} \vec{u}$.
6. El producto tiene neutro. $1 \cdot_{\mathbb{K}} \vec{u} = \vec{u}$.
7. El producto es distributivo respecto la suma vectorial.
8. El producto es distributivo respecto la suma escalar.

A los elementos de V se les llamarán vectores. De aquí en adelante, al producto por un escalar de $\mathbb{K}$, $\cdot_{\mathbb{K}}$, lo denotaremos por $\cdot$ o, simplemente, por yuxtaposición.

EJEMPLO 13. Los siguientes ejemplos conformarán los espacios vectoriales más recurrentes durante la asignatura, que no los únicos:

1. El plano real $\mathbb{R}^2$. El conjunto de puntos de dos coordenadas:

$$\mathbb{R}^2 := \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x, y \in \mathbb{R} \right\},$$

   con las operaciones definidas como

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} := \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}, \qquad \text{y} \qquad \lambda \begin{bmatrix} x \\ y \end{bmatrix} := \begin{bmatrix} \lambda x \\ \lambda y \end{bmatrix},$$

   con $\lambda \in \mathbb{R}$, forma un $\mathbb{R}$-espacio vectorial.
2. Sea $n > 1$ un número natural. $\mathbb{R}^n$, el conjunto de puntos de n coordenadas:

$$\mathbb{R}^n := \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mid x_1, x_2, ..., x_n \in \mathbb{R} \right\},$$

con las operaciones definidas como

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} := \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}, \quad \text{y} \quad \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} := \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix},$$

con $\lambda \in \mathbb{R}$, forman un $\mathbb{R}$-espacio vectorial.

3. Sea $n > 1$ un número natural. El conjunto de polinomios de grado menor o igual que n

$$\mathbb{P}_n := \left\{ \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + \cdots + a_n x^n \mid a_i \in \mathbb{R}, i = 0, ..., n \right\},$$

con operaciones definidas como

$$\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i := \sum_{i=0}^{n} (a_i + b_i) x^i, \quad \text{y} \quad \lambda \sum_{i=0}^{n} a_i x^i := \sum_{i=0}^{n} (\lambda a_i) x^i,$$

con $\lambda \in \mathbb{R}$, forman un $\mathbb{R}$-espacio vectorial.

4. Sean $m, n > 1$ dos números naturales. El conjunto de matrices de tamaño $m \times n$

$$\mathcal{M}_{m\times n}(\mathbb{R}) := \left\{ A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \mid a_{ij} \in \mathbb{R}, \text{ con } i = 1, ..., m \text{ y } j = 1, ..., n \right\},$$

con operaciones definidas como

$$(a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij}) \quad \text{y} \quad \lambda(a_{ij}) := (\lambda a_{ij}),$$

con $\lambda \in \mathbb{R}$, forman un $\mathbb{R}$-espacio vectorial.

Definición. Sea V un $\mathbb{K}$-espacio vectorial. A la expresión

$$\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \cdots + \lambda_n \vec{u}_n$$

la llamaremos **combinación lineal** de los vectores $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n \in V$ con escalares $\lambda_1, ..., \lambda_n \in \mathbb{K}$.

Ejemplo 14.

1. En $\mathbb{R}^2$, una combinación lineal de los vectores $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} 4 \\ -6 \end{bmatrix}$ puede ser

$$2 \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} + (-5) \cdot \begin{bmatrix} 4 \\ -6 \end{bmatrix} = \begin{bmatrix} -18 \\ 36 \end{bmatrix}.$$

También se dirá que el vector $\begin{bmatrix} -18 \\ 36 \end{bmatrix}$ es combinación lineal de $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} 4 \\ -6 \end{bmatrix}$. Puede entenderse que hemos *generado* el vector $\begin{bmatrix} -18 \\ 36 \end{bmatrix}$ a partir de los iniciales. Por supuesto, si cambiamos los escalares de la combinación lineal, podemos generar nuevos vectores.

2. Sin embargo, considerar nuevos escalares no necesariamente significa generar nuevos vectores. Por ejemplo,
$$\begin{bmatrix}1\\1\\3\end{bmatrix}+\begin{bmatrix}5\\-2\\1\end{bmatrix}+\begin{bmatrix}-4\\1\\-2\end{bmatrix}=\begin{bmatrix}2\\0\\2\end{bmatrix}=4\begin{bmatrix}1\\1\\3\end{bmatrix}+6\begin{bmatrix}5\\-2\\1\end{bmatrix}+8\begin{bmatrix}-4\\1\\-2\end{bmatrix}.$$
Es decir, hemos dado dos alternativas para generar el vector $\begin{bmatrix}2\\0\\2\end{bmatrix}$ a partir de los otros tres. Esto, en general, es algo indeseable, pues no sabremos cuál de las alternativas se ha considerado para generar el nuevo vector.
3. Recíprocamente, podemos estudiar si un vector es combinación lineal de otros dados. Por ejemplo, ¿es $\begin{bmatrix}2\\1\\1\end{bmatrix}$ combinación lineal de los vectores $\begin{bmatrix}1\\1\\3\end{bmatrix}$, $\begin{bmatrix}3\\1\\0\end{bmatrix}$ y $\begin{bmatrix}1\\0\\-2\end{bmatrix}$? Por la definición de combinación lineal, esta pregunta es equivalente a estudiar si existen $x_1, x_2, x_3 \in \mathbb{R}$ tales que se verifica
$$x_1\begin{bmatrix}1\\1\\3\end{bmatrix}+x_2\begin{bmatrix}3\\1\\0\end{bmatrix}+x_3\begin{bmatrix}1\\0\\-2\end{bmatrix}=\begin{bmatrix}2\\1\\1\end{bmatrix},$$
que, tal y como se presentó en el `Ejemplo` 2, la expresión de la izquierda se puede reescribir como un producto matricial
$$\begin{pmatrix}1&3&1\\1&1&0\\3&0&-2\end{pmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}2\\1\\1\end{bmatrix}$$
que a su vez denota un sistema de ecuaciones. Así, la respuesta será afirmativa si el sistema anterior es incompatible, y negativa si es incompatible. Para estudiar la compatibilidad del sistema, basta con aplicar el Teorema de Rouché-Frobenius, y para ello necesitamos hallar los rangos de la matriz de coeficientes y ampliada del sistema.
$$\left(\begin{array}{ccc|c}\boxed{1}&3&1&2\\1&1&0&1\\3&0&-2&1\end{array}\right)\to\left(\begin{array}{ccc|c}\boxed{1}&3&1&2\\0&\boxed{-2}&-1&-1\\0&-9&-5&-5\end{array}\right)\to\left(\begin{array}{ccc|c}\boxed{1}&3&1&2\\0&\boxed{-2}&-1&-1\\0&0&\boxed{-1}&-1\end{array}\right).$$
Como el rango de la matriz de coeficientes y ampliada coinciden, el sistema es compatible y por tanto el vector $\begin{bmatrix}2\\1\\1\end{bmatrix}$ es combinación lineal los otros tres. Además, podemos hallar cuál es la combinación lineal resolviendo el sistema. La solución al sistema es $x_3 = 1$, $x_2 = 0$ y $x_1 = 1$, que verifica:
$$1\cdot\begin{bmatrix}1\\1\\3\end{bmatrix}+0\cdot\begin{bmatrix}3\\1\\0\end{bmatrix}+1\cdot\begin{bmatrix}1\\0\\-2\end{bmatrix}=\begin{bmatrix}2\\1\\1\end{bmatrix}.$$

Definición. Sea V un $\mathbb{K}$-espacio vectorial. Diremos que un conjunto de vectores $\{\vec{u}_1, ..., \vec{u}_n\} \subset V$ es **linealmente independiente** si la ecuación vectorial

$$x_1\vec{u}_1 + \cdots + x_n\vec{u}_n = \vec{0}$$

tiene solución única $x_1 = ... = x_n = 0$. En caso contrario, diremos que el conjunto de vectores es **linealmente dependiente**.

Ejemplo 15.

1. Estudiemos la dependencia lineal del conjunto $\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}-1\\2\\1\end{bmatrix}\right\}$. Por definición, debemos estudiar las soluciones de la ecuación vectorial
$$x_1\begin{bmatrix}1\\1\\0\end{bmatrix} + x_2\begin{bmatrix}0\\1\\1\end{bmatrix} + x_3\begin{bmatrix}-1\\2\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$
Se puede reescribir matricialmente como
$$\begin{pmatrix}1 & 0 & -1\\1 & 1 & 2\\0 & 1 & 1\end{pmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$
Es decir, estudiar la dependencia lineal de esos tres vectores es equivalente a estudiar las soluciones del sistema de ecuaciones lineales homogéneo con matriz de coeficientes $\begin{pmatrix}1 & 0 & -1\\1 & 1 & 2\\0 & 1 & 1\end{pmatrix}$. Para discutir el sistema hallemos su rango utilizando el método de Gauss (podríamos hacerlo perfectamente con determinantes)
$$\begin{pmatrix}\boxed{1} & 0 & -1\\1 & 1 & 2\\0 & 1 & 1\end{pmatrix} \to \begin{pmatrix}\boxed{1} & 0 & -1\\0 & \boxed{1} & 3\\0 & 1 & 1\end{pmatrix} \to \begin{pmatrix}\boxed{1} & 0 & -1\\0 & \boxed{1} & 3\\0 & 0 & \boxed{-2}\end{pmatrix},$$
teniendo que, por el teorema de Rouché-Frobenius, es un sistema compatible determinado y por tanto la única solución que tiene el sistema, y la ecuación vectorial, es la nula. Así, el conjunto es linealmente independiente.
2. Podemos estudiar la dependencia en otros conjuntos de otros espacios vectoriales. Por ejemplo del conjunto $\{1+x,\ x+x^2,\ -1+2x+x^2\}$. De nuevo, por definición debemos estudiar las soluciones de la ecuación vectorial
$$y_1(1+x) + y_2(x+x^2) + y_3(-1+2x+x^2) = 0.$$
Donde hemos denotado a las incógnitas por y_i para evitar confusiones con la indeterminada x de los polinomios. De la ecuación, podemos plantear el siguiente sistema de ecuaciones lineales homogéneo, donde igualamos término a término
$$\left\{\begin{array}{cccc} y_1 & & -y_3 & = & 0\\ y_1 & +y_2 & +2y_3 & = & 0\\ & y_2 & +y_3 & = & 0\end{array}\right.,$$

que matricialmente podemos escribirlo como el sistema homogéneo $\begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}$, coincidiendo con el sistema del ejemplo anterior y concluyendo que también es un conjunto linealmente independiente.

Como se ha visto en este ejemplo, de manera general, vamos a identificar la ecuación vectorial $x_1\vec{u}_1 + \cdots + x_n\vec{u}_n = \vec{0}$ con un sistema de ecuaciones homogéneo, conformado por los "*vectores en columnas*". Entonces, estudiar si un conjunto de vectores es linealmente independiente se reducirá a estudiar si un sistema homogéneo es determinado, y en caso de ser indeterminado, el conjunto será dependiente.

Otra definición equivalente de conjunto linealmente dependiente es la siguiente.

TEOREMA 2.1. *Sea V un $\mathbb{K}$-espacio vectorial. El conjunto de vectores $\{\vec{u}_1, ..., \vec{u}_n\} \subset V$ es linealmente dependiente si, y sólo si, al menos uno de los vectores pueda expresarse como combinación lineal de los demás.*

DEMOSTRACIÓN. Supongamos que $\{\vec{u}_1, ..., \vec{u}_n\}$ es un conjunto linealmente dependiente, es decir, la ecuación vectorial

$$x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots x_n\vec{u}_n = \vec{0}$$

tiene alguna solución no nula. Sin pérdida de generalidad, supongamos que $x_1 \neq 0$. Así, como $x_1 \neq 0$, podemos despejar entonces $\vec{u}_1$ en términos de los demás:

$$\vec{u}_1 = -\frac{x_2}{x_1}\vec{u}_2 - \cdots - \frac{x_n}{x_1}\vec{u}_n,$$

o lo que es lo mismo, $\vec{u}_1$ se puede expresar como combinación lineal de los demás.

Recíprocamente, supongamos que al menos uno de ellos se puede expresar como combinación lineal de los demás. De nuevo, sin pérdida de generalidad, supongamos que $\vec{u}_1 = \lambda_2\vec{u}_2 + \cdots + \lambda_n\vec{u}_n \neq \vec{0}$, con $\lambda_2, ..., \lambda_n \in \mathbb{R}$, no todos nulos. Entonces, pasando todo al lado izquierdo tenemos que

$$\vec{u}_1 - \lambda_2\vec{u}_2 - \cdots - \lambda_n\vec{u}_n = \vec{0},$$

es decir, la ecuación vectorial $x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots x_n\vec{u}_n = \vec{0}$ admite una solución distinta a la nula y, por tanto, es un conjunto linealmente dependiente. □

Esta otra definición de dependencia lineal pone de manifiesto uno de los defectos de estos conjuntos. Suponiendo que nuestra única herramienta para *generar* nuevos vectores serán las combinaciones lineales, un conjunto dependiente nos indica que hay información redudante dentro de nuestro conjunto. Esto atenta contra la optimalidad del número de vectores para *generar* un nuevo espacio vectorial.

Un conjunto dependiente no sólo ofrece información redudante, sino que también provoca problemas de no unicidad a la hora de *generar* nuevos vectores.

EJEMPLO 16.

1. El conjunto $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$ es claramente linealmente dependiente. Esto se debe a que el sistema de ecuaciones homogéneo que debemos estudiar tiene tres incógnitas (igual al número de vectores), pero tan sólo dos ecuaciones (igual al número de

componentes del vector), haciendo imposible que el rango de la matriz de coeficientes coincida con el número de incógnitas, pues el rango será como máximo dos. De todas maneras, vamos a realizar un estudio exhaustivo,

Para estudiar la dependencia del conjunto, necesitamos estudiar la solución de la ecuación vectorial

$$x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

que tiene asociado el sistema de ecuaciones homogéneo $\begin{pmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix}$. Resolviendo este sistema mediante Gauss

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

tenemos que la solución es $x_3 = \lambda$, $x_2 = \lambda$ y $x_1 = -\lambda$, con $\lambda \in \mathbb{R}$, que podemos reescribir vectorialmente como

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \lambda \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}.$$

En particular, para $\lambda = 1$, podemos observar que efectivamente al menos uno de los vectores se puede expresar como combinación lineal del resto:

$$(-1) \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Por otro lado, vamos a comprobar que podemos expresar un nuevo vector $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ como combinación lineal de los vectores de $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$. Para ello, debemos resolver la ecuación vectorial

$$x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix},$$

que es equivalente a resolver el sistema de ecuaciones con matriz ampliada asociada $\left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 2 & 1 & 1 & 3 \end{array} \right)$.

$$\left(\begin{array}{ccc|c} \boxed{1} & -1 & 2 & 4 \\ 2 & 1 & 1 & 3 \end{array} \right) \to \left(\begin{array}{ccc|c} \boxed{1} & -1 & 2 & 4 \\ 0 & \boxed{3} & -3 & -5 \end{array} \right),$$

que tiene como solución $x_3 = \lambda$, $x_2 = -\frac{5}{3} + \lambda$ y $x_1 = \frac{7}{3} - \lambda$. Que podemos reescribir vectorialmente como

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \lambda \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} \frac{7}{3} \\ -\frac{5}{3} \\ 0 \end{bmatrix}.$$

Así, por ejemplo, para $\lambda = 0$ y $\lambda = 1$ se tiene que

$$\frac{7}{3}\begin{bmatrix}1\\2\end{bmatrix} - \frac{5}{3}\begin{bmatrix}-1\\1\end{bmatrix} + 0\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}4\\3\end{bmatrix} = \frac{4}{3}\begin{bmatrix}1\\2\end{bmatrix} - \frac{2}{3}\begin{bmatrix}-1\\1\end{bmatrix} + 1\begin{bmatrix}2\\1\end{bmatrix}.$$

Es decir, podemos generar el mismo vector de infinitas maneras.

2. Sin embargo, esto no ocurre con los conjuntos linealmente independientes. Por ejemplo, podemos considerar el conjunto $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}-1\\1\end{bmatrix}\right\}$ que es linealmente independiente. Esto se puede argumentar aprovechando el ejemplo previo, donde se ha eliminado el vector donde no se ha pivotado, consiguiendo así que en el estudio que se realice posteriormente, el sistema de ecuaciones homogéneo sea determinado, y por tanto, se tenga la independencia lineal.

   Visto que el conjunto es independiente, resolvamos ahora la ecuación vectorial

   $$x_1\begin{bmatrix}1\\2\end{bmatrix} + x_2\begin{bmatrix}-1\\1\end{bmatrix} = \begin{bmatrix}4\\3\end{bmatrix},$$

   o equivalentemente, el sistema de ecuaciones completo

   $$\left(\begin{array}{cc|c} \boxed{1} & -1 & 4 \\ 2 & 1 & 3 \end{array}\right) \to \left(\begin{array}{cc|c} \boxed{1} & -1 & 4 \\ 0 & \boxed{3} & --5 \end{array}\right),$$

   que tiene como única solución $x_2 = -\frac{5}{3}$ y $x_1 = \frac{7}{3}$, es decir,

   $$\frac{7}{3}\begin{bmatrix}1\\2\end{bmatrix} - \frac{5}{3}\begin{bmatrix}-1\\1\end{bmatrix} = \begin{bmatrix}4\\3\end{bmatrix}.$$

   Estos coeficientes $\frac{7}{3}$ y $-\frac{5}{3}$ que nos han permitido *generar* el vector $\begin{bmatrix}4\\3\end{bmatrix}$ es lo que llamaremos *coordenadas*.

Una observación muy importante pasa por ver que la solución del sistema homogéneo y del sistema completo comparten el término con el parámetro. Esto nos deja intuir que la solución de un sistema de ecuaciones completo se construye con una solución particular del sistema sumada a la solución del sistema homogéneo asociado.

Definición. Diremos que un conjunto de vectores $S = \{\vec{u}_1, ..., \vec{u}_n\}$ es un **sistema generador de** V si para todo $\vec{v} \in V$ existen $\lambda_1, ..., \lambda_n \in \mathbb{K}$ tales que

$$\vec{v} = \lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + \cdots + \lambda_n\vec{u}_n,$$

y denotaremos como $V = \mathrm{Gen}(S)$.

Decir que un conjunto es un sistema generador de un espacio vectorial es equivalente a decir que la ecuación vectorial

$$\vec{v} = x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n$$

tiene solución para todo $\vec{v} \in V$.

Ejemplo 17.

1. Veamos que el conjunto $S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right\}$ es un sistema generador de $\mathbb{R}^2$.

   Necesitamos demostrar que para todo $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$ la ecuación vectorial
$$x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$
   tiene solución. Esta ecuación es equivalente al sistema de ecuaciones con matriz ampliada $\left(\begin{array}{ccc|c} 1 & 2 & 2 & x \\ 2 & -1 & 4 & y \end{array} \right)$. Observar que en este sistema las letras x e y denotan parámetros, no incógnitas. Así, como la matriz ampliada tiene rango 2, por ejemplo porque $\begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix} = -5 \neq 0$, el sistema es compatible y por tanto tiene solución demostrando que $\mathbb{R}^2 = \mathrm{Gen}(S)$.
2. Por otro lado, veamos que el conjunto $S = \left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ no genera $\mathbb{R}^3$. Para ello, estudiemos las soluciones de la ecuación vectorial
$$x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$
   equivalente a un sistema, al cual le aplicamos el método de Gauss a su matriz ampliada para estudiar el rango.
$$\left(\begin{array}{ccc|c} \boxed{1} & 0 & 1 & x \\ -1 & 1 & 0 & y \\ 0 & 1 & 1 & z \end{array} \right) \to \left(\begin{array}{ccc|c} \boxed{1} & 0 & 1 & x \\ 0 & \boxed{1} & 1 & x+y \\ 0 & 1 & 1 & z \end{array} \right) \to \left(\begin{array}{ccc|c} \boxed{1} & 0 & 1 & x \\ 0 & \boxed{1} & 1 & x+y \\ 0 & 0 & 0 & z-x-y \end{array} \right)$$
   Así, la ecuación sólo tiene solución si $-x - y + z = 0$, por lo tanto, por ejemplo, el vector $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ no puede ser generado a partir de S, demostrando así que S no genera $\mathbb{R}^3$.

   Sin embargo, S sí que sigue generando nuevos vectores, en particular todos los que verifiquen la ecuación homogénea $-x - y + z = 0$. Ya veremos más adelante que todos los vectores que verifican la ecuación anterior formarán un plano. De esta manera, podemos decir que S genera un plano. Observar también que S es un conjunto linealmente dependiente y un sistema generador del plano con ecuación $-x-y+z = 0$.
3. Sea V un $\mathbb{K}$-espacio vectorial. Supongamos que $S = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es un sistema generador de V. Demostremos que los vectores
$$\vec{v}_1 := \vec{u}_1 + \vec{u}_2 - \vec{u}_3, \quad \vec{v}_2 := \vec{u}_1 + 2\vec{u}_2 + 3\vec{u}_3, \quad \vec{v}_3 := 2\vec{u}_1 + \vec{u}_2 + \vec{u}_3$$
   forman también un sistema generador de V. Debemos estudiar si para todo $\vec{w} \in V$ la ecuación vectorial
$$x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 = \vec{w}$$

tiene solución. La ecuación anterior es equivalente a

$$x_1(\vec{u}_1 + \vec{u}_2 - \vec{u}_3) + x_2(\vec{u}_1 + 2\vec{u}_2 + 3\vec{u}_3) + x_3(2\vec{u}_1 + \vec{u}_2 + \vec{u}_3) = \vec{w}$$

sacando factor común los vectores

$$(x_1 + x_2 + 2x_3)\vec{u}_1 + (x_1 + 2x_2 + x_3)\vec{u}_2 + (-x_1 + 3x_2 + x_3)\vec{u}_3 = \vec{w}.$$

Como S es un sistema generador, la última ecuación vectorial tiene solución, es decir, existirán $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{K}$ tales que

$$\lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + \lambda_3\vec{u}_3 = \vec{w},$$

es decir,

$$\left\{\begin{array}{rcl} x_1 + x_2 + 2x_3 & = & \lambda_1 \\ x_1 + 2x_2 + x_3 & = & \lambda_2 \\ -x_1 + 3x_2 + x_3 & = & \lambda_3 \end{array}\right.$$

Ahora bien, necesitamos ver si existe solución a este último sistema, pues esto aseguraría que hemos encontrado solución a nuestra ecuación vectorial inicial. Estudiamos el sistema mediante el rango de la matriz ampliada:

$$\left(\begin{array}{ccc|c} \boxed{1} & 1 & 2 & \lambda_1 \\ 1 & 2 & 1 & \lambda_2 \\ -1 & 3 & 1 & \lambda_3 \end{array}\right) \to \left(\begin{array}{ccc|c} \boxed{1} & 1 & 2 & \lambda_1 \\ 0 & \boxed{1} & -1 & \lambda_2 - \lambda_1 \\ 0 & 4 & 3 & \lambda_3 + \lambda_1 \end{array}\right) \to \left(\begin{array}{ccc|c} \boxed{1} & 1 & 2 & \lambda_1 \\ 0 & \boxed{1} & -1 & \lambda_2 - \lambda_1 \\ 0 & 0 & \boxed{7} & \lambda_3 + 5\lambda_1 - 4\lambda_2 \end{array}\right)$$

donde obtenemos que el sistema de ecuaciones es compatible, y por tanto, los vectores $\vec{v}_1, \vec{v}_2$ y $\vec{v}_3$ también generan a V.

DEFINICIÓN. Sea V un $\mathbb{K}$-espacio vectorial. Diremos que un conjunto $\mathcal{B} = \{\vec{b}_1, ..., \vec{b}_n\}$ es una **base de** V si $\mathcal{B}$ es linealmente independiente y un sistema generador de V. El número de vectores de una base es único, y llamaremos **dimensión de** V a dicho número, que denotaremos por $\dim(V)$.

EJEMPLO 18.

1. **Base canónica de** $\mathbb{R}^3$ es $B_c = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$, así $\dim(\mathbb{R}^3) = 3$.
2. **Base estándar/canónica de** $\mathbb{P}_n$ es $B_c = \{1, x, x^2, ..., x^n\}$ teniendo que $\dim(\mathbb{P}_n) = n + 1$.
3. **Base canónica de** $\mathcal{M}_{2\times 2}(\mathbb{R})$ es $B_c = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ y entonces $\dim(M_2(\mathbb{R})) = 4$.

Estas bases serán importantes pues las *coordenadas* serán inmediatas, como veremos en el siguiente ejemplo.

4. Vimos en el ejemplo anterior que $S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right\}$ era un sistema generador de $\mathbb{R}^2$, pero no es una base, ya que es un conjunto linealmente dependiente. Veamos cómo

extraer un subconjunto que siga siendo generador pero linealmente independiente. Para ello estudiemos la dependencia que existe:

$$x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

con sistema homogéneo asociado

$$\begin{pmatrix} \boxed{1} & 2 & 2 \\ 2 & -1 & 4 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & 2 & 2 \\ 0 & \boxed{\text{-5}} & 0 \end{pmatrix}$$

que tiene como solución $x_3 = \lambda, x_2 = 0$ y $x_1 = -2\lambda$, con $\lambda \in \mathbb{R}$. Teniendo que existe una dependencia (que era obvia) entre el primer y tercer vector.

Importante: Ser un sistema generador de un espacio vectorial significa que todo vector se puede expresar como combinación lineal del sistema, por eso, si existe dependencia en el conjunto, hablando de los vectores generadores, podremos eliminar todos aquellos que puedan ser generado a partir de los restantes.

En nuestro caso, el segundo vector no puede ser generado a partir de los restantes, por lo que no puede ser eliminado. El primero y el tercero son proporcionales, así que cualquiera puede generar al otro. Como regla general para considerar un subconjunto linealmente independiente eliminaremos aquellos en los que no se haya pivotado. Por tanto, $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}$ es un conjunto independiente y sigue siendo sistema generador de $\mathbb{R}^2$, es decir, es base.

5. Siguiendo el ejemplo anterior. Sea V un $\mathbb{K}$-espacio vectorial. Supongamos que $\mathcal{B} = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es una base de V. Entonces $\mathcal{C} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ donde

$$\vec{v}_1 := \vec{u}_1 + \vec{u}_2 - \vec{u}_3, \quad \vec{v}_2 := \vec{u}_1 + 2\vec{u}_2 + 3\vec{u}_3, \quad \vec{v}_3 := 2\vec{u}_1 + \vec{u}_2 + \vec{u}_3$$

es también una base de V. Ya se ha justificado que es un sistema generador. De la misma manera, se puede comprobar que es un conjunto linealmente independiente y por tanto es base.

Definición. Sea V un $\mathbb{K}$-espacio vectorial y $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, ..., \vec{b}_n\}$ una base. Para cualquier $\vec{u} \in V$, definimos las **coordenadas de $\vec{u}$ respecto $\mathcal{B}$** como los escalares $\lambda_1, ..., \lambda_n$ que verifican

$$\vec{u} = \lambda_1 \vec{b}_1 + \lambda_2 \vec{b}_2 + \cdots + \lambda_n \vec{b}_n,$$

y denotaremos por $[\vec{u}]_{\mathcal{B}} := \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} \in \mathbb{R}^n$. Además, si $S \subset V$ entonces escribiremos

$$[S]_{\mathcal{B}} := \{[\vec{s}]_{\mathcal{B}} \mid \vec{s} \in S\} \subset \mathbb{R}^n.$$

Las coordenadas permiten identificar cualquier vector, de cualquier espacio vectorial, con un elemento de $\mathbb{R}^n$. Esto nos permitirá trabajar matricialmente se manera consistente y sin depender de la propia estructura del espacio vectorial.

Ejemplo 19.

1. En $\mathbb{P}_2$, el cálculo de las coordenadas en base canónica es inmediato, basta con listar, en orden creciente de grado, los coeficientes del polinomio: $[5-3x]_{\mathcal{B}_c} = \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix}$. Y, por ejemplo, en $\mathbb{P}_3$, $[5-3x]_{\mathcal{B}_c} = \begin{bmatrix} 5 \\ -3 \\ 0 \\ 0 \end{bmatrix}$.
2. En $M(\mathbb{R})_2$, las coordendas en base canónica es equivalente a leer los coeficientes de la matriz por columna: $\left[\begin{pmatrix} 3 & 2 \\ 7 & -1 \end{pmatrix}\right]_{\mathcal{B}_c} = \begin{bmatrix} 3 \\ 7 \\ 2 \\ -1 \end{bmatrix}$.
3. En $\mathbb{R}^3$, si queremos hallar las coordenadas del vector $\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ en base $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ necesitamos resolver la ecuación vectorial
$$x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$$
o lo que es equivalente, resolver el sistema de ecuaciones
$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \end{array}\right)$$
teniendo como solución $x_1 = 1, x_2 = 2, x_3 = -1$, es decir, $\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$.
4. De manera idéntica, si queremos hallar las coordenadas de $\begin{pmatrix} 1 & 4 \\ 6 & -2 \end{pmatrix}$ en base $\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ tenemos que resolver la ecuación vectorial
$$x_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 6 & -2 \end{pmatrix}$$
que plantea el siguiente sistema de ecuaciones
$$\left(\begin{array}{cccc|c} \boxed{1} & 0 & 0 & 0 & 1 \\ 0 & \boxed{1} & -1 & 0 & 6 \\ 0 & 1 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & -2 \end{array}\right) \to \left(\begin{array}{cccc|c} \boxed{1} & 0 & 0 & 0 & 1 \\ 0 & \boxed{1} & -1 & 0 & 6 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 1 & -2 \end{array}\right) \to \left(\begin{array}{cccc|c} \boxed{1} & 0 & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 0 & 5 \\ 0 & 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & 0 & \boxed{1} & -2 \end{array}\right)$$

teniendo como solución $x_1 = 1, x_2 = 5, x_3 = -1, x_4 = -2$, y consecuentemente $\left[\begin{pmatrix} 1 & 4 \\ 6 & -2 \end{pmatrix}\right]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 5 \\ -1 \\ -2 \end{bmatrix}$

5. Recíprocamente, si queremos hallar un vector $\vec{u} \in \mathbb{R}^3$ tal que sus coordenadas en $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} \right\}$ son $\begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix}$ entonces, por dejar clara la relación entre vector y coordenadas,
$$[\vec{u}]_{\mathcal{B}} = \begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix} \iff \vec{u} = 4 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + (-1) \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \\ -2 \end{bmatrix}.$$

6. De igual manera, en $\mathbb{P}_3$, el polinomio $p(x)$ con coordenadas $\begin{bmatrix} 4 \\ 2 \\ -1 \\ 3 \end{bmatrix}$ en la base de polinomios $\mathcal{B} = \{1 + x - 3x^2, 4x + 2x^3, x^2 - x^3, 1 + x^2\}$, es decir, $[p(x)]_{\mathcal{B}} = \begin{bmatrix} 4 \\ 2 \\ -1 \\ 3 \end{bmatrix}$, es
$$\begin{aligned} p(x) &= 4(1 + x - 3x^2) + 2(4x + 2x^3) + (-1)(x^2 - x^3) + 3(1 + x^2) \\ &= 7 + 12x - 10x^2 + 5x^3 \end{aligned}$$

El siguiente teorema nos dice que la operación de tomar coordenadas de un vector es un operador lineal. Esta propiedad nos permitirá identificar combinaciones lineales de vectores con combinaciones lineales de coordenadas y, tal y como hemos visto, toda la teoría hasta ahora está basada en combinaciones lineales de vectores. A partir de esta sección, podremos traducirlo a combinaciones lineales de coordenadas (o vectores de $\mathbb{R}^n$) pudiendo construir siempre las matrices donde hemos discutido los resultados.

Teorema 2.2. *Sea V un $\mathbb{K}$-espacio vectorial y $\mathcal{B}$ una base. Entonces, para todo $\vec{u}, \vec{v} \in V$ y $\lambda \in \mathbb{K}$,*
$$[\vec{u} + \vec{v}]_{\mathcal{B}} = [\vec{u}]_{\mathcal{B}} + [\vec{v}]_{\mathcal{B}}, \quad y \quad [\lambda\vec{u}]_{\mathcal{B}} = \lambda[\vec{u}]_{\mathcal{B}}.$$

Demostración. Supongamos que $[\vec{u}]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ y $[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$. Esto, por definición, significa que
$$\vec{u} = x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_n\vec{b}_n, \quad \text{y} \quad \vec{v} = y_1\vec{b}_1 + y_2\vec{b}_2 + \cdots + y_n\vec{b}_n.$$
Entonces la suma
$$\begin{aligned} \vec{u} + \vec{v} &= (x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_n\vec{b}_n) + (y_1\vec{b}_1 + y_2\vec{b}_2 + \cdots + y_n\vec{b}_n) \\ &= (x_1 + y_1)\vec{b}_1 + (x_2 + y_2)\vec{b}_2 + \cdots (x_n + y_n)\vec{b}_n, \end{aligned}$$

es decir, $[\vec{u}+\vec{v}]_{\mathcal{B}} = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = [\vec{u}]_{\mathcal{B}} + [\vec{v}]_{\mathcal{B}}$.

Se deja como ejercicio demostrar que $[\lambda\vec{u}]_{\mathcal{B}} = \lambda[\vec{u}]_{\mathcal{B}}$. □

Las coordenadas nos permiten establecer la conexión, que en el siguiente tema formalizaremos, entre cualquier espacio vectorial de dimensión n y $\mathbb{R}^n$. De esta forma, a partir de ahora, podremos trabajar sobre cualquier espacio vectorial únicamente desde la perspectiva de $\mathbb{R}^n$.

Más concretamente, dado un espacio vectorial V de dimensión n con base $\mathcal{B}$, una combinación lineal

$$\lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + \cdots + \lambda_m\vec{u}_m = \vec{v}$$

se puede expresar, gracias al Teorema 2.2, en términos de coordenadas como

$$[\lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + \cdots + \lambda_m\vec{u}_m]_{\mathcal{B}} = [\vec{v}]_{\mathcal{B}} \iff \lambda_1[\vec{u}_1]_{\mathcal{B}} + \lambda_2[\vec{u}_2]_{\mathcal{B}} + \cdots + \lambda_m[\vec{u}_m]_{\mathcal{B}} = [\vec{v}]_{\mathcal{B}}$$

que, si $[\vec{u}_i]_{\mathcal{B}} = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix}$ y $[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ entonces podríamos escribirlo como

$$\lambda_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + \lambda_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \cdots + \lambda_m \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

que a su vez se reescribir matricialmente como

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Esta identificación de una combinación lineal con la expresión matricial anterior resultará muy útil y se utilizará de manera reiterada durante muchas argumentaciones. También pone de manifiesto que el cálculo de coordenadas es equivalente a resolver un sistema de ecuaciones lineales.

Definición. Sea V un $\mathbb{K}$-espacio vectorial de dimensión n y sea $\mathcal{B}$ una base. Consideremos $S = \{\vec{u}_1, ..., \vec{u}_m\} \subset V$. Llamaremos **matriz asociada a S respecto de $\mathcal{B}$** a la matriz construida por columnas $M(S)_{\mathcal{B}} = \left(\ [\vec{u}_1]_{\mathcal{B}} \mid [\vec{u}_2]_{\mathcal{B}} \mid \cdots \mid [\vec{u}_m]_{\mathcal{B}} \ \right)_{n\times m}$.

Merece la pena reincidir en que la matriz asociada a un conjunto de vectores respecto de una base de un espacio V tiene tantas filas como dimensión de V y tiene tantas columnas como vectores tenga el conjunto.

Ejemplo 20. Hallemos la matriz asociada de distintos conjuntos respecto de una base.

1. En $\mathbb{R}^3$, la matriz asociada al conjunto $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -3 \\ 2 \end{bmatrix} \right\}$ en base canónica es $M(S)_{\mathcal{B}_c} = \begin{pmatrix} 1 & 0 \\ 2 & -3 \\ -1 & 2 \end{pmatrix}$.
2. En $\mathbb{P}_3$, la matriz asociada al conjunto $S = \{x - x^2, 1 + x + x^3, 3 - 4x + x^2, 2x + x^3\}$ en base canónica es $M(S)_{\mathcal{B}_c} = \begin{pmatrix} 0 & 1 & 3 & 0 \\ 1 & 1 & -4 & 2 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$.
3. En $\mathbb{R}^2$, consideremos la base $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ y el subconjunto $S = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right\}$.
Para hallar la matriz asociada a S en $\mathcal{B}$ debemos calcular las coordenadas de los vectores de S en $\mathcal{B}$:
$$M(S)_{\mathcal{B}} = \left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\mathcal{B}} \middle| \begin{bmatrix} 1 \\ 3 \end{bmatrix}_{\mathcal{B}} \middle| \begin{bmatrix} 3 \\ -2 \end{bmatrix}_{\mathcal{B}} \right).$$
Para calcular las coordenadas tendremos que resolver los siguientes sistemas de ecuaciones.
$$\begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \iff x_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \iff \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 \\ 3 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \iff x_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \iff \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$
$$\begin{bmatrix} 3 \\ -2 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \iff x_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \end{bmatrix} \iff \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$
que tienen la misma matriz de coeficientes, tan sólo cambiando los términos independientes. Entonces, observar, que la resolución del sistema sólo dependerá de las transformaciones filas necesarias para hallar la inversa de la matriz de coeficientes:
$$\left(\begin{array}{cc|ccc} 1 & 1 & 2 & 1 & 3 \\ -1 & 1 & 1 & 3 & -2 \end{array} \right) \to \left(\begin{array}{cc|ccc} 1 & 1 & 2 & 1 & 3 \\ 0 & 2 & 3 & 4 & 1 \end{array} \right) \to \left(\begin{array}{cc|ccc} 1 & 0 & \frac{1}{2} & -1 & \frac{5}{2} \\ 0 & 1 & \frac{3}{2} & 2 & \frac{1}{2} \end{array} \right)$$
teniendo que
$$\begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} \frac{1}{2} \\ \frac{3}{2} \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 3 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ -2 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} \frac{5}{2} \\ \frac{1}{2} \end{bmatrix},$$
y, por tanto,
$$M(S)_{\mathcal{B}} = \left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\mathcal{B}} \middle| \begin{bmatrix} 1 \\ 3 \end{bmatrix}_{\mathcal{B}} \middle| \begin{bmatrix} 3 \\ -2 \end{bmatrix}_{\mathcal{B}} \right) = \begin{pmatrix} \frac{1}{2} & -1 & \frac{5}{2} \\ \frac{3}{2} & 2 & \frac{1}{2} \end{pmatrix},$$
que es distinta de la matriz $M(S)_{\mathcal{B}_c} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 3 & -2 \end{pmatrix}$.

Importante observar cómo, mediante el método de Gauss-Jordan, ha sido posible resolver de manera simultánea más de un sistema de ecuaciones que comparten matriz de coeficientes.

4. En $\mathbb{R}^2$, consideremos la base $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$. Entonces
$$M(\mathcal{B})_{\mathcal{B}} = \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}_{\mathcal{B}} \middle| \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{B}} \right) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

TEOREMA 2.3. *Sea V un $\mathbb{K}$-espacio vectorial con $\mathcal{B} = \{\vec{b}_1, ..., \vec{b}_n\}$ base de V. Sea $S = \{\vec{u}_1, ..., \vec{u}_m\} \subset V$ y $M(S)_{\mathcal{B}}$ su matriz asociada respecto de $\mathcal{B}$. Entonces*

1. *S es linealmente independiente si, y solo si, $\mathrm{rg}(M(S)_{\mathcal{B}}) = m =$ nº columnas.*
2. *S es un sistema generador de V si, y solo si, $\mathrm{rg}(M(S)_{\mathcal{B}}) = n =$ nº filas.*
3. *S es base de V si, y solo si, $\mathrm{rg}(M(S)_{\mathcal{B}}) = m = n$, o dicho de otra forma, la matriz es cuadrada de rango máximo.*

DEMOSTRACIÓN.

1. $\boxed{\Rightarrow}$ Supongamos que S es linealmente independiente. Esto implica que la ecuación vectorial
$$x_1\vec{u}_1 + \cdots + x_m\vec{u}_m = \vec{0}$$
tiene la solución nula como la única. Consideremos coordenadas en $\mathcal{B}$ en la expresión anterior:
$$[x_1\vec{u}_1 + \cdots + x_m\vec{u}_m]_{\mathcal{B}} = [\vec{0}]_{\mathcal{B}} \quad \Longleftrightarrow \quad x_1[\vec{u}_1]_{\mathcal{B}} + \cdots + x_m[\vec{u}_m]_{\mathcal{B}} = [\vec{0}]_{\mathcal{B}}$$
donde hemos utilizado la linealidad del operador coordenadas. A su vez, la última expresión
$$\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = [\vec{0}]_{\mathcal{B}} = x_1[\vec{u}_1]_{\mathcal{B}} + \cdots + x_m[\vec{u}_m]_{\mathcal{B}} = \left([\vec{u}_1]_{\mathcal{B}} \mid \cdots \mid [\vec{u}_m]_{\mathcal{B}} \right) \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = M(S)_{\mathcal{B}} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}.$$
Como la solución es única, esto significa que este sistema de ecuaciones homogéneos es compatible determinado, es decir, $\mathrm{rg}(M(S)_{\mathcal{B}}) =$ nº columnas $= m$.

$\boxed{\Leftarrow}$ Recíprocamente, supongamos que $\mathrm{rg}(M(S)_{\mathcal{B}}) = m$. Queremos estudiar las soluciones de la ecuación vectorial
$$x_1\vec{u}_1 + \cdots + x_m\vec{u}_m = \vec{0}.$$
De nuevo, esta ecuación se puede identificar con el sistema de ecuaciones homogéneo
$$M(S)_{\mathcal{B}} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$
Como $\mathrm{rg}(M(S)_{\mathcal{B}}) = m =$ número de incógnitas, por el teorema de Rouché-Frobenius, el sistema es determinado y tiene solución única, así, la ecuación vectorial también tiene solución única, teniendo que el conjunto S es linealmente independiente.

2. Siguiendo un razonamiento similar al anterior.

3. Es inmediato de los puntos anteriores.

□

Otras lecturas que se pueden hacer al teorema son: Para que un conjunto de vectores sea linealmente independiente el número de vectores que contiene debe ser *como máximo igual a la dimensión*, por otro lado, para que sea sistema generador este número debe ser *como mínimo igual a la dimensión*. De modo que un conjunto será base si se encuentra el equilibrio entre independencia y generador.

EJEMPLO 21. Gracias al teorema anterior, *si conocemos una base del espacio vectorial*, estudiar las características de un conjunto se traduce en estudiar el rango de una matriz.

1. En $\mathbb{P}_2$, estudiemos el conjunto $S = \{1-x^2,\, 2x+3x^2,\, 1+x+x^2,\, 1-x\}$. Consideramos la matriz asociada a S en $\mathcal{B}_c$, $M(S)_{\mathcal{B}_c} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & -1 \\ -1 & 3 & 1 & 0 \end{pmatrix}$. Como la matriz no es cuadrada ya sabemos que S no puede ser una base. En particular, no puede ser base porque S no puede ser linealmente independiente, ya que $\mathrm{rg}(M(S)_{\mathcal{B}_c}) \leqslant 3 \neq 4$. Por otro lado, como $\begin{vmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ -1 & 3 & 1 \end{vmatrix} = 2-(-2+3) = 1 \neq 0$, entonces $\mathrm{rg}(M(S)_{\mathcal{B}_c}) = 3 = \dim(\mathbb{P}_2)$, teniendo que S es un sistema generador de $\mathbb{P}_2$, pero no una base.
2. Sea V un $\mathbb{K}$-espacio vectorial y $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ una base. Demostremos que $\mathcal{C} = \{-2\vec{b}_1 + 3\vec{b}_2, \vec{b}_1 + 4\vec{b}_2\}$ es una base de V. Bastará con estudiar el rango de la matriz

$$M(\mathcal{C})_{\mathcal{B}} = \left(\, [-2\vec{b}_1 + \vec{b}_2]_{\mathcal{B}} \;\middle|\; [\vec{b}_1 + 4\vec{b}_2]_{\mathcal{B}} \, \right) = \begin{pmatrix} -2 & 1 \\ 3 & 4 \end{pmatrix}.$$

Que tiene rango dos ya que, por ejemplo, $\begin{vmatrix} -2 & 1 \\ 3 & 4 \end{vmatrix} = -11 \neq 0$. Demostrando así que $\mathcal{C}$ también es base de V.

2. Matriz de cambio de base

Para un mismo espacio vectorial pueden considerarse más de una base. Por tanto, podemos preguntarnos que relación existirá entre las coordenadas de un mismo vector en distintas bases.

Sea V un $\mathbb{K}$-espacio vectorial y sean $\mathcal{B} = \{\vec{b}_1, ..., \vec{b}_n\}$ y $\mathcal{C} = \{\vec{c}_1, ..., \vec{c}_n\}$ dos bases de V. Si $\vec{u} \in V$, entonces

$$\vec{u} = x_1\vec{b}_1 + x_2\vec{b}_2 + ... + x_n\vec{b}_n, \text{ es decir, } [\vec{u}]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

y

$$\vec{u} = y_1\vec{c}_1 + y_2\vec{c}_2 + ... + y_n\vec{c}_n, \text{ es decir, } [\vec{u}]_{\mathcal{C}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Reescribiendo la primera ecuación vectorial en coordenadas respecto de C tenemos:

$$[\vec{u}]_{\mathcal{C}} = x_1[\vec{b}_1]_{\mathcal{C}} + x_2[\vec{b}_2]_{\mathcal{C}} + ... + x_n[\vec{b}_n]_{\mathcal{C}} = \left(\ [\vec{b}_1]_{\mathcal{C}} \ \middle| \ [\vec{b}_2]_{\mathcal{C}} \ \middle| \ \cdots \ \middle| \ [\vec{b}_n]_{\mathcal{C}} \ \right) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = M(\mathcal{B})_{\mathcal{C}}[\vec{u}]_{\mathcal{B}}.$$

Así, a la matriz asociada a $\mathcal{B}$ en $\mathcal{C}$, la llamaremos **matriz de cambio de base $\mathcal{B}$ a $\mathcal{C}$** y la denotaremos por $P_{\mathcal{C}\leftarrow\mathcal{B}}$. La expresión anterior se puede entonces escribir como

$$P_{\mathcal{C}\leftarrow\mathcal{B}}[\vec{u}]_{\mathcal{B}} = [\vec{u}]_{\mathcal{C}}.$$

Además, por el Teorema 2.3 la matriz de cambio de base es invertible y verifica que

$$(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}.$$

Por último, si tenemos una tercera base $\mathcal{D}$, entonces se verifica que

$$P_{\mathcal{D}\leftarrow\mathcal{B}} = P_{\mathcal{D}\leftarrow\mathcal{C}}P_{\mathcal{C}\leftarrow\mathcal{B}}.$$

Ejemplo 22.

1. En $\mathbb{R}^3$, hallemos $P_{\mathcal{B}\leftarrow\mathcal{B}_c}$ y $P_{\mathcal{B}_c\leftarrow\mathcal{B}}$ donde $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$. La matriz de cambio de base *a la base canónica* siempre es inmediata de calcular.

$$P_{\mathcal{B}_c\leftarrow\mathcal{B}} = M(\mathcal{B})_{\mathcal{B}_c} = \left(\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{B}_c} \ \middle| \ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}_{\mathcal{B}_c} \ \middle| \ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{B}_c} \ \right) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Para calcular $P_{\mathcal{B}\leftarrow\mathcal{B}_c}$ tendremos en cuenta que

$$P_{\mathcal{B}\leftarrow\mathcal{B}_c} = (P_{\mathcal{B}_c\leftarrow\mathcal{B}})^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. En $\mathbb{P}_2$, hallemos $P_{\mathcal{C}\leftarrow\mathcal{B}}$ para $\mathcal{B} = \{1,\ 1+x,\ 1+x+x^2\}$ y $\mathcal{C} = \{1+2x,\ -x+3x^2,\ 1-2x^2\}$. Para hallar la matriz tenemos que calcular varias coordenadas, más concretamente

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = M(\mathcal{B})_{\mathcal{C}} = \left(\ [1]_{\mathcal{C}} \ \middle| \ [1+x]_{\mathcal{C}} \ \middle| \ [1+x+x^2]_{\mathcal{C}} \ \right).$$

Ya vimos en el `Ejemplo` 20.3 que podemos hallar todas estas coordenadas de manera simultánea aplicando Gauss-Jordan a la siguiente matriz

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 1 & 1 \\ 2 & -1 & 0 & 0 & 1 & 1 \\ 0 & 3 & -2 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -2 & -1 & -1 \\ 0 & 3 & -2 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -2 & -1 & -1 \\ 0 & 0 & -8 & -6 & -3 & -2 \end{array}\right)$$
$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -2 & -1 & -1 \\ 0 & 0 & 1 & \frac{3}{4} & \frac{3}{8} & \frac{1}{4} \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{4} & \frac{5}{8} & \frac{3}{4} \\ 0 & 1 & 0 & -\frac{1}{2} & -\frac{1}{4} & -\frac{1}{2} \\ 0 & 0 & 1 & \frac{3}{4} & \frac{3}{8} & \frac{1}{4} \end{array}\right),$$

teniendo que

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{pmatrix} \frac{1}{4} & \frac{5}{8} & \frac{3}{4} \\ -\frac{1}{2} & -\frac{1}{4} & -\frac{1}{2} \\ \frac{3}{4} & \frac{3}{8} & \frac{1}{4} \end{pmatrix}.$$

Desde otro punto de vista, podemos descomponer la matriz de cambio de base como $P_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{C}\leftarrow\mathcal{B}_c}P_{\mathcal{B}_c\leftarrow\mathcal{B}}$, y sabiendo que

$$P_{\mathcal{B}_c\leftarrow\mathcal{B}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{y} \quad P_{\mathcal{C}\leftarrow\mathcal{B}_c} = (P_{\mathcal{C}\leftarrow\mathcal{B}_c})^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \\ 0 & 3 & -2 \end{pmatrix}^{-1},$$

bastaría con calcular la inversa de la segunda matriz y multiplicarla por la primera (por la izquierda).

Observar que justo esto es lo que se ha realizado en el proceso de Gauss-Jordan. Visto como producto de matrices, el método de Gauss-Jordan anterior equivale a lo siguiente:

$$\begin{aligned} (P_{\mathcal{B}_c\leftarrow\mathcal{C}})^{-1}\left(P_{\mathcal{B}_c\leftarrow\mathcal{C}} \mid P_{\mathcal{B}_c\leftarrow\mathcal{B}} \right) &= \left((P_{\mathcal{B}_c\leftarrow\mathcal{C}})^{-1}P_{\mathcal{B}_c\leftarrow\mathcal{C}} \mid (P_{\mathcal{B}_c\leftarrow\mathcal{C}})^{-1}P_{\mathcal{B}_c\leftarrow\mathcal{B}} \right) \\ &= \left(I_3 \mid P_{\mathcal{C}\leftarrow\mathcal{B}_c}P_{\mathcal{B}_c\leftarrow\mathcal{B}} \right) = \left(I_3 \mid P_{\mathcal{C}\leftarrow\mathcal{B}} \right) \end{aligned}.$$

Teorema 2.4. *Sea V un $\mathbb{K}$-espacio vectorial y sean $\mathcal{B}$ y $\mathcal{C}$ dos bases. Demostrar que para todo $S \subset V$ las matrices $M(S)_\mathcal{B}$ y $M(S)_\mathcal{C}$ son equivalentes por filas.*

Demostración. Sea $S = \{\vec{u}_1, \dots, \vec{u}_m\} \subset V$ y sea $P_{\mathcal{C}\leftarrow\mathcal{B}}$ la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ que verifica $P_{\mathcal{C}\leftarrow\mathcal{B}}[\vec{v}]_\mathcal{B} = [\vec{v}]_\mathcal{C}$ para todo $\vec{v} \in V$. Entonces

$$\begin{aligned} P_{\mathcal{C}\leftarrow\mathcal{B}}M(S)_\mathcal{B} &= P_{\mathcal{C}\leftarrow\mathcal{B}}\left([\vec{u}_1]_\mathcal{B} \mid \cdots \mid [\vec{u}_m]_\mathcal{B} \right) \\ &= \left(P_{\mathcal{C}\leftarrow\mathcal{B}}[\vec{u}_1]_\mathcal{B} \mid \cdots \mid P_{\mathcal{C}\leftarrow\mathcal{B}}[\vec{u}_m]_\mathcal{B} \right) = \left([\vec{u}_1]_\mathcal{C} \mid \cdots \mid [\vec{u}_m]_\mathcal{C} \right) = M(S)_\mathcal{C}. \end{aligned}$$

□

3. Subespacios vectoriales

$\boxed{\text{Definición}}$. Sea V un $\mathbb{K}$-espacio vectorial. Diremos que un subconjunto $E \subset V$ es un **subespacio vectorial de** V si para todo $\vec{e}, \vec{f} \in E$ y $\lambda, \mu \in \mathbb{K}$ entonces $\lambda\vec{e} + \mu\vec{f} \in E$.

Un subespacio vectorial es de nuevo un espacio vectorial que se ve “sumergido” en otro más “grande”.

Ejemplo 23.

1. En $\mathbb{R}^2$, el subconjunto $E = \left\{ \begin{bmatrix} x \\ x \end{bmatrix} \mid x \in \mathbb{R} \right\}$ sí es un subespacio vectorial. Para demostrarlo es suficiente con entender que los elementos del conjunto son aquellos que su primera y segunda componente coincide. Entonces, para dos elementos cualesquiera y $\lambda, \mu \in \mathbb{R}$,
$$\lambda \begin{bmatrix} x \\ x \end{bmatrix} + \mu \begin{bmatrix} y \\ y \end{bmatrix} = \begin{bmatrix} \lambda x + \mu y \\ \lambda x + \mu y \end{bmatrix} \in E.$$
2. Sin embargo, el subconjunto $S = \left\{ \begin{bmatrix} x \\ x \end{bmatrix}, \begin{bmatrix} x \\ -x \end{bmatrix} \mid x \in \mathbb{R} \right\}$ no es un subespacio vectorial. Por ejemplo,
$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \notin S.$$
3. Sea $A \in M(\mathbb{K})_{m\times n}$ y consideremos el sistema de ecuaciones lineales homogéneo $AX = 0$. Entonces $E = \{X \in \mathbb{K}^n \mid AX = 0\}$ es un subespacio vectorial de $\mathbb{K}^n$. Si tomamos $X, Y \in E$ dos soluciones cualesquiera del sistema y $\lambda, \mu \in \mathbb{K}$, entonces debemos demostrar que la combinación lineal $\lambda X + \mu Y \in E$, es decir, vuelve a ser una solución del sistema de ecuaciones. En efecto:
$$A(\lambda X + \mu Y) = A(\lambda X) + A(\mu Y) = \lambda(AX) + \mu(AY) = \lambda 0 + \mu 0 = 0.$$
4. En $M(\mathbb{K})_n$, el subconjunto $E = \{A \in M(\mathbb{K})_n \mid tr(A) = 0\}$ es un subespacio vectorial.

Nuestro primer objetivo es cómo describir los subespacios vectoriales, para luego saber cómo manipularlos y combinarlos entre ellos.

1. **Sistemas generadores de un subespacio.** Podemos referirnos a cualquier espacio vectorial mediante un sistema generador. En particular, en muchas ocasiones *definiremos* el subespacio directamente dando su sistema generador.

   Sea V un $\mathbb{K}$-espacio vectorial y sea $S = \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_m\} \subset V$. Definimos el **subespacio vectorial** E **generado por** S como el espacio vectorial $E = \text{Gen}(S)$. Es decir, para todo $\vec{e} \in E$ existen $\lambda_1, \lambda_2..., \lambda_m \in \mathbb{K}$ tales que
$$\lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + \cdots + \lambda_m\vec{u}_m = \vec{e}.$$
   Por definición, S es un sistema generador de E, pero no necesariamente linealmente independiente. Pero ya hemos visto cómo extraer un subconjunto que siga manteniendo el carácter generador y que sea linealmente independiente (véase `Ejemplo` 18.4), pudiendo considerar una base de E, que denotaremos por $\mathcal{B}_E$.

   EJEMPLO 24.

   a) En $\mathbb{R}^3$, hallemos un sistema generador del subespacio vectorial
$$E = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \mid x + 4y - 2z = 0 \right\},$$
   que por el `Ejemplo` 23.3 ya sabemos que es un subespacio vectorial. En particular, el conjunto de E está formado por las soluciones de la ecuación lineal homogénea

$x + 4y - 2z = 0$. Resolviendo el sistema obtenemos, para $\lambda, \mu \in \mathbb{R}$,

$$\left\{\begin{array}{ccrr} x & = & -4\lambda & +2\mu \\ y & = & \lambda & \\ z & = & & \mu \end{array}\right., \text{ o vectorialmente, } \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

Pudiendo así reescribir el conjunto E como

$$E = \left\{ \lambda \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^3 \mid \lambda, \mu \in \mathbb{R} \right\}.$$

Donde claramente podemos afirmar que cualquier vector de E es combinación lineal de los vectores $S = \left\{ \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$, es decir, $E = \mathrm{Gen}(S)$.

b) En $\mathbb{P}_3$, vamos a comprobar si los subespacios generados por $S_E = \{1 - x + x^3,\ 1 + x + 2x^2,\ 1 + x^2 - x^3\}$ y $S_F = \{2 + 2x^2 + x^3,\ 2 - x + x^2,\ x + x^2 + x^3\}$ son iguales. Podemos comprobar si los vectores de S_F se pueden generar a partir de S_E y viceversa. Empecemos por generar S_F:

$$\begin{array}{rcl} \lambda_1(1 - x + x^3) + \lambda_2(1 + x + 2x^2) + \lambda_3(1 + x^2 - x^3) & = & 2 + 2x^2 + x^3 \\ \lambda_1(1 - x + x^3) + \lambda_2(1 + x + 2x^2) + \lambda_3(1 + x^2 - x^3) & = & 2 - x + x^2 \\ \lambda_1(1 - x + x^3) + \lambda_2(1 + x + 2x^2) + \lambda_3(1 + x^2 - x^3) & = & x + x^2 + x^3 \end{array}$$

que es equivalente a estudiar, de manera simultánea, los tres sistemas de ecuaciones mediante la matriz

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 2 & 2 & 0 \\ -1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 2 & 1 & 2 & 1 & 1 \\ 1 & 0 & -1 & 1 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 2 & 2 & 0 \\ 0 & 2 & 1 & 2 & 1 & 1 \\ 0 & 2 & 1 & 2 & 1 & 1 \\ 0 & -1 & -2 & -1 & -2 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 2 & 2 & 0 \\ 0 & 2 & 1 & 2 & 1 & 1 \\ 0 & 0 & -3 & 0 & -3 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Obteniendo que los tres sistemas son compatibles y pudiendo afirmar así que $\mathrm{Gen}(S_F) \subset \mathrm{Gen}(S_E)$. Sin embargo, no hemos demostrado aún si son o no iguales, sólo que el subespacio generado por S_F está contenido en el de S_E.
Recíprocamente, estudiemos ahora si los vectores de S_E se pueden generar a partir de S_F:

$$\begin{array}{rcl} \lambda_1(2 + 2x^2 + x^3) + \lambda_2(2 - x + x^2) + \lambda_3(x + x^2 + x^3) & = & 1 - x + x^3 \\ \lambda_1(2 + 2x^2 + x^3) + \lambda_2(2 - x + x^2) + \lambda_3(x + x^2 + x^3) & = & 1 + x + 2x^2 \\ \lambda_1(2 + 2x^2 + x^3) + \lambda_2(2 - x + x^2) + \lambda_3(x + x^2 + x^3) & = & 1 + x^2 - x^3 \end{array},$$

teniendo así que estudiar la matriz

$$\left(\begin{array}{ccc|ccc} 2 & 2 & 0 & 1 & 1 & 1 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 2 & 1 & 1 & 0 & 2 & 1 \\ 1 & 0 & 1 & 1 & 0 & -1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 2 & 2 & 0 & 1 & 1 & 1 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & -2 & 2 & 1 & -1 & -3 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 2 & 2 & 0 & 1 & 1 & 1 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & -3 & -3 \end{array}\right).$$

Obteniendo que ningún sistema de ecuaciones es compatible y, por tanto que $\mathrm{Gen}(S_E) \not\subset \mathrm{Gen}(S_F)$. Concluimos que $\mathrm{Gen}(S_E) \neq \mathrm{Gen}(S_F)$, pero $\mathrm{Gen}(S_F)$ es un subespacio de $\mathrm{Gen}(S_E)$.

Podríamos haberlo estudiado de otra manera. Sabiendo que S_E es un conjunto linealmente independiente, forma una base de $\mathrm{Gen}(S_E)$, que denotaremos por $\mathcal{B}_E$. De esta forma podríamos estudiar si S_F es también base del mismo subespacio mediante el rango de la matriz $M(S_F)_{\mathcal{B}_E}$. Para hallar esta matriz retomaremos los cálculos previos y terminaremos de realizar Gauss-Jordan:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 2 & 2 & 0 \\ 0 & 2 & 1 & 2 & 1 & 1 \\ 0 & 0 & -3 & 0 & -3 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Obteniendo $M(S_F)_{\mathcal{B}_E} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$, que tiene claramente rango 2. Así, por el Teorema 2.3, S_F no es ni linealmente independiente ni sistema generador de $\mathrm{Gen}(S_E)$.

2. **Ecuaciones paramétricas en $\mathcal{B}$.** Consideremos una base $\mathcal{B}$ de V y $\mathcal{B}_E = \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_m\}$ una base del subespacio E. Ya sabemos que todo $\vec{e} \in E$ se puede expresar como

$$\vec{e} = \lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \cdots + \lambda_m \vec{u}_m,$$

que tomando coordenadas en $\mathcal{B}$, obtenemos que

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [\vec{e}]_{\mathcal{B}} = \lambda_1 [\vec{u}_1]_{\mathcal{B}} + \lambda_2 [\vec{u}_2]_{\mathcal{B}} + \cdots + \lambda_r [\vec{u}_m]_{\mathcal{B}} = \left([\vec{u}_1]_{\mathcal{B}} \mid [\vec{u}_2]_{\mathcal{B}} \mid \cdots \mid [\vec{u}_m]_{\mathcal{B}} \right) \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix},$$

es decir, a la expresión

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = M(\mathcal{B}_E)_{\mathcal{B}} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix}$$

es a la que llamaremos **ecuaciones parámetricas de E en $\mathcal{B}$,** ya que según vayamos cambiando los parámetros λ_i obtendremos distintas coordenadas en $\mathcal{B}$ de vectores de E. Se puede observar que el número de parámetros depende del tamaño del sistema generador y, sin embargo, el vector de E se expresa en coordenadas en $\mathcal{B}$, es decir, hay tantas componentes como dimensión tenga V, no E.

Importante, para hablar de ecuaciones paramétricas debemos trabajar con una base de E y otra base de V. No es suficiente un sistema generador.

Por otro lado, las ecuaciones paramétricas, por la estructura que tienen, pueden interpretarse como la solución de un sistema de ecuaciones lineales homogéneo.

Ejemplo 25. En $\mathbb{R}^3$, hallemos las ecuaciones paramétricas del subespacio

$$E = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \mid x + 4y - 2z = 0 \right\}$$

en $\mathcal{B}_c$ y en $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$. Ya vimos en el ejemplo anterior que $\mathcal{B}_E = \left\{ \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$ es una base de E, entonces basta con calcular $M(\mathcal{B}_E)_{\mathcal{B}_c}$, que es inmediata:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{pmatrix} -4 & 2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} \quad \Longleftrightarrow \quad \left\{ \begin{array}{ccrr} x & = & -4\lambda & +2\mu \\ y & = & \lambda & \\ z & = & & \mu \end{array} \right.$$

con $[\vec{e}]_{\mathcal{B}_c} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ son las ecuaciones paramétricas de E en $\mathcal{B}_c$.

Por otro lado, para hallar las ecuaciones paramétricas de E en $\mathcal{B}$ debemos calcular primero $M(\mathcal{B}_E)_{\mathcal{B}}$

$$\left(\begin{array}{ccc|cc} 1 & 1 & 1 & -4 & 2 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{array} \right) \to \left(\begin{array}{ccc|cc} 1 & 1 & 0 & -4 & 1 \\ 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 1 \end{array} \right) \to \left(\begin{array}{ccc|cc} 1 & 0 & 0 & -5 & 2 \\ 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 1 \end{array} \right),$$

teniendo que:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{pmatrix} -5 & 2 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} \quad \Longleftrightarrow \quad \left\{ \begin{array}{ccrr} x' & = & -5\lambda & +2\mu \\ y' & = & \lambda & -\mu \\ z' & = & & \mu \end{array} \right. ,$$

con $[\vec{e}]_{\mathcal{B}} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$ son las ecuaciones paramétricas de E en $\mathcal{B}$.

Observar que las ecuaciones paramétricas en $\mathcal{B}_c$ coinciden con la solución de $x + 4y - 2z = 0$, pero las ecuaciones paramétricas en $\mathcal{B}$ no son solución. Esto se debe a que están expresadas en bases distintas.

3. **Ecuaciones implícitas en $\mathcal{B}$.** Como las ecuaciones paramétricas de un subespacio E en $\mathcal{B}$ pueden interpretarse como la solución de un sistema de ecuaciones lineales homogéneo, podemos hallar uno que tenga como solución las ecuaciones paramétricas de E en $\mathcal{B}$. A este sistema lo llamaremos **ecuaciones implícitas/cartesianas de E en $\mathcal{B}$.**

Propiedades de las ecuaciones implícitas: Partamos de las ecuaciones paramétricas de E en $\mathcal{B}$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = M(\mathcal{B}_E)_{\mathcal{B}} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix}.$$

Cualquier sistema de ecuaciones homogéneo, $AX = 0$, que las tenga como solución debe verificar que la matriz A debe tener $\dim(V) = n$ columnas. Por otro lado, la solución cuenta con m parámetros (o grados de libertad), es decir, siguiendo el

Teorema 1.1, $n - \operatorname{rg}(A) = m$. Como m coincide con la dimensión de E y n con la de V, podemos reescribir la expresión anterior como

$$\dim(V) - \operatorname{rg}(A) = \dim(E).$$

En general, reservamos el nombre de ecuaciones implícitas al mínimo número de ecuaciones necesarias para describir el sistema de ecuaciones. Matricialmente esto equivale a que $\operatorname{rg}(A) =$ número de ecuaciones implícitas.

EJEMPLO 26. *a*) Para hallar las ecuaciones implícitas de un subespacio eliminaremos los parámetros de las ecuaciones paramétricas.
En $\mathbb{R}^3$, supongamos que las ecuaciones paramétricas del subespacio E en $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ son

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{pmatrix} -5 & 2 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} \quad \Longleftrightarrow \quad \left\{ \begin{array}{rcrr} x' & = & -5\lambda & +2\mu \\ y' & = & \lambda & -\mu \\ z' & = & & \mu \end{array} \right. ,$$

con $[\vec{e}]_{\mathcal{B}} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$. Si pensamos las ecuaciones como un sistema de ecuaciones con incógnitas λ, μ, sabemos que resultaría en un sistema compatible. Imponiendo esta condición sobre el sistema:

$$\left(\begin{array}{cc|c} -5 & 2 & x' \\ \boxed{1} & -1 & y' \\ 0 & 1 & z' \end{array} \right) \to \left(\begin{array}{cc|c} \boxed{1} & -1 & y' \\ 0 & -3 & x' + 5y' \\ 0 & \boxed{1} & z' \end{array} \right) \to \left(\begin{array}{cc|c} \boxed{1} & -1 & y' \\ 0 & \boxed{1} & z' \\ 0 & 0 & x' + 5y' + 3z' \end{array} \right)$$

obtenemos la condición $x' + 5y' + 3z' = 0$ para que el sistema sea compatible. Así, diremos que $x' + 5y' + 3z' = 0$ es la ecuación implícita de E en $\mathcal{B}$.
Observar que las ecuaciones paramétricas en $\mathcal{B}$ son solución de la ecuación.
Y, siguiendo el ejemplo anterior, ya sabemos que la ecuación implícita en $\mathcal{B}_c$ es $x + 4y - 2z = 0$, que no coincide con la anterior, pero siguen siendo sendas ecuaciones implícitas del mismo subespacio, sólo que en distintas bases.

b) Supongamos ahora que tenemos el siguiente subespacio de $\mathbb{R}^4$

$$E \equiv \left\{ \begin{array}{rrrrcl} x' & -y' & +z' & +2t' & = & 0 \\ x' & +y' & & +t' & = & 0 \end{array} \right. \qquad \text{en } \mathcal{B},$$

donde $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$. Si queremos hallar una base de F, resolvamos el sistema

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

teniendo como solución

$$\left\{\begin{array}{rcrr} x' & = & -\lambda & -3\mu \\ y' & = & \lambda & +\mu \\ z' & = & 2\lambda & \\ t' & = & & 2\mu \end{array}\right., \text{ o vectorialmente, } \begin{bmatrix} x' \\ y' \\ z' \\ t' \end{bmatrix} = \lambda \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} -3 \\ 1 \\ 0 \\ 2 \end{bmatrix}.$$

Pero observamos que hemos calculado las coordenadas de un vector en $\mathcal{B}$, así que hemos obtenido que $[\mathcal{B}_F]_{\mathcal{B}} = \left\{ \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \\ 0 \\ 2 \end{bmatrix} \right\}$ y, por tanto, $\mathcal{B}_F = \left\{ \begin{bmatrix} 2 \\ 3 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 2 \\ 2 \end{bmatrix} \right\}$.

Teorema 2.5. *Sea V un $\mathbb{K}$-espacio vectorial y $\mathcal{B}$ una base de V. Son equivalentes*

1. *$E \subset V$ es un subespacio vectorial,*
2. *existe $A \in M(\mathbb{K})_{r\times \dim(V)}$ de rango r tal que $[E]_{\mathcal{B}} = \{X \in \mathbb{R}^{\dim(V)} \mid AX = 0\}$.*

Demostración.

$\boxed{\text{a)} \Rightarrow \text{b)}}$ Esta implicación es la definición de las ecuaciones implícitas de E en $\mathcal{B}$.

$\boxed{\text{b)} \Rightarrow \text{a)}}$ Recíprocamente, supongamos que existe una matriz $A \in M(\mathbb{K})_{r\times \dim(V)}$ de rango r tal que las coordenadas del subconjunto E se pueden escribir como $[E]_{\mathcal{B}} = \{X \in \mathbb{R}^{\dim(V)} \mid AX = 0\}$. Como es un sistema homogéneo será compatible, y tendrá $\dim(V) - r = m$ grados de libertad. Si resolvemos el sistema $AX = 0$, tendremos que cualquier solución X se puede expresar como

$$X = \lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_m X_m.$$

En particular, $X_1, X_2, ..., X_m$ son también soluciones de $AX = 0$. Como $X, X_1, ..., X_m$ son a su vez las coordenadas de vectores de E en $\mathcal{B}$, existirán $\vec{v}, \vec{u}_1, ..., \vec{u}_m \in E$ tales que $[\vec{v}]_{\mathcal{B}} = X$, $[\vec{u}_1]_{\mathcal{B}} = X_1$, ..., $[\vec{u}_m]_{\mathcal{B}} = X_m$. Es decir, podemos reescribir la expresión anterior vectorialmente

$$\vec{v} = \lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \cdots + \lambda_m \vec{u}_m.$$

Obteniendo que todo vector $\vec{v} \in E$ se puede expresar como combinación lineal de $\vec{u}_1, ..., \vec{u}_m$ o, en otras palabras, $E = \text{Gen}(\{\vec{u}_1, ..., \vec{u}_m\})$ siendo así que E es un subespacio vectorial generado por $\{\vec{u}_1, ..., \vec{u}_m\}$. □

El teorema entonces nos permite justificar que un subconjunto es un subespacio vectorial si podemos describir las coordenadas de sus vectores como solución de un sistema de ecuaciones lineales homogéneo. Simplificando así el estudio a una mera observación.

Teorema 2.6. *Sea V un $\mathbb{K}$-espacio vectorial y E un subespacio vectorial. Consideremos $\mathcal{B}$ y $\mathcal{C}$ dos bases de V y supongamos que $A_{\mathcal{B}}, A_{\mathcal{C}} \in M_{r\times \dim(V)}(\mathbb{K})$ son las matrices de coeficientes de las ecuaciones implícitas de E en $\mathcal{B}$ y $\mathcal{C}$ respectivamente. Demuestra que $A_{\mathcal{B}}$ y $A_{\mathcal{C}}$ son matrices equivalentes por columnas.*

Demostración. Por definición, las ecuaciones implícitas en $\mathcal{B}$ de un subespacio E son el sistema de ecuaciones homogéneo que tiene a las ecuaciones paramétricas en $\mathcal{B}$ $M(\mathcal{B}_E)_{\mathcal{B}}$ como solución, donde $\mathcal{B}_E$ es una base de E. Esto, matricialmente se puede expresar como $A_{\mathcal{B}} M(\mathcal{B}_E)_{\mathcal{B}} = 0$.

Por otro lado, sabemos que las matrices $M(\mathcal{B}_E)_\mathcal{B}$ y $M(\mathcal{B}_E)_\mathcal{C}$ son equivalentes por filas. En particular,

$$P_{\mathcal{B}\leftarrow\mathcal{C}}M(\mathcal{B}_E)_\mathcal{C} = M(\mathcal{B}_E)_\mathcal{B}.$$

Entonces, sustituyendo

$$0 = A_\mathcal{B}M(\mathcal{B}_E)_\mathcal{B} = (A_\mathcal{B}P_{\mathcal{B}\leftarrow\mathcal{C}})M(\mathcal{B}_E)_\mathcal{C},$$

obteniendo que $A_\mathcal{B}P_{\mathcal{B}\leftarrow\mathcal{C}}$ son ecuaciones implícitas de E en $\mathcal{C}$, es decir, $A_\mathcal{C} = A_\mathcal{B}P_{\mathcal{B}\leftarrow\mathcal{C}}$. Teniendo así que ecuaciones implícitas de E en distintas bases son equivalentes por columnas. □

Ejemplo 27. A toda matriz $A = (A_1|A_2|\cdots|A_m)$ de tamaño $n \times m$ se le pueden asociar dos importantes subespacios vectoriales:

1. El **subespacio columna de** A, aquel generado por las columnas de A, es decir
$$\mathrm{Col}(A) = \mathrm{Gen}(\{A_1, ..., A_m\}) \subset \mathbb{R}^n.$$
Por definición, las columnas ya forman un sistema generador. Para comprobar si las columnas forman una base de este subespacio bastaría con comprobar si son linealmente independientes, y en caso de que no lo fuesen, considerar un subconjunto de columnas que sí lo sean.
2. El **subespacio nulo de** A, aquel que coincide con las soluciones del sistema de ecuaciones homogéneo $AX = 0$, es decir,
$$\mathrm{Nul}(A) = \{X \in \mathbb{R}^m \mid AX = 0\} \subset \mathbb{R}^m.$$
Para hallar una base, bastaría con resolver el sistema de ecuaciones homogéneo.

Estos dos ejemplos son, en otras palabras, un resumen de cómo describir un subespacio vectorial con una matriz. O mediante una matriz que indique sus ecuaciones paramétricas coincidiendo con la matriz del subespacio columna, o que indique sus ecuaciones implícitas coincidiendo con la matriz del subespacio nulo.

Ejercicio. En $M(\mathbb{R})_2$, consideremos el subespacio vectorial $E = \{A \mid tr(A) = 0\}$. Hallar dos matrices N_E y C_E tales que $\mathrm{Col}(C_E) = \mathrm{Nul}(N_E) = [E]_{\mathcal{B}_c}$.

Definición. Sea V un $\mathbb{K}$-espacio vectorial y sean E y F dos subespacios vectoriales. Se definen los subespacios:

1. **Intersección de** E **y** F
$$E \cap F := \{\vec{u} \in V \mid \vec{u} \in E \text{ y } \vec{u} \in F\}.$$
2. **Suma de** E **y** F
$$E + F := \{\vec{e} + \vec{f} \in V \mid \vec{e} \in E,\ \vec{f} \in F\}.$$

En el caso de que $E \cap F = \{\vec{0}\}$, denotaremos al subespacio suma por $E \oplus F$ y diremos que la **suma es directa**. Además, diremos que F **es el complementario de** E y lo denotaremos por $E^c = F$.

La idea de suma directa es una manera de "trocear" un espacio vectorial de manera "limpia", con la intención de que cada trozo tenga un interés independiente. Esto se puede ver claramente por ejemplo si queremos realizar una rotación en el espacio, el eje y el plano de rotación jugarán papeles distintos.

Ejemplo 28. En $\mathbb{R}^4$, consideremos los subespacios E con base $\mathcal{B}_E = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix} \right\}$ y

$$F \equiv \left\{ \begin{array}{rrrrcl} x & -y & & +t & = & 0 \\ x & & +z & -t & = & 0 \end{array} \right. \quad \text{en } \mathcal{B}_c.$$

1. Para hallar una base de la intersección $E \cap F$, primero hallemos sus ecuaciones implícitas, y es suficiente con juntar las ecuaciones implícitas de ambos subespacios en la misma base. Ya que tenemos las ecuaciones de F en $\mathcal{B}_c$, hallemos las de E también en $\mathcal{B}_c$:

$$\left(\begin{array}{rrr|c} \boxed{1} & 0 & 0 & x \\ 0 & 1 & 2 & y \\ 1 & 1 & -1 & z \\ -1 & 1 & 1 & t \end{array} \right) \rightarrow \left(\begin{array}{rrr|c} \boxed{1} & 0 & 0 & x \\ 0 & \boxed{1} & 2 & y \\ 0 & 1 & -1 & z-x \\ 0 & 1 & 1 & x+t \end{array} \right) \rightarrow \left(\begin{array}{rrr|c} \boxed{1} & 0 & 0 & x \\ 0 & \boxed{1} & 2 & y \\ 0 & 0 & \boxed{-3} & z-x-y \\ 0 & 0 & -1 & x+t-y \end{array} \right)$$

$$\rightarrow \left(\begin{array}{rrr|c} \boxed{1} & 0 & 0 & x \\ 0 & \boxed{1} & 2 & y \\ 0 & 0 & \boxed{-3} & z-x-y \\ 0 & 0 & 0 & -4x+2y+z-3t \end{array} \right)$$

teniendo que la ecuación implícita de E en $\mathcal{B}_c$

$$E \equiv \{4x - 2y - z + 3t = 0 \quad \text{en } \mathcal{B}_c.$$

Entonces las ecuaciones implícitas de $E \cap F$ en $\mathcal{B}_c$.

$$E \cap F \equiv \left\{ \begin{array}{rrrrcl} 4x & -2y & -z & +3t & = & 0 \\ x & -y & & +t & = & 0 \\ x & & +z & -t & = & 0 \end{array} \right. \quad \text{en } \mathcal{B}_c,$$

que resolviendo el sistema

$$\begin{pmatrix} 4 & -2 & -1 & 3 \\ \boxed{1} & -1 & 0 & 1 \\ 1 & 0 & 1 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} \boxed{1} & -1 & 0 & 1 \\ 0 & 2 & -1 & -1 \\ 0 & \boxed{1} & 1 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} \boxed{1} & -1 & 0 & 1 \\ 0 & \boxed{1} & 1 & -2 \\ 0 & 0 & \boxed{-3} & 3 \end{pmatrix},$$

teniendo como solución, expresada vectorialmente,

$$\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \lambda \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

y por tanto $[\mathcal{B}_{E\cap F}]_{\mathcal{B}_c} = \mathcal{B}_{E\cap F} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$.

2. Hallemos una base de $E+F$. Sólo necesitamos saber que $E+F$ está generado por la unión de bases de ambos subespacios. Como ya tenemos una base de E, hallemos una base de F. Tan sólo tenemos que resolver las ecuaciones implícitas de F

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

teniendo que $[\mathcal{B}_F]_{\mathcal{B}_c} = \mathcal{B}_F = \left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}$. De este modo, un *sistema generador* de $E+F$ es $S_{E+F} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}$. Basta con obtener un subconjunto linealmente independiente, que siguiendo el razonamiento del `Ejemplo` 18.4,

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

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

Obteniendo que $\mathcal{B}_{E+F} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} \right\}$ es una base de $E+F$. Teniendo entonces que $E+F = \mathbb{R}^4$.

3. Por último, hallemos el subespacio complementario de F, F^c. La idea es muy sencilla: partiendo de una base del subespacio F debemos añadir vectores hasta conseguir una base de $V = \mathbb{R}^4$. Hay muchas formas de añadir estos vectores. Veamos dos:

   a) Añadiendo vectores que no verifiquen sus ecuaciones implícitas. Por ejemplo, $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ no satisfacen claramente las ecuaciones implícitas de F en $\mathcal{B}_c$. Además

$\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$ forma claramente una base de $\mathbb{R}^4$, así que $\mathcal{B}_{F^c} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$ es una base del subespacio F^c, que es complementario a F.

b) Añadiendo vectores tras estudiar la independencia de una base de F. La única particularidad que tiene este método es que trabajaremos sobre la matriz traspuesta $(M(\mathcal{B}_F)_{\mathcal{B}_c})^t$:

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

Y añadiremos vectores de la base canónica tales que no coincidan con los pivotes. En este caso $\mathcal{B}_{F^c} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ es una base de otro subespacio F^c, que es complementario a F.

4.

Teorema 2.7. *Sea V un $\mathbb{K}$-espacio vectorial y $\mathcal{B}$ una base de V. Consideremos E y F dos subespacios vectoriales y N_E, N_F, C_E, C_F matrices tales que $[E]_{\mathcal{B}} = Nul(N_E) = Col(C_E)$ y $[F]_{\mathcal{B}} = Nul(N_F) = Col(C_F)$. Entonces*

1. $[E \cap F]_{\mathcal{B}} = Nul\left(\dfrac{N_E}{N_F} \right)$ *y,*
2. $[E + F]_{\mathcal{B}} = Col\left(\begin{array}{c|c} C_E & C_F \end{array} \right)$.

Demostración. Queda como ejercicio. □

Ejemplo 29.

Teorema 2.8. *Sea V un $\mathbb{K}$-espacio vectorial con E y F dos subespacios vectoriales. Entonces*

$$\dim(E + F) = \dim(E) + \dim(F) - \dim(E \cap F).$$

Demostración. Por comodidad, escribamos $\dim(E) = d_e$, $\dim(F) = d_f$ y $\dim(E \cap F) = d_i$. Consideremos $\mathcal{B}_{E\cap F} = \{\vec{i}_1, \ldots, \vec{i}_{d_i}\}$ una base de $E \cap F$. Ahora, consideremos dos bases de los respectivos subespacios complementarios a E y F, $\{\vec{e}_{d_i+1}, \ldots, \vec{e}_{d_e}\}$ y $\{\vec{f}_{d_i+1}, \ldots, \vec{f}_{d_f}\}$. Entonces debemos demostrar que

$$\mathcal{B} = \{\vec{i}_1, \ldots, \vec{i}_{d_i}, \vec{e}_{d_i+1}, \ldots, \vec{e}_{d_e}, \vec{f}_{d_i+1}, \ldots, \vec{f}_{d_f}\}$$

es una base de $E + F$.

- $\mathcal{B}$ es un sistema generador de $E + F$ porque $\mathcal{B}_E = \{\vec{i}_1, \ldots, \vec{i}_{d_i}, \vec{e}_{d_i+1}, \ldots, \vec{e}_{d_e}\}$ es base de E y $\mathcal{B}_F = \{\vec{i}_1, \ldots, \vec{i}_{d_i}, \vec{f}_{d_i+1}, \ldots, \vec{f}_{d_f}\}$ es base de F, y $\mathcal{B} = \mathcal{B}_E \cup \mathcal{B}_F$.

- Para ver que son linealmente independientes estudiemos la ecuación vectorial

$$(x_1\vec{i}_1 + \cdots + x_{d_i}\vec{i}_{d_i}) + (x_{d_i+1}\vec{e}_{d_i+1} + \cdots + x_{d_e}\vec{e}_{d_e}) + (x_{d_e+1}\vec{f}_{d_i+1} + \cdots + x_{d_f}\vec{f}_{d_f}) = \vec{0}.$$

  Si llamamos $\vec{u} = x_{d_e+1}\vec{f}_{d_i+1} + \cdots + x_{d_f}\vec{f}_{d_f} \in F$, tenemos que

$$\vec{u} = -(x_1\vec{i}_1 + \cdots + x_{d_i}\vec{i}_{d_i}) - (x_{d_i+1}\vec{e}_{d_i+1} + \cdots + x_{d_e}\vec{e}_{d_e}) \in E,$$

  es decir, $\vec{u} \in E \cap F$, así que podremos expresarlo únicamente como combinación lineal de la base $\mathcal{B}_{E\cap F}$:

$$\vec{u} = \lambda_1\vec{i}_1 + \cdots \lambda_{d_i}\vec{i}_{d_i}.$$

  Sustituyendo esta expresión en la ecuación anterior tenemos:

$$\lambda_1\vec{i}_1 + \cdots \lambda_{d_i}\vec{i}_{d_i} = -(x_1\vec{i}_1 + \cdots + x_{d_i}\vec{i}_{d_i}) - (x_{d_i+1}\vec{e}_{d_i+1} + \cdots + x_{d_e}\vec{e}_{d_e}),$$

  o lo que es lo mismo,

$$((x_1 + \lambda_1)\vec{i}_1 + \cdots + (x_{d_i} + \lambda_{d_i})\vec{i}_{d_i}) + (x_{d_i+1}\vec{e}_{d_i+1} + \cdots + x_{d_e}\vec{e}_{d_e}) = \vec{0}$$

  Como es una ecuación expresada en términos de la base $\mathcal{B}_E$, que es linealmente independiente, tenemos que en particular $x_{d_i+1} = \cdots = x_{d_e} = 0$. Pudiendo reescribir nuestra primera ecuación vectorial como:

$$(x_1\vec{i}_1 + \cdots + x_{d_i}\vec{i}_{d_i}) + (x_{d_e+1}\vec{f}_{d_i+1} + \cdots + x_{d_f}\vec{f}_{d_f}) = \vec{0},$$

  que es una ecuación expresada en términos de la base $\mathcal{B}_F$, y de nuevo, como es linealmente independiente, tenemos que $x_1 = \cdots = x_{d_i} = x_{d_e+1} = \cdots = x_{d_f} = 0$. Concluyendo que $\mathcal{B}$ es linealmente independiente, y por tanto base de $E + F$.

Por último, sólo falta contar cuántos vectores tiene la base $\mathcal{B}$, y tiene exáctamente $d_e + d_f - d_i = \dim(E) + \dim(F) - \dim(E \cap F)$. □

TEOREMA 2.9. *Sea V un $\mathbb{K}$-espacio vectorial y $\mathcal{B}$ una base. $\mathcal{B}_E, \mathcal{B}_F \subset V$ son base del mismo subespacio vectorial si, y solo si, $M(\mathcal{B}_E)_{\mathcal{B}}$ y $M(\mathcal{B}_F)_{\mathcal{B}}$ son equivalentes por columnas.*

DEMOSTRACIÓN.

$\boxed{\Rightarrow}$ Supongamos que $\mathcal{B}_E = \{\vec{e}_1, \ldots, \vec{e}_m\}$ y $\mathcal{B}_F = \{\vec{f}_1, \ldots, \vec{f}_m\}$ son base del mismo subespacio vectorial. Entonces, en particular, cada vector de $\mathcal{B}_E$ se puede expresar como combinación lineal de los vectores de $\mathcal{B}_F$, obteniendo así las siguientes m expresiones vectoriales:

$$\begin{array}{rcl} \vec{e}_1 &=& p_{11}\vec{f}_1 + \cdots + p_{m1}\vec{f}_m \\ \vec{e}_2 &=& p_{12}\vec{f}_1 + \cdots + p_{m2}\vec{f}_m \\ \vdots & & \vdots \\ \vec{e}_m &=& p_{1m}\vec{f}_1 + \cdots + p_{mm}\vec{f}_m \end{array} \iff \begin{array}{rcl} [\vec{e}_1]_{\mathcal{B}_F} &=& \begin{bmatrix} p_{11} \\ \vdots \\ p_{m1} \end{bmatrix} \\ [\vec{e}_2]_{\mathcal{B}_F} &=& \begin{bmatrix} p_{12} \\ \vdots \\ p_{m2} \end{bmatrix} \\ \vdots & & \vdots \\ [\vec{e}_m]_{\mathcal{B}_F} &=& \begin{bmatrix} p_{1m} \\ \vdots \\ p_{mm} \end{bmatrix} \end{array}$$

Por otro lado, tomando coordenadas en $\mathcal{B}$ en cada expresión, tenemos que

$$\begin{aligned}
[\vec{e}_1]_{\mathcal{B}} &= p_{11}[\vec{f}_1]_{\mathcal{B}} + \cdots + p_{m1}[\vec{f}_m]_{\mathcal{B}} &= M(\mathcal{B}_F)_{\mathcal{B}} \begin{bmatrix} p_{11} \\ \vdots \\ p_{m1} \end{bmatrix} \\
[\vec{e}_2]_{\mathcal{B}} &= p_{12}[\vec{f}_1]_{\mathcal{B}} + \cdots + p_{m2}[\vec{f}_m]_{\mathcal{B}} &= M(\mathcal{B}_F)_{\mathcal{B}} \begin{bmatrix} p_{12} \\ \vdots \\ p_{m2} \end{bmatrix} \\
\vdots & \qquad\qquad \vdots & \vdots \\
[\vec{e}_m]_{\mathcal{B}} &= p_{1m}[\vec{f}_1]_{\mathcal{B}} + \cdots + p_{mm}[\vec{f}_m]_{\mathcal{B}} &= M(\mathcal{B}_F)_{\mathcal{B}} \begin{bmatrix} p_{1m} \\ \vdots \\ p_{mm} \end{bmatrix}.
\end{aligned}$$

Entonces, columna a columna:

$$M(\mathcal{B}_E)_{\mathcal{B}} = \left(\; [\vec{e}_1]_{\mathcal{B}} \mid \cdots \mid [\vec{e}_m]_{\mathcal{B}} \; \right) = \left(M(\mathcal{B}_F)_{\mathcal{B}} \begin{bmatrix} p_{11} \\ \vdots \\ p_{m1} \end{bmatrix} \middle| \cdots \middle| M(\mathcal{B}_F)_{\mathcal{B}} \begin{bmatrix} p_{1m} \\ \vdots \\ p_{mm} \end{bmatrix} \right) = M(\mathcal{B}_F)_{\mathcal{B}} P,$$

donde

$$P = \left(\begin{bmatrix} p_{11} \\ \vdots \\ p_{m1} \end{bmatrix} \middle| \cdots \middle| \begin{bmatrix} p_{1m} \\ \vdots \\ p_{mm} \end{bmatrix} \right) = \left(\; [\vec{e}_1]_{\mathcal{B}_F} \mid \cdots \mid [\vec{e}_m]_{\mathcal{B}_F} \; \right) = M(\mathcal{B}_E)_{\mathcal{B}_F} = P_{\mathcal{B}_F \leftarrow \mathcal{B}_E}.$$

Siendo entonces P invertible ya que coincide con la matriz de cambio de base de $\mathcal{B}_E$ a $\mathcal{B}_F$.

$\boxed{\Leftarrow}$ Recíprocamente, supongamos que existe una matriz invertible P tal que $M(\mathcal{B}_E)_{\mathcal{B}} = M(\mathcal{B}_F)_{\mathcal{B}} P$. Entonces, observando las columnas de las matrices tenemos que

$$\begin{aligned}
[\vec{e}_1]_{\mathcal{B}} &= M(S_F)_{\mathcal{B}} \begin{bmatrix} p_{11} \\ \vdots \\ p_{m1} \end{bmatrix} &= p_{11}[\vec{f}_1]_{\mathcal{B}} + \cdots + p_{m1}[\vec{f}_m]_{\mathcal{B}} \\
[\vec{e}_2]_{\mathcal{B}} &= M(S_F)_{\mathcal{B}} \begin{bmatrix} p_{12} \\ \vdots \\ p_{m2} \end{bmatrix} &= p_{12}[\vec{f}_1]_{\mathcal{B}} + \cdots + p_{m2}[\vec{f}_m]_{\mathcal{B}} \\
\vdots & \qquad \vdots & \vdots \\
[\vec{e}_m]_{\mathcal{B}} &= M(S_F)_{\mathcal{B}} \begin{bmatrix} p_{1m} \\ \vdots \\ p_{mm} \end{bmatrix} &= p_{1m}[\vec{f}_1]_{\mathcal{B}} + \cdots + p_{mm}[\vec{f}_m]_{\mathcal{B}}.
\end{aligned}$$

Demostrando que $[\mathrm{Gen}(\mathcal{B}_E)]_{\mathcal{B}} \subset [\mathrm{Gen}(\mathcal{B}_F)]_{\mathcal{B}}$. Como P es invertible, si trabajamos con la expresión matricial $M(\mathcal{B}_E)_{\mathcal{B}}P^{-1} = M(\mathcal{B}_F)_{\mathcal{B}}$, obtendremos que $[\mathrm{Gen}(\mathcal{B}_F)]_{\mathcal{B}} \subset [\mathrm{Gen}(\mathcal{B}_E)]_{\mathcal{B}}$, y por tanto $[\mathrm{Gen}(\mathcal{B}_E)]_{\mathcal{B}} = [\mathrm{Gen}(\mathcal{B}_F)]_{\mathcal{B}}$.

$\square$

TEMA 3

Transformaciones lineales

En Matemáticas, es básico relacionar distintos conjuntos de números, de vectores, de soluciones de una ecuación diferencial, etc. Y para ello utilizamos el concepto de *función*. En particular, para dos conjuntos X e Y, queremos asociar un elemento $x \in X$ con un único elemento $y \in Y$. Esta relación la denotaremos como $f\colon X \to Y$ y $f(x) = y$. Al conjunto X se le llama **dominio de** f y a Y **codominio de** f. Para cada $x \in X$ llamaremos **imagen de** x **por** f a $f(x) \in Y$, y denotamos al conjunto de todas las imágenes por f como

$$\mathrm{Im}(f) := \{f(x) \mid x \in X\} \subset Y,$$

que llamaremos **imagen de** f. En general $\mathrm{Im}(f)$ no tiene por qué coincidir con Y.

1. Definiciones básicas

DEFINICIÓN. Sean V y W dos $\mathbb{K}$-espacios vectoriales. Diremos que $T\colon V \to W$ es una transformación lineal si verifica para todo $\vec{u}, \vec{v} \in V$ y $\lambda \in \mathbb{K}$

1. $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$ y,
2. $T(\lambda\vec{u}) = \lambda T(\vec{u})$.

Si $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\} \subset V$ entonces definimos $T(S) := \{T(\vec{u}_1), T(\vec{u}_2), \ldots, T(\vec{u}_n)\} \subset W$.

EJEMPLO 30. Sean $T\colon V \to W$ y $S\colon V \to W$ dos transformaciones lineales. Podemos definir las siguientes nuevas transformaciones:

1. $T + S\colon V \to W$ definida como $(T + S)(\vec{u}) := T(\vec{u}) + S(\vec{u})$. Que resulta ser de nuevo una transformación lineal. Por la definición de $T + S$ y sabiendo que T y S son transformaciones lineales:

   a)
$$\begin{aligned}(T + S)(\vec{u} + \vec{v}) &= T(\vec{u} + \vec{v}) + S(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) + S(\vec{u}) + S(\vec{v}) \\ &= (T + S)(\vec{u}) + (T + S)(\vec{v})\end{aligned}$$

   b)
$$\begin{aligned}(T + S)(\lambda\vec{u}) &= T(\lambda\vec{u}) + S(\lambda\vec{u}) = \lambda T(\vec{u}) + \lambda S(\vec{u}) = \lambda(T(\vec{u}) + S(\vec{u})) \\ &= \lambda(T + S)(\vec{u})\end{aligned}$$

2. $\lambda T : V \to W$ definida como $(\lambda T)(\vec{u}) := \lambda\vec{u}$ también es una transformación lineal.

Con esto hemos demostrado que el conjunto

$$\mathrm{Hom}(V, W) := \{T\colon V \to W \mid T \text{ es lineal}\}$$

es un $\mathbb{K}-$espacio vectorial que se denomina el **espacio vectorial de homomorfismos entre** V **y** W.

$\boxed{\text{DEFINICIÓN}}$. Sea $T\colon V \to W$ una transformación lineal. Diremos que

1. T es inyectiva si $\vec{u}, \vec{v} \in V$ tales que $T(\vec{u}) = T(\vec{v})$ entonces $\vec{u} = \vec{v}$
2. T es sobreyectiva si para todo $\vec{w} \in W$ existe $\vec{u} \in V$ tal que $T(\vec{u}) = \vec{w}$.
3. T es biyectiva si es inyectiva y sobreyectiva.

El siguiente teorema nos indica que para comprobar la inyectividad de una transformación lineal basta con estudiarlo únicamente para el vector nulo.

TEOREMA 3.1. *Sea $T\colon V \to W$ una transformación lineal. Entonces T es inyectiva si, y solo si, $\vec{u} \in V$ tal que $T(\vec{u}) = \vec{0}$ entonces $\vec{u} = \vec{0}$.*

DEMOSTRACIÓN. Supongamos primero que T es inyectiva. Como T es lineal verifica $T(\vec{0}) = T(0 \cdot \vec{0}) = 0 \cdot T(\vec{0}) = \vec{0}$. Ahora, si $\vec{u} \in V$ tal que $T(\vec{u}) = \vec{0} = T(\vec{0})$, por la inyectividad de T, $\vec{u} = \vec{0}$.

Recíprocamente, supongamos que nuestra transformación lineal T verifica que si $\vec{u} \in V$ tal que $T(\vec{u}) = \vec{0}$ entonces $\vec{u} = \vec{0}$. Para demostrar que es inyectiva, consideremos $\vec{u}, \vec{v} \in V$ tales que $T(\vec{u}) = T(\vec{v})$. Como T es lineal

$$T(\vec{u}) = T(\vec{v}) \quad \Longleftrightarrow \quad T(\vec{u}) - T(\vec{v}) = \vec{0} \quad \Longleftrightarrow \quad T(\vec{u} - \vec{v}) = \vec{0},$$

y por tanto $\vec{u} - \vec{v} = \vec{0}$, es decir, $\vec{u} = \vec{v}$, demostrando que T es inyectiva. □

Hemos estudiado en el tema anterior la importancia de las combinaciones lineales de un conjunto finito de vectores, al igual que las propiedades de independencia y sistema generador. Más concretamente, si V es un $\mathbb{K}-$espacio vectorial y $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ una base de V, entonces todo $\vec{u} \in V$ lo podemos expresar como la siguiente combinación lineal

$$\vec{u} = \lambda_1 \vec{b}_1 + \lambda_2 \vec{b}_2 + \cdots + \lambda_n \vec{b}_n \in V.$$

Ahora, si consideramos la transformación lineal $T\colon V \to W$, como tenemos dos espacios vectoriales relacionados mediante T, queremos saber cuándo estas propiedades de independencia y sistema generador se conservan, es decir, aplicando la transformación T a $\vec{u}$, y utilizando la linealidad:

$$T(\vec{u}) = \lambda_1 T(\vec{b}_1) + \lambda_2 T(\vec{b}_2) + \cdots + \lambda_n T(\vec{b}_n) \in W,$$

queremos saber cuándo $T(\mathcal{B}) = \{T(\vec{b}_1), T(\vec{b}_2), \ldots, T(\vec{b}_n)\}$ es también independiente y sistema generador de W. El siguiente teorema resuelve este problema.

TEOREMA 3.2. *Sea $T\colon V \to W$ una transformación lineal.*

1. *T es inyectiva si, y solo si, para todo $S \subset V$ linealmente independiente entonces $T(S)$ es linealmente independiente.*
2. *T es sobreyectiva si, y solo si, para todo $S \subset V$ sistema generador de V entonces $T(S)$ es sistema generador de W.*
3. *T es biyectiva si, y solo si, para toda base $\mathcal{B}$ de V entonces $T(\mathcal{B})$ es base de W.*

DEMOSTRACIÓN.

1. $\boxed{\Rightarrow}$ Supongamos que T es inyectiva. Queremos ver que si $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m\}$ es linealmente independiente, entonces $T(S) = \{T(\vec{u}_1), T(\vec{u}_2), \ldots, T(\vec{u}_m)\}$ es linealmente independiente. Para estudiar la independencia, debemos estudiar las soluciones de la ecuación vectorial

$$x_1T(\vec{u}_1) + x_2T(\vec{u}_2) + \cdots + x_nT(\vec{u}_n) = \vec{0}.$$

Aplicando la linealidad de T, la ecuación vectorial anterior la podemos reescribir como

$$T(x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n) = \vec{0}.$$

Ahora, como T es inyectiva, por el Teorema 3.1, el argumento de T coincide con $\vec{0}$, es decir,

$$x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n = \vec{0},$$

y como S es linealmente independiente, tenemos que $x_1 = x_2 = \cdots = x_n = 0$, demostrando así que $T(S)$ es linealmente independiente.

$\boxed{\Leftarrow}$ Supongamos que T verifica la propiedad: Si $S \subset V$ es linealmente independiente, entonces $T(S)$ es linealmente independiente. Queremos demostrar que T es inyectiva. Por el Teorema 3.1, es suficiente con demostrar que T verifica la propiedad: Si $T(\vec{u}) = \vec{0}$ entonces $\vec{u} = \vec{0}$, que es equivalente a: si $\vec{u} \neq \vec{0}$ entonces $T(\vec{u}) \neq \vec{0}$. Entonces, si tomamos $\vec{u} \neq \vec{0}$, el conjunto $\{\vec{u}\}$ es linealmente independiente, y por la propiedad de T, sabemos que $\{T(\vec{u})\}$ es linealmente independiente, que implica necesariamente $T(\vec{u}) \neq \vec{0}$. Demostrando que T es inyectiva.

$\square$

2. Matriz asociada a una transformación lineal

Sean V y W dos $\mathbb{K}$-espacios vectoriales con bases $\mathcal{B}_V = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ y $\mathcal{B}_W$ respectivamente. Queremos estudiar una transformación lineal $T: V \to W$ y para ello vamos a estudiar el transformado de un vector cualquiera $\vec{u} \in V$. Primero, las coordenadas de $\vec{u}$ en $\mathcal{B}_V$ son $[\vec{u}]_{\mathcal{B}_V} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, es decir, $\vec{u} = x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots x_n\vec{b}_n$. Así, aplicando la transformación a $\vec{u}$

$$T(\vec{u}) = T(x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots x_n\vec{b}_n) = x_1T(\vec{b}_1) + x_2T(\vec{b}_2) + \cdots + x_nT(\vec{b}_n),$$

y tomando coordenadas en $\mathcal{B}_W$:

$$[T(\vec{u})]_{\mathcal{B}_W} = x_1[T(\vec{b}_1)]_{\mathcal{B}_W} + x_2[T(\vec{b}_2)]_{\mathcal{B}_W} + \cdots + x_n[T(\vec{b}_n)]_{\mathcal{B}_W} = M(T(\mathcal{B}_V))_{\mathcal{B}_W} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

que denotanto a la matriz $M(T(\mathcal{B}_V))_{\mathcal{B}_W} = M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V}$, verifica que

$$[T(\vec{u})]_{\mathcal{B}_W} = M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V}[\vec{u}]_{\mathcal{B}_V}.$$

Es decir, la matriz $M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V}$ está actuando como la transformación lineal T.

Utilizando la matriz asociada a una transformación lineal podemos reescribir el Teorema 3.2 teniendo en cuenta el Teorema 2.3:

TEOREMA 3.3. *Sea $T\colon V \to W$ una transformación lineal con $\mathcal{B}_V$ y $\mathcal{B}_W$ bases de V y W respectivamente.*

1. *T es inyectiva si, y solo si, $\operatorname{rg}(M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V}) = \dim V =$ nº columnas,*
2. *T es sobreyectiva si, y solo si, $\operatorname{rg}(M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V}) = \dim W =$ nº filas,*
3. *T es biyectiva si, y solo si, $\operatorname{rg}(M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V}) = \dim V = \dim W =$ nº columnas = nº filas.*

TEOREMA 3.4. *Sea $T\colon V \to W$ una transformación lineal y sean $\mathcal{B}_V, \mathcal{C}_V$ bases de V y $\mathcal{B}_W, \mathcal{C}_W$ bases de W. Entonces*

$$M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V} = P_{\mathcal{B}_W \leftarrow \mathcal{C}_W} M(T)_{\mathcal{C}_W \leftarrow \mathcal{C}_V} P_{\mathcal{C}_V \leftarrow \mathcal{B}_V}.$$

En particular, si $V = W$ con $\mathcal{B} = \mathcal{B}_V = \mathcal{B}_W$ y $\mathcal{C} = \mathcal{C}_V = \mathcal{C}_W$, entonces

$$\begin{aligned} M(T)_{\mathcal{B} \leftarrow \mathcal{B}} &= P_{\mathcal{B} \leftarrow \mathcal{C}} M(T)_{\mathcal{C} \leftarrow \mathcal{C}} P_{\mathcal{C} \leftarrow \mathcal{B}} \\ &= P_{\mathcal{B} \leftarrow \mathcal{C}} M(T)_{\mathcal{C} \leftarrow \mathcal{C}} \left(P_{\mathcal{B} \leftarrow \mathcal{C}}\right)^{-1} \\ &= \left(P_{\mathcal{C} \leftarrow \mathcal{B}}\right)^{-1} M(T)_{\mathcal{C} \leftarrow \mathcal{C}} P_{\mathcal{C} \leftarrow \mathcal{B}}. \end{aligned}$$

DEMOSTRACIÓN. □

DEFINICIÓN. Sean A y B dos matrices cuadradas del mismo tamaño. Decimos que A y B son **matrices semejantes** si existe P invertible tal que $A = PBP^{-1}$.

Según el Teorema 3.4, distintas matrices asociadas a una misma transformación lineal $T\colon V \to V$ en mismas bases son matrices semejantes.

TEOREMA 3.5. *Sean $T\colon V \to W$ y $S\colon V \to W$ dos transformaciones lineales y $\mathcal{B}_V$ y $\mathcal{B}_W$ bases de V y W respectivamente. Entonces*

$$M(T+S)_{\mathcal{B}_W \leftarrow \mathcal{B}_V} = M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V} + M(S)_{\mathcal{B}_W \leftarrow \mathcal{B}_V}, \quad M(\lambda \cdot T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V} = \lambda \cdot M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V}.$$

TEOREMA 3.6. *Sea $Hom(V, W)$ el espacio vectorial de homomorfismos entre V y W con $\mathcal{B}_V$ y $\mathcal{B}_W$ bases de V y W respectivamente. Entonces*

$$M(\,\cdot\,)_{\mathcal{B}_W \leftarrow \mathcal{B}_V} \colon Hom(V, W) \to M_{\dim(W) \times \dim(V)}(\mathbb{K})$$

es una transformación lineal biyectiva.

Adicionalmente, el siguiente Teorema justifica la definición del producto matricial. El producto entre dos matrices se define para que se verifique que la matriz asociada a la composición coincida con el producto de las matrices asociadas. Primero recordamos que dadas dos transformaciones lineales $T\colon U \to V$ y $S\colon V \to W$, se define la composición de T con S como la nueva transformación $S \circ T\colon U \to W$ tal que $(S \circ T)(\vec{u}) := S(T(\vec{u}))$, que resulta ser también lineal.

TEOREMA 3.7. *Sean $T\colon U \to V$ y $S\colon V \to W$ dos transformaciones lineales y $\mathcal{B}_U, \mathcal{B}_V$ y $\mathcal{B}_W$ bases de U, V y W respectivamente. Entonces*

$$M(S \circ T)_{\mathcal{B}_W \leftarrow \mathcal{B}_U} = M(S)_{\mathcal{B}_W \leftarrow \mathcal{B}_V} \cdot M(T)_{\mathcal{B}_V \leftarrow \mathcal{B}_U}.$$

DEMOSTRACIÓN. Sea $\mathcal{B}_U = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ la base de U. Entonces, para todo $i \in \{1, \dots, n\}$,

$$\begin{aligned} M(S \circ T)_{\mathcal{B}_W \leftarrow \mathcal{B}_U}[\vec{b}_i]_{\mathcal{B}_U} &= [(S \circ T)(\vec{b}_i)]_{\mathcal{B}_W} = [S(T(\vec{b}_i))]_{\mathcal{B}_W} \\ &= M(S)_{\mathcal{B}_W \leftarrow \mathcal{B}_V}[T(\vec{b}_i)]_{\mathcal{B}_V} \\ &= M(S)_{\mathcal{B}_W \leftarrow \mathcal{B}_V} M(T)_{\mathcal{B}_V \leftarrow \mathcal{B}_U}[\vec{b}_i]_{\mathcal{B}_U}. \end{aligned}$$

Concluyendo que todas las columnas de $M(S \circ T)_{\mathcal{B}_W \leftarrow \mathcal{B}_U}$ y $M(S)_{\mathcal{B}_W \leftarrow \mathcal{B}_V} \cdot M(T)_{\mathcal{B}_V \leftarrow \mathcal{B}_U}$ son iguales y, consecuentemente, las matrices también lo son. □

3. Subespacios asociados a una transformación lineal

DEFINICIÓN. Sea $T \colon V \to W$ una transformación lineal. Se pueden definir los siguientes conjuntos:

1. El **núcleo de** T, $\mathrm{Nuc}(T) := \{\vec{u} \in V \mid T(\vec{u}) = \vec{0}\} \subset V$.
2. La **imagen de** T, $\mathrm{Im}(T) = \{\vec{w} \in W \mid \text{ existe } \vec{u} \in V \text{ tal que } T(\vec{u}) = \vec{w}\} \subset W$.

EJERCICIO. Sea $T \colon V \to W$ una transformación lineal. Demostrar que los conjuntos $\mathrm{Nuc}(T)$ e $\mathrm{Im}(T)$ son subespacios vectoriales de V y W respectivamente.

TEOREMA 3.8. *Sea $T \colon V \to W$ una transformación lineal.*

1. T es inyectiva si, y solo si, $\dim(\mathrm{Nuc}(T)) = 0$.
2. T es sobreyectiva si, y solo si, $\dim(\mathrm{Im}(T)) = \dim(W)$.

TEOREMA 3.9. *Sea $T \colon V \to W$ una transformación lineal con $\mathcal{B}_V$ y $\mathcal{B}_W$ bases de V y W respectivamente. Entonces*

$$[\mathrm{Nuc}(T)]_{\mathcal{B}_V} = \mathrm{Nul}\left(M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V}\right), \quad [\mathrm{Im}(T)]_{\mathcal{B}_W} = \mathrm{Col}\left(M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V}\right).$$

En particular, $\dim(\mathrm{Im}(T)) = \mathrm{rg}(M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V})$ *y* $\dim(\mathrm{Nuc}(T)) = \dim(V) - \mathrm{rg}(M(T)_{\mathcal{B}_W \leftarrow \mathcal{B}_V})$.

TEOREMA 3.10. *Sea $T \colon V \to W$ una transformación lineal. Entonces*

$$\dim(V) = \dim(\mathrm{Nuc}(T)) + \dim(\mathrm{Im}(T)).$$

4. Diagonalización por semejanza

En este tema vamos a trabajaremos con transformaciones lineales con un único espacio vectorial, $T \colon V \to V$. Estas transformaciones lineales también reciben el nombre de **endomorfismos**.

Ya hemos estudiado que, dada una base $\mathcal{D} = \{\vec{d}_1, \vec{d}_2, \dots, \vec{d}_n\}$ de V, podemos considerar la matriz asociada en $\mathcal{D}$ de T, y que hemos denotado por $M(T)_{\mathcal{D} \leftarrow \mathcal{D}}$. El pregunta que queremos resolver en este tema es:

¿Qué debe verificar la base $\mathcal{D}$ para que $M(T)_{\mathcal{D} \leftarrow \mathcal{D}}$ sea diagonal?

Sabiendo que $M(T)_{\mathcal{D} \leftarrow \mathcal{D}} = M(T(\mathcal{D}))_{\mathcal{D}} = \left([T(\vec{d}_1)]_{\mathcal{D}} \mid [T(\vec{d}_2)]_{\mathcal{D}} \mid \cdots \mid [T(\vec{d}_n)]_{\mathcal{D}} \right)$, si la matriz fuese diagonal significaría que las columnas serían de la forma

$$[T(\vec{d}_1)]_{\mathcal{D}} = \begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad [T(\vec{d}_2)]_{\mathcal{D}} = \begin{bmatrix} 0 \\ \lambda_2 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad [T(\vec{d}_n)]_{\mathcal{D}} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \lambda_n \end{bmatrix},$$

que vectorialmente es equivalente a

$$T(\vec{d_1}) = \lambda_1\vec{d_1}, \quad T(\vec{d_2}) = \lambda_2\vec{d_2}, \quad \ldots, \quad T(\vec{d_n}) = \lambda_n\vec{d_n}.$$

Entonces debemos hallar una base $\mathcal{D} = \{\vec{d_1}, \vec{d_2}, \ldots, \vec{d_n}\}$ cuyos vectores verifiquen la propiedad anterior $T(\vec{d_i}) = \lambda_i\vec{d_i}$ para todo $i \in \{1, \ldots, n\}$.

Ejemplo 31. Sea $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ definida como $T\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}y\\x\end{bmatrix}$. Hallar vectores $\vec{u} \in \mathbb{R}^2$ que verifiquen la propiedad $T(\vec{u}) = \lambda\vec{u}$ no es una tarea que deba subestimarse. Por ejemplo, ya que no tenemos ninguna condición sobre λ, ¿existe $\vec{u} = \begin{bmatrix}x\\y\end{bmatrix}$ tal que $T\begin{bmatrix}x\\y\end{bmatrix} = 2\begin{bmatrix}x\\y\end{bmatrix}$?, que, aplicando la definición de T, es equivalente a, ¿existe $\vec{u} = \begin{bmatrix}x\\y\end{bmatrix}$ tal que $\begin{bmatrix}y\\x\end{bmatrix} = 2\begin{bmatrix}x\\y\end{bmatrix}$? Proporcionando el siguiente sistema de ecuaciones

$$\left\{\begin{array}{rrcl} -2x & +y & = & 0 \\ x & -2y & = & 0 \end{array}\right.$$

que tiene como única solución $x = y = 0$, es decir, el único vector $\vec{u}$ que verifica $T(\vec{u}) = 2\vec{u}$ es $\vec{u} = \vec{0}$ y, como queremos que $\vec{u}$ forme parte de una base, queremos soluciones no nulas. Así que para $\lambda = 2$ no es posible. ¿Existirá algún valor de λ para el que podamos encontrar un vector no nulo?, es decir, ¿existe λ para el cual existe $\vec{u} = \begin{bmatrix}x\\y\end{bmatrix} \neq \vec{0}$ tal que $T\begin{bmatrix}x\\y\end{bmatrix} = \lambda\begin{bmatrix}x\\y\end{bmatrix}$? Al igual que antes, esto es equivalente a, ¿existe λ para el cual existe $\vec{u} = \begin{bmatrix}x\\y\end{bmatrix} \neq \vec{0}$ tal que $\begin{bmatrix}y\\x\end{bmatrix} = \lambda\begin{bmatrix}x\\y\end{bmatrix}$? Del cual obtenemos el siguiente sistema de ecuaciones lineales homogéneo:

$$\left\{\begin{array}{rrcl} -\lambda x & +y & = & 0 \\ x & -\lambda y & = & 0 \end{array}\right.,$$

que no tendrá solución única si el rango de la matriz de coeficientes es menor que dos. En particular, si su determinante es igual a 0:

$$\det\begin{pmatrix}-\lambda & 1\\ 1 & -\lambda\end{pmatrix} = \lambda^2 - 1 = 0,$$

teniendo que el sistema será indeterminado solo si $\lambda = 1$ o $\lambda = -1$. Entonces, para $\lambda = 1$ tenemos una posible solución $\begin{bmatrix}1\\1\end{bmatrix}$, y para $\lambda = -1$ otra solución $\begin{bmatrix}-1\\1\end{bmatrix}$. Es decir,

$$T\begin{bmatrix}1\\1\end{bmatrix} = 1\cdot\begin{bmatrix}1\\1\end{bmatrix}, \quad T\begin{bmatrix}-1\\1\end{bmatrix} = (-1)\cdot\begin{bmatrix}-1\\1\end{bmatrix}.$$

Y, por tanto, la base $\mathcal{D} = \left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}-1\\1\end{bmatrix}\right\}$ verifica que

$$M(T)_{\mathcal{D}\leftarrow\mathcal{D}} = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

Así, diseñaremos una estrategia general para hallar primero esos escalares $\lambda_i \in \mathbb{K}$ y, después, hallaremos los vectores de la base tan sólo resolviendo un sistema de ecuaciones lineales. Para ello, denotemos por $A = M(T)_{\mathcal{B}\leftarrow\mathcal{B}}$ la matriz asociada a T en una base cualquiera $\mathcal{B}$ de V y por $X = [\vec{u}]_{\mathcal{B}}$.

Primero, hallemos los escalares $\lambda \in \mathbb{K}$ aptos para resolver el problema $T(\vec{u}) = \lambda\vec{u}$, con $\vec{u} \neq 0$:

$$T(\vec{u}) = \lambda\vec{u} \iff T(\vec{u}) - \lambda\vec{u} = \vec{0} \iff T(\vec{u}) - (\lambda \mathrm{Id}_V)(\vec{u}) \iff (T - \lambda\mathrm{Id}_V)(\vec{u}) = \vec{0},$$

es decir, $\vec{u} \in \mathrm{Nuc}(T - \lambda\mathrm{Id}) \neq \{\vec{0}\}$. Que matricialmente podemos reescribir como:

$$AX = \lambda X \iff AX - \lambda X = 0 \iff AX - \lambda IX = 0 \iff (A - \lambda I)X = 0,$$

en este caso, $X \in \mathrm{Nul}(A - \lambda I) \neq \{0\}$. Que tendrá solución no nula si la matriz de coeficientes $A - \lambda I$ no tiene rango máximo, es decir, para aquellos valores $\lambda \in \mathbb{K}$ tales que $\det(A - \lambda I) = 0$.

Después, una vez hallado los valores de λ aptos, basta con hallar una solución no nula de los respectivos sistemas de ecuaciones homogéneos con matriz de coeficientes $A - \lambda I$.

EJERCICIO. Repetir la argumentación si la matriz asociada a T está expresada en bases distintas, es decir, $A = M(T)_{\mathcal{C}\leftarrow\mathcal{B}}$.

DEFINICIÓN. Sea $T\colon V \to V$ una transformación lineal y sea $\mathcal{B}$ una base de V.

1. $\lambda \in \mathbb{K}$ es un **autovalor de** T si existe $\vec{0} \neq \vec{u} \in V$ tal que $T(\vec{u}) = \lambda\vec{u}$, y se dirá que λ **es un autovalor asociado a** $\vec{u}$.
2. $\vec{0} \neq \vec{u} \in V$ es un **autovector de** T si existe $\lambda \in \mathbb{K}$ tal que $T(\vec{u}) = \lambda\vec{u}$, y se dirá que $\vec{u}$ **es un autovector asociado a** λ.
3. Para cada autovalor λ de T definimos el **subespacio propio asociado a** λ como $V_\lambda := \mathrm{Nuc}(T - \lambda\mathrm{Id}_V)$.
4. Se define el **polinomio característico** de T como $p(\lambda) := \det(M(T)_{\mathcal{B}\leftarrow\mathcal{B}} - \lambda I)$.

PROPIEDADES.

1. *λ es autovalor si, y solo si, es raíz del polinomio característico, es decir, $p(\lambda) = 0$.*
2. *$V_\lambda \neq \{\vec{0}\}$ si, y solo si, λ es autovalor.*
3. *$\vec{u}$ es autovector asociado a λ si, y solo si, $\vec{0} \neq \vec{u} \in V_\lambda$.*

TEOREMA 3.11. *Sea $T\colon V \to V$ una transformación lineal y sean λ y μ dos autovalores distintos. Entonces $V_\lambda \cap V_\mu = \{\vec{0}\}$.*

DEMOSTRACIÓN. Si $\vec{u} \in V_\lambda \cap V_\mu$, entonces por un lado $\vec{u} \in V_\lambda$ teniendo que $T(\vec{u}) = \lambda\vec{u}$, y por el otro lado, $\vec{u} \in V_\mu$ teniendo además que $T(\vec{u}) = \mu\vec{u}$. Así, juntando ambas igualdades tenemos que

$$\lambda\vec{u} = \mu\vec{u} \iff (\lambda - \mu)\vec{u} = \vec{0} \iff \vec{u} = \vec{0}$$

ya que $\lambda - \mu \neq 0$, pues $\lambda \neq \mu$. Demostrando que si un vector está en $V_\lambda \cap V_\mu$ debe ser el vector nulo. □

Este teorema nos permite asegurar entonces que a distintos autovalores, los autovectores asociados serán linealmente independientes.

Obteniendo así la estrategia para tratar de hallar una base $\mathcal{D}$ tal que $M(T)_{\mathcal{D}\leftarrow\mathcal{D}}$ sea diagonal.

EJEMPLO 32. Sea $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ la transformación definida como $T\begin{bmatrix} x\\ y\\ z\end{bmatrix} = \begin{bmatrix} 3x & +3y & -9z\\ -x & +7y & +z\\ -5x & +5y & -z\end{bmatrix}$. Hallemos, si es posible, una base de autovectores de T.

Primero, debemos hallar el polinomio característico. Para ello necesitaremos trabajar sobre una matriz asociada, en particular, en las bases canónicas

$$A = M(T)_{\mathcal{B}_c\leftarrow\mathcal{B}_c} = \begin{pmatrix} 3 & 3 & -9\\ -1 & 7 & 1\\ -5 & 5 & -1\end{pmatrix}.$$

Entonces, su polinomio característico es:

$$\begin{aligned}
p(\lambda) = \det(A-\lambda I_3) \quad &= \det\left(\begin{pmatrix} 3 & 3 & -9\\ -1 & 7 & 1\\ -5 & 5 & -1\end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\right) = \det\begin{pmatrix} 3-\lambda & 3 & -9\\ -1 & 7-\lambda & 1\\ -5 & 5 & -1-\lambda\end{pmatrix}\\
&\xlongequal{C_3^N = C_3 + C_1} \det\begin{pmatrix} 3-\lambda & 3 & -6-\lambda\\ -1 & 7-\lambda & 0\\ -5 & 5 & -6-\lambda\end{pmatrix}\\
&\xlongequal{F_1^N = F_1 - F_3} \det\begin{pmatrix} 8-\lambda & -2 & 0\\ -1 & 7-\lambda & 0\\ -5 & 5 & -6-\lambda\end{pmatrix} = (-6-\lambda)\det\begin{pmatrix} 8-\lambda & -2\\ -1 & 7-\lambda\end{pmatrix}\\
&= (-6-\lambda)((8-\lambda)(7-\lambda)-2) = (-6-\lambda)(\lambda^2-15\lambda+54)\\
&= (-6-\lambda)(6-\lambda)(9-\lambda)
\end{aligned}$$

Como los autovalores de T son las raíces del polinomio característico, tenemos que $\lambda_1 = -6$, $\lambda_2 = 6$ y $\lambda_3 = 9$ son los tres autovalores. Ahora, para cada uno de ellos, calculemos su subespacio propio para así poder determinar tres autovectores asociados a cada uno de ellos:

$\boxed{V_{-6}}$ Como $V_{-6} = \mathrm{Nuc}(T - (-6)Id_{\mathbb{R}^3})$, sabemos por el Teorema 3.9, que $[V_{-6}]_{\mathcal{B}_c} = \mathrm{Nul}(A + 6I_3)$. Entonces basta con resolver el sistema de ecuaciones con matriz asociada $A+6I_3$:

$$A + 6I_3 = \begin{pmatrix} 3 & 3 & -9\\ -1 & 7 & 1\\ -5 & 5 & -1\end{pmatrix} + \begin{pmatrix} 6 & 0 & 0\\ 0 & 6 & 0\\ 0 & 0 & 6\end{pmatrix} = \begin{pmatrix} 9 & 3 & -9\\ -1 & 13 & 1\\ -5 & 5 & 5\end{pmatrix}$$

$$\begin{pmatrix} 9 & 3 & -9\\ -1 & 13 & 1\\ \boxed{-5} & 5 & 5\end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -1 & -1\\ 0 & \boxed{12} & 0\\ 0 & 12 & 0\end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -1 & -1\\ 0 & \boxed{1} & 0\\ 0 & 0 & 0\end{pmatrix},$$

teniendo que $\mathcal{B}_{[V_{-6}]_{\mathcal{B}_c}} = \left\{\begin{bmatrix} 1\\ 0\\ 1\end{bmatrix}\right\}$.

$\boxed{V_6}$ Repitiendo la argumentación anterior:

$$A - 6I_3 = \begin{pmatrix} 3 & 3 & -9 \\ -1 & 7 & 1 \\ -5 & 5 & -1 \end{pmatrix} - \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} = \begin{pmatrix} -3 & 3 & -9 \\ -1 & 1 & 1 \\ -5 & 5 & -7 \end{pmatrix}$$

$$\begin{pmatrix} -3 & 3 & -9 \\ \boxed{\text{-1}} & 1 & 1 \\ -5 & 5 & -7 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{\text{-1}} & 1 & 1 \\ 0 & 0 & \boxed{\text{-12}} \\ 0 & 0 & -12 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{\text{-1}} & 1 & 1 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \end{pmatrix},$$

teniendo que $\mathcal{B}_{[V_6]_{\mathcal{B}_c}} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$.

$\boxed{V_9}$

$$A - 9I_3 = \begin{pmatrix} 3 & 3 & -9 \\ -1 & 7 & 1 \\ -5 & 5 & -1 \end{pmatrix} - \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} = \begin{pmatrix} -6 & 3 & -9 \\ -1 & -2 & 1 \\ -5 & 5 & -10 \end{pmatrix}$$

$$\begin{pmatrix} -6 & 3 & -9 \\ -1 & -2 & 1 \\ \boxed{\text{-5}} & 5 & -10 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -1 & 2 \\ 0 & \boxed{\text{-3}} & 3 \\ 0 & -3 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -1 & 2 \\ 0 & \boxed{1} & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

teniendo que $\mathcal{B}_{[V_9]_{\mathcal{B}_c}} = \left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$.

Aunque se pueda comprobar, al ser todos los autovalores distintos, por el Teorema 3.11, estos subespacios son linealmente independientes, y por tanto $\mathcal{D} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$ forman una base de autovectores de $\mathbb{R}^3$, es decir,

$$T(\mathcal{D}) = \left\{ -6 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, 6 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, 9 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\},$$

y por tanto

$$D = M(T)_{\mathcal{D} \leftarrow \mathcal{D}} = \begin{pmatrix} -6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

Adicionalmente, si quisiéramos relacionar la matriz $A = M(T)_{\mathcal{B}_c \leftarrow \mathcal{B}_c}$ con $D = M(T)_{\mathcal{D} \leftarrow \mathcal{D}}$, basta con utilizar las matrices de cambio de bases:

$$A = P_{\mathcal{B}_c \leftarrow \mathcal{D}} D (P_{\mathcal{B}_c \leftarrow \mathcal{D}})^{-1},$$

donde

$$P = P_{\mathcal{B}_c \leftarrow \mathcal{D}} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

DEFINICIÓN. Sea $T\colon V \to V$ una transformación lineal. Diremos que **una transformación T es diagonalizable por semejanza** si existe una base de autovectores de T en V.

A pesar de que hemos introducido el problema de diagonalización por semejanza a través de las transformaciones lineales, lo más habitual es trabajar directamente matricialmente. Por ello, todas las definiciones anteriores se pueden reescribir matricialmente y, por no repetir en exceso, sólo escribiremos la siguiente definición:

DEFINICIÓN. Sea $A \in M_n(\mathbb{K})$. Diremos que **una matriz A es diagonalizable por semejanza** si es semejante a una matriz diagonal D, es decir, si existe P invertible tal que $A = PDP^{-1}$.

Además, es importante observar que la matriz diagonal D está formada por los autovalores de A y la matriz P por sus autovectores asociados, en el mismo orden.

EJEMPLO 33. Estudiemos ahora si es posible diagonalizar por semejanza la matriz $A = \begin{pmatrix} 12 & 4 & -2 \\ 1 & 12 & -1 \\ -1 & -2 & 11 \end{pmatrix}$, es decir, hallemos D diagonal y P invertible tales que $A = PDP^{-1}$.

Primero, hallemos su polinomio característico:

$$
\begin{aligned}
p(\lambda) = \det(A - \lambda I_3) &= \det \begin{pmatrix} 12-\lambda & 4 & -2 \\ 1 & 12-\lambda & -1 \\ -1 & -2 & 11-\lambda \end{pmatrix} \\
&\overset{C_3^N = C_3 + C_1}{=\!=\!=} \det \begin{pmatrix} 12-\lambda & 4 & 10-\lambda \\ 1 & 12-\lambda & 0 \\ -1 & -2 & 10-\lambda \end{pmatrix} \\
&\overset{F_1^N = F_1 - F_3}{=\!=\!=} \det \begin{pmatrix} 13-\lambda & 6 & 0 \\ 1 & 12-\lambda & 0 \\ -1 & -2 & 10-\lambda \end{pmatrix} = (10-\lambda) \det \begin{pmatrix} 13-\lambda & 6 \\ 1 & 12-\lambda \end{pmatrix} \\
&= (10-\lambda)((13-\lambda)(12-\lambda) - 6) = (10-\lambda)(\lambda^2 - 25\lambda + 150) \\
&= (10-\lambda)^2(15-\lambda).
\end{aligned}
$$

Teniendo como autovalores $\lambda_1 = 10$ (doble) y $\lambda_2 = 15$ (simple). Ahora sólo contamos con dos subespacios propios, pero necesitamos tres autovectores linealmente independientes. Parece razonable que del autovalor $\lambda_1 = 10$ doble sea posible obtener dos autovectores linealmente independientes.

$\boxed{V_{10}}$ Siguiendo el razonamiento del ejemplo anterior

$$
A - 10I_3 = \begin{pmatrix} 12 & 4 & -2 \\ 1 & 12 & -1 \\ -1 & -2 & 11 \end{pmatrix} - \begin{pmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \end{pmatrix} = \begin{pmatrix} 2 & 4 & -2 \\ 1 & 2 & -1 \\ -1 & -2 & 1 \end{pmatrix}
$$

$$\begin{pmatrix} 2 & 4 & -2 \\ \boxed{1} & 2 & -1 \\ -1 & -2 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} \boxed{1} & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

teniendo que $\mathcal{B}_{V_{10}} = \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$, consiguiendo así dos autovectores linealmente independientes.

$\boxed{V_{15}}$

$$A - 15I_3 = \begin{pmatrix} 12 & 4 & -2 \\ 1 & 12 & -1 \\ -1 & -2 & 11 \end{pmatrix} - \begin{pmatrix} 15 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 15 \end{pmatrix} = \begin{pmatrix} -3 & 4 & -2 \\ 1 & -3 & -1 \\ -1 & -2 & -4 \end{pmatrix}$$

$$\begin{pmatrix} -3 & 4 & -2 \\ \boxed{1} & -3 & -1 \\ -1 & -2 & -4 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -3 & -1 \\ 0 & \boxed{\text{-5}} & -5 \\ 0 & -5 & -5 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -3 & -1 \\ 0 & \boxed{1} & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

teniendo que $\mathcal{B}_{V_{15}} = \left\{ \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \right\}$.

Construyendo así la base de autovectores $\mathcal{D} = \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \right\}$, que nos proporciona las matrices

$$D = \begin{pmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 15 \end{pmatrix} \quad \text{y} \quad P = \begin{pmatrix} -2 & 1 & -2 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix},$$

que verifican $A = PDP^{-1}$.

En este último ejemplo, nos hemos encontrado con la situación de que uno de los autovalores haya aparecido de manera repetida, y como esto podría haber ocasionado problemas para hallar la cantidad necesaria de autovectores. Como sí ocurrirá en el siguiente ejemplo:

Ejemplo 34. Estudiemos si $A = \begin{pmatrix} 8 & -2 & 1 \\ 0 & 6 & 3 \\ -1 & 1 & 7 \end{pmatrix}$ es diagonalizable por semejanza. Como va ser habitual, hallemos el polinomio característico:

$$\begin{aligned} p(\lambda) = \det(A - \lambda I_3) &= \det \begin{pmatrix} 8-\lambda & -2 & 1 \\ 0 & 6-\lambda & 3 \\ -1 & 1 & 7-\lambda \end{pmatrix} \\ &\overset{C_2^N = C_2 + C_1}{=\!=\!=} \det \begin{pmatrix} 8-\lambda & 6-\lambda & 1 \\ 0 & 6-\lambda & 3 \\ -1 & 0 & 7-\lambda \end{pmatrix} \\ &\overset{F_1^N = F_1 - F_2}{=\!=\!=} \det \begin{pmatrix} 8-\lambda & 0 & -2 \\ 0 & 6-\lambda & 3 \\ -1 & 0 & 7-\lambda \end{pmatrix} = (6-\lambda) \det \begin{pmatrix} 8-\lambda & -2 \\ -1 & 7-\lambda \end{pmatrix} \\ &= (6-\lambda)((8-\lambda)(7-\lambda) - 2) = (6-\lambda)(\lambda^2 - 15\lambda + 54) \\ &= (6-\lambda)^2(9-\lambda). \end{aligned}$$

Teniendo ahora dos autovalores, $\lambda_1 = 6$ (doble) y $\lambda_2 = 9$ (simple). Estudiemos si del autovalor múltiple, es posible hallar dos autovectores linealmente independientes:

$\boxed{V_6}$

$$A - 6I_3 = \begin{pmatrix} 8 & -2 & 1 \\ 0 & 6 & 3 \\ -1 & 1 & 7 \end{pmatrix} - \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 1 \\ 0 & 0 & 3 \\ -1 & 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & -2 & 1 \\ 0 & 0 & 3 \\ \boxed{\text{-1}} & 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{\text{-1}} & 1 & 1 \\ 0 & 0 & \boxed{3} \\ 0 & 0 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{\text{-1}} & 1 & 1 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \end{pmatrix}$$

teniendo entonces que $\dim(V_6) = 1$ y una base $\mathcal{B}_{V_6} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$, de forma que no existen dos autovectores asociados a $\lambda_1 = 6$ que sean linealmente independientes. Demostrando así que A no es diagonalizable por semejanza.

Resulta que no todas las transformaciones lineales (o matrices) son diagonalizables por semejanza. De hecho, hemos podido observar en este último ejemplo que hay veces que el número de autovectores es insuficiente cuando un autovalor se ha repetido, o lo que es lo mismo, ha sido una raíz múltiple del polinomio característico. Por ello, a estos valores le daremos un nombre específico, pues resultarán cruciales para estudiar cuando una transformación lineal (o matriz) es diagonalizable por semejanza.

$\boxed{\textbf{Definición}}$. Sea $T\colon V \to W$ una transformación lineal con λ autovalor de T.

1. Se define la **multiplicidad algebraica de** λ como su multiplicidad algebraica como raíz del polinomio característico. Se denotará por m.a.(λ)
2. Se define la **multiplicidad geométrica de** λ como la dimensión de su subespacio propio asociado V_λ. Se denotará por m.g.$(\lambda) := \dim(V_\lambda)$.

En el siguiente teorema se demuestra que el número de autovectores linealmente independientes asociados a un mismo autovalor está limitado por la multiplicidad algebraica, es decir, el número de veces que el autovalor ha aparecido como raíz del polinomio característico. Indicando que, si un subespacio propio no ha generado los suficientes autovectores, otro distinto no podrá suplir las carencias. Y, por otra parte, el número de autovalores estará limitado por la dimensión del espacio, o en términos matriciales, por el tamaño de la matriz.

TEOREMA 3.12. *Sea* $T\colon V \to V$ *una transformación lineal con* $\lambda_1, \ldots, \lambda_r$ *autovalores. Entonces* $1 \leqslant \text{m.g.}(\lambda_i) \leqslant \text{m.a.}(\lambda_i)$ *y* $\sum_{i=1}^r \text{m.a.}(\lambda_i) \leqslant \dim(V)$.

DEMOSTRACIÓN. Llamemos $d = \text{m.g.}(\lambda_i) = \dim(V_{\lambda_i})$ por comodidad. Como λ_i es autovalor, sabemos que $V_{\lambda_i} \neq \{\vec{0}\}$ y por tanto $1 \leqslant d$. Consideremos ahora $\mathcal{B}_{V_{\lambda_i}} = \{\vec{b}_1, \ldots, \vec{b}_d\}$ una base de V_{λ_i} y, consideremos $\mathcal{C}_{(V_{\lambda_i})^c} = \{\vec{c}_{d+1}, \ldots, \vec{c}_n\}$ una base del subespacio complementario $(V_{\lambda_i})^c$. Podemos entonces construir $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_d, \vec{c}_{d+1}, \ldots, \vec{c}_n\}$ una base de V. Ahora, la matriz asociada a T en $\mathcal{B}$ será de la forma:

$$
\begin{aligned}
M(T)_{\mathcal{B}\leftarrow\mathcal{B}} &= \left(\begin{array}{c|c|c|c|c|c} [T(\vec{b}_1)]_{\mathcal{B}} & \cdots & [T(\vec{b}_d)]_{\mathcal{B}} & [T(\vec{c}_{d+1})]_{\mathcal{B}} & \cdots & [T(\vec{c}_n)]_{\mathcal{B}} \end{array} \right) \\
&= \left(\begin{array}{cccc|c|c|c}
\lambda_i & 0 & \cdots & 0 & & & \\
0 & \lambda_i & \cdots & 0 & & & \\
\vdots & \vdots & \ddots & \vdots & & & \\
0 & 0 & \cdots & \lambda_i & [T(\vec{c}_{d+1})]_{\mathcal{B}} & \cdots & [T(\vec{c}_n)]_{\mathcal{B}} \\
0 & 0 & \cdots & 0 & & & \\
\vdots & \vdots & \vdots & \vdots & & & \\
0 & 0 & \cdots & 0 & & &
\end{array} \right)
\end{aligned}
$$

Entonces su polinomio característico $p(\lambda) = \det(M(T)_{\mathcal{B}\leftarrow\mathcal{B}} - \lambda I_n) = (\lambda_i - \lambda)^d q(\lambda)$. Por tanto m.g.$(\lambda_i) \leqslant$ m.a.(λ_i), y se tiene la igualdad si λ_i no es raíz del polinomio $q(\lambda)$.

Repitiendo este proceso para cada $i \in \{1, \ldots, r\}$, podremos expresar $p(\lambda) = (\lambda_1 - \lambda)^{\text{m.g.}(\lambda_1)} \cdots (\lambda_r - \lambda)^{\text{m.g.}(\lambda_r)} q(\lambda)$ y como $\deg(p(\lambda)) = \dim(V)$, tenemos que $\sum_{i=1}^r \text{m.a.}(\lambda_i) \leqslant \dim(V)$ $\square$

El número de raíces que tiene un polinomio es un tema de estudio a parte. Aquí, sólo recordaremos que las raíces de un polinomio pueden ser reales, o complejas. En general, si admitimos raíces complejas, todo polinomio de grado n tendrá n raíces complejas. Pero esto no es cierto si sólo queremos raíces reales. Por ejemplo, el polinomio $p(x) = 1 + x^2$ no tiene raíces reales.

De aquí en adelante, sólo nos interesará resolver el problema de la diagonalización por semejanza para matrices reales.

DEFINICIÓN. Sea $A \in M_n(\mathbb{R})$. Diremos que A es **diagonalizable por semejanza real** si existe $D \in M_n(\mathbb{R})$ diagonal y $P \in M_n(\mathbb{R})$ invertible tal que $A = PDP^{-1}$. Cuando no se especifique lo contrario, diagonalizar por semejanza se entenderá como diagonalizar por semejanza real.

EJEMPLO 35. Hallemos los autovalores de $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

$$p(\lambda) = \det \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = (-i - \lambda)(i - \lambda),$$

que tiene autovalores $\lambda_1 = -i$ y $\lambda_2 = i$. Como no son autovalores reales, la matriz A no es diagonalizable por semejanza real, ya que la matriz $D = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \in M_2(\mathbb{C})$. No obstante, podríamos seguir calculando los subespacios propios:

$\boxed{V_{-i}}$

$$A - (-i)I_2 = \begin{pmatrix} i & -1 \\ \boxed{1} & i \end{pmatrix} \to \begin{pmatrix} \boxed{1} & i \\ 0 & 0 \end{pmatrix},$$

donde se ha utilizado que $i^2 = -1$. Teniendo que $\mathcal{B}_{V_{-i}} = \left\{ \begin{bmatrix} -i \\ 1 \end{bmatrix} \right\}$.

$\boxed{V_i}$

$$A - iI_2 = \begin{pmatrix} -i & -1 \\ \boxed{1} & -i \end{pmatrix} \to \begin{pmatrix} \boxed{1} & -i \\ 0 & 0 \end{pmatrix},$$

Teniendo que $\mathcal{B}_{V_i} = \left\{ \begin{bmatrix} i \\ 1 \end{bmatrix} \right\}$.

Construyendo así las matrices

$$D = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \in M_2(\mathbb{C}) \quad \text{y} \quad P = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} \in M_2(\mathbb{C}),$$

que verifican $A = PDP^{-1}$. Podemos decir que A es diagonalizable por semejanza compleja, pero no es diagonalizable por semejanza real.

TEOREMA 3.13. *Sea $T\colon V \to V$ una transformación lineal con $\lambda_1, \lambda_2, \ldots, \lambda_r$ autovalores de T. Entonces T es diagonalizable por semejanza si, y solo si,*

1. $\sum_{i=1}^r \mathrm{m.a.}(\lambda_i) = \dim(V)$,
2. $\mathrm{m.a.}(\lambda_i) = \mathrm{m.g.}(\lambda_i)$ *para todo* $i \in \{1, \ldots, r\}$.

En particular, si todos los autovalores son distintos, es decir, $\mathrm{m.a.}(\lambda_i) = 1$ *para todo* $i \in \{1, \ldots, r\}$*, entonces T es diagonalizable por semejanza.*

DEMOSTRACIÓN.

$\boxed{\Rightarrow}$ Supongamos que T es diagonalizable. Entonces existe $\mathcal{B}$ base de autovectores y, por el Teorema 3.11, $V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r}$, con lo que, también por el Teorema 2.8 y Teorema 3.12, $\dim(V) = \sum_{i=1}^r \dim(V_{\lambda_i}) = \sum_{i=1}^r \mathrm{m.g.}(\lambda_i) \leqslant \sum_{i=1}^r \mathrm{m.a.}(\lambda_i) \leqslant \dim(V)$, es decir, $\sum_{i=1}^r \mathrm{m.a.}(\lambda_i) = \dim(V)$. Por el mismo motivo, $\mathrm{m.g.}(\lambda_i) = \mathrm{m.a.}(\lambda_i)$ para todo $i \in \{1, \ldots, n\}$.

$\boxed{\Leftarrow}$ Recíprocamente, si

$\square$

Ejemplo 36. Estudiemos, en función de $a \in \mathbb{R}$, cuando $A = \begin{pmatrix} 0 & 1 & 1 \\ -a & a-1 & a \\ 1 & 1 & 0 \end{pmatrix}$ es diagonalizable por semejanza. Aplicando el Teorema 3.13, estudiaremos las multiplicidades algebraicas y geométricas de los autovalores.

$$\begin{aligned} p(\lambda) &= \det \begin{pmatrix} -\lambda & 1 & 1 \\ -a & a-1-\lambda & a \\ 1 & 1 & -\lambda \end{pmatrix} \\ &\xlongequal{C_3^N = C_3 + C_1} \det \begin{pmatrix} -\lambda & 1 & 1-\lambda \\ -a & a-1-\lambda & 0 \\ 1 & 1 & 1-\lambda \end{pmatrix} \\ &\xlongequal{F_1^N = F_1 - F_3} \det \begin{pmatrix} -1-\lambda & 0 & 0 \\ -a & a-1-\lambda & 0 \\ 1 & 1 & 1-\lambda \end{pmatrix} = (-1-\lambda)(1-\lambda)(a-1-\lambda) \end{aligned}$$

Entonces tenemos tres casos, según la multiplicidad algebraica de los autovalores, que son si $a-1 \neq -1, 1$, $a-1=-1$ y $a-1=1$, que son equivalentes a $a \neq 0, 2$, $a=0$ y $a=2$.

$\boxed{a \neq 0, 2}$ Como tenemos tres autovalores distintos con multiplicidad algebraica 1. Por el Teorema 3.13, A es diagonalizable por semejanza.

$\boxed{a = 0}$ Entonces $\lambda_1 = -1$ con $\text{m.a.}(-1) = 2$, $\lambda_2 = 1$ y $\text{m.a.}(1) = 1$. Tan solo debemos estudiar $\text{m.g.}(-1)$, ya que $1 \leqslant \text{m.g.}(1) \leqslant \text{m.a.}(1) = 1$, es decir, $\text{m.g.}(1) = \text{m.a.}(1) = 1$. En general, sólo es suficiente estudiar los autovalores múltiples.

$\boxed{V_{-1}}$

$$A - (-1)I_3 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \boxed{1} & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

teniendo que $\dim(V_{-1}) = 3 - \text{rg}(A + I_3) = 3 - 1 = 2 = \text{m.g.}(-1) = \text{m.a.}(-1)$. Y por tanto A es diagonalizable por semejanza.

$\boxed{a = 2}$ Entonces $\lambda_1 = -1$ con $\text{m.a.}(-1) = 1$, $\lambda_2 = 1$ y $\text{m.a.}(1) = 2$. Solo es necesario estudiar $\text{m.g.}(1)$:

$\boxed{V_1}$

$$A - I_3 = \begin{pmatrix} 0 & 1 & 1 \\ -2 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \boxed{1} & 1 & 1 \\ -2 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & 1 & 1 \\ 0 & \boxed{4} & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

teniendo que $\dim(V_1) = 3 - \text{rg}(A - I_3) = 3 - 2 = 1 = \text{m.g.}(1) \neq 2 = \text{m.a.}(1)$, y por tanto A no es diagonalizable por semejanza.

Teorema 3.14. *Sea $A \in M_{n \times n}(\mathbb{K})$ con autovalores $\lambda_1, \ldots, \lambda_r$.*

1. $\det(A) = \prod_{i=1}^{r} \lambda_i^{\text{m.a.}(\lambda_i)}$,
2. $\text{tr}(A) = \sum_{i=1}^{r} \text{m.a.}(\lambda_i)\lambda_i$.

Demostración.

1. Sea $p(\lambda) = \det(A - \lambda I)$ el polinomio característico de A. Como los autovalores de A son las raíces del polinomio característico, entonces lo podemos expresar como
$$\det(A - \lambda I) = p(\lambda) = (\lambda_1 - \lambda)^{\mathrm{m.a.}(\lambda_1)}(\lambda_2 - \lambda)^{\mathrm{m.a.}(\lambda_2)} \cdots (\lambda_r - \lambda)^{\mathrm{m.a.}(\lambda_r)}.$$
Por tanto, evaluando en $\lambda = 0$:
$$\det(A) = \det(A - 0I) = p(0) = \lambda_1^{\mathrm{m.a.}(\lambda_1)}\lambda_2^{\mathrm{m.a.}(\lambda_2)} \cdots \lambda_r^{\mathrm{m.a.}(\lambda_r)} = \prod_{i=1}^{r} \lambda_i^{\mathrm{m.a.}(\lambda_i)}.$$
2. Sólo lo demostraremos para el caso en que A sea diagonalizable por semejanza, aunque sea cierto en general. Supongamos que $A = PDP^{-1}$ con D matriz diagonal con los autovalores en la diagonal y P la matriz con los autovectores en sus columnas. Entonces
$$\mathrm{tr}(A) = \mathrm{tr}(P(DP^{-1})) = \mathrm{tr}((DP^{-1})P) = \mathrm{tr}(D) = \sum_{i=1}^{r} \mathrm{m.a.}(\lambda_i)\lambda_i.$$

□

5. Algunas aplicaciones de la diagonalización por semejanza

Una primera, y buena pregunta, puede ser: ¿Qué ventajas tiene la diagonalización por semejanza respecto a la diagonalización por equivalencia por filas y columnas?

Ya hemos estudiado que si realizamos el algoritmo de Gauss-Jordan por filas a una matriz A obtendremos una matriz equivalente por filas FA que será triangular superior. A su vez, si aplicamos Gauss-Jordan por columnas a esta última matriz, obtendremos otra matriz equivalente por columnas FAC que será diagonal. Sin embargo, afirmar que la matriz A es diagonalizable por semejanza implica que $P^{-1}AP$ es diagonal, que para coincidir con la diagonalización por equivalencias significaría que $F = C^{-1}$, una propiedad que casi nunca se obtendría de manera “sencilla”. Entonces,

¿cómo podemos aprovechar la expresión $A = PDP^{-1}$?

5.1. La exponencial de una matriz.

Teorema 3.15. *Sea $A \in M_n(\mathbb{K})$ con autovector $\vec{u} \in \mathbb{K}^n$ asociado al autovalor $\lambda \in \mathbb{K}$.*

1. *Para todo $k \in \mathbb{N} \cup \{-1\}$, $\vec{u}$ es un autovector de A^k asociado al autovalor λ^k,*
2. *λ es autovalor de A^t, aunque generalmente $\vec{u}$ no será autovector de A^t.*
3. *Si A es diagonalizable por semejanza, es decir, existe una matriz D diagonal y P invertible tales que*
$$A = PDP^{-1} \quad \text{entonces} \quad A^k = PD^kP^{-1}.$$

Demostración.

1. Sea $k \in \mathbb{N}$. Vamos a demostrar por inducción que $A^k\vec{u} = \lambda^k\vec{u}$. Claramente, para $k = 0$ es cierto, ya que $A^0 = I_n$, que tiene a $\lambda^0 = 1$ como autovalor, y $\vec{u}$ es un autovector, pues en particular, $\vec{u} \in \mathrm{Nul}(A^0 - 1 \cdot I_n) = \mathrm{Nul}(O_n) = \mathbb{K}^n$. Supongamos que es cierto para $k - 1$, es decir, supongamos que la hipótesis de inducción es:
$$\text{Hipótesis de Inducción (H.I.)} \quad A^{k-1}\vec{u} = \lambda^{k-1}\vec{u}$$

y veamos que también lo es para k:

$$A^k\vec{u} = A(A^{k-1}\vec{u}) \overset{\text{(H.I.)}}{=\!=\!=} A(\lambda^{k-1}\vec{u}) = \lambda^{k-1}A\vec{u} \overset{\text{def. de } \vec{u}}{=\!=\!=\!=} \lambda^{k-1}\lambda\vec{u} = \lambda^k\vec{u}.$$

□

Ejemplo 37. Sea $A = \begin{pmatrix} -7 & 3 & 3 \\ -9 & 5 & 3 \\ -9 & 3 & 5 \end{pmatrix}$, hallemos la expresión general para A^k. Utilizando el Teorema 3.15(3), diagonalicemos por semejanza la matriz A. Primero, hallemos su polinomio característico:

$$\begin{aligned}
p(\lambda) &= \det(A - \lambda I_3) = \det \begin{pmatrix} -7-\lambda & 3 & 3 \\ -9 & 5-\lambda & 3 \\ -9 & 3 & 5-\lambda \end{pmatrix} \\
&\overset{F_1^N = F_1 - F_3}{=\!=\!=\!=} \det \begin{pmatrix} 2-\lambda & 0 & -2+\lambda \\ -9 & 5-\lambda & 3 \\ -9 & 3 & 5-\lambda \end{pmatrix} \overset{C_3^N = C_3 + C_1}{=\!=\!=\!=} \det \begin{pmatrix} 2-\lambda & 0 & 0 \\ -9 & 5-\lambda & -6 \\ -9 & 3 & -4-\lambda \end{pmatrix} \\
&= (2-\lambda)\det \begin{pmatrix} 5-\lambda & -6 \\ 3 & -4-\lambda \end{pmatrix} = (2-\lambda)((5-\lambda)(-4-\lambda)+18) \\
&= (2-\lambda)(\lambda^2-\lambda-2) = (2-\lambda)^2(-1-\lambda).
\end{aligned}$$

Teniendo que $\lambda_1 = 2$ con m.a.$(2) = 2$ y $\lambda_2 = -1$ con m.a.$(-1) = 1$. Hallemos ahora sus subespacios propios:

$\boxed{V_2}$

$$A - 2I_3 = \begin{pmatrix} -7 & 3 & 3 \\ -9 & 5 & 3 \\ -9 & 3 & 5 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \boxed{-9} & 3 & 3 \\ -9 & 3 & 3 \\ -9 & 3 & 3 \end{pmatrix} \to \begin{pmatrix} \boxed{3} & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

teniendo que $\mathcal{B}_{V_2} = \left\{ \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \right\}$ es una base de autovectores asociados a $\lambda_1 = 2$.

$\boxed{V_{-1}}$

$$A-(-1)I_3 = \begin{pmatrix} -7 & 3 & 3 \\ -9 & 5 & 3 \\ -9 & 3 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \boxed{-6} & 3 & 3 \\ -9 & 6 & 3 \\ -9 & 3 & 6 \end{pmatrix} \to \begin{pmatrix} \boxed{2} & -1 & -1 \\ 0 & \boxed{3} & -3 \\ 0 & -3 & 3 \end{pmatrix} \to \begin{pmatrix} \boxed{2} & -1 & -1 \\ 0 & \boxed{1} & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

teniendo que $\mathcal{B}_{V_{-1}} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ es una base de autovectores asociados a $\lambda_2 = -1$.

Entonces, las matrices

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{y} \quad P = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix}$$

verifican $A = PDP^{-1}$. Entonces, por el Teorema 3.15(3), las matrices

$$D^k = \begin{pmatrix} 2^k & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & (-1)^k \end{pmatrix} \quad \text{y} \quad P = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix}$$

verifican $A^k = PD^kP^{-1}$.

Definición. Sea A una matriz cuadrada. Definimos la **exponencial de** A como la siguiente serie matricial

$$e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!} = I + A + \frac{A^2}{2} + \frac{A^3}{6} + \dots$$

Teorema 3.16. *Sean $A, B, P \in M_n(\mathbb{R})$. Entonces*

1. *Si $AB = BA$ entonces $e^{A+B} = e^A e^B$,*
2. *Si P es invertible entonces $e^{PAP^{-1}} = Pe^AP^{-1}$,*
3. *Si λ es autovalor de A con autovector asociado $\vec{u}$ entonces e^λ es autovalor de e^A con el mismo autovector asociado $\vec{u}$,*
4. $\det(e^A) = e^{\operatorname{tr}(A)}$.

Ejemplo 38. Hallemos la matriz exponencial de $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Al ser D diagonal, $D^k = \begin{pmatrix} 2^k & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & (-1)^k \end{pmatrix}$. Por definición

$$e^D := \sum_{k=0}^{\infty} \frac{D^k}{k!} = \sum_{k=0}^{\infty} \begin{pmatrix} \frac{2^k}{k!} & 0 & 0 \\ 0 & \frac{2^k}{k!} & 0 \\ 0 & 0 & \frac{(-1)^k}{k!} \end{pmatrix} = \begin{pmatrix} \sum_{k=0}^{\infty} \frac{2^k}{k!} & 0 & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{2^k}{k!} & 0 \\ 0 & 0 & \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \end{pmatrix} = \begin{pmatrix} e^2 & 0 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & e^{-1} \end{pmatrix}.$$

Entonces, hallar la matriz exponencial de una matriz diagonal es inmediata.

Ejemplo 39. Consideremos la matriz del ejemplo anterior $A = \begin{pmatrix} -7 & 3 & 3 \\ -9 & 5 & 3 \\ -9 & 3 & 5 \end{pmatrix}$. Hallemos su matriz exponencial e^A. Como ya hemos visto en el `Ejemplo` 37, podemos descomponer $A = PDP^{-1}$ con

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{y} \quad P = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix},$$

y utilizando el Teorema 3.16(2),

$$e^A = Pe^DP^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} e^2 & 0 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & e^{-1} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 3 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 3e^{-1} - 2e^2 & -e^{-1} + e^2 & -e^{-1} + e^2 \\ 3e^{-1} - 3e^2 & -e^{-1} + 2e^2 & -e^{-1} + e^2 \\ 3e^{-1} - 3e^2 & -e^{-1} + e^2 & -e^{-1} + 2e^2 \end{pmatrix},$$

ya que, tal y como hemos visto en el `Ejemplo` 38, la matriz exponencial de una matriz diagonal es inmediata de calcular $e^D = \begin{pmatrix} e^2 & 0 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & e^{-1} \end{pmatrix}$.

TEMA 4

Espacios vectoriales euclídeos reales

En los temas anteriores, hemos estudiado con detalle la noción de espacio vectorial. También cómo transformar sus vectores, o cómo relacionarlos entre ellos mediante las transformaciones lineales. Finalmente, en este tema, añadiremos un nuevo concepto, el de *producto escalar*, que nos permitirá estudiar una nueva característica de los vectores, en concreto la longitud de un vector y el ángulo que forman entre ellos. Así, podemos entender que el *producto escalar* es una herramienta de *medición*, y podríamos referinos a ella como la *métrica* del espacio vectorial.

1. Definiciones básicas

DEFINICIÓN. Sea V un $\mathbb{R}$-espacio vectorial. Diremos que la función

$$\begin{array}{rccc} \langle \cdot\,,\cdot \rangle\colon & V \times V & \to & \mathbb{R} \\ & (\vec{u},\vec{v}) & \mapsto & \langle \vec{u}\,,\vec{v} \rangle \end{array}$$

es un **producto escalar de** V si cumple las siguientes propiedades, para todo $\vec{u},\vec{v},\vec{w} \in V$ y $\lambda \in \mathbb{R}$:

1. Es bilineal:

   Primera componente: $\langle \vec{u}+\vec{w},\, \vec{v}\rangle = \langle \vec{u},\, \vec{v}\rangle + \langle \vec{w},\, \vec{v}\rangle$, y $\langle \lambda\vec{u},\, \vec{v}\rangle = \lambda\langle \vec{u},\, \vec{v}\rangle$,

   Segunda componente: $\langle \vec{u},\, \vec{v}+\vec{w}\rangle = \langle \vec{u},\, \vec{v}\rangle + \langle \vec{u},\, \vec{w}\rangle$, y $\langle \vec{u},\, \lambda\vec{v}\rangle = \lambda\langle \vec{u},\, \vec{v}\rangle$.
2. Es simétrica: $\langle \vec{u}\,,\vec{v}\rangle = \langle \vec{v}\,,\vec{u}\rangle$.
3. Es definida positiva: $\langle \vec{u}\,,\vec{u}\rangle \geqslant 0$ y $\langle \vec{u}\,,\vec{u}\rangle = 0$ si, y solo si, $\vec{u} = \vec{0}$.

Diremos que $(V, \langle \cdot\,,\cdot\rangle)$ es un **$\mathbb{R}$-espacio vectorial euclídeo.**

Aunque se puede dar una definición para espacios vectoriales complejos, hay varios detalles que se deberían cubrir con cuidado. Por falta de tiempo, tan solo estudiaremos espacios vectoriales euclídeos reales.

EJEMPLO 40.

1. El **producto escalar usual en** $\mathbb{R}^2$ se define como

$$\left\langle \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\rangle = x_1y_1 + x_2y_2.$$

   Que dota a $\mathbb{R}^2$ de estructura de espacio vectorial euclídeo real. Una observación muy importante es que

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1y_1 + x_2y_2 = \left\langle \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\rangle.$$

Por ejemplo, $\left\langle \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix} \right\rangle = 1\cdot 3 + 2\cdot 4 = 11.$

2. En general, el **producto escalar usual en** $\mathbb{R}^n$ se define como

$$\left\langle \begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}, \begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix} \right\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Siguiendo la observación anterior, si $\vec{u} = \begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$ y $\vec{v} = \begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}$, entonces

$$\langle \vec{u},\, \vec{v}\rangle = \vec{u}^{\,t}\,\vec{v} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

3. Sin embargo, podemos definir nuevos productos escalares en $\mathbb{R}^n$ satisfaciendo reglas distintas. Por ejemplo, para $n = 2$,

$$\left\langle \begin{bmatrix}x_1\\x_2\end{bmatrix}, \begin{bmatrix}y_1\\y_2\end{bmatrix} \right\rangle = 2x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2.$$

es también otro producto escalar (hay que comprobar que verifica las tres propiedades). Por ejemplo, $\left\langle \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix} \right\rangle = 2\cdot 1\cdot 3 - 1\cdot 4 - 2\cdot 3 + 2\cdot 4 = 6 - 4 - 6 + 8 = 4$. No coincidiendo con el valor obtenido con el producto escalar usual.

4. En $M_{m\times n}(\mathbb{R})$, se puede definir el siguiente producto escalar:

$$\langle A,\, B\rangle = \mathrm{tr}(A^tB).$$

5. En $\mathbb{P}_n$, para $a, b \in \mathbb{R}$, se puede definir

$$\langle p(x),\, q(x)\rangle = \int_a^b p(x)q(x)\mathrm{d}x.$$

1.1. Matriz de Gram. Sea $(V, \langle\cdot\,,\cdot\rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y sea $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ una base de V. Consideremos $\vec{u}, \vec{v} \in V$ con coordenadas en $\mathcal{B}$

$$\vec{u} = x_1\vec{b}_1 + \cdots + x_n\vec{b}_n = \sum_{i=1}^n x_i\vec{b}_i, \quad \vec{v} = y_1\vec{b}_1 + \cdots + y_n\vec{b}_n = \sum_{j=1}^n y_j\vec{b}_j.$$

Entonces, si queremos hallar $\langle \vec{u}\,, \vec{v}\rangle$, utilizando la bilinealidad:

$$\begin{aligned}\langle \vec{u}\,, \vec{v}\rangle &= \langle \textstyle\sum_{i=1}^n x_i\vec{b}_i\,, \sum_{j=1}^n y_j\vec{b}_j\rangle \xlongequal{\text{Lin. 1ª comp.}} \sum_{i=1}^n x_i\langle \vec{b}_i\,, \sum_{j=1}^n y_j\vec{b}_j\rangle \\ &\xlongequal{\text{Lin. 2ª comp.}} \textstyle\sum_{i=1}^n x_i \sum_{j=1}^n y_j\langle \vec{b}_i\,, \vec{b}_j\rangle = \sum_{i,j=1}^n x_iy_j\langle \vec{b}_i\,, \vec{b}_j\rangle\end{aligned}$$

Por otro lado,

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{pmatrix} \langle \vec{b}_1 , \vec{b}_1 \rangle & \langle \vec{b}_1 , \vec{b}_2 \rangle & \cdots & \langle \vec{b}_1 , \vec{b}_n \rangle \\ \langle \vec{b}_2 , \vec{b}_1 \rangle & \langle \vec{b}_2 , \vec{b}_2 \rangle & \cdots & \langle \vec{b}_2 , \vec{b}_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{b}_n , \vec{b}_1 \rangle & \langle \vec{b}_n , \vec{b}_2 \rangle & \cdots & \langle \vec{b}_n , \vec{b}_n \rangle \end{pmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \sum_{j=1}^n \langle \vec{b}_1 , \vec{b}_j \rangle y_j \\ \sum_{j=1}^n \langle \vec{b}_2 , \vec{b}_j \rangle y_j \\ \vdots \\ \sum_{j=1}^n \langle \vec{b}_n , \vec{b}_j \rangle y_j \end{bmatrix}$$

$$= \sum_{i=1}^n x_i \sum_{j=1}^n \langle \vec{b}_i , \vec{b}_j \rangle y_j = \sum_{i,j=1}^n x_i y_j \langle \vec{b}_i , \vec{b}_j \rangle = \langle \vec{u} , \vec{v} \rangle.$$

Entonces, si llamamos $G_{\mathcal{B}} = (g_{ij} = \langle \vec{b}_i , \vec{b}_j \rangle)$ entonces se verifica:

$$\langle \vec{u}, \vec{v} \rangle = [\vec{u}]_{\mathcal{B}}^t \, G_{\mathcal{B}} \, [\vec{v}]_{\mathcal{B}}.$$

A la matriz $G_{\mathcal{B}}$ la llamaremos **matriz de Gram de $\langle \cdot , \cdot \rangle$ en $\mathcal{B}$**. Como consecuencia de la definición del producto escalar, por ser simétrico, la matriz de Gram es una matriz simétrica, es decir, $G_{\mathcal{B}}^t = G_{\mathcal{B}}$.

Si tuviéramos otra base $\mathcal{C}$ de V, nos podemos plantear la relación entre $G_{\mathcal{B}}$ y $G_{\mathcal{C}}$. Para ello únicamente recordemos las propiedades de la trasposición del producto de matrices $(AB)^t = B^t A^t$:

$$[\vec{u}]_{\mathcal{C}}^t \, G_{\mathcal{C}} \, [\vec{v}]_{\mathcal{C}} = \langle \vec{u}, \vec{v} \rangle = [\vec{u}]_{\mathcal{B}}^t \, G_{\mathcal{B}} \, [\vec{v}]_{\mathcal{B}} = \left(P_{\mathcal{B}\leftarrow\mathcal{C}}[\vec{u}]_{\mathcal{C}}\right)^t \, G_{\mathcal{B}} \, P_{\mathcal{B}\leftarrow\mathcal{C}}[\vec{v}]_{\mathcal{C}} = [\vec{u}]_{\mathcal{C}}^t \left(\left(P_{\mathcal{B}\leftarrow\mathcal{C}}\right)^t \, G_{\mathcal{B}} \, P_{\mathcal{B}\leftarrow\mathcal{C}}\right) [\vec{v}]_{\mathcal{C}},$$

de donde podemos concluir que

$$G_{\mathcal{C}} = \left(P_{\mathcal{B}\leftarrow\mathcal{C}}\right)^t \, G_{\mathcal{B}} \, P_{\mathcal{B}\leftarrow\mathcal{C}}.$$

Definición. Sean $A, B \in M_n(\mathbb{K})$. Diremos que A y B son **matrices congruentes** si existe $P \in M_n(\mathbb{K})$ invertible tal que $A = P^t B P$.

Entonces hemos visto que las matrices de Gram en bases distintas son matrices congruentes.

Es importante también notar que si dos matrices congruentes, una de ellas es simétrica, entonces la otra también debe serlo: Si B es simétrica, entonces $A^t = (P^t B P)^t = P^t B^t (P^t)^t = P^t B P = A$. Y en particular, si B es diagonal, es simétrica.

Definición. Sea $(V, \langle \cdot , \cdot \rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo. Definimos la **norma de** $\vec{u} \in V$ como

$$\|\vec{u}\| := \sqrt{\langle \vec{u} , \vec{u} \rangle}.$$

También diremos que $\vec{u} \in V$ es un **vector unitario o normalizado** si $\|\vec{u}\| = 1$, y diremos que $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m\} \subset V$ es un **conjunto unitario o normalizado** si $\|\vec{u}_i\| = 1$ para todo $i \in \{1, \ldots, m\}$.

Debido a que el producto escalar es definido positivo, el radicando siempre será positivo o igual a cero, y no hay ningún problema en la definición.

La norma de un vector se puede interpretar intuitivamente como la *longitud* del vector.

PROPIEDADES. *Sea* $(V, \langle \cdot\,,\cdot\rangle)$ *un* $\mathbb{R}$*-espacio vectorial euclídeo. Para todo* $\vec{u},\vec{v}\in V$ *y* $\lambda\in\mathbb{R}$*:*

1. $\|\vec{u}\|\geqslant 0$ *y* $\|\vec{u}\|=0$ *si, y solo si,* $\vec{u}=\vec{0}$.
2. $\|\lambda\vec{u}\|=|\lambda|\|\vec{u}\|$.
3. *Si* $\vec{u}\neq\vec{0}$*, el vector* $\dfrac{\vec{u}}{\|\vec{u}\|}$ *es unitario, que llamaremos el* ***vector normalizado de*** $\vec{u}$.
4. *(Desigualdad de Cauchy-Schwarz)* $|\langle\vec{u}\,,\vec{v}\rangle|\leqslant\|\vec{u}\|\|\vec{v}\|$.
5. *(Desigualdad triangular)* $\|\vec{u}+\vec{v}\|\leqslant\|\vec{u}\|+\|\vec{v}\|$.

DEMOSTRACIÓN.

1. Por definición: $\|\vec{u}\|=\sqrt{\langle\vec{u},\,\vec{u}\rangle}\geqslant 0$. Además, si $0=\|\vec{u}\|=\sqrt{\langle\vec{u},\,\vec{u}\rangle}$, entonces $\langle\vec{u},\,\vec{u}\rangle=0$, que por ser definido positivo el producto escalar, $\vec{u}=\vec{0}$.
2. $\|\lambda\vec{u}\|=\sqrt{\langle\lambda\vec{u},\,\lambda\vec{u}\rangle}=\sqrt{\lambda^2\langle\vec{u},\,\vec{u}\rangle}=|\lambda|\sqrt{\langle\vec{u},\,\vec{u}\rangle}$.
3. $\|\frac{\vec{u}}{\|\vec{u}\|}\|=\frac{1}{\|\vec{u}\|}\|\vec{u}\|=1$.
4. Si alguno de los vectores es nulo, la desigualdad se cumple trivialmente. Supongamos que $\vec{u},\vec{v}\neq\vec{0}$ y sea $\lambda\in\mathbb{R}$:
$$0\leqslant\langle\vec{u}-\lambda\vec{v},\,\vec{u}-\lambda\vec{v}\rangle=\|\vec{u}\|^2-2\lambda\langle\vec{u},\,\vec{v}\rangle+\lambda^2\|\vec{v}\|^2,$$
en particular, si $\lambda=\dfrac{\langle\vec{u},\,\vec{v}\rangle}{\|\vec{v}\|^2}$ tenemos la desigualdad
$$0\leqslant\|\vec{u}\|^2-2\frac{\langle\vec{u},\,\vec{v}\rangle^2}{\|\vec{v}\|^2}+\frac{\langle\vec{u},\,\vec{v}\rangle^2}{\|\vec{v}\|^2}=\|\vec{u}\|^2-\frac{\langle\vec{u},\,\vec{v}\rangle^2}{\|\vec{v}\|^2}\quad\Longleftrightarrow\quad\langle\vec{u},\,\vec{v}\rangle^2\leqslant\|\vec{u}\|^2\|\vec{v}\|^2,$$
que tomando raíz cuadrada en la última desigualdad:
$$|\langle\vec{u},\,\vec{v}\rangle|=\sqrt{\langle\vec{u},\,\vec{v}\rangle^2}\leqslant\sqrt{\|\vec{u}\|^2\|\vec{v}\|^2}=\|\vec{u}\|\|\vec{v}\|$$
5. Teniendo en cuenta la desigualdad de Cauchy-Schwarz:
$$\begin{aligned}\|\vec{u}+\vec{v}\|^2 &= \langle\vec{u}+\vec{v},\,\vec{u}+\vec{v}\rangle=\|\vec{u}\|^2+2\,\langle\vec{u},\,\vec{v}\rangle+\|\vec{v}\|^2\\ &\leqslant\|\vec{u}\|^2+2|\langle\vec{u},\,\vec{v}\rangle|+\|\vec{v}\|^2\\ &\leqslant\|\vec{u}\|^2+2\|\vec{u}\|\|\vec{v}\|+\|\vec{v}\|^2\\ &=(\|\vec{u}\|+\|\vec{v}\|)^2.\end{aligned}$$
Tomando raíz cuadrada, tenemos la desigualdad triangular:
$$\|\vec{u}+\vec{v}\|\leqslant\|\vec{u}\|+\|\vec{v}\|.$$

$\square$

La desigualdad de Cauchy-Schwarz nos permite tener intuición del ángulo que forman dos vectores. Sean $\vec{u},\vec{v}\neq\vec{0}$:
$$|\langle\vec{u}\,,\vec{v}\rangle|\leqslant\|\vec{u}\|\|\vec{v}\|\quad\Longleftrightarrow\quad-\|\vec{u}\|\|\vec{v}\|\leqslant\langle\vec{u}\,,\vec{v}\rangle\leqslant\|\vec{u}\|\|\vec{v}\|\quad\Longleftrightarrow\quad-1\leqslant\frac{\langle\vec{u}\,,\vec{v}\rangle}{\|\vec{u}\|\|\vec{v}\|}\leqslant 1,$$

es decir, $\dfrac{\langle \vec{u}\,, \vec{v}\rangle}{\|\vec{u}\|\|\vec{v}\|} \in [-1,\, 1]$. Ya que la función trigonométrica $\cos\colon [0,\, \pi] \to [-1,\, 1]$ es biyectiva, existirá $\alpha \in [0,\, \pi]$ tal que $\cos(\alpha) = \dfrac{\langle \vec{u}\,, \vec{v}\rangle}{\|\vec{u}\|\|\vec{v}\|}$. A este número $\alpha \in [0,\, \pi]$ lo llamaremos **ángulo entre $\vec{u}$ y $\vec{v}$** y lo representaremos por $\angle(\vec{u}, \vec{v})$.

DEFINICIÓN. Sea $(V, \langle \cdot\,, \cdot \rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo. Diremos que $\vec{u} \in V$ y $\vec{v} \in V$ son **vectores ortogonales** si $\langle \vec{u},\, \vec{v}\rangle = 0$, y lo denotaremos por $\vec{u} \perp \vec{v}$. Además, diremos que $S = \{\vec{u}_1,\, \vec{u}_2,\, \ldots,\, \vec{u}_m\} \subset V$ es un **conjunto ortogonal** si $\vec{u}_i \perp \vec{u}_j$ para todo $i \neq j$.

EJEMPLO 41. En $\mathbb{R}^3$ con el producto escalar usual, dado $\vec{u} = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix}$ y $\vec{v} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$, definimos el **producto vectorial de $\vec{u}$ y $\vec{v}$** como

$$\vec{u} \times \vec{v} := \begin{bmatrix} u_y v_z - u_z v_y \\ u_z v_x - u_x v_z \\ u_x v_y - u_y v_x \end{bmatrix} = \det\begin{pmatrix} u_y & u_z \\ v_y & v_z \end{pmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \det\begin{pmatrix} u_x & u_z \\ v_x & v_z \end{pmatrix}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \det\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

que habitualmente suele representarse mediante el determinante *formal*, y utilizando la notación clásica de Física:

$$\vec{u} \times \vec{v} = \det\begin{pmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{pmatrix}$$

donde $\vec{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ y $\vec{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. El producto vectorial de $\vec{u}$ y $\vec{v}$ cumple las siguientes propiedades

1. $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$. Esta propiedad se ve claramente de las propiedades del determinante, al permutar dos filas.
2. $\vec{u} \times \vec{v} \perp \vec{u}$ y $\vec{u} \times \vec{v} \perp \vec{v}$. Efectivamente,

$$\begin{aligned} \langle \vec{u} \times \vec{v},\, \vec{u}\rangle &= (u_y v_z - u_z v_y)u_x + (u_z v_x - u_x v_z)u_y + (u_x v_y - u_y v_x)u_z \\ &= u_y v_z u_x - u_z v_y u_x + u_z v_x u_y - u_x v_z u_y + u_x v_y u_z - u_y v_x u_z = 0 \end{aligned}.$$

   Y el mismo cálculo para $\vec{v}$.

Por ejemplo, si $\vec{u} = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$ y $\vec{v} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$, entonces $\vec{u} \times \vec{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

TEOREMA 4.1 (Teorema de Pitágoras). *Sea $(V, \langle \cdot\,, \cdot \rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y sean $\vec{u}, \vec{v} \in V$*

$$\vec{u} \perp \vec{v} \iff \|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$$

DEMOSTRACIÓN.

$\boxed{\Rightarrow}$ Supongamos que $\vec{u} \perp \vec{v}$, es decir, $\langle \vec{u},\, \vec{v}\rangle = 0$. Entonces

$$\|\vec{u} + \vec{v}\|^2 = \langle \vec{u} + \vec{v},\, \vec{u} + \vec{v}\rangle = \langle \vec{u},\, \vec{u}\rangle + \langle \vec{u},\, \vec{v}\rangle + \langle \vec{v},\, \vec{u}\rangle + \langle \vec{v},\, \vec{v}\rangle = \|\vec{u}\|^2 + \|\vec{v}\|^2.$$

$\boxed{\Leftarrow}$ Supongamos que $\|\vec{u}+\vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$. Entonces, por una parte,

$$\|\vec{u}+\vec{v}\|^2 = \langle \vec{u},\, \vec{u}\rangle + \langle \vec{u},\, \vec{v}\rangle + \langle \vec{v},\, \vec{u}\rangle + \langle \vec{v},\, \vec{v}\rangle,$$

y por otra parte,

$$\|\vec{u}\|^2 + \|\vec{v}\|^2 = \langle \vec{u},\, \vec{u}\rangle + \langle \vec{v},\, \vec{v}\rangle,$$

que igualando ambas expresiones, tenemos que

$$\langle \vec{u},\, \vec{v}\rangle + \langle \vec{v},\, \vec{u}\rangle = 0 \quad\Longleftrightarrow\quad 2\langle \vec{u},\, \vec{v}\rangle = 0 \quad\Longleftrightarrow\quad \langle \vec{u},\, \vec{v}\rangle = 0 \quad\Longleftrightarrow\quad \vec{u}\perp\vec{v}.$$

□

A su vez, hablaremos de **bases ortogonales** cuando la base es un conjunto ortogonal. También, si la base es un conjunto ortogonal y unitario diremos que es una **base ortonormal**.

En general, sea $(V,\, \langle\cdot\,,\cdot\rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y $\mathcal{B} = \{\vec{b}_1,\, \vec{b}_2,\, \ldots,\, \vec{b}_n\}$ una base ortogonal de V, si se normalizan todos sus vectores, se obtendrá una base ortonormal, ya que se tiene la propiedad para cualesquiera vectores $\vec{u},\vec{v}\in V$:

$$\vec{u}\perp\vec{v} \quad\Longleftrightarrow\quad \frac{\vec{u}}{\|\vec{u}\|}\perp\frac{\vec{v}}{\|\vec{v}\|}$$

Teorema 4.2 (Teorema de Gram-Schmidt). *Sea $(V,\, \langle\cdot\,,\cdot\rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y sea E un subespacio vectorial con $\mathcal{B}_E = \{\vec{b}_1,\, \vec{b}_2,\, \ldots,\, \vec{b}_m\}$ una base de E. Entonces $\mathcal{C}_E = \{\vec{c}_1,\, \vec{c}_2,\, \ldots,\, \vec{c}_m\}$ donde*

$$\vec{c}_i := \vec{b}_i - \sum_{j=1}^{i-1}\frac{\langle \vec{b}_i,\, \vec{c}_j\rangle}{\langle \vec{c}_j,\, \vec{c}_j\rangle}\vec{c}_j$$

es una base ortogonal de E.

El siguiente teorema pone de manifiesto algunas de las innumerables virtudes de las bases ortogonales y ortonormales.

Teorema 4.3. *Sea $(V,\, \langle\cdot\,,\cdot\rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y sean $\mathcal{B} = \{\vec{b}_1,\, \vec{b}_2,\, \ldots,\, \vec{b}_n\}$ y $\mathcal{C}$ dos bases ortogonales de V.*

1. *$\mathcal{B}$ es ortogonal si, y solo si, $G_{\mathcal{B}}$ es diagonal. Además, $\mathcal{B}$ es ortonormal si, y solo si, $G_{\mathcal{B}} = I_{\dim(V)}$.*
2. *Si $\mathcal{B}$ y $\mathcal{C}$ son ortonormales entonces $(P_{\mathcal{B}\leftarrow\mathcal{C}})^t = P_{\mathcal{C}\leftarrow\mathcal{B}}$.*
3. *Todo $\vec{u}\in V$ se puede expresar como*

$$\vec{u} = \frac{\langle \vec{u},\, \vec{b}_1\rangle}{\langle \vec{b}_1,\, \vec{b}_1\rangle}\vec{b}_1 + \frac{\langle \vec{u},\, \vec{b}_2\rangle}{\langle \vec{b}_2,\, \vec{b}_2\rangle}\vec{b}_2 + \cdots + \frac{\langle \vec{u},\, \vec{b}_n\rangle}{\langle \vec{b}_n,\, \vec{b}_n\rangle}\vec{b}_n,$$

*a $\dfrac{\langle \vec{u},\, \vec{b}_i\rangle}{\langle \vec{b}_i,\, \vec{b}_i\rangle}$ se les llaman los **coeficientes de Fourier de $\vec{u}$ en $\mathcal{B}$**.*

Demostración.

1. Suponiendo que $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ es ortogonal, por la definición de la matriz de Gram
$$G_{\mathcal{B}} = \begin{pmatrix} \langle\vec{b}_1, \vec{b}_1\rangle & \langle\vec{b}_1, \vec{b}_2\rangle & \cdots & \langle\vec{b}_1, \vec{b}_n\rangle \\ \langle\vec{b}_2, \vec{b}_1\rangle & \langle\vec{b}_2, \vec{b}_2\rangle & \cdots & \langle\vec{b}_2, \vec{b}_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\vec{b}_n, \vec{b}_1\rangle & \langle\vec{b}_n, \vec{b}_2\rangle & \cdots & \langle\vec{b}_n, \vec{b}_n\rangle \end{pmatrix} = \begin{pmatrix} \langle\vec{b}_1, \vec{b}_1\rangle & 0 & \cdots & 0 \\ 0 & \langle\vec{b}_2, \vec{b}_2\rangle & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \langle\vec{b}_n, \vec{b}_n\rangle \end{pmatrix},$$
ya que $\langle\vec{b}_i, \vec{b}_j\rangle = 0$ si $i \neq j$ y, si $\mathcal{B}$ es ortonormal, entonces $\langle\vec{b}_i, \vec{b}_j\rangle = 1$ para todo $i \in \{1, \ldots, n\}$.
2. Por el apartado 1., si $\mathcal{C}$ es una base ortonormal entonces $G_{\mathcal{C}}$ es la matriz identidad. Por otro lado, si $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ es una base ortonormal, entonces
$$\langle\vec{b}_i, \vec{b}_j\rangle = [\vec{b}_i]^t_{\mathcal{C}} G_{\mathcal{C}} [\vec{b}_j]_{\mathcal{C}} = [\vec{b}_i]^t_{\mathcal{C}} [\vec{b}_j]_{\mathcal{C}} = 0, \text{ si } i \neq j, \quad \text{y} \quad \langle\vec{b}_i, \vec{b}_i\rangle = [\vec{b}_i]^t_{\mathcal{C}} [\vec{b}_i]_{\mathcal{C}} = 1.$$
Sabemos que $P_{\mathcal{C}\leftarrow\mathcal{B}} = \left([\vec{b}_1]_{\mathcal{C}} \,\middle|\, [\vec{b}_2]_{\mathcal{C}} \,\middle|\, \cdots \,\middle|\, [\vec{b}_n]_{\mathcal{C}} \right)$ y por tanto $(P_{\mathcal{C}\leftarrow\mathcal{B}})^t = \begin{pmatrix} [\vec{b}_1]^t_{\mathcal{C}} \\ \hline [\vec{b}_2]^t_{\mathcal{C}} \\ \hline \vdots \\ \hline [\vec{b}_n]^t_{\mathcal{C}} \end{pmatrix}$.

Así que, utilizando las cuentas anteriores:
$$(P_{\mathcal{C}\leftarrow\mathcal{B}})^t P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{pmatrix} [\vec{b}_1]^t_{\mathcal{C}} [\vec{b}_1]_{\mathcal{C}} & [\vec{b}_1]^t_{\mathcal{C}} [\vec{b}_2]_{\mathcal{C}} & \cdots & [\vec{b}_1]^t_{\mathcal{C}} [\vec{b}_n]_{\mathcal{C}} \\ [\vec{b}_2]^t_{\mathcal{C}} [\vec{b}_1]_{\mathcal{C}} & [\vec{b}_2]^t_{\mathcal{C}} [\vec{b}_2]_{\mathcal{C}} & \cdots & [\vec{b}_2]^t_{\mathcal{C}} [\vec{b}_n]_{\mathcal{C}} \\ \vdots & \vdots & \ddots & \vdots \\ [\vec{b}_n]^t_{\mathcal{C}} [\vec{b}_1]_{\mathcal{C}} & [\vec{b}_n]^t_{\mathcal{C}} [\vec{b}_2]_{\mathcal{C}} & \cdots & [\vec{b}_n]^t_{\mathcal{C}} [\vec{b}_n]_{\mathcal{C}} \end{pmatrix} = I_{\dim(V)},$$
es decir, $(P_{\mathcal{C}\leftarrow\mathcal{B}})^t = (P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}})$.
3. Sea $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ es una base ortogonal y $\vec{u} \in V$. Entonces $\vec{u} = \lambda_1\vec{b}_1 + \lambda_2\vec{b}_2 + \cdots + \lambda_n\vec{b}_n$. Tenemos que demostrar que $\lambda_i = \dfrac{\langle\vec{u}, \vec{b}_i\rangle}{\langle\vec{b}_i, \vec{b}_i\rangle}$. Para ello, realicemos el siguiente cálculo:
$$\langle\vec{u}, \vec{b}_1\rangle = \langle\lambda_1\vec{b}_1 + \lambda_2\vec{b}_2 + \cdots + \lambda_n\vec{b}_n, \vec{b}_1\rangle = \lambda_1\langle\vec{b}_1, \vec{b}_1\rangle + \lambda_2\langle\vec{b}_2, \vec{b}_1\rangle + \cdots + \lambda_n\langle\vec{b}_n, \vec{b}_1\rangle = \lambda_1\langle\vec{b}_1, \vec{b}_1\rangle,$$
que despejando, tenemos $\lambda_1 = \dfrac{\langle\vec{u}, \vec{b}_1\rangle}{\langle\vec{b}_1, \vec{b}_1\rangle}$. Repitiendo este argumento para cada $i \in \{1, \ldots, n\}$ se tiene el resultado.

$\square$

Definición. Sea $P \in M_n(\mathbb{R})$. Diremos que P es una **matriz ortogonal** si $P^{-1} = P^t$.

Según el Teorema 4.3, la matriz de cambio de bases entre bases ortonormales es una matriz ortogonal.

Definición. Sea $(V, \langle\cdot\,,\cdot\rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y sean E y F dos subespacios vectoriales. Diremos que E y F son **subespacios vectoriales ortogonales** si $\vec{e} \perp \vec{f}$ para todo $\vec{e} \in E$ y $\vec{f} \in F$, y denotaremos por $E \perp F$.

TEOREMA 4.4. *Sea $(V, \langle \cdot\,, \cdot\rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y sean E y F dos subespacios vectoriales con $S_E = \{\vec{e}_1, \dots, \vec{e}_r\}$ y $S_F = \{\vec{f}_1, \dots, \vec{f}_s\}$ sistemas generadores de E y F respectivamente. Entonces $E \perp F$ si, y solo si, $\vec{e}_i \perp \vec{f}_j$ para todo $i \in \{1, \dots, r\}$ y $j \in \{1, \dots, s\}$.*

DEMOSTRACIÓN.

$\boxed{\Rightarrow}$ Por la propia definición de subespacios vectoriales ortogonales.

$\boxed{\Leftarrow}$ Supongamos que $\vec{e}_i \perp \vec{f}_j$ para todos los posibles i y j, es decir, $\langle \vec{e}_i, \vec{f}_j\rangle = 0$. Sean $\vec{e} \in E$ y $\vec{f} \in F$, queremos demostrar que $\vec{e} \perp \vec{f}$. Como S_E y S_F son sistemas generadores de E y F respectivamente,

$$\vec{e} = \lambda_1\vec{e}_1 + \lambda_2\vec{e}_2 + \cdots + \lambda_r\vec{e}_r \quad \text{y} \quad \vec{f} = \mu_1\vec{f}_1 + \mu_2\vec{f}_2 + \cdots + \mu_s\vec{f}_s.$$

Entonces

$$\begin{aligned}\langle \vec{e}, \vec{f}\rangle &= \langle \lambda_1\vec{e}_1 + \lambda_2\vec{e}_2 + \cdots + \lambda_r\vec{e}_r, \mu_1\vec{f}_1 + \mu_2\vec{f}_2 + \cdots + \mu_s\vec{f}_s\rangle \\ &= \textstyle\sum_{i=1}^{r}\sum_{j=1}^{s} \lambda_i\mu_j\langle \vec{e}_i, \vec{f}_j\rangle = 0.\end{aligned}$$

Demostrando que $E \perp F$.

□

TEOREMA 4.5. *Sea $(V, \langle \cdot\,, \cdot\rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y sean E y F dos subespacios vectoriales de V. Si $E \perp F$ entonces $E \cap F = \{\vec{0}\}$. Por tanto, si $E \perp F$ entonces $E \oplus F$ la suma siempre es directa.*

DEMOSTRACIÓN. Supongamos que $E \perp F$ y sea $\vec{u} \in E \cap F$. Entonces queremos demostrar que $\vec{u} = \vec{0}$. Observar que se tiene que $\vec{u} \perp \vec{u}$, ya que $\vec{u} \in E$ y $\vec{u} \in F$ y $E \perp F$. Así, como $\vec{u} \perp \vec{u}$ significa que $\langle \vec{u}, \vec{u}\rangle = 0$, por ser el producto escalar definido positivo, teniendo que $\vec{u} = \vec{0}$. □

DEFINICIÓN. Sea $(V, \langle \cdot\,, \cdot\rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y sea E un subespacio vectorial. Definimos el **complemento ortogonal de** E como aquel subespacio vectorial F que verifica $E \oplus F = V$ y $E \perp F$, que denotaremos por $E^{\perp}$

A diferencia de un complemento cualquiera de un subespacio vectorial, el complemento ortogonal es único y además verifica la propiedad $(E^{\perp})^{\perp} = E$.

El cálculo del complemento ortogonal es sencillo. Supongamos que $\mathcal{B}_E = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m\}$ es una base de E. Entonces, por el Teorema 4.4, un vector $\vec{e}^{\perp} \in E^{\perp}$ verifica $\vec{e}^{\perp} \perp \vec{e}_i$ para todo $i \in \{1, \dots, m\}$, es decir, debe verificar

$$\left\{\begin{array}{rcl} \langle \vec{e}^{\perp}, \vec{e}_1\rangle &=& 0 \\ \langle \vec{e}^{\perp}, \vec{e}_2\rangle &=& 0 \\ \vdots & & \vdots \\ \langle \vec{e}^{\perp}, \vec{e}_m\rangle &=& 0 \end{array}\right. \iff \left\{\begin{array}{rcl} [\vec{e}^{\perp}]^t_{\mathcal{B}}\, G_{\mathcal{B}}\, [\vec{e}_1]_{\mathcal{B}} &=& 0 \\ [\vec{e}^{\perp}]^t_{\mathcal{B}}\, G_{\mathcal{B}}\, [\vec{e}_2]_{\mathcal{B}} &=& 0 \\ \vdots & & \vdots \\ [\vec{e}^{\perp}]^t_{\mathcal{B}}\, G_{\mathcal{B}}\, [\vec{e}_m]_{\mathcal{B}} &=& 0 \end{array}\right.$$

donde hemos expresado los productos escalares en coordenadas $\mathcal{B}$ utilizando la matriz de Gram $G_{\mathcal{B}}$, proporcionándonos un sistema de ecuaciones lineales homogéneo, que son las ecuaciones implícitas de $E^{\perp}$ en $\mathcal{B}$.

EJEMPLO 42. En $\mathbb{R}^3$, consideremos el plano π con ecuación implicita en $\mathcal{B}_c$ $\pi \equiv \{x-2y+z = 0$. Según el razonamiento anterior, para hallar el complemento ortogonal, que en este caso $\dim(\pi^{\perp}) = \dim(\mathbb{R}^3) - \dim(\pi) = 3 - 2 = 1$, tendremos que hallar primero una base de π.

Como ya viene siendo habitual, una base de $\mathcal{B}_\pi = \left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$. Entonces, si $\vec{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ pertenece a $\pi^\perp$ si, y solo si, es solución de

$$\pi^\perp \equiv \left\{ \begin{array}{rcl} \left\langle \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \right\rangle & = & 0 \\ \left\langle \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\rangle & = & 0 \end{array} \right. \iff \pi^\perp \equiv \left\{ \begin{array}{rrrcl} 2x & +y & & = & 0 \\ -x & & +z & = & 0 \end{array} \right. \begin{array}{l} \text{Ecuaciones} \\ \text{implícitas en } \mathcal{B}_c \end{array}$$

que tiene como base $\mathcal{B}_{\pi^\perp} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}$.

Aprovechando que estamos trabajando en $\mathbb{R}^3$, y queremos calcular un vector ortogonal a un plano, podríamos haber calculado el producto vectorial de dos vectores, no proporcionales, del plano:

$$\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \times \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \det \begin{pmatrix} \vec{i} & \vec{j} & \vec{k} \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \vec{i} - 2\vec{j} + \vec{k} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

Además, posiblemente se haya estudiado que el *vector normal* del plano coincide con los coeficientes de la ecuación del plano. Esto, en general, es cierto para todo subespacio vectorial de dimensión $n-1$. Para verlo, basta con saber que $(\pi^\perp)^\perp = \pi$. Efectivamente, si $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ es el *vector normal* de π, entonces el plano tendrá como ecuación implícita en $\mathcal{B}_c$

$$\left\langle \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \begin{bmatrix} a \\ b \\ c \end{bmatrix} \right\rangle = 0 \iff ax + by + cz = 0,$$

coincidiendo así las coordenadas en $\mathcal{B}_c$ del *vector normal* con los coeficientes de la ecuación implícita de π.

2. Proyección ortogonal

DEFINICIÓN. Sea $(V, \langle \cdot, \cdot \rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y sea E un subespacio vectorial y $E^\perp$ su complemento ortogonal. Como $V = E \oplus E^\perp$, todo $\vec{u} \in V$ se puede expresar de manera única como

$$\vec{u} = \vec{e} + \vec{e}^\perp$$

con $\vec{e} \in E$ y $\vec{e}^\perp \in E^\perp$. Al vector $\vec{e}$ de esta descomposición lo llamaremos **proyección ortogonal de $\vec{u}$ sobre E** y lo denotaremos por $p_E(\vec{u})$. Entonces $p_{E^\perp}(\vec{u}) = \vec{e}^\perp$.

TEOREMA 4.6. *Sea $(V, \langle \cdot\,, \cdot \rangle)$ un $\mathbb{K}$-espacio vectorial euclídeo y sea E un subespacio vectorial con $\mathcal{B}_E = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m\}$ una base ortogonal de E. Entonces*

$$p_E(\vec{u}) = \frac{\langle \vec{u}, \vec{e}_1\rangle}{\langle \vec{e}_1, \vec{e}_1\rangle}\vec{e}_1 + \frac{\langle \vec{u}, \vec{e}_2\rangle}{\langle \vec{e}_2, \vec{e}_2\rangle}\vec{e}_2 + \cdots + \frac{\langle \vec{u}, \vec{e}_m\rangle}{\langle \vec{e}_m, \vec{e}_m\rangle}\vec{e}_m.$$

DEMOSTRACIÓN. $\square$

EJEMPLO 43. En $\mathbb{R}^4$ con el producto escalar usual, consideremos el subespacio vectorial E y $\mathcal{B}_E = \left\{\begin{bmatrix}1\\0\\1\\-1\end{bmatrix}, \begin{bmatrix}1\\-1\\2\\0\end{bmatrix}\right\}$ una base de E. Para hallar la proyección ortogonal de $\vec{u} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$ en E tenemos dos alternativas:

1. Por definición. Para ello debemos calcular primero $E^\perp$, y luego descomponer $\vec{u} = \vec{e} + \vec{e}^\perp$ con $\vec{e} \in E$ y $\vec{e}^\perp \in E^\perp$, y sabiendo que $\vec{e} = p_E(\vec{u})$.

   Primero, calculemos $E^\perp$ mediante sus ecuaciones implícitas en $\mathcal{B}_c$:

$$\left\{\begin{array}{l} \left\langle \begin{bmatrix}x\\y\\z\\t\end{bmatrix}, \begin{bmatrix}1\\0\\1\\-1\end{bmatrix}\right\rangle = 0 \\[2ex] \left\langle \begin{bmatrix}x\\y\\z\\t\end{bmatrix}, \begin{bmatrix}1\\-1\\2\\0\end{bmatrix}\right\rangle = 0 \end{array}\right. \iff \left\{\begin{array}{rrrrl} x & & +z & -t & = 0 \\ x & -y & +2z & & = 0 \end{array}\right.$$

   teniendo $\mathcal{B}_{E^\perp} = \left\{\begin{bmatrix}-1\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\\1\end{bmatrix}\right\}$ es base de $E^\perp$.

   Segundo, hallemos las coordenadas de $\vec{u}$ en la nueva base $\mathcal{B} := \mathcal{B}_E \cup \mathcal{B}_{E^\perp} = \left\{\begin{bmatrix}1\\0\\1\\-1\end{bmatrix}, \begin{bmatrix}1\\-1\\2\\0\end{bmatrix}, \begin{bmatrix}-1\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\\1\end{bmatrix}\right\}$:

$$\left(\begin{array}{cccc|c} \boxed{1} & 1 & -1 & 1 & 1 \\ 0 & -1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 0 & 1 \\ -1 & 0 & 0 & 1 & 1 \end{array}\right) \to \left(\begin{array}{cccc|c} \boxed{1} & 1 & -1 & 1 & 1 \\ 0 & \boxed{-1} & 1 & 1 & 1 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 1 & -1 & 2 & 2 \end{array}\right) \to \left(\begin{array}{cccc|c} \boxed{1} & 1 & -1 & 1 & 1 \\ 0 & \boxed{-1} & 1 & 1 & 1 \\ 0 & 0 & \boxed{3} & 0 & 1 \\ 0 & 0 & 0 & \boxed{3} & 3 \end{array}\right)$$

$$\left(\begin{array}{cccc|c} \boxed{1} & 1 & -1 & 1 & 1 \\ 0 & \boxed{-1} & 1 & 1 & 1 \\ 0 & 0 & \boxed{1} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \boxed{1} & 1 \end{array}\right) \to \left(\begin{array}{cccc|c} \boxed{1} & 1 & -1 & 0 & 0 \\ 0 & \boxed{-1} & 1 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \boxed{1} & 1 \end{array}\right) \to \left(\begin{array}{cccc|c} \boxed{1} & 1 & 0 & 0 & \frac{1}{3} \\ 0 & \boxed{-1} & 0 & 0 & -\frac{1}{3} \\ 0 & 0 & \boxed{1} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \boxed{1} & 1 \end{array}\right)$$

$$\left(\begin{array}{cccc|c} \boxed{1} & 0 & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \boxed{1} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \boxed{1} & 1 \end{array}\right).$$

Donde hemos obtenido que

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix},$$

y por tanto $p_E(\vec{u}) = 0 \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}$. Adicionalmente, hemos hallado $p_{E^\perp}(\vec{u}) = \frac{1}{3} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 \\ 4 \\ 1 \\ 3 \end{bmatrix}$.

2. Por el Teorema 4.6. Necesitaremos una base ortogonal de E.

   Primero, hallemos una base ortogonal de E utilizando el Teorema 4.2:

   a)

$$\vec{c}_1 := \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

   b)

$$\vec{c}_2 := \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} - \frac{\left\langle \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right\rangle} \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} - \frac{3}{3} \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix}$$

Obteniendo que $\mathcal{C}_E = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix} \right\}$ es una base ortogonal de E. Ahora podemos aplicar el Teorema 4.6:

$$\begin{aligned} p_E(\vec{u}) &= \frac{\langle \vec{u},\, \vec{c}_1 \rangle}{\langle \vec{c}_1,\, \vec{c}_1 \rangle}\vec{c}_1 + \frac{\langle \vec{u},\, \vec{c}_2 \rangle}{\langle \vec{c}_2,\, \vec{c}_2 \rangle}\vec{c}_2 = \frac{\left\langle \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right\rangle} \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} + \frac{\left\langle \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix} \right\rangle} \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}. \end{aligned}$$

Ejemplo 44. Sea $(V, \langle \cdot,\, \cdot \rangle)$ un espacio vectorial euclídeo real. Consideremos un subespacio vectorial E con base $\mathcal{B}_E = \{\vec{e}_1,\, \vec{e}_2,\, \ldots,\, \vec{e}_m\}$, y su complemento ortogonal $E^\perp$ con base $\mathcal{B}_{E^\perp} = \{\vec{e}^{\perp}_{m+1},\, \vec{e}^{\perp}_{m+2},\, \ldots,\, \vec{e}^{\perp}_n\}$. Podemos considerar la proyección ortogonal $p_E \colon V \to V$ como transformación lineal. Entonces podemos hallar la matriz asociada a p_E en $\mathcal{B} = \{\vec{e}_1,\, \vec{e}_2,\, \ldots,\, \vec{e}_m,\, \vec{e}^{\perp}_{m+1},\, \vec{e}^{\perp}_{m+2},\, \ldots,\, \vec{e}^{\perp}_n\}$. Para hallar $M(p_E)_{\mathcal{B}\leftarrow\mathcal{B}} = M(p_E(\mathcal{B}))_{\mathcal{B}}$ calculemos primero

$$\begin{aligned} p_E(\mathcal{B}) &= \{p_E(\vec{e}_1),\, p_E(\vec{e}_2),\, \ldots,\, p_E(\vec{e}_m),\, p_E(\vec{e}^{\perp}_{m+1}),\, p_E(\vec{e}^{\perp}_{m+2}),\, \ldots,\, p_E(\vec{e}^{\perp}_n)\} \\ &= \{\vec{e}_1,\, \vec{e}_2,\, \ldots,\, \vec{e}_m,\, \vec{0},\, \vec{0},\, \ldots,\, \vec{0}\}, \end{aligned}$$

entonces

$$M(p_E)_{\mathcal{B}\leftarrow\mathcal{B}} = \left(\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right) = \left(\begin{array}{c|c} I_m & O_{m\times(n-m)} \\ \hline O_{(n-m)\times m} & O_m \end{array}\right)$$

2.1. Solución de mínimos cuadrados. El siguiente teorema nos asegura que dado un vector $\vec{u} \in V$ y un subespacio vectorial E, el vector más "cercano" a $\vec{u}$ en E es la proyección ortogonal.

Teorema 4.7. *Sea $(V, \langle \cdot\,, \cdot \rangle)$ un $\mathbb{K}$-espacio vectorial euclídeo y sea E un subespacio vectorial. Si $\vec{u} \in V$ entonces*

$$\|\vec{u} - p_E(\vec{u})\| \leqslant \|\vec{u} - \vec{e}\|$$

para todo $\vec{e} \in E$.

DEMOSTRACIÓN. $\square$

En virtud de este teorema, podremos proponer la idea de resolver un sistema de ecuaciones lineales incompatible.

Sean $A \in M_{m\times n}(\mathbb{K})$, $X \in \mathbb{R}^n$ y $B \in \mathbb{R}^m$. Supongamos que el sistema de ecuaciones $AX = B$ es incompatible. Esto se puede interpretar de manera equivalente como $B \notin \mathrm{Col}(A)$. De hecho, $AX = B$ sólo tiene solución si $B \in \mathrm{Col}(A)$. Sin embargo, puede seguir siendo interesante tener "algo" que sea lo más parecido a una solución. Por ello, si pensamos en el espacio vectorial euclídeo $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$, la proyección ortogonal de B sobre el subespacio vectorial $\mathrm{Col}(A)$ es la solución del sistema más "cercana" a B. A esta solución la llamaremos la **solución de mínimos cuadrados**. Es decir, si $AX = B$ es un sistema incompatible, entonces $p_{\mathrm{Col}(A)}(B)$ es la solución de mínimos cuadrados del sistema. Llamaremos el **error cuadrático** a $\|B - p_{\mathrm{Col}(A)}(B)\|^2$.

Aunque ya hemos visto como calcular la proyección ortogonal de un vector sobre un subespacio, este problema se puede resolver de una forma alternativa. Esta otra forma parte del siguiente hecho: $(B - p_{\mathrm{Col}(A)}(B)) \perp \mathrm{Col}(A)$.

Sea $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ el $\mathbb{R}$-espacio vectorial euclídeo y $\mathcal{B}$ una base. Como cualquier vector de $\mathrm{Col}(A) \subset \mathbb{R}^n$ se puede expresar como AY para algún $Y \in \mathbb{R}^m$ y, por otro lado, existe $X_p \in \mathbb{R}^m$ tal que $AX_p = p_{\mathrm{Col}(A)}(B)$, entonces:

$$0 = \langle AY, B - AX_p \rangle = (AY)^t G_{\mathcal{B}} (B - AX_p) = Y^t A^t G_{\mathcal{B}} B - Y^t A^t G_{\mathcal{B}} A X_p,$$

es decir, para todo $Y \in \mathbb{R}^m$,

$$Y^t A^t G_{\mathcal{B}} B = Y^t A^t G_{\mathcal{B}} A X_p.$$

Como es cierto para cualquier vector $Y \in \mathbb{R}^m$, podemos ver X_p como la solución al sistema

$$A^t G_{\mathcal{B}} B = A^t G_{\mathcal{B}} A X.$$

EJEMPLO 45. Hallemos la parábola $y = c + bx + ax^2$ que "mejor" se aproxime a los puntos del plano $A = (0, 1)$, $B = (1, 2)$, $C = (2, 0)$ y $D = (3, -1)$. Lo ideal sería que existiera tal parábola, pero en caso de que no, calcularemos la parábola que minimice el error cuadrático respecto al producto escalar usual. Para ello, se plantea el sistema:

$$\begin{array}{r} A\colon x = 0,\ y = 1 \\ B\colon x = 1,\ y = 2 \\ C\colon x = 2,\ y = 0 \\ D\colon x = 3,\ y = -1 \end{array} \left\{ \begin{array}{rrrcr} & & c & = & 1 \\ a & +b & +c & = & 2 \\ 4a & +2b & +c & = & 0 \\ 9a & +3b & +c & = & -1 \end{array} \right. \iff \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix},$$

que se puede comprobar que este sistema es incompatible. Entonces, planteemos el nuevo sistema:

$$\begin{pmatrix} 0 & 1 & 4 & 9 \\ 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{pmatrix} 0 & 1 & 4 & 9 \\ 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix} \iff \begin{pmatrix} 98 & 36 & 14 \\ 36 & 14 & 6 \\ 14 & 6 & 4 \end{pmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -7 \\ -1 \\ 2 \end{bmatrix}$$

que tiene como solución $a = -\dfrac{1}{2}$, $b = \dfrac{7}{10}$ y $c = \dfrac{6}{5}$. Entonces la parábola $y = \frac{6}{5} + \frac{7}{10}x - \frac{1}{2}x^2$ es la parábola que minimiza el error cuadrático para los puntos anteriores. El error que se ha cometido es

$$\varepsilon^2 = \left\| \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} \begin{bmatrix} -\frac{1}{2} \\ \frac{7}{10} \\ \frac{6}{5} \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} \frac{6}{5} \\ \frac{7}{5} \\ \frac{3}{5} \\ -\frac{6}{5} \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix} \right\|^2 = \frac{1}{5}^2 + \left(-\frac{3}{5}\right)^2 + \frac{3}{5}^2 + \left(-\frac{1}{5}\right)^2,$$

es decir, $\varepsilon^2 = \frac{20}{25} = \frac{4}{5}$.

3. Diagonalización por congruencia

Dada $A \in M_n(\mathbb{R})$ queremos estudiar cuando es congruente con una matriz $D \in M_n(\mathbb{R})$ diagonal, es decir, cuando existe $P \in M_n(\mathbb{R})$ invertible tal que $A = P^tDP$. Como D es diagonal, en particular es simétrica, y esto implica que A necesariamente debe ser simétrica, ya que $A^t = (P^tDP)^t = P^tD^t(P^t)^t = P^tDP = A$.

A diferencia de la diagonalización por semejanza, este problema no resulta complicado de resolver. Tan solo es suficiente con recordar que realizar transformaciones elementales por filas a A es multiplicar por la izquierda por sus respectivas matrices elementales, y realizar las mismas transformaciones elementales por columnas corresponde con multiplicar por la derecha por la traspuesta de las matrices elementales. Veamos esto en un ejemplo:

Ejemplo 46. Sea $A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 1 \end{pmatrix}$, entonces

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 4 & 1 & 0 & 0 \\ 2 & 1 & 2 & 0 & 1 & 0 \\ 4 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow[F_3^N = F_3 - 4F_1]{F_2^N = F_2 - 2F_1} \left(\begin{array}{ccc|ccc} 1 & 2 & 4 & 1 & 0 & 0 \\ 0 & -3 & -6 & -2 & 1 & 0 \\ 0 & -6 & -15 & -4 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[C_3^N = C_3 - 4C_1]{C_2^N = C_2 - 2C_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -3 & -6 & -2 & 1 & 0 \\ 0 & -6 & -15 & -4 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{F_3^N = F_3 - 2F_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -3 & -6 & -2 & 1 & 0 \\ 0 & 0 & -3 & 0 & -2 & 1 \end{array}\right)$$

$$\xrightarrow{C_3^N = C_3 - 2C_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -3 & 0 & -2 & 1 & 0 \\ 0 & 0 & -3 & 0 & -2 & 1 \end{array}\right)$$

teniendo que

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

es decir, A es congruente con la matriz diagonal $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}$.

Esta diagonalización resulta interesante para la clasificación de matrices (semi)definidas positivas, negativas o indefinidas, que a su vez son utilizadas para clasificar puntos críticos de funciones en varias variables como máximos o mínimos.

4. Diagonalización por congruencia ortogonal

Por último, estudiemos cuándo $A \in M_n(\mathbb{R})$ es congruente con $D \in M_n(\mathbb{R})$ donde, adicionalmente, $P \in M_n(\mathbb{R})$ tal que $A = P^t DP$ es ortogonal, es decir,

$$A = P^t DP \quad \text{con} \quad P^{-1} = P^t.$$

Como ya hemos visto antes, A necesariamente debe ser simétrica. Sin embargo, este problema resulta más complicado que el caso no ortogonal. Esto sucede por su parecido con el problema de la diagonalización por semejanza, donde se tuvo que resolver un problema de autovalores y autovectores.

A pesar de ello, ya hemos visto que para tener la diagonalización por semejanza, necesitamos una base de autovectores, y por otro lado, para que la matriz de cambio de base sea una matriz ortogonal, la matriz de cambio de bases debe ser entre bases ortonormales. Dando pie a la siguiente definición.

Definición. Sea $(V, \langle\cdot\,,\cdot\rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y $T\colon V \to V$ una transformación lineal. Diremos que T es **diagonalizable por congruencia ortogonal** si existe una base de autovectores ortonormal.

De esta forma, partiremos de una matriz asociada a $T\colon V \to V$ en una base ortonormal $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$, para poder asegurar que la matriz de cambio de base sea ortogonal, que tendrá que ser también simétrica, como condición necesaria para ser congruente con una matriz diagonal. Ahora, utilizando los coeficientes de Fourier:

$$M(T)_{\mathcal{B}\leftarrow\mathcal{B}} = \left(\ [T(\vec{b}_1)]_{\mathcal{B}} \;\middle|\; [T(\vec{b}_2)]_{\mathcal{B}} \;\middle|\; \cdots \;\middle|\; [T(\vec{b}_n]_{\mathcal{B}}\ \right) = \begin{pmatrix} \langle T(\vec{b}_1), \vec{b}_1\rangle & \langle T(\vec{b}_2), \vec{b}_1\rangle & \cdots & \langle T(\vec{b}_n), \vec{b}_1\rangle \\ \langle T(\vec{b}_1), \vec{b}_2\rangle & \langle T(\vec{b}_2), \vec{b}_2\rangle & \cdots & \langle T(\vec{b}_n), \vec{b}_2\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle T(\vec{b}_1), \vec{b}_n\rangle & \langle T(\vec{b}_2), \vec{b}_n\rangle & \cdots & \langle T(\vec{b}_n), \vec{b}_n\rangle \end{pmatrix},$$

y sabiendo que tanto la matriz es simétrica como el producto escalar, tendremos que, para todo $i, j \in \{1, \ldots, n\}$.

$$\langle T(\vec{b}_i), \vec{b}_j\rangle = \langle \vec{b}_i, T(\vec{b}_j)\rangle.$$

De hecho, es una condición suficiente y necesaria para que $M(T)_{\mathcal{B}\leftarrow\mathcal{B}}$ sea simétrica. Si T verifica la propiedad anterior para todo $\vec{u}, \vec{v} \in V$, se dirá que es una **transformación autoadjunta**.

Teorema 4.8. *Sea $(V, \langle\cdot\,,\cdot\rangle)$ un $\mathbb{R}$-espacio vectorial euclídeo y $T\colon V \to V$ una transformación lineal autoadjunta. Si λ y μ son dos autovalores distintos de T entonces $V_\lambda \perp V_\mu$.*

Demostración. Supongamos que $\lambda \neq \mu$. Si $\vec{u} \in V_\lambda$ y $\vec{v} \in V_\mu$ entonces

$$\lambda\langle \vec{u}, \vec{v}\rangle = \langle \lambda\vec{u}, \vec{v}\rangle = \langle T(\vec{u}), \vec{v}\rangle = \langle \vec{u}, T(\vec{v})\rangle = \langle \vec{u}, \mu\vec{v}\rangle = \mu\langle \vec{u}, \vec{v}\rangle,$$

que es equivalente a

$$(\lambda - \mu)\langle \vec{u}, \vec{v}\rangle = 0 \quad\Longleftrightarrow\quad \langle \vec{u}, \vec{v}\rangle = 0,$$

por que $\lambda - \mu \neq 0$. Teniendo que $V_\lambda \perp V_\mu$. □

Este teorema nos asegura entonces que si una transformación lineal autoadjunta $T\colon V \to V$ es diagonalizable por semejanza, entonces, junto con el Teorema 4.2, se puede hallar una base de autovectores ortonormal. Teniendo como resultado que si una transformación es diagonalizable por semejanza entonces es diagonalizable por congruencia ortogonal, y esta sería una manera de hallar dicha base de autovectores ortonormal.

EJEMPLO 47. Consideremos $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ y diagonalicemosla por congruencia ortogonal, es decir, hallemos $P \in M_3(\mathbb{R})$ ortogonal y $D \in M_3(\mathbb{R})$ tales que $A = PDP^t$ con $P^t = P^{-1}$.

Como ya hemos razonado previamente, debemos hallar una base ortonormal de autovectores. El primer paso será hallar la base de autovectores, y luego, por el Teorema 4.8 y el Teorema 4.2, podremos ortonormalizarla asegurando que no pierdan la propiedad de ser autovectores.

Primero, hallemos el polinomio característico de A:

$$\begin{aligned} p(\lambda) &= \det \begin{pmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix} \overset{C_3^N = C_3 - C_1}{=\!=\!=\!=} \det \begin{pmatrix} -\lambda & 1 & 1+\lambda \\ 1 & -\lambda & 0 \\ 1 & 1 & -1-\lambda \end{pmatrix} \\ \overset{F_1^N = F_1 + F_3}{=\!=\!=\!=} & \det \begin{pmatrix} 1-\lambda & 2 & 0 \\ 1 & -\lambda & 0 \\ 1 & 1 & -1-\lambda \end{pmatrix} = (-1-\lambda) \det \begin{pmatrix} 1-\lambda & 2 \\ 1 & -\lambda \end{pmatrix} \\ &= (-1-\lambda)((1-\lambda)(-\lambda) - 2) = (-1-\lambda)(\lambda^2 - \lambda - 2) \\ &= (-1-\lambda)^2(2-\lambda). \end{aligned}$$

Los autovalores son $\lambda_1 = -1$ con m.a.$(-1) = 2$ y $\lambda_2 = 2$ con m.a.$(2) = 1$. Hallemos los subespacios propios

$\boxed{V_{-1}}$

$$A - (-1)I_3 = A + I_3 = \begin{pmatrix} \boxed{1} & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Y por tanto $\mathcal{B}_{V_{-1}} = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ es una base de autovectores asociados a $\lambda_1 = -1$. Aprovechemos para hallar una base ortogonal aplicando el Teorema 4.2:

1.
$$\vec{c}_1 := \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

2.

$$\vec{c}_2 := \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{\left\langle \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\rangle} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

teniendo que $\mathcal{C}_{V_{-1}} = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \right\}$ es una base de autovectores ortonormal de V_{-1}.

$\boxed{V_2}$

$$A - 2I_3 = \begin{pmatrix} -2 & 1 & 1 \\ \boxed{1} & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & -2 & 1 \\ 0 & \boxed{-3} & 3 \\ 0 & 3 & -3 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & -2 & 1 \\ 0 & \boxed{1} & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

teniendo que $\mathcal{B}_{V_2} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ es una base (trivialmente ortogonal) de V_2. Y, normalizándola, $\mathcal{C}_{V_2} = \left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ es una base ortonormal de V_2.

Juntando ambas bases de autovectores, y en virtud del Teorema 4.8,

$$\mathcal{D} = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

es una base ortonormal de autovectores de A. Construyendo entonces las matrices

$$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{y} \quad P = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix},$$

que verifican

$$A = PDP^{-1} \quad \text{con} \quad P^t = P^{-1}.$$

Ya hemos visto que si una matriz simétrica es diagonalizable por semejanza, entonces es diagonalizable por congruencia ortogonal, pues siempre podremos asegurar hallar una base ortonormal de autovectores. Pero podemos decir más, y es que resulta que toda transformación lineal autoadjunta $T: V \to V$ es *siempre* diagonalizable por congruencia ortogonal.

TEOREMA 4.9 (Teorema espectral para transformaciones). *Sea* $(V, \langle \cdot, \cdot \rangle)$ *un* $\mathbb{R}$*-espacio vectorial euclídeo y* $T: V \to V$ *una transformación lineal autoadjunta. Entonces* T *es diagonalizable por congruencia ortogonal.*

TEOREMA 4.10 (Teorema espectral para matrices). *Sea* $A \in M_n(\mathbb{R})$ *simétrica. Entonces existe* $D \in M_n(\mathbb{R})$ *diagonal y* $P \in M_n(\mathbb{R})$ *ortogonal tal que* $A = PDP^t$.