.UBA fiuba FACULTAD DE INGENIERÍA

150 ING el país celebra su ingeniería 1870-2020

Apuntes para la práctica de Álgebra II

Sugerencias, desarrollos teóricos y ejercicios resueltos

Maximiliano Contino

2021

Resumen

Este apunte tiene por objetivo cubrir todos los temas de la práctica correspondiente a la asignatura Álgebra II de la Facultad de Ingeniería de la Universidad de Buenos Aires. En cada capítulo, se hace un repaso de los contenidos teorícos mínimos para resolver los distintos ejercicios haciendo hincapié en las respectivas demostraciones. En varios casos, se agregaron aplicaciones teóricas que amplian el alcance de algunos conceptos de la materia.

Los ejercicios específicos que corresponden a la guía de Álgebra II tienen la misma numeración que en dicha guía. Como referencia se tomó la versión de la guía de marzo de 2021 aunque con el paso de los cuatrimestres dichos ejercicios podrían cambiar de numeración o desaparecer. También, se agregaron ejercicios similares a los que figuran en la guía o ejercicios de parciales y finales con el objetivo de profundizar en algunos conceptos determinados, dichos ejercicios aparecen ordenados con letras mayúsculas.

En el Capítulo 0 se comienza con las definiciones básicas y la notación que se usará a lo largo de todo el apunte. En el Capítulo 1, estudiamos las nociones de espacio vectorial, subespacios, sistema de generadores e independencia lineal. En el Capítulo 2, se definen y estudian las transformaciones lineales y se comienza con el estudio de ecuaciones diferenciales a coeficientes constantes. En el Capítulo 3, seguimos con las nociones de espacios con producto interno reales y complejos y las proyecciones ortogonales. La diagonalización de matrices, las formas de Jordan en espacios de dimensión finita y los sistemas de ecuaciones diferenciales se desarrollan a continuación en el Capítulo 4 seguidas por el estudio de matrices hermíticas, matrices simétricas y la descomposición en valores singulares de matrices en el Capítulo 5. Concluimos con el estudio de formas cuadráticas en el Capítulo 6.

Siempre serán bienvenidas todas las correcciones y aportes que permitan mejorar este apunte.

Maximiliano Contino

Índice general

CAPÍTULO 0

Introducción

En todo el apunte usaremos la siguiente notación y propiedades básicas:

Conjuntos de números

$\mathbb{N}$ denota el conjunto de todos los números naturales, $\mathbb{N} = \{1, 2, \cdots\}$, $\mathbb{Z}$ el conjunto de los números enteros, $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$, $\mathbb{Q}$ denota el conjunto de números racionales, es decir

$$\mathbb{Q} = \{\frac{m}{n} : m \in \mathbb{Z}, n \in \mathbb{N}\}.$$

Finalmente $\mathbb{R}$ denota el conjunto de números reales y $\mathbb{C}$ el conjunto de los números complejos.

Por definición $z \in \mathbb{C}$ si $z = a + ib$ con $a, b \in \mathbb{R}$ e $i^2 = -1$. Si $z = a + ib \in \mathbb{C}$ entonces $\overline{z} := a - ib$ es el conjugado de z. Observar que si $z, y \in \mathbb{C}$ entonces

$$\overline{\overline{z}} = z, \quad \overline{z+y} = \overline{z} + \overline{y}, \quad \overline{zy} = \overline{z}\ \overline{y} \text{ y}$$

$$\overline{z} = z, \text{ si y sólo si } z \in \mathbb{R}.$$

Además, si $z = a + ib \in \mathbb{C}$ tenemos que

$$z + \overline{z} = 2a = 2\operatorname{Re}(z) \text{ y } z - \overline{z} = i2b = 2i\operatorname{Im}(z).$$

La *fórmula de Euler* establece que para cada $z = a + ib \in \mathbb{C}$ vale que

$$e^z = e^{a+ib} = e^a(\cos(b) + i\sin(b)).$$

Matrices, vectores, funciones continuas y polinomios

Sea $\mathbb{K} = \mathbb{C}$ ó $\mathbb{K} = \mathbb{R}$. Entonces $\mathbb{K}^{m\times n}$ es el conjunto de todas las matrices de m filas y n columnas con entradas (coeficientes) en $\mathbb{K}$. Usaremos el símbolo I_n (ó cuando se infiera el contexto simplemente I) para denotar a la matriz *identidad* de $n \times n$. Dicha matriz $I_n \in \mathbb{K}^{n\times n}$ se define como

$$I_n := \begin{bmatrix} 1 & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & 1 \end{bmatrix}.$$

Dada $A \in \mathbb{K}^{m\times n}$ se denota por $A^T \in \mathbb{K}^{n\times m}$ a su matriz transpuesta. Si

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

con $a_{ij} \in \mathbb{K}$ para $i \in \{1, 2, \cdots, m\}$ y $j \in \{1, 2, \cdots, n\}$. Entonces

$$A^T := \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}.$$

Dada $A \in \mathbb{K}^{n\times n}$, su *traza* se define como

$$\operatorname{tr}(A) := a_{11} + \cdots + a_{nn} \in \mathbb{K},$$

donde $a_{ii} \in \mathbb{K}$ son los elementos de la diagonal de A.

Dada $A \in \mathbb{K}^{m\times n}$, $A^* \in \mathbb{K}^{n\times n}$ es la matriz traspuesta conjugada de A. Es decir, $A^* := \overline{A^T}$.

$\mathbb{K}^n$ denota el conjunto de las matrices de $n \times 1$ o de vectores (columnas) de $\mathbb{K}^n$. Entonces si $x \in \mathbb{K}^n$, tenemos que

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1 \ x_2 \ \cdots \ x_n]^T$$

con $x_1, x_2, \cdots, x_n \in \mathbb{K}$.

El conjunto $\mathcal{E} := \{e_1, \cdots, e_n\} \subseteq \mathbb{K}^n$ cuyos elementos se definen por $e_1 = \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{bmatrix}, \cdots, e_n = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{bmatrix}$, se denomina la *base canónica* de $\mathbb{K}^n$.

Dado un intervalo $I \subseteq \mathbb{R}$,

$$C(I) := \{f : I \to \mathbb{R} : f \text{ es continua}\},$$

y para cada $n \in \mathbb{N}$, $C^n(I)$ denota el conjunto de las funciones de $C(I)$ que son n veces derivables y cuyas derivadas son continuas. $C^\infty(I)$ denota el conjunto de todas las funciones infinitamente derivables.

$\mathbb{K}[x]$ denota el conjunto de polinomios con coeficientes en $\mathbb{K}$. Es decir,

$$\mathbb{K}[x] := \{p : p(x) = a_n x^n + \cdots + a_1 x + a_0 : n \in \mathbb{N},\ a_0, a_1, \cdots, a_n \in \mathbb{K}\}.$$

Finalmente, dado $n \in \mathbb{N}$, $\mathbb{K}_n[x]$ denota el conjunto de polinomios de $\mathbb{K}$ con grado menor o igual a n unión el polinomio nulo $0(x) := 0$. Es decir,

$$\mathbb{K}_n[x] = \{p \in \mathbb{K}[x] : \operatorname{grado}(p) \leq n\} \cup \{0\}.$$

Conjuntos en general

Sean A, B dos conjuntos, las siguientes son algunas operaciones básicas que usaremos:

- $A \cap B := \{x : x \in A \text{ y } x \in B\}$.
- $A \cup B := \{x : x \in A \text{ ó } x \in B\}$.
- $A \setminus B := \{x : x \in A \text{ y } x \notin B\}$.
- $A \times B := \{(a, b) : a \in A, b \in B\}$.

Sean A, B dos conjuntos, las siguientes son algunas relaciones básicas entres A y B

- $A \subseteq B$, si para cada $x \in A$ vale que $x \in B$.
- $A \subsetneq B$, si para cada $x \in A$ vale que $x \in B$ y existe $z \in B$ tal que $z \notin A$.
- $A = B$, si $A \subseteq B$ y $B \subseteq A$.

Observar que $A \subseteq B$ si y sólo si $A \cap B = A$ y, si y sólo si $A \cup B = B$.

Denotaremos por $\emptyset$ al conjunto que no tiene ningún elemento. Entonces, es fácil ver que $A \cap \emptyset = \emptyset$ y $A \cup \emptyset = A$.

Funciones en general

Sean X, Y dos conjuntos. Una *función* $f : X \to Y$ (de X a Y), es una relación entre X e Y que a cada elemento $x \in X$ le asigna un único elemento del conjunto Y que se denota $f(x)$.

Un ejemplo de función, es la *función identidad*, definida por $id_X : X \to X$,

$$id_X(x) := x.$$

Sea $f : X \to Y$ una función de X a Y. Entonces, diremos que:

- f es *inyectiva* si $f(x_1) = f(x_2)$ implica que $x_1 = x_2$ para cualquier $x_1, x_2 \in X$.
- f es *sobreyectiva* si para todo $y \in Y$ existe $x \in X$ tal que $y = f(x)$.
- f es *biyectiva* si f es inyectiva y sobreyectiva.

Observar que si f es biyectiva, entonces podemos definir la relación $f^{-1} : Y \to X$, como $f^{-1}(y) := x$, donde para cada $y \in Y$, x es el único elemento de X tal que $f(x) = y$. En este caso, f^{-1} es una función, pues como f es sobreyectiva, para cada elemento de Y existe un elemento de $x \in X$ tal que $f(x) = y$. Además como f es inyectiva, el elemento $x \in X$ es único.

Sean X, Y, Z conjuntos. $f : X \to Y$ y $g : Y \to Z$ dos funciones. La *composición* de g con f es la relación $g \circ f : X \to Z$, definida por

$$g \circ f(x) := g(f(x)),$$

para cada $x \in X$. Es fácil ver que, la composición de funciones, también es una función.

CAPÍTULO 1

Espacios Vectoriales

1.1. Espacios vectoriales

En este capítulo vamos a estudiar un tipo de conjunto que nos acompañará a lo largo de toda la matería. Estos conjuntos que llamaremos espacios vectoriales tiene la característica de que en ellos podemos definir dos operaciones: la suma y el producto por escalares. Veremos varios resultados relevantes y resolveremos ejercicios correspondientes a la Guía 1, así como también ejercicios de examenes y ejercicios relacionados.

Para definir el concepto de espacio vectorial, primero vamos a definir lo que entendemos por cuerpo.

Definición. Sea $\mathbb{K}$ un conjunto no vacío y sean $+$ y $\cdot$ dos operaciones definidas en $\mathbb{K}$. Se dice que $(\mathbb{K}, +, \cdot)$ es un *cuerpo* si:

1. $\mathbb{K}$ es *cerrado* para las operaciones $+$ y $\cdot$, es decir $a + b \in \mathbb{K}$ y $a \cdot b \in \mathbb{K}$ para todo $a, b \in \mathbb{K}$.

2. Las operaciones $+$ y $\cdot$ son *asociativas*, es decir $(a + b) + c = a + (b + c)$ y $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ para todo $a, b, c \in \mathbb{K}$.

3. Las operaciones $+$ y $\cdot$ son *conmutativas*, es decir $a + b = b + a$ y $a \cdot b = b \cdot a$ para todo $a, b \in \mathbb{K}$.

4. Las operaciones $+$ y $\cdot$ tienen *elemento neutro*, es decir existe un elemento que llamaremos $0 \in \mathbb{K}$ tal que $a + 0 = 0 + a = a$ para todo $a \in \mathbb{K}$ y existe un elemento que llamaremos $1 \in \mathbb{K}$ tal que $a \cdot 1 = 1 \cdot a = a$ para todo $a \in \mathbb{K}$.

5. Cada $a \in \mathbb{K}$ tiene un elemento inverso respecto de la operación $+$ y cada $a \in \mathbb{K} \setminus \{0\}$ tiene un elemento inverso respecto de la operación $\cdot$. Es decir, para cada $a \in \mathbb{K}$ existe un elemento que llamaeremos $-a \in \mathbb{K}$ tal que $a + (-a) = (-a) + a = 0$, donde 0 es el elemento neutro para $+$ y para cada $a \in \mathbb{K} \setminus \{0\}$ existe un elemento neutro que llamaremos $a^{-1} \in \mathbb{K}$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = 1$, donde 1 es el elemento neutro para $\cdot$.

6. Valen las propiedades distributivas: $a \cdot (b + c) = a \cdot b + a \cdot c$ y $(b + c) \cdot a = b \cdot a + c \cdot a$ para todo $a, b, c \in \mathbb{K}$.

Ejemplo 1.a.

- Si $+$ y $\cdot$ representan la suma y multiplicación usual de números, entonces los conjuntos $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ y $(\mathbb{C}, +, \cdot)$, son cuerpos.

- Consideremos el conjunto de dos elementos $\{0, 1\}$, definamos $+$ como $0 + 0 = 0, 0 + 1 = 1$, $1 + 0 = 1$ y $1 + 1 = 0$ y definamos $\cdot$ como $0 \cdot 0 = 0, 0 \cdot 1 = 0$, $1 \cdot 0 = 0$ y $1 \cdot 1 = 1$. Entonces $(\{0, 1\}, +, \cdot)$ es un cuerpo. **Si no queda claro, demostrarlo como ejercicio**.

Ejercicio 1.A. Sean $+$ y $\cdot$ la suma y multiplicación usual de números y sea $F := \{r + s\sqrt{3} : r, s \in \mathbb{Q}\}$. Entonces $(F, +, \cdot)$ es un cuerpo.

Dem. Probemos los axiomas de cuerpo. Claramente, F es un conjunto no vacío. Por otra parte, si $a, b \in F$ entonces $a = r_1 + s_1\sqrt{3}$ y $b = r_2 + s_2\sqrt{3}$ para ciertos $r_1, s_1, r_2, s_2 \in \mathbb{Q}$. Entonces

$$a + b = r_1 + s_1\sqrt{3} + r_2 + s_2\sqrt{3} = (r_1 + s_1) + \sqrt{3}(r_2 + s_2) \in F,$$

pues $r_1 + s_1 \in Q$ y $r_2 + s_2 \in \mathbb{Q}$ y además,

$$\begin{aligned} a \cdot b &= (r_1 + s_1\sqrt{3}) \cdot (r_2 + s_2\sqrt{3}) = r_1 \cdot r_2 + 3s_1 \cdot s_2 + r_1 \cdot s_2\sqrt{3} + s_1 \cdot r_2\sqrt{3} \\ &= (r_1 \cdot r_2 + 3s_1 \cdot s_2) + (r_1 \cdot s_2 + r_1 \cdot r_2)\sqrt{3} \in F, \end{aligned}$$

pues $= r_1 \cdot r_2 + 3s_1 \cdot s_2 \in \mathbb{Q}$ y $r_1 \cdot s_2 + r_1 \cdot r_2 \in \mathbb{Q}$. Por lo tanto, $\mathbb{K}$ es cerrado para las operaciones $+$ y $\cdot$

Como $F \subseteq \mathbb{R}$, $0 = 0 + \sqrt{3}\,0 \in F$ y $1 = 1 + 0\,\sqrt{3} \in F$. Es claro que las operaciones $+$ y $\cdot$ son asociativas, conmutativas, 0 es un elemento neutro para $+$, 1 es un elemento neutro para $\cdot$ y valen las propiedades distributivas. Por otra parte, si $a \in F$, entonces $-a := (-1)a \in F$ y $-a$ es un inverso para $+$. De hecho, si $a \in F$ entonces $a = r + s\sqrt{3}$ para ciertos $r, s \in Q$. Entonces, $-a = (-1) \cdot a = (-r) + (-s)\sqrt{3} \in F$, pues $-r, -s \in \mathbb{Q}$ y claramente $a + -a = 0$.

Sólo nos resta probar que todo elemento no nulo de F tiene inverso para $\cdot$. Sea $a \in F \setminus \{0\}$. Entonces $a = r + s\sqrt{3}$, para cierto $r, s \in \mathbb{Q}$. Entonces $(r + s\sqrt{3}) \cdot (r - s\sqrt{3}) = r^2 - 3s^2 \neq 0$. De hecho, si fuera que $r^2 - 3s^2 = 0$, entonces $|r| = \sqrt{3}|s|$. Pero como $s \in \mathbb{Q}$, entonces $r \notin \mathbb{Q}$ (número irracional multiplicado por número racional) lo cual es absurdo pues $r \in \mathbb{Q}$. Entonces, como $r^2 - 3s^2 \neq 0$ se sigue que $(r + s\sqrt{3})^{-1} = \frac{r}{r^2-3s^2} - \frac{s}{r^2-3s^2}\sqrt{3} \in F$ porque $\frac{r}{r^2-3s^2} \in \mathbb{Q}$ y $\frac{s}{r^2-3s^2} \in \mathbb{Q}$. Claramente $a^{-1} := (r + s\sqrt{3})^{-1} = \frac{r}{r^2-3s^2} - \frac{s}{r^2-3s^2}\sqrt{3} \in F$ es un inverso de a para $\cdot$. ■

De ahora en más, usaremos los cuerpos de escalares $(\mathbb{R}, +, \cdot)$ y $(\mathbb{C}, +, \cdot)$ con las suma y producto usuales. En general, cuando no queramos diferenciar entre el cuerpo real o complejo, directamente usaremos la notación $\mathbb{K}$ para referirnos a los mismos.

Definición. Sea $\mathbb{K}$ un cuerpo ($\mathbb{K} = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$). Sea $\mathbb{V}$ un conjunto no vacío donde se define una operación $+$ llamada *suma* entre dos elementos de $\mathbb{V}$ y una acción $\cdot$ llamada *producto por escalares* entre un elemento de $\mathbb{V}$ y un escalar del cuerpo $\mathbb{K}$. Diremos que $(\mathbb{V}, +, \mathbb{K}, \cdot)$ es un $\mathbb{K}$-*espacio vectorial* si:

1. La suma $+$ es una *ley de composición interna*, es decir $u + v \in \mathbb{V}$ para todo $u, v \in \mathbb{V}$.
2. La suma $+$ es *conmutativa*, es decir $u + v = v + u$ para todo $u, v \in \mathbb{V}$.
3. La suma $+$ es *asociativa*, es decir $(u + v) + w = u + (v + w)$ para todo $u, v, w \in \mathbb{V}$.
4. La suma $+$ tiene *elemento neutro*, es decir existe un elemento que llamaremos $0_{\mathbb{V}} \in \mathbb{V}$ tal que $u + 0_{\mathbb{V}} = 0_{\mathbb{V}} + u = u$ para todo $u \in \mathbb{V}$.
5. Para cada $u \in \mathbb{V}$ existe un elemento inverso respecto de la suma $+$. Es decir, para cada $u \in \mathbb{V}$ existe un elemento que llamaremos $-u \in \mathbb{V}$ tal que $u + (-u) = (-u) + u = 0_{\mathbb{V}}$, donde $0_{\mathbb{V}}$ es un elemento neutro de la suma.
6. El producto por escalares $\cdot$ es una *ley de composición externa*, es decir $\alpha \cdot v \in \mathbb{V}$ para todo $\alpha \in \mathbb{K}$ y para todo $v \in \mathbb{V}$.
7. El producto por escalares $\cdot$ es *distributivo respecto de los escalares*, es decir $(\alpha+\beta)\cdot u = \alpha\cdot u+\beta\cdot u$ para todo $\alpha, \beta \in \mathbb{K}$ y para todo $u \in \mathbb{V}$.
8. El producto por escalares $\cdot$ es *distributivo respecto de la suma de vectores*, es decir $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$ para todo $\alpha \in \mathbb{K}$ y para todo $u, v \in \mathbb{V}$.
9. El producto por escalares $\cdot$ es *asociativo respecto de los escalares*, es decir $(\alpha\beta) \cdot u = \alpha \cdot (\beta \cdot v)$ para todo $\alpha, \beta \in \mathbb{K}$ y para todo $u \in \mathbb{V}$.

10. El escalar $1 \in \mathbb{K}$ cumple que $1 \cdot u = u$ para todo $u \in \mathbb{V}$.

Ejercicio 1.B. Sea $I \subseteq \mathbb{R}$ un intervalo y $C(I) = \{f : I \to \mathbb{R} : f \text{ es continua}\}$. En $C(I)$ definimos la operación suma $+$ como

$$(f+g)(x) = f(x) + g(x) \text{ para cada } f, g \in C(I) \text{ y } x \in I.$$

Para el cuerpo de escalares $\mathbb{K} = \mathbb{R}$ definimos la acción producto por escalares de $\mathbb{R}$ en $C(I)$ como

$$(\alpha \cdot f)(x) = \alpha f(x) \text{ para cada } f \in C(I), \alpha \in \mathbb{K} \text{ y } x \in I.$$

Demostar que $(C(I), +, \mathbb{R}, \cdot)$ es un $\mathbb{R}$-espacio vectorial.

Antes de comenzar el ejercicio, dado que vamos a trabajar con igualdad de funciones, repasemos qué significa eso. Supongamos que tenemos $f, g : I \to \mathbb{R}$ (es decir dos funciones definidas en un intervalo $I \subseteq \mathbb{R}$ que toman valores en $\mathbb{R}$). Entonces

$$f = g \text{ si y sólo si } f(x) = g(x) \text{ para todo } x \in I. \tag{1.1}$$

En palabras, dos funciones son iguales si y sólo si tienen el mismo dominio y coinciden en todo punto de su dominio.

Por ejemplo, consideremos $f, g : \mathbb{R} \to \mathbb{R}$ con $f(x) = \begin{cases} 13 \text{ si } x = 0 \\ x \text{ si } x \neq 0 \end{cases}$ y $g(x) = x$. Entonces $f \neq g$ simplemente porque $f(0) \neq g(0)$.

Dem. Vamos a verificar que se cumplen los axiomas de los espacios vectoriales.

Observemos que $C(I)$ es un conjunto no vacío. Por ejemplo si definimos la función $\mathbf{0} : I \to \mathbb{R}$ como $\mathbf{0}(x) = 0,$ para todo $x \in I$. Claramente $\mathbf{0} \in C(I)$.

Ax. 1: Veamos que $+$ es una ley de composición interna. Sean $f, g \in C(I)$ entonces $f+g \in C(I)$, esto es así debido a que la suma usual (tal como se definió en este caso) de funciones continuas es también una función continua. Esta propiedad se vio en el curso de Análisis Matemático del CBC.

Ax. 2: La suma $+$ es conmutativa en $C(I)$. Queremos probar que $f + g = g + f$ para todo $f, g \in C(I)$.

Vamos a probar que la igualdad anterior ocurre para todo $x \in I$. Sean $f, g \in C(I)$ y $x \in I$, entonces $(f+g)(x) = f(x) + g(x) = g(x) + f(x) = (g+f)(x)$, donde usamos que la suma de números reales es conmutativa. Entonces como $(f+g)(x) = (g+f)(x)$, para todo $x \in I$ (fijarse que la igualdad anterior la obtuvimos para cualquier x), entonces (aplicando la definición de igualdad de funciones) $f + g = g + f$.

Ax. 3: La suma $+$ es asociativa en $C(I)$. Queremos probar que $(f+g)+h = f+(g+h)$ para todo $f, g, h \in C(I)$. Vamos a operar de manera similar a lo que hicimos en la prueba del axioma 2. Sean $f, g, h \in C(I)$ y $x \in I$, entonces $((f+g)+h)(x) = (f+g)(x) + h(x) = f(x) + g(x) + h(x) = f(x) + (g(x) + h(x)) = f(x) + (g+h)(x) = (f + (g+h))(x)$, donde usamos que la suma de números reales es asociativa. Como la igualdad anterior vale para todo $x \in I$, tenemos lo que queríamos probar.

Ax. 4: Existe un elemento neutro en $C(I)$ respecto de la suma. En este ítem, tenemos que proponer un elemento de $C(I)$ que sea elemento neutro para la suma y verificar que dicho elemento cumple. Por ejemplo, podemos proponer como elemento neutro la función nula $\mathbf{0} \in C(I)$ que definimos arriba. Veamos que $f + \mathbf{0} = f$ para todo $f \in C(I)$. Sean $f \in C(I)$ y $x \in I$, entonces $(f + \mathbf{0})(x) = f(x) + \mathbf{0}(x) = f(x) + 0 = f(x)$. Como la igualdad anterior vale para todo $x \in I$, tenemos lo que queríamos probar.

Observación: Se puede probar que el elemento neutro de un espacio vectorial es único. Por lo tanto, en este caso, la función nula $\mathbf{0} \in C(I)$, no sólo es "un" elemento neutro sino que es "el" elemento neutro. Meditar sobre la diferencia de estas expresiones.

Ax. 5: Para cada $f \in C(I)$ existe opuesto (inverso aditivo) en $C(I)$ respecto de la suma. En este ítem, tenemos que proponer un elemento de $C(I)$ que sea un inverso aditivo para la suma y verificar que dicho elemento cumple, es decir, queremos proponer una función $g \in C(I)$ tal que para cada $f \in CC(I)$ se cumple $f + g = \mathbf{0}$ (observar que ya verificamos que $\mathbf{0}$ es el elemento neutro para la suma). Sean $f \in C(I)$ y $x \in I$. Se define la función $g : I \to \mathbb{R}$ por $g(x) = -f(x)$. Entonces veamos que g es un inverso aditivo de f. Claramente $g \in C(I)$ porque el producto de una función continua por el escalar -1 es también una función continua. Por otra parte $(f + g)(x) = f(x) + g(x) = f(x) + -f(x) = 0 = \mathbf{0}(x)$ y como la igualdad anterior vale para todo $x \in I$, tenemos lo que queríamos probar.

Observación: Se puede probar que el inverso aditivo de cada vector de un espacio vectorial es único.

Ax. 6: Veamos que el producto por escalares $\cdot$ es una ley de composición externa. Sean $f \in C(I)$ y $\alpha \in \mathbb{R}$ entonces $\alpha \cdot f \in C(I)$; esto es así debido a que en primer lugar, $\alpha f(x) \in \mathbb{R}$ para todo $x \in I$ y además, porque el producto de un número real por una función continua (tal como se definió) es también una función continua. Esta propiedad se vio en el curso de Análisis Matemático del CBC.

Ax. 7: El producto por escalares $\cdot$ es distributivo respecto de la suma de escalares. Queremos ver que $(\alpha+\beta)\cdot f = \alpha \cdot f + \beta \cdot f$, para todo $f \in \mathbb{C}(I)$ y $\alpha, \beta \in \mathbb{R}$. Sean $f \in C(I)$, $x \in I$ y $\alpha, \beta \in \mathbb{R}$, entonces $((\alpha+\beta)\cdot f)(x) = (\alpha+\beta)f(x) = \alpha f(x)+\beta f(x) = (\alpha \cdot f)(x)+(\beta \cdot f)(x) = (\alpha \cdot f+\beta \cdot f)(x)$, donde usamos la propiedad distributiva de los números reales. Finalmente, como la igualdad anterior vale para todo $x \in I$, tenemos lo que queríamos probar.

Ax. 8: El producto por escalares $\cdot$ es distributivo respecto de la suma de vectores. Queremos ver que $\alpha \cdot (f + g) = \alpha \cdot f + \alpha \cdot g$, para todo $f, g \in C(I)$ y $\alpha \in \mathbb{R}$. Sean $f, g \in C(I)$, $x \in I$ y $\alpha \in \mathbb{R}$, entonces $(\alpha\cdot(f+g))(x) = \alpha((f+g)(x)) = \alpha(f(x)+g(x)) = \alpha f(x)+\alpha g(x) = (\alpha\cdot f)(x)+(\alpha\cdot g)(x) = (\alpha \cdot f + \alpha \cdot g)(x)$, donde usamos la propiedad distributiva de los números reales. Finalmente, como la igualdad anterior vale para todo $x \in I$, tenemos lo que queríamos probar.

Ax. 9: El producto por escalares $\cdot$ es asociativo respecto de los escalares. Queremos ver que $(\alpha\beta) \cdot f = \alpha \cdot (\beta \cdot f)$, para todo $f \in C(I)$ y $\alpha, \beta \in \mathbb{R}$. Sean $f \in C(I)$, $x \in I$ y $\alpha, \beta \in \mathbb{R}$, entonces $((\alpha\beta) \cdot f)(x) = (\alpha\beta)f(x) = \alpha(\beta f(x)) = \alpha((\beta \cdot f)(x)) = (\alpha \cdot (\beta \cdot f))(x)$, donde usamos la propiedad asociativa de los números reales. Finalmente, como la igualdad anterior vale para todo $x \in I$, tenemos lo que queríamos probar.

Ax. 10: El escalar $1 \in \mathbb{R}$ cumple que $1 \cdot f = f$, para todo $f \in C(I)$. Sean $f \in C(I)$ y $x \in I$, entonces $(1 \cdot f)(x) = 1f(x) = f(x)$ y como la igualdad anterior vale para todo $x \in I$, tenemos lo que queríamos probar. ■

El siguiente ejercicio recopila algunas propiedades importantes respecto del elemento neutro y el inverso aditivo de un espacio vectorial.

Ejercicio 1.C. Sea $\mathbb{K}$ un cuerpo y $(\mathbb{V}, +, \mathbb{K}, \cdot)$ un $\mathbb{K}$-espacio vectorial. Sea $0_{\mathbb{V}}$ el elemento neutro de $\mathbb{V}$ y dado $v \in \mathbb{V}$ notaremos como $-v$ al inverso aditivo de v. Demostar que, para todo $v \in \mathbb{V}$,

$$0 \cdot v = 0_{\mathbb{V}} \text{ y } -v = (-1) \cdot v.$$

Para demostrar estas propieades usaremos los axiomas que verifican los espacios vectoriales.

Dem. Por el Ax. 7, vale que $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v$. Sea w el inverso aditivo de $0 \cdot v$, que existe por el Ax.5. Entonces:

$$0_{\mathbb{V}} = 0 \cdot v + w = (0 \cdot v + 0 \cdot v) + w = 0 \cdot v + (0 \cdot v + w) = 0 \cdot v + 0_{\mathbb{V}} = 0 \cdot v.$$

Donde se usaron los Ax. 3, 4 y 5.

Finalmente,

$$v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 - 1) \cdot v = 0 \cdot v = 0_{\mathbb{V}}.$$

Donde se usaron los Ax. 7 y 10 y la igualdad anterior. Por lo tanto $(-1) \cdot v$ es el inverso aditivo de v, es decir $-v = (-1) \cdot v$. ■

A continuación veremos un ejemplo (delirante pero) interesante de espacios vectoriales. Este ejemplo es similar al **Ejercicio 1.1**.

Ejemplo 1.b. Sean $\mathbb{V} := \{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 : x_1, x_2 > 0\}$ y $\mathbb{K} = \mathbb{R}$. Se define

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \oplus \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} := \begin{bmatrix} x_1 y_1 \\ x_2 y_2 \end{bmatrix} \text{ para cada } x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{V}, y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{V} \text{ y}$$

$$\alpha \odot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} := \begin{bmatrix} x_1^\alpha \\ x_2^\alpha \end{bmatrix} \text{ para cada } x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{V}, \alpha \in \mathbb{R}.$$

Demostrar que $(\mathbb{V}, \oplus, \mathbb{R}, \odot)$ es un $\mathbb{R}$-espacio vectorial.

Dem. Vamos a verificar que se cumplen los axiomas de los espacios vectoriales.

Observemos que $\mathbb{V}$ es un conjunto no vacío. Por ejemplo $(1, 1) \in \mathbb{V}$.

Ax. 1: Veamos que $\oplus$ es una ley de composición interna. Sean $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{V}$ entonces

$$x \oplus y = \begin{bmatrix} x_1 y_1 \\ x_2 y_2 \end{bmatrix} \in \mathbb{V},$$

esto es así debido a que como $x, y \in \mathbb{V}$, tenemos que $x_1, x_2 > 0$ e $y_1, y_2 > 0$. Entonces $x_1 x_2 > 0, y_1 y_2 > 0$.

Ax. 2: La suma $\oplus$ es conmutativa en $\mathbb{V}$. Sean $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{V}$ entonces

$$x \oplus y = \begin{bmatrix} x_1 y_1 \\ x_2 y_2 \end{bmatrix} = \begin{bmatrix} y_1 x_1 \\ y_2 x_2 \end{bmatrix} = y \oplus x,$$

donde usamos que el producto de números reales es conmutativo.

Ax. 3: La suma $\oplus$ es asociativa en $\mathbb{V}$. Sean $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, $z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \mathbb{V}$ entonces

$$(x \oplus y) \oplus z = \begin{bmatrix} x_1 y_1 \\ x_2 y_2 \end{bmatrix} \oplus z = \begin{bmatrix} (x_1 y_1) z_1 \\ (x_2 y_2) z_2 \end{bmatrix} = \begin{bmatrix} x_1 (y_1 z_1) \\ x_2 (y_2 z_2) \end{bmatrix} = x \oplus (y \oplus z),$$

donde usamos que el producto de números reales es asociativo.

Ax. 4: Existe un elemento neutro en $\mathbb{V}$ respecto de la suma. Vamos proponer como elemento neutro al vector $0_{\mathbb{V}} := \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbb{V}$. Sea $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{V}$, entonces

$$x \oplus 0_{\mathbb{V}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \oplus \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1\, 1 \\ x_2\, 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x.$$

Por lo tanto, $0_{\mathbb{V}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbb{V}$ es el elemento neutro para la suma.

Ax. 5: Para cada $x \in \mathbb{V}$ existe opuesto (inverso aditivo) en $\mathbb{V}$ respecto de la suma. Sea $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{V}$, vamos proponer como inverso aditivo al vector $-x := \begin{bmatrix} \frac{1}{x_1} \\ \frac{1}{x_2} \end{bmatrix} \in \mathbb{V}$. El vector $-x$ que definimos es un elemento de $\mathbb{V}$ pues como $x \in \mathbb{V}$ tenemos que $x_1, x_2 > 0$ y el inverso de un número real positivo es otro número real positivo. Entonces

$$x \oplus -x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \oplus \begin{bmatrix} \frac{1}{x_1} \\ \frac{1}{x_2} \end{bmatrix} = \begin{bmatrix} \frac{x_1}{x_1} \\ \frac{x_2}{x_2} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0_{\mathbb{V}}.$$

Por lo tanto, $-x := \begin{bmatrix} \frac{1}{x_1} \\ \frac{1}{x_2} \end{bmatrix} \in \mathbb{V}$ es el inverso aditivo de x.

Ax. 6: Veamos que $\odot$ es una ley de composición externa. Sean $x \in \mathbb{V}$ y $\alpha \in \mathbb{R}$ entonces

$$\alpha \odot x = \begin{bmatrix} x_1^{\alpha} \\ x_2^{\alpha} \end{bmatrix} \in \mathbb{V},$$

esto es así debido a que como $x \in \mathbb{V}$, tenemos que $x_1, x_2 > 0$ y entonces $x_1^{\alpha}, x_2^{\alpha} > 0$.

Ax. 7: El producto por escalares $\odot$ es distributivo respecto de la suma de escalares. Sean $x \in \mathbb{V}$ y $\alpha, \beta \in \mathbb{R}$, entonces

$$(\alpha + \beta) \odot x = \begin{bmatrix} x_1^{\alpha+\beta} \\ x_2^{\alpha+\beta} \end{bmatrix} = \begin{bmatrix} x_1^{\alpha} x_1^{\beta} \\ x_2^{\alpha} x_2^{\beta} \end{bmatrix} = \begin{bmatrix} x_1^{\alpha} \\ x_2^{\alpha} \end{bmatrix} \oplus \begin{bmatrix} x_1^{\beta} \\ x_2^{\beta} \end{bmatrix} = \alpha \odot x \oplus \beta \odot x.$$

donde usamos propiedades de potenciación.

Ax. 8: El producto por escalares $\odot$ es distributivo respecto de la suma de vectores. Sean $x,\ y \in \mathbb{V}$ y $\alpha \in \mathbb{R}$, entonces

$$\alpha \odot (x \oplus y) = \alpha \odot \begin{bmatrix} x_1 y_1 \\ x_2 y_2 \end{bmatrix} = \begin{bmatrix} (x_1 y_1)^{\alpha} \\ (x_2 y_2)^{\alpha} \end{bmatrix} = \begin{bmatrix} x_1^{\alpha} x_2^{\alpha} \\ y_1^{\alpha} y_2^{\alpha} \end{bmatrix} = \begin{bmatrix} x_1^{\alpha} \\ x_2^{\alpha} \end{bmatrix} \oplus \begin{bmatrix} y_1^{\alpha} \\ y_2^{\alpha} \end{bmatrix} = \alpha \odot x \oplus \alpha \odot y,$$

donde usamos propiedades de potenciación.

Ax. 9: El producto por escalares $\odot$ es asociativo respecto de los escalares. Sean $x \in \mathbb{V}$ y $\alpha, \beta \in \mathbb{R}$, entonces

$$(\alpha\beta) \odot x = \begin{bmatrix} x_1^{\alpha\beta} \\ x_2^{\alpha\beta} \end{bmatrix} = \begin{bmatrix} (x_1^{\beta})^{\alpha} \\ (x_2^{\beta})^{\alpha} \end{bmatrix} = \alpha \odot \begin{bmatrix} x_1^{\beta} \\ x_2^{\beta} \end{bmatrix} = \alpha \odot (\beta \odot x),$$

donde usamos propiedades de potenciación.

Ax. 10: El escalar $1 \in \mathbb{R}$ cumple que $1 \odot x = x$, para todo $x \in \mathbb{V}$. Sea $x \in \mathbb{V}$, entonces

$$1 \odot x = \begin{bmatrix} x_1^{1} \\ x_2^{1} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x.$$

■

Los siguientes ejemplos son algunos de los $\mathbb{K}$-espacios vectoriales con los que trabajaremos con más frecuencia:

Ejemplo 1.c.

- Sea $\mathbb{K} = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$. En $\mathbb{C}^n$ definimos, para $x, y \in \mathbb{C}^n$ y $\alpha \in \mathbb{K}$ la suma como

$$x + y = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

  y el producto por escalares como

$$\alpha \cdot x = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix}.$$

  Entonces: $(\mathbb{C}^n, +, \mathbb{C}, \cdot)$ es un $\mathbb{C}$-espacio vectorial y $(\mathbb{C}^n, +, \mathbb{R}, \cdot)$ es un $\mathbb{R}$-espacio vectorial. Con la misma suma y producto por escalares también tenemos que $(\mathbb{R}^n, +, \mathbb{R}, \cdot)$ es un $\mathbb{R}$-espacio vectorial. Notar que si estamos tabajando en $\mathbb{R}^n$ y tomamos como cuerpo $\mathbb{C}$ entonces el producto por escalares que definimos NO es una ley de composición externa, por lo que $(\mathbb{R}^n, +, \mathbb{C}, \cdot)$ NO es un $\mathbb{C}$-espacio vectorial.

- Sea $\mathbb{K} = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$. En $\mathbb{C}^{m\times n}$ definimos, para $A, B \in \mathbb{C}^{m\times n}$ y $\alpha \in \mathbb{K}$ la suma como

$$A + B = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

y el producto por escalares como

$$\alpha \cdot A = \begin{bmatrix} \alpha a_{11} & \cdots & \alpha a_{1n} \\ \vdots & \ddots & \vdots \\ \alpha a_{m1} & \cdots & \alpha a_{mn} \end{bmatrix}.$$

Entonces: $(\mathbb{C}^{m\times n}, +, \mathbb{C}, \cdot)$ es un $\mathbb{C}$-espacio vectorial y $(\mathbb{C}^{m\times n}, +, \mathbb{R}, \cdot)$ es un $\mathbb{R}$-espacio vectorial. Con la misma suma y producto por escalares también tenemos que $(\mathbb{R}^{m\times n}, +, \mathbb{R}, \cdot)$ es un $\mathbb{R}$-espacio vectorial. Notar que si estamos tabajando en $\mathbb{R}^{m\times n}$ y tomamos como cuerpo $\mathbb{C}$ entonces el producto por escalares que definimos NO es una ley de composición externa, por lo que $(\mathbb{R}^{m\times n}, +, \mathbb{C}, \cdot)$ NO es un $\mathbb{C}$-espacio vectorial.

- Sea $\mathbb{K} = \mathbb{R}$ y

$$\mathbb{R}[x] := \{p : p(x) = a_n x^n + \cdots + a_1 x + a_0 : n \in \mathbb{N} \cup \{0\},\ a_0, a_1, \cdots, a_n \in \mathbb{R}\}.$$

Para cada $p, q \in \mathbb{R}[x]$ y $\alpha \in \mathbb{R}$ definimos la suma y el producto por escalares de la misma manera que lo hicimos en el ejemplo de arriba para las funciones continuas. Entonces $(\mathbb{R}[x], +, \mathbb{R}, \cdot)$ es un $\mathbb{R}$-espacio vectorial.

1.2. Subespacios

Cuando un subconjunto $\mathcal{S}$ de un $\mathbb{K}$-espacio vectorial $\mathbb{V}$ es en sí mismo un $\mathbb{K}$-espacio vectorial con las mismas operaciones definidas en $\mathbb{V}$ diremos que $\mathcal{S}$ es un *subespacio*. Es decir, si $\mathbb{V}$ es un $\mathbb{K}$-espacio vectorial, un subespacio de $\mathbb{V}$ es un subconjunto de $\mathbb{V}$ que es un $\mathbb{K}$-espacio vectorial con las mismas operaciones $+$ y $\cdot$ restringidas a dicho subconjunto.

Definición. Sea $\mathbb{K}$ un cuerpo y $(\mathbb{V}, +, \mathbb{K}, \cdot)$ un $\mathbb{K}$-espacio vectorial. Un subconjunto $\mathcal{S} \subseteq \mathbb{V}$ es un *subespacio* de $\mathbb{V}$ si $(\mathcal{S}, +, \mathbb{K}, \cdot)$ es un $\mathbb{K}$-espacio vectorial.

Teorema 1.2.1. *Sea $\mathbb{K}$ un cuerpo y $(\mathbb{V}, +, \mathbb{K}, \cdot)$ un $\mathbb{K}$-espacio vectorial. Un subconjunto $\mathcal{S} \subseteq \mathbb{V}$ es un subespacio de $\mathbb{V}$ si y sólo si:*

i) $\mathcal{S} \neq \emptyset$,

ii) *si $u, v \in \mathcal{S}$ entonces $u + v \in \mathcal{S}$,*

iii) *si $\alpha \in \mathbb{K}$ y $v \in \mathcal{S}$ entonces $\alpha \cdot v \in \mathcal{S}$.*

Dem. Supongamos que $\mathcal{S}$ es un subespacio de $\mathbb{V}$. Entonces por definición $(\mathcal{S}, +, \mathbb{K}, \cdot)$ es un $\mathbb{K}$-espacio vectorial. Por lo tanto, $\mathcal{S} \neq \emptyset$ y se cumplen los 10 axiomas de la definición de espacio vectorial. En particular, valen $i), ii)$ y $iii)$.

Recíprocamente, si valen $i), ii)$ y $iii)$ entonces $\mathcal{S} \neq \emptyset$. Por otra parte, si restringimos la suma $+$ al subconjunto $\mathcal{S}$ entonces, por $ii)$, la suma $+$ es una ley de composición interna (en $\mathcal{S}$) y vale el axioma 1. De la misma manera, si restringimos el producto por escalares $\cdot$ al subconjunto $\mathcal{S}$ entonces, por $iii)$, el producto por escalares $\cdot$ es una ley de composición externa (en $\mathcal{S}$) y se cumple el axioma 6. Finalmente, como $\mathcal{S} \subseteq \mathbb{V}$ y acabamos de ver que la suma $+$ y el producto por escalares $\cdot$ restringidos a $\mathcal{S}$ quedan bien definidos (porque son ley de composición interna y externa respectivamente) entonces se siguen de manera inmediata los axiomas $2, 3, 8, 9$ y 10. Sólo nos resta probar los axiomas 4 y 5. Primero, veamos que cada elemento de $\mathcal{S}$ tiene un inverso aditivo en

$\mathcal{S}$. De hecho, sea $v \in \mathcal{S}$ (existe pues $\mathcal{S}$ es no vacío), entonces $-v = (-1) \cdot v \in \mathcal{S}$, donde usamos el Ejercicio 1.C y que vale *iii*). Por lo tanto, cada elemento de $\mathcal{S}$ tiene un inverso aditivo en $\mathcal{S}$ y vale el axioma 5. Finalmente, sea $v \in \mathcal{S}$, si $-v$ es el inverso aditivo de v (que ya vimos que es un elemento de $\mathcal{S}$), tenemos que $0_{\mathbb{V}} = v + -v \in \mathcal{S}$, donde usamos que vale *ii*) y entonces tenemos el axioma 4.

Por lo tanto $(\mathcal{S}, +, \mathbb{K}, \cdot)$ es un $\mathbb{K}$-espacio vectorial y entonces, por definición, $\mathcal{S}$ es un subespacio de $\mathbb{V}$. ■

El item *i*) del teorema anterior se puede reemplazar por la condición *i'*) $0_{\mathbb{V}} \in \mathcal{S}$.

De hecho, si el item *i*) del teorema anterior se reemplaza por la condición *i'*), claramente tenemos que $\mathcal{S}$ es no vacío (pues $0_{\mathbb{V}} \in \mathcal{S}$) y recuperamos *i*) del teorema anterior.

Recíprocamente, en la demostración del teorema anterior, vimos que que si valen *i*), *ii*), *iii*) entonces $0_{\mathbb{V}} \in \mathcal{S}$.

Ejercicio 1.2 b, d). Verificar las siguientes afirmaciones:

- El conjunto $\mathcal{S} := \{\alpha[1\ 0\ 0]^T + \beta[1\ 0\ 1]^T : \alpha, \beta \in \mathbb{R}\}$ es un subespacio de $\mathbb{R}^3$.
- Para cada $n \in \mathbb{N}$, el conjunto $\mathcal{S} := \{\sum_{k=0}^{n} a_k x^k : a_0, a_1, \cdots, a_n \in \mathbb{R}\}$ es un subespacio de $\mathbb{R}[x]$.

Dem. Vamos a usar el teorema anterior para probar esta afirmación. Entonces, para probar que los conjuntos en cuestión son subespacios, basta ver los items *i*), *ii*), *iii*) del Teorema 1.2.1.

1.2 (b).

i) El vector $[0\ 0\ 0]^T = 0[1\ 0\ 0]^T + 0[1\ 0\ 1]^T \in \mathcal{S}$ con $\alpha = \beta = 0$.

ii) Si $u, v \in \mathcal{S}$, entonces $u = \alpha_1[1\ 0\ 0]^T + \beta_1[1\ 0\ 1]^T$ para ciertos $\alpha_1, \beta_1 \in \mathbb{R}$ y $v = \alpha_2[1\ 0\ 0]^T + \beta_2[1\ 0\ 1]^T$ para ciertos $\alpha_2, \beta_2 \in \mathbb{R}$. Entonces $u + v = (\alpha_1 + \alpha_2)[1\ 0\ 0]^T + (\beta_1 + \beta_2)[1\ 0\ 1]^T \in \mathcal{S}$ con $\alpha = \alpha_1 + \alpha_2 \in \mathbb{R}$ y $\beta = \beta_1 + \beta_2 \in \mathbb{R}$.

iii) Si $u \in \mathcal{S}$, $u = \alpha_1[1\ 0\ 0]^T + \beta_1[1\ 0\ 1]^T$ para ciertos $\alpha_1, \beta_1 \in \mathbb{R}$. Sea $c \in \mathbb{R}$, entonces $cu = (c\alpha_1)[1\ 0\ 0]^T + (c\beta_1)[1\ 0\ 1]^T \in \mathcal{S}$, con $\alpha = c\alpha_1 \in \mathbb{R}$ y $\beta = c\beta_1 \in \mathbb{R}$.

Como se cumple *i*), *ii*), *iii*), por Teorema 1.2.1, $\mathcal{S}$ es un subespacio de $\mathbb{R}^3$.

Al conjunto $\mathcal{S}$ lo notamos como $\mathcal{S} = gen\{[1\ 0\ 0]^T, [1\ 0\ 1]^T\}$.

1.2 (d). Sea $n \in \mathbb{N}$. Entonces:

i) El polinomio nulo $\mathbf{0}(x) = 0 = \sum_{k=0}^{n} 0x^k$. En este caso $a_0 = a_1 = \cdots = a_n = 0$. Entonces $\mathbf{0} \in \mathcal{S}$.

ii) Si $p, q \in \mathcal{S}$, entonces $p(x) = \sum_{k=0}^{n} b_k x^k$, para ciertos $b_0, b_1, \cdots b_n \in \mathbb{R}$ y $q(x) = \sum_{k=0}^{n} c_k x^k$, para ciertos $c_0, c_1, \cdots c_n \in \mathbb{R}$. Entonces $p(x) + q(x) = \sum_{k=0}^{n} b_k x^k + \sum_{k=0}^{n} c_k x^k = \sum_{k=0}^{n} (b_k + c_k) x^k$. Entonces $p + q \in \mathcal{S}$, con $a_i = b_i + c_i \in \mathbb{R}$ para cada $i \in \{0, 1, \cdots, n\}$.

iii) Si $p \in \mathcal{S}$, entonces $p(x) = \sum_{k=0}^{n} b_k x^k$, para ciertos $b_0, b_1, \cdots b_n \in \mathbb{R}$. Sea $\alpha \in \mathbb{R}$, entonces $\alpha p(x) = \alpha \sum_{k=0}^{n} b_k x^k = \sum_{k=0}^{n} (\alpha b_k) x^k$. Entonces $\alpha p \in \mathcal{S}$, con $a_i = \alpha b_i \in \mathbb{R}$ para cada $i \in \{0, 1, \cdots, n\}$.

Como se cumple *i*), *ii*), *iii*), por el Teorema 1.2.1, $\mathcal{S}$ es un subespacio de $\mathbb{R}[x]$. ■

1.3. Combinaciones lineales y sistema de generadores

Sea $\mathbb{K}$ un cuerpo y $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial. Sean $a_1, a_2, \cdots, a_r \in \mathbb{K}$ y $v_1, v_2, \cdots, v_r \in \mathbb{V}$ una *combinación lineal* de los vectores $v_1, v_2, \cdots, v_r$ es un elemento $v \in \mathbb{V}$ tal que

$$v = a_1v_1 + a_2v_2 + \cdots + a_rv_r.$$

Notaremos como $gen\{v_1, \cdots, v_r\}$ al conjunto de todas las combinaciones lineales de los vectores $v_1, \cdots, v_r$. Es decir,

$$gen\{v_1, v_2, \cdots, v_r\} := \{v : v = a_1v_1 + a_2v_2 + \cdots + a_rv_r : a_1, a_2, \cdots, a_r \in \mathbb{K}\}.$$

Usando las ideas del Ejercicio 1.2 b, d), es claro que el conjunto $gen\{v_1, v_2, \cdots, v_r\}$ es un subespacio de $\mathbb{V}$. Por otra parte, sea $G := \{v_1, v_2, \cdots, v_r\}$ un subconjunto de r elementos de $\mathbb{V}$. Entonces,

$$gen\{G\} := gen\{v_1, v_2, \cdots, v_r\}.$$

Definición. Sea $\mathbb{K}$ un cuerpo, $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y $\mathcal{S} \subseteq \mathbb{V}$ un subespacio de $\mathbb{V}$. Diremos que el conjunto $G = \{v_1, \cdots, v_r\} \subseteq \mathbb{V}$ es un *sistema de generadores* de $\mathcal{S}$ si

$$\mathcal{S} = gen\{G\} = gen\{v_1, \cdots, v_r\}.$$

Es decir, todo elemento de $\mathcal{S}$ es una combinación lineal de elementos de G.

Finalmente enunciaremos la definición de *sistema de generadores minimal*.

Definición. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y $\mathcal{S} \subseteq \mathbb{V}$ un subespacio. Un conjunto G es un sistema de generadores minimal de $\mathcal{S}$ si:

i) $\mathcal{S} = gen\{G\}$ (es decir G es un sistema de generadores de $\mathcal{S}$),

ii) si $G' \subsetneq G$ (es decir G' es un subconjunto propio de G) entonces G' no es sistema generador de $\mathcal{S}$.

A partir de la definición anterior podemos decir que si G es un sistema de generadores minimal de $\mathcal{S}$ entonces si algún conjunto G'' cumple que $gen\{G''\} = \mathcal{S}$ es porque $G \subseteq G''$. De esta propiedad viene la idea de "minimal" (minimal respecto de la inclusión de conjuntos).

Ejercicio 1.6 c, d). En cada uno de los siguientes casos describir el subespacio $\mathcal{S}$ mediante un sistema de generadores minimal.

- $\mathcal{S} = \{X \in \mathbb{R}^{2\times 2} : X \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} X\}$.
- $\mathcal{S} = \{p \in \mathbb{R}_3[x] : \int_{-1}^{1} p(x)dx = \int_{-1}^{1} xp(x)dx = 0\}$.

Dem. **1.6 (c).** Si $X \in \mathcal{S}$ entonces $X \in \mathbb{R}^{2\times 2}$, es decir $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ con $a, b, c, d \in \mathbb{R}$ y además

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Operando, nos queda que

$$\begin{bmatrix} 2a & 3b \\ 2c & 3d \end{bmatrix} = \begin{bmatrix} 2a & 2b \\ 3c & 3d \end{bmatrix}.$$

Entonces nos quedó el siguiente sistema:

$$\begin{cases} 2a = 2a \\ 3b = 2b \\ 2c = 3c \\ 3d = 3d. \end{cases}$$

La solución del sistema es $b = c = 0$. Entonces, volviendo a la expresión de X nos queda $X = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ con $a, d \in \mathbb{R}$. Entonces

$$\begin{aligned} \mathcal{S} &= \{X \in \mathbb{R}^{2\times 2} : X \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} X\} \\ &= \{X \in \mathbb{R}^{2\times 2} : X = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} : a, d \in \mathbb{R}\} \\ &= gen\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\}. \end{aligned}$$

Además el conjunto $G := \{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\}$ es un sistema de generadores minimal de $\mathcal{S}$. Cualquier subconjunto propio de G por ejemplo $G' := \{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\} \subsetneq G$ (¿qué otros subconjuntos propios tiene G?) no es un sistema de generadores de $\mathcal{S}$ (¿por qué?).

1.6 (d). Sea $p \in \mathcal{S}$ entoces $p \in \mathbb{R}_3[x]$, es decir $p(x) = ax^3 + bx^2 + cx + d$ con $a, b, c, d \in \mathbb{R}$ y $\int_{-1}^{1} p(x)dx = \int_{-1}^{1} xp(x)dx = 0$. Entonces

$$0 = \int_{-1}^{1} (ax^3 + bx^2 + cx + d)dx = \frac{a}{4}x^4 + \frac{b}{3}x^3 + \frac{c}{2}x^2 + dx|_{-1}^{1} = \frac{2}{3}b + 2d$$

y

$$0 = \int_{-1}^{1} x(ax^3 + bx^2 + cx + d)dx = \frac{a}{5}x^5 + \frac{b}{4}x^4 + \frac{c}{3}x^3 + \frac{d}{2}x^2|_{-1}^{1} = \frac{2}{5}a + \frac{2}{3}c.$$

Es decir nos quedó el siguiente sistema:

$$\begin{cases} \frac{2}{3}b + 2d = 0 \\ \frac{2}{5}a + \frac{2}{3}c = 0. \end{cases}$$

La solución del sistema es $b = -3d$ y $a = -\frac{5}{3}c$. Entonces volviendo a la expresión de p nos queda que $p(x) = ax^3 + bx^2 + cx + d = -\frac{5}{3}cx^3 - 3dx^2 + cx + d$ con $c, d \in \mathbb{R}$. Entonces

$$\mathcal{S} = \{p \in \mathbb{R}_3[x] : p(x) = c(-\frac{5}{3}x^3 + x) + d(-3x^2 + 1) : c, d \in \mathbb{R}\} = gen\{-\frac{5}{3}x^3 + x, -3x^2 + 1\}.$$

Un sistema de generadores minimal de $\mathcal{S}$ puede ser $G := \{-\frac{5}{3}x^3 + x, -3x^2 + 1\}$; claramente $\mathcal{S} = gen\{G\}$, además G es minimal porque si G' es un subconjunto propio de G (¿cuántos subconjuntos propios tiene G?) no genera $\mathcal{S}$. Por ejemplo, el subconjunto $G' := \{-3x^2 + 1\} \subsetneq G$ claramente no es un sistema de generadores de $\mathcal{S}$ (¿por qué?). De la misma manera se razona con los demás subconjuntos propios de G. ■

Ejercicio 1.5. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y $G = \{v_1, v_2, v_3, v_4\}$ un conjunto de vectores de $\mathbb{V}$. Sabiendo que $2v_1 - v_2 + v_3 = 0$ y $2v_1 + v_3 - v_4 = 0$. Hallar todos los subconjuntos de G que pueden ser un sistema de generadores minimal del subespacio generado por G.

Para simplificar un poco la notación, vamos a llamar

$$\mathcal{S} := gen\{G\} = gen\{v_1, v_2, v_3, v_4\}$$

es decir $\mathcal{S}$ es el subespacio generado por lo elementos de G.

Antes de empezar es crucial entender la diferencia entre G que es un conjunto de 4 elementos (que en general no es un subespacio) y $\mathcal{S}$ que es el subespacio generado por los elementos de G (que en general tiene infinitos elementos). Entendido eso, sigamos.

Dem. Por lo que nos dice el ejercicio, vale que $v_2 = 2v_1 + v_3$ y $v_4 = 2v_1 + v_3$. Entonces, tenemos que:

$$G = \{v_1, v_2, v_3, v_4\} = \{v_1, 2v_1 + v_3, v_3, 2v_1 + v_3\} = \{v_1, v_3, 2v_1 + v_3\},$$

(la igualdad de conjuntos $\{v_1, 2v_1 + v_3, v_3, 2v_1 + v_3\} = \{v_1, v_3, 2v_1 + v_3\}$ se prueba fácilmente probando doble inclusión). Y por otra parte

$$\mathcal{S} = gen\{G\} = gen\{v_1, 2v_1 + v_3, v_3, 2v_1 + v_3\} = gen\{v_1, v_3\}.$$

Para obtener un sistema de generadores minimal del subespacio generado por G (que llamamos $\mathcal{S}$), tenemos que pensar en 6 casos:

Caso 1: v_1 y v_3 son LI. En este caso, los subconjuntos de G que pueden ser un sistema de generadores minimal de $\mathcal{S}$ son: $\{v_1, v_3\}$, $\{v_1, 2v_1 + v_3\}$ y $\{v_3, 2v_1 + v_3\}$. Veamos por qué:

a) Los conjuntos $\{v_1, v_3\}, \{v_1, 2v_1 + v_3\}, \{v_3, 2v_1 + v_3\}$ son subconjuntos (todos distintos entre sí) de G.

b) $\mathcal{S} = gen\{v_1, v_3\} = gen\{v_1, 2v_1 + v_3\} = gen\{v_3, 2v_1 + v_3\}$. Entonces los conjuntos $\{v_1, v_3\}$, $\{v_1, 2v_1 + v_3\}$ y $\{v_3, 2v_1 + v_3\}$ son sistemas de generadores de $\mathcal{S}$. Observar que los tres conjuntos son distintos entre sí pero generan el mismo subespacio.

c) Los subconjuntos propios de $\{v_1, v_3\}$ son: $\{v_1\}$, $\{v_3\}$ y $\emptyset$. Claramente, ninguno de los tres subconjuntos propios de $\{v_1, v_3\}$ generan $\mathcal{S}$ (por ejemplo, el conjunto $\{v_1\}$ no genera $\mathcal{S}$ pues $v_3 \in \mathcal{S}$ pero $v_3 \notin gen\{v_1\}$ porque estamos suponiendo que v_1 y v_3 son LI, la misma idea se puede usar para ver que los otros subconjuntos propios de $\{v_1, v_3\}$ no generan $\mathcal{S}$). Por lo tanto $\{v_1, v_3\}$ no sólo es un sistema de generadores de $\mathcal{S}$ sino que también es minimal. De la misma manera podemos que ver que los conjuntos $\{v_1, 2v_1 + v_3\}$ y $\{v_3, 2v_1 + v_3\}$ también son sistemas de generadores minimales de $\mathcal{S}$.

Por $a), b)$ y $c)$, los conjuntos $\{v_1, v_3\}, \{v_1, 2v_1 + v_3\}, \{v_3, 2v_1 + v_3\}$ son los subconjuntos de G que pueden ser un sistema de generadores minimal del subespacio generado por G.

Por si alguien se hizo esta pregunta: Observar que el conjunto $\{15v_1, 16v_3\}$ es un sistema de generadores minimal de $\mathcal{S}$ (además es distinto a los 3 que ya enumeramos en este caso) pero como NO es un subconjunto de G, no aparece como uno de los sistema de generadores minimales de $\mathcal{S}$ que nos pide el ejercicio.

Caso 2: v_1 y v_3 son LD (y además $v_1 \neq 0$, $v_3 \neq 0$ y $2v_1 + v_3 \neq 0$). En este caso, los subconjuntos de G que pueden ser un sistema de generadores minimal de $\mathcal{S}$ son: $\{v_1\}$, $\{v_3\}$ y $\{2v_1 + v_3\}$. Veamos por qué:

a) Los conjuntos $\{v_1\}, \{v_3\}, \{2v_1 + v_3\}$ son subconjuntos de G.

b) $\mathcal{S} = gen\{v_1, v_3\} = gen\{v_1\} = gen\{v_3\} = gen\{2v_1 + v_3\}$. Entonces los conjuntos $\{v_1\}$, $\{v_3\}$ y $\{2v_1 + v_3\}$ son sistemas de generadores de $\mathcal{S}$.

c) El único subconjunto propio de $\{v_1\}$, $\{v_3\}$ y $\{2v_1 + v_3\}$ es $\emptyset$ que claramente no genera $\mathcal{S}$. Por lo tanto $\{v_1\}$, $\{v_3\}$ y $\{2v_1 + v_3\}$ no sólo son sistemas de generadores de $\mathcal{S}$ sino que también son todos minimales.

Por $a), b)$ y $c)$, los conjuntos $\{v_1\}$, $\{v_3\}$ y $\{2v_1 + v_3\}$ son los subconjuntos de G que pueden ser un sistema de generadores minimal del subespacio generado por G.

Acá hacemos la siguiente aclaración. Si en este caso, tenemos la suerte de que $v_1 = v_3$. Entonces en realidad tenemos sólo dos conjuntos (distintos entre sí) que pueden ser un sistema de generadores minimal del subespacio generado por G, serían los conjuntos $\{v_3\}$ y $\{2v_1 + v_3\} = \{3v_3\}$.

Si en este caso, tenemos la suerte de que $v_1 = -v_3$ entonces tenemos sólo dos conjuntos (distintos entre sí) que pueden ser un sistema de generadores minimal del subespacio generado por $G : \{v_1\}$ y $\{v_3\} = \{-v_1\}$.

Por último, con las hipótesis del caso 2, si $v_1 \neq v_3$ y $v_1 \neq -v_3$, los tres conjuntos $\{v_1\}$, $\{v_3\}$ y $\{2v_1 + v_3\}$ son todos distintos entre sí.

Caso 3: v_1 y v_3 son LD ($v_1 \neq 0$, $v_3 \neq 0$ y $2v_1 + v_3 = 0$). En este caso, $v_3 = -2v_1 \neq 0$ y $G = \{v_1, -2v_1, 0\}$. Entonces, los subconjuntos de G que pueden ser un sistema de generadores minimal de $\mathcal{S}$ son: $\{v_1\}$, $\{-2v_1\}$. Veamos por qué:

a) Los conjuntos $\{v_1\}, \{-2v_1\}$ son subconjuntos (distintos entre sí) de G.

b) $\mathcal{S} = gen\{v_1, v_3\} = gen\{v_1\} = gen\{-2v_1\}$. Entonces los conjuntos $\{v_1\}$, $\{-2v_1\}$ son sistemas de generadores de $\mathcal{S}$.

c) El único subconjunto propio de $\{v_1\}$, $\{-2v_1\}$ es $\emptyset$ que claramente no genera $\mathcal{S}$. Por lo tanto $\{v_1\}$, $\{-2v_1\}$ no sólo son sistemas de generadores de $\mathcal{S}$ sino que también son todos minimales.

Por $a), b)$ y $c)$, los conjuntos $\{v_1\}$, $\{-2v_1\}$ son los subconjuntos de G que pueden ser un sistema de generadores minimal del subespacio generado por G.

Caso 4: $v_1 = 0$ y $v_3 \neq 0$. Entonces $2v_1 + v_3 = v_3$ y $G = \{0, v_3, v_3\} = \{0, v_3\}$.

En este caso, el único subconjunto de G que puede ser un sistema de generadores minimal de $\mathcal{S}$ es $\{v_3\}$. Veamos por qué:

a) $\{v_3\}$ es subconjunto de G.

b) $\mathcal{S} = gen\{v_1, v_3\} = gen\{v_3\}$. Entonces $\{v_3\}$ es un sistema de generadores de $\mathcal{S}$.

c) El único subconjunto propio de $\{v_3\}$ es $\emptyset$ que claramente no genera $\mathcal{S}$. Por lo tanto $\{v_3\}$ no sólo es un sistema de generadores de $\mathcal{S}$ sino que también es minimal.

Por $a), b)$ y $c)$, el conjunto $\{v_3\}$ es el único subconjunto de G que puede ser un sistema de generadores minimal de $\mathcal{S}$.

Caso 5: $v_1 \neq 0$ y $v_3 = 0$. Entonces, $2v_1 + v_3 = 2v_1 \neq 0$ y $G = \{v_1, 0, 2v_1\}$. En este caso, los subconjuntos de G que pueden ser un sistema de generadores minimal de S son: $\{v_1\}$ y $\{2v_1\}$. Veamos por qué:

a) $\{v_1\}$ y $\{2v_1\}$ son subconjuntos (distintos entre sí) de G.

b) $\mathcal{S} = gen\{v_1, v_3\} = gen\{v_1\} = gen\{2v_1\}$. Entonces $\{v_1\}$ y $\{2v_1\}$ son sistemas de generadores de $\mathcal{S}$.

c) El único subconjunto propio de $\{v_1\}$ y $\{2v_1\}$ es $\emptyset$ que claramente no genera $\mathcal{S}$. Por lo tanto $\{v_1\}$ y $\{2v_1\}$ no sólo son sistemas de generadores de $\mathcal{S}$ sino que también son minimal.

Por $a), b)$ y $c)$, los conjuntos $\{v_1\}$ y $\{2v_1\}$ son los subconjuntos de G que puede ser un sistema de generadores minimal de $\mathcal{S}$

El último caso excede los alcances de la materia pero lo vamos a ver igual para aquellos que tengan curiosidad.

Caso 6: $v_1 = 0$ y $v_3 = 0$. Entonces $2v_1 + v_3 = 0$ y $G = \{0, 0, 0\} = \{0\}$.

En este caso (veremos a continuación que) el único subconjunto de G que puede ser un sistema de generadores minimal de $\mathcal{S}$ es $\emptyset$. Veamos por qué:

a) $\emptyset$ es un subconjunto de G.

b) $\mathcal{S} = gen\{0\} = \{0\} = gen\{\emptyset\}$. Para probar esto vamos a tener que definir qué entendemos por "$gen\{\emptyset\}$". Ver más abajo.

c) El conjunto $\emptyset$ no tiene subconjuntos propios.

Por *a*), *b*) y *c*), el conjunto $\emptyset$ es el único subconjunto de G que puede ser un sistema de generadores minimal de $\mathcal{S}$. ■

Veamos por qué podemos decir que vale que $\{0\} = gen\{\emptyset\}$

Antes de contestar esta pregunta, vamos a probar una propiedad que nos será muy útil.

Proposición 1.3.1. *Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial. Consideremos $H = \{v_1, v_2, \cdots, v_r\}$ donde $v_1, v_2, \cdots, v_r \in \mathbb{V}$. Entonces, si $\mathcal{S}$ es cualquier subespacio de $\mathbb{V}$, vale que*

$$gen\{H\} = \bigcap_{H \subseteq \mathcal{S}} \mathcal{S}.$$

En palabras, afirmamos que el subespacio $gen\{H\}$ es igual a la intersección de todos los subespacios $\mathcal{S}$ de $\mathbb{V}$ que contienen al conjunto H.

Dem. Probemos la doble inclusión. Supongamos que

$$v \in gen\{H\} := \{a_1v_1 + a_2v_2 + \cdots + a_rv_r : \ a_1, a_2, \cdots, a_r \in \mathbb{K}\}$$

entonces,

$$v = a_1v_1 + a_2v_2 + \cdots + a_rv_r,$$

para ciertos $a_1, a_2, \cdots, a_r \in \mathbb{K}$. Entonces, sea $\mathcal{S}$ cualquier subespacio de $\mathbb{V}$ que contiene al conjunto H. Como $v_1, v_2, \cdots, v_r \in H$, tenemos que $v_1, v_2, \cdots, v_r \in \mathcal{S}$ y como $\mathcal{S}$ es un subespacio, se sigue que $a_1v_1 + a_2v_2 + \cdots + a_rv_r \in \mathcal{S}$. Como lo anterior vale para cualquier $\mathcal{S}$ subespacio que contenga a H, entonces $v \in \bigcap_{H \subseteq \mathcal{S}} \mathcal{S}$.

Recíprocamente, si $v \in \bigcap_{H \subseteq \mathcal{S}} \mathcal{S}$, entonces $v \in \mathcal{S}$ para todo $\mathcal{S}$ subespacio de $\mathbb{V}$ que contiene a H. Como (ya vimos que) el conjunto $gen\{H\} = \{a_1v_1 + a_2v_2 + \cdots + a_rv_r : a_1, a_2, \cdots, a_r \in \mathbb{K}\}$ es un subespacio de $\mathbb{V}$ y claramente

$$H \subseteq \{a_1v_1 + a_2v_2 + \cdots + a_rv_r : a_1, a_2, \cdots, a_r \in \mathbb{K}\}$$

(verificarlo). Entonces, en particular, tenemos que $v \in gen\{H\} = \{a_1v_1 + a_2v_2 + \cdots + a_rv_r : a_1, a_2, \cdots, a_r \in \mathbb{K}\}$ y probamos la doble inclusión. Es decir, efectivamente tenemos que $gen\{H\} = \bigcap_{H \subseteq \mathcal{S}} \mathcal{S}$. ■

Volviendo al **Caso 6** del **Ejercicio 1.5**, veamos que

$$\{0\} = gen\{\emptyset\}.$$

Para probar esto último, es necesario definir qué entendemos por $gen\{\emptyset\}$. Hasta ahora, sólo vimos qué entendemos por $gen\{H\}$ cuando H es un conjunto (finito) de vectores (de hecho definimos $gen\{H\}$ como todas las combinaciones lineales de los elementos de H) y con esa definición podemos hacer prácticamente todos los ejercicios de la Guía (salvo quizás el Caso 6) del Ejercicio 1.5).

Vamos a dar una definición (razonable) de $gen\{\emptyset\}$ que a su vez extienda la definición para el caso en que H sea no vacío (que fue la que ya vimos). Entonces, por lo que acabamos de probar en la Proposición 1.3.1, resulta natural definir:

$$gen\{\emptyset\} := \bigcap_{\emptyset \subseteq \mathcal{S}} \mathcal{S} = \bigcap_{\mathcal{S}} \mathcal{S},$$

donde la última igualdad vale porque todo conjunto (en particular todo subespacio) contiene al conjunto vacío.

Veamos entonces, que $\bigcap_{\mathcal{S}} \mathcal{S} = \{0\}$. De hecho, es claro que $0 \in \mathcal{S}$ para todo $\mathcal{S}$ subespacio de $\mathbb{V}$, entonces

$$\{0\} \subseteq \bigcap_{\mathcal{S}} \mathcal{S}.$$

Por otra parte, como en particular $\{0\}$ es un subespacio entonces,

$$\bigcap_{\mathcal{S}} \mathcal{S} \subseteq \{0\}.$$

Probamos la doble inclusión y entonces se sigue que

$$gen\{\emptyset\} = \bigcap_{\mathcal{S}} \mathcal{S} = \{0\}.$$

Con estas ideas y la prueba del **Ejercicio 1.7** (a continuación) se puede demostrar también que el conjunto $\emptyset$ es (la única) base del subespacio $\{0\}$.

1.4. Conjuntos linealmente independientes

Definición. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial. El conjunto $\{v_1, v_2, \cdots, v_r\} \subseteq \mathbb{V}$ es linealmente independiente (de ahora en más abreviamos LI) si dados $a_1, a_2, \cdots, a_r \in \mathbb{K}$ tales que

$$a_1 v_1 + a_2 v_2 + \cdots + a_r v_r = 0_{\mathbb{V}}$$

entonces $a_1 = a_2 = \cdots = a_r = 0$. Cuando un conjunto no es LI diremos que es linealmente dependiente (de ahora en más abreviamos LD).

Recordar que $0_{\mathbb{V}}$ denota el elemento neutro del $\mathbb{K}$-espacio vectorial $\mathbb{V}$.

El siguiente ejemplo demuestra que la independencia lineal de vectores depende del $\mathbb{K}$-espacio vectorial considerado.

Ejemplo 1.d. Sea $(\mathbb{V}, \oplus, \mathbb{R}, \odot)$ el $\mathbb{R}$-espacio vectorial del Ejemplo 1.b y consideremos $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$ con la suma y producto por escalares usuales en $\mathbb{R}^2$. Entonces el conjunto $\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \end{bmatrix}\}$ es LD en $(\mathbb{V}, \oplus, \mathbb{R}, \odot)$ pero es LI en $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$.

Dem. Observar que

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 2^2 \\ 1^2 \end{bmatrix} = 2 \odot \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Entonces

$$1 \odot \begin{bmatrix} 4 \\ 1 \end{bmatrix} \oplus (-2) \odot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \oplus \begin{bmatrix} 2^{-2} \\ 1^{-2} \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \oplus \begin{bmatrix} \frac{1}{4} \\ 1 \end{bmatrix} = \begin{bmatrix} 4\frac{1}{4} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0_{\mathbb{V}}.$$

Tenemos una combinación lineal no nula de los vectores $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ igualada a $0_{\mathbb{V}}$. Por lo tanto, concluimos que el conjunto $\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \end{bmatrix}\}$ es LD en $(\mathbb{V}, \oplus, \mathbb{R}, \odot)$.

Por otra parte, supongamos que existen $a, b \in \mathbb{R}$ tales que

$$a \begin{bmatrix} 2 \\ 1 \end{bmatrix} + b \begin{bmatrix} 4 \\ 1 \end{bmatrix} = 0_{\mathbb{R}^2} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Entonces $2a + 4b = 0$ y $a + b = 0$. Ese sistema sólo tiene como solución $a = b = 0$. Por lo tanto, concluimos que el conjunto $\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \end{bmatrix} \}$ es LI en $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$. ∎

Para resolver el próximo ejercicio, vamos a usar la ayuda que nos dan. La ayuda nos pregunta qué significa que G sea un sistema de generadores que no es minimal. Si lo hacemos de esta manera estaremos probando el contrarrecíproco del ejercicio, es decir probaremos que vale "no p si y sólo si no q" que es equivalante a "p si y sólo si q".

Ejercicio 1.7. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y $G = \{v_1, v_2, v_3\}$ un sistema de generadores de $\mathbb{V}$. Probar que G es un sistema de generadores minimal de $\mathbb{V}$ si y sólo si G es LI.

Dem. Supongamos que G es un sistema de generadores de $\mathbb{V}$ que no es minimal. Entonces, por definición, existe un subconjunto (propio) $G' \subsetneq G$ no vacío tal que $\mathbb{V} = gen\{G'\}$. Sabemos por hipótesis que $\mathbb{V} = gen\{G\} = gen\{v_1, v_2, v_3\}$. Entonces $gen\{G'\} = gen\{v_1, v_2, v_3\}$. Para continuar, observemos que todos los subconjuntos propios (no vacíos) de G posibles son: $\{v_1\}$, $\{v_2\}$, $\{v_3\}$, $\{v_1, v_2\}$, $\{v_1, v_3\}$ y $\{v_2, v_3\}$. Supongamos que $G' = \{v_1\}$ (la misma idea vale si G' es cualquiera de los otros subconjunto propios de G posibles). Entonces, como $gen\{v_1\} = gen\{G'\} = gen\{v_1, v_2, v_3\}$, tenemos que (por ejemplo) $v_2 \in gen\{v_1\}$ entonces existe $a \in \mathbb{K}$ tal que $v_2 = av_1$ y entonces el conjunto $\{v_1, v_2, v_3\}$ no es LI. Finalmente, observar que con lo que vimos en el Caso 6) del Ejercicio 1.5), si G' es el conjunto vacío (que es un subconjunto propio de G), tendríamos que $gen\{\emptyset\} = \{0\} = gen\{G\} = gen\{v_1, v_2, v_3\}$. Entonces, claramente $\{v_1, v_2, v_3\}$ no es LI y arribamos a la misma conclusión.

Recíprocamente, si G no es un conjunto linealmente independiente de $\mathbb{V}$, entonces existen $a_1, a_2, a_3 \in \mathbb{K}$ (no todos nulos) tales que $0_{\mathbb{V}} = a_1v_1 + a_2v_2 + a_3v_3$. Supongamos que $a_1 \neq 0$ (es lo mismo si suponemos que $a_2 \neq 0$ o que $a_3 \neq 0$). Entonces, $v_1 = -\frac{a_2}{a_1}v_2 - \frac{a_3}{a_1}v_3$, por lo tanto $\mathbb{V} = gen\{v_1, v_2, v_3\} = gen\{v_2, v_3\}$. Entonces, si llamamos $G' := \{v_2, v_3\}$, tenemos que G' es un subconjunto propio de G (eso es así porque G' está contenido en G y además v_1 pertenece a G y no pertenece a G') y que $\mathbb{V} = gen\{G'\}$. Entonces G no puede ser un sistema de generadores minimal de $\mathbb{V}$. ∎

Ejercicio 1.8. Sean $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial, $\{v_1, v_2, \cdots, v_2\} \subseteq \mathbb{V}$ un conjunto LI, y $A = [a_{ij}] \in \mathbb{K}^{n\times n}$. Para cada $j \in \{1, 2, \cdots, n\}$ se definen los vectores $w_j := \sum_{i=1}^{n} a_{ij}v_i$. Mostrar que $\{w_1, w_2, \cdots, w_n\}$ es LI si y sólo si $det(A) \neq 0$.

Dem. Vamos a probar el ejercicio usando la definición de independencia lineal. Sean $\alpha_1, \alpha_2, \cdots \alpha_n \in \mathbb{K}$ tales que

$$\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n = 0_{\mathbb{V}}. \tag{1.2}$$

Entonces, el conjunto $\{w_1, w_2, \cdots, w_n\}$ es LI si y sólo si $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

Veamos entonces, qué pinta tiene la expresión (1.2) desarrollándola. De hecho, sabiendo que $w_1 = a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n,\ w_2 = a_{12}v_1 + a_{22}v_2 + \cdots a_{n2}v_n$, etc., tenemos que

$$\begin{aligned}&\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n = \alpha_1(a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n) + \\ &\alpha_2(a_{12}v_1 + a_{22}v_2 + \cdots a_{n2}v_n) + \cdots + \alpha_n(a_{1n}v_1 + a_{2n}v_2 + \cdots a_{nn}v_n) = 0_{\mathbb{V}}.\end{aligned}$$

Si sacamos factor común los vectores $v_1, v_2, \cdots, v_n$ en la expresión de arriba, nos queda que:

$$\begin{aligned}&v_1(\alpha_1 a_{11} + \alpha_2 a_{12} + \cdots + \alpha_n a_{1n}) + v_2(\alpha_1 a_{21} + \alpha_2 a_{22} + \cdots + \alpha_n a_{2n}) + \cdots + \\ &v_n(\alpha_1 a_{n1} + \alpha_2 a_{n2} + \cdots + \alpha_n a_{nn}) = 0_{\mathbb{V}}.\end{aligned}$$

Como el conjunto $\{v_1, v_2, \cdots, v_n\}$ es LI (por hipótesis) y arriba tenemos una CL de dichos vectores igualada a $0_{\mathbb{V}}$ tenemos que cada escalar que multiplica a dichos vectores es nulo, es decir:

$$\begin{cases} \alpha_1 a_{11} + \alpha_2 a_{12} + \cdots + \alpha_n a_{1n} = 0 \\ \alpha_1 a_{21} + \alpha_2 a_{22} + \cdots + \alpha_n a_{2n} = 0 \\ \cdots \\ \alpha_1 a_{n1} + \alpha_2 a_{n2} + \cdots + \alpha_n a_{nn} = 0. \end{cases}$$

Como $A = [a_{ij}] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in \mathbb{K}^{n\times n}$, el sistema de ecuaciones anterior es equivalente a

$$A \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{bmatrix} = 0. \tag{1.3}$$

Por lo tanto $\{w_1, w_2, \cdots, w_n\}$ es LI si y sólo si $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, si y sólo si la única solución del sistema (1.3) es la trivial, si y sólo si A es inversible (acá usamos que A es cuadrada), si y sólo si $\det(A) \neq 0$. ■

Independencia lineal de conjuntos en $\mathbb{K}^n$

A continuación, vamos a justificar por qué cuando se tiene una cantidad (finita) de vectores de $\mathbb{K}^n$ ($\mathbb{K} = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$) para verificar su independencia lineal basta con colocarlos en filas (o en columnas) en una matriz, triangular y ver si una (o más) fila/s (columna/s) se anulan. La siguiente proposición también resultará útil para entender los algoritmos del **Ejercicio 1.13**.

Proposición 1.4.1. *Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y $\{v_1, v_2, \cdots, v_r\} \subseteq \mathbb{V}$ un conjunto de vectores. Entonces:*

1. $\{v_1, v_2, \cdots, v_i, \cdots, v_j, \cdots, v_r\}$ *es LI y sólo si* $\{v_1, v_2, \cdots, v_j, \cdots, v_i, \cdots, v_r\}$ *es LI. Es decir la independencia lineal de un conjunto no cambia si cambiamos el orden de los elementos del conjunto.*

2. $\{v_1, v_2, \cdots, v_i, \cdots, v_r\}$ *es LI y sólo si* $\{v_1, v_2, \cdots, cv_i, \cdots v_r\}$ *es LI, para todo* $c \in \mathbb{K} \setminus \{0\}$. *Es decir la independencia lineal de un conjunto no cambia si multiplicamos por un escalar no nulo alguno de los vectores.*

3. $\{v_1, v_2, \cdots, v_i, \cdots, v_j, \cdots, v_r\}$ *es LI y sólo si* $\{v_1, v_2, \cdots, v_i, \cdots, v_j + \lambda vi, \cdots, v_r\}$ *es LI para todo* $\lambda \in \mathbb{K}$ *(acá λ puede ser 0). Es decir la independencia lineal de un conjunto no cambia si a un vector lo reemplazamos por ese mismo vector sumado un multiplo de otro vector.*

Dem. 1. : Es obvio.
2. : Se deja como ejercicio.
3. : Supongamos que el conjunto $\{v_1, v_2, \cdots, v_i, \cdots, v_j, \cdots, v_r\}$ es LI.

Queremos ver que $\{v_1, v_2, \cdots, v_i, \cdots, v_j + \lambda vi, \cdots, v_r\}$ con $\lambda \in \mathbb{K}$ es LI. Lo haremos por definición. Sean $a_1, a_2, \cdots, a_r \in \mathbb{K}$ tales que

$$a_1 v_1 + a_2 v_2 + \cdots + a_i v_i + \cdots + a_j(v_j + \lambda vi) + \cdots + a_r v_r = 0_{\mathbb{V}}.$$

Operando y sacando factor común $v_1, v_2, \cdots, v_r$, tenemos que:

$$a_1 v_1 + a_2 v_2 + \cdots + (a_i + a_j\lambda) v_i + \cdots + a_j v_j + \cdots + a_r v_r = 0_{\mathbb{V}}.$$

Nos quedó una CL de los vectores $v_1, v_2, \cdots, v_r$ igualada a $0_{\mathbb{V}}$. Entonces, usando la hipótesis (es decir que $\{v_1, v_2, \cdots, v_i, \cdots, v_j, \cdots, v_r\}$ es LI), tenemos que, $a_1 = a_2 = \cdots = (a_i + a_j\lambda) = \cdots = a_j = \cdots = a_r = 0$. Entonces, $a_i = -\lambda a_j = -\lambda 0 = 0$. Entonces, como concluimos que $a_1 = a_2 = \cdots = a_i = \cdots = a_j = \cdots = a_r = 0$ tenemos que $\{v_1, v_2, \cdots, v_i, \cdots, v_j + \lambda vi, \cdots, v_r\}$ es LI y probamos lo que queríamos (observar que λ era cualquier elemento del cuerpo $\mathbb{K}$, por lo tanto, lo que probamos vale para todo $\lambda \in \mathbb{K}$).

Recíprocamente, supongamos que $\{v_1, v_2, \cdots, v_i, \cdots, v_j + \lambda vi, \cdots v_r\}$ es LI para todo $\lambda \in \mathbb{K}$. Queremos ver que el conjunto $\{v_1, v_2, \cdots, v_i, \cdots, v_j, \cdots v_r\}$ es LI. Observar que

$$\{v_1, v_2, \cdots, v_i, \cdots, v_j, \cdots v_r\} = \{v_1, v_2, \cdots, v_i, \cdots, v_j + \lambda v_i + (-\lambda)v_i, \cdots v_r\}$$

que es LI por lo que acabamos de ver. ■

Recordemos que cuando triangulamos una matriz, podemos realizar las siguientes operaciones entre las filas:

1. intercambiar filas,
2. multiplicar una fila por un escalar no nulo,
3. reemplazar una fila por esa fila sumada a un multiplo de otra.

Por lo tanto, a partir de la Proposición 1.4.1, para decidir si un conjunto de vectores $\{v_1, v_2, \cdots, v_r\}$ de $\mathbb{K}^n$ es linealmente idependiente, basta con colocar dichos vectores en filas (o en columnas) en una matriz y luego triangular. Si una (o más) fila/s (columna/s) se anulan, el conjunto es LD. De lo contrario, el conjunto es LI.

La utilidad de este "método" (que sólo sirve para vectores de $\mathbb{K}^n$) está en que si al triangular la matriz (que resulta de colocar los vectores en filas), una fila hubiera dado nula, entonces el vector que originalmente estaba en dicha fila es LD con el resto de los vectores. La desventaja de este método es que (obviamente) sólo se puede aplicar a una cantidad finita de vectores de $\mathbb{K}^n$ (con la suma y el producto por escalares usuales de $\mathbb{K}^n$).

Ejercicio 1.10 a). Determinar cuáles de los siguientes subconjuntos son LI en su correspondiente espacio vectorial:

- El subconjunto $\{\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ -6 \end{bmatrix}, \begin{bmatrix} 0 \\ -5 \\ 6 \end{bmatrix}\}$.

Dem. **1.10 (a).** Vamos a resolver el problema de dos maneras.

Por definición: Sean $a, b, c \in \mathbb{R}$ tales que

$$a \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + b \begin{bmatrix} 3 \\ 5 \\ -6 \end{bmatrix} + c \begin{bmatrix} 0 \\ -5 \\ 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Operando, nos queda el siguiente sistema a resolver, $\begin{bmatrix} 1 & 3 & 0 \\ 3 & 5 & -5 \\ -2 & -6 & 6 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$. Resolviendo el sistema, obtenemos que $a = b = c = 0$. Entonces el conjunto es LI.

Triangulando la matriz que resulta de colocar los vectores en fila:

Llamemos $v_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, v_2 = \begin{bmatrix} 3 \\ 5 \\ -6 \end{bmatrix}, v_3 \begin{bmatrix} 0 \\ -5 \\ 6 \end{bmatrix}$, coloquemos dichos vectores en las filas de una matriz y triangulemos.

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -2 \\ 3 & 5 & -6 \\ 0 & -5 & 6 \end{bmatrix} \underset{v_2 - 3v_1 \to v_2'}{\Rightarrow} \begin{bmatrix} 1 & 3 & -2 \\ 0 & -4 & 0 \\ 0 & -5 & 6 \end{bmatrix} \underset{v_3 - \frac{5}{4}v_2' \to v_3'}{\Rightarrow} \begin{bmatrix} 1 & 3 & -2 \\ 0 & -4 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

En la Proposición 1.4.1, vimos que las operaciones que realizamos entre los vectores (colocados en las filas de una matriz) cuando triangulamos dicha matriz no modifican su independencia lineal. Es decir, si llamamos $v_1' = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$, $v_2' = \begin{bmatrix} 0 \\ -4 \\ 0 \end{bmatrix}$ y $v_3' = \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix}$, entonces claramente $\{v_1', v_2', v_3'\}$ es un conjunto LI. Entonces, por la Proposición 1.4.1, el conjunto original $\{v_1, v_2, v_3\}$ era LI. ■

Ejercicio 1.12 a). Hallar todos los valores de $a \in \mathbb{R}$ para los cuales los siguientes subconjuntos son linealmente dependientes (LD) en su correspondiente espacio vectorial.

- El subconjunto $\{1 + ax + 3x^2, 4 + 5x + 7x^2, a + x + x^2\}$ en $\mathbb{R}_2[x]$.

Dem. Para este ejercicio no podemos usar la Proposición 1.4.1 (no tenemos vectores de $\mathbb{R}^n$ sino polinomios) así que lo haremos por definición. Sean $\alpha, \beta, \gamma \in \mathbb{R}$ tales que

$$\alpha(1 + ax + 3x^2) + \beta(4 + 5x + 7x^2) + \gamma(a + x + x^2) = 0 \text{ para todo } x \in \mathbb{R}.$$

Buscaremos para qué valores de $a \in \mathbb{R}$, obtenemos que $\alpha = \beta = \gamma = 0$; para esos valores de a el conjunto será LI, o equivalentemente, cuando a NO adquiera eso valores el conjunto será LD.

Operando y sacando factor común $1, x, x^2$ en la ecuación anterior, tenemos que:

$$1(\alpha + 4\beta + a\gamma) + x(a\alpha + 5\beta + \gamma) + x^2(3\alpha + 7\beta + \gamma) = 0 = \mathbf{0}(x) \text{ para todo } x \in \mathbb{R},$$

donde $\mathbf{0}$ denota al polinomio nulo (es decir al elemento neutro de $\mathbb{R}_2[x]$). Como sabemos que el conjunto $\{1, x, x^2\}$ es LI en $\mathbb{R}_2[x]$ y tenemos una CL de dichos vectores igualada al elemento neutro de $\mathbb{R}_2[x]$ vale que $\alpha + 4\beta + a\gamma = a\alpha + 5\beta + \gamma = 3\alpha + 7\beta + \gamma = 0$. Escrito de manera matricial, el sistema anterior es equivalente a

$$\begin{bmatrix} 1 & 4 & a \\ a & 5 & 1 \\ 3 & 7 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Como tenemos una matriz cuadrada de 3×3, sabemos que el determinante de dicha matriz es no nulo si y sólo si la solución del sistema anterior es la trivial ($\alpha = \beta = \gamma = 0$) si y sólo si el subconjunto de vectores iniciales es LI. Equivalentemente, el determinante de la matriz en cuestión es nulo si y sólo si la solución del sistema es no trivial si y sólo si el subconjunto de vectores iniciales es LD. Calculemos el determinante:

$\det(\begin{bmatrix} 1 & 4 & a \\ a & 5 & 1 \\ 3 & 7 & 1 \end{bmatrix}) = 7a^2 - 19a + 10$. Entonces $7a^2 - 19a + 10 = 0$ si y sólo si $a = \frac{5}{7}$ ó $a = 2$. Para eso valores de a el subconjunto $\{1+ax+3x^2, 4+5x+7x^2, a+x+x^2\}$ es LD o equivalentemente si $a \in \mathbb{R} \setminus \{\frac{5}{7}, 2\}$ el subconjunto $\{1 + ax + 3x^2, 4 + 5x + 7x^2, a + x + x^2\}$ es LI. ■

Observar que podemos usar el "método" del determinante (tal como hicimos en el ejercicio anterior) cuando el sistema a resolver tiene la misma cantidad de incógnitas que de ecuaciones (es decir nos quedan matrices cuadradas). Si ese no es el caso, el determinante no está definido y hay que buscar otras maneras de resolver el problema. Por ejemplo, triangulando la matriz que nos quedó (al plantear la independencia lineal por definición) y viendo para que valores de a una (o más) filas se anulan.

1.5. Wronskiano

En este sección vamos a estudiar una condición suficiente para determinar la independencia lineal de un conjunto de funciones. Sea $I \subseteq \mathbb{R}$ un intervalo y sean $\{f_1, f_2, \cdots, f_n\}$ un conjunto de

n funciones pertenecientes a $C^{n-1}(I)$ (es decir funciones continuas con $n-1$ derivadas continuas). Se define el *wronskiano* de $\{f_1, f_2, \cdots, f_n\}$ en $x_0 \in I$ como

$$W(f_1, \cdots, f_n)(x_0) = \det\left(\begin{bmatrix} f_1(x_0) & f_2(x_0) & \cdots & f_n(x_0) \\ f_1'(x_0) & f_2'(x_0) & \cdots & f_n'(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{n-1}(x_0) & f_2^{n-1}(x_0) & \cdots & f_n^{n-1}(x_0) \end{bmatrix}\right).$$

En este sección $\mathbf{0}$ denotará a la función nula (es decir al elemento neutro de $C(I)$).

Veamos un ejemplo. Sean $f_1(x) = x^2, f_2(x) = x, f_3(x) = 1$. Claramente, $\{f_1, f_2, f_3\}$ es un conjunto de 3 funciones pertenecientes a $C^2(\mathbb{R})$ (en realidad son infinitamente derivables con derivadas continuas). Entonces

$$W(f_1, f_2, f_3)(x) = \det(\begin{bmatrix} x^2 & x & 1 \\ 2x & 1 & 0 \\ 2 & 0 & 0 \end{bmatrix}) = -2.$$

El siguiente resultado nos da una condición suficiente para determinar la independencia lineal de un conjunto de n funciones en $C^{n-1}(I)$.

Teorema 1.5.1. *Sean $\{f_1, f_2, \cdots, f_n\}$ un conjunto de funciones pertenecientes a $C^{n-1}(I)$ (es decir funciones continuas con $n-1$ derivadas continuas) tal que existe $x_0 \in I$ tal que $W(f_1, \cdots, f_n)(x_0) \neq 0$. Entonces el conjunto $\{f_1, f_2, \cdots, f_n\}$ es LI.*

Dem. Vamos a probar el resultado usando la definición de independencia lineal. Sean $a_1, \cdots, a_n \in \mathbb{K}$ tales que

$$a_1 f_1 + a_2 f_2 + \cdots + a_n f_n = \mathbf{0}.$$

Queremos ver que $a_1 = a_2 = \cdots = a_n = 0$. Para probar esto, derivemos la ecuación anterior $n-1$ veces (lo podemos hacer porque las funciones involucradas pertenecen todas a $C^{n-1}(I)$). Nos queda el siguiente sistema de n ecuaciones y n incógnitas.

$$\begin{cases} a_1 f_1 + a_2 f_2 + \cdots + a_n f_n = \mathbf{0} \\ a_1 f_1' + a_2 f_2' + \cdots + a_n f_n' = \mathbf{0} \\ \vdots \\ a_1 f_1^{n-1} + a_2 f_2^{n-1} + \cdots + a_n f_n^{n-1} = \mathbf{0}. \end{cases}.$$

En particular, si evaluamos el sistema de ecuaciones anterior en x_0 tenemos que

$$\begin{cases} a_1 f_1(x_0) + a_2 f_2(x_0) + \cdots + a_n f_n(x_0) = \mathbf{0}(x_0) = 0 \\ a_1 f_1'(x_0) + a_2 f_2'(x_0) + \cdots + a_n f_n'(x_0) = \mathbf{0}(x_0) = 0 \\ \vdots \\ a_1 f_1^{n-1}(x_0) + a_2 f_2^{n-1}(x_0) + \cdots + a_n f_n^{n-1}(x_0) = \mathbf{0}(x_0) = 0. \end{cases}.$$

Sea

$$A := \begin{bmatrix} f_1(x_0) & f_2(x_0) & \cdots & f_n(x_0) \\ f_1'(x_0) & f_2'(x_0) & \cdots & f_n'(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{n-1}(x_0) & f_2^{n-1}(x_0) & \cdots & f_n^{n-1}(x_0) \end{bmatrix}.$$

Entonces, el sistema anterior, nos queda escrito matricialmente como

$$A \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Como $A \in \mathbb{R}^{n\times n}$ (está definido el determinante de A) y

$$\det(A) = \det\left(\begin{bmatrix} f_1(x_0) & f_2(x_0) & \cdots & f_n(x_0) \\ f_1'(x_0) & f_2'(x_0) & \cdots & f_n'(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{n-1}(x_0) & f_2^{n-1}(x_0) & \cdots & f_n^{n-1}(x_0) \end{bmatrix}\right) = W(f_1, \cdots, f_n)(x_0) \neq 0,$$

donde para la última igualdad usamos la hipótesis. Entonces, $\text{rg}(A) = n$ y por lo tanto, la solución del sistema anterior es la trivial. Es decir, $a_1 = a_2 = \cdots = a_n = 0$. Entonces, el conjunto $\{f_1, f_2, \cdots, f_n\}$ es LI. ■

Observar que el Teorema 1.5.1 NO es un “si y sólo si”. Es un teorema del tipo “si p entonces q”. Es decir, con ese teorema podemos decir que si el conjunto de funciones es LD, entonces el Wronskiano se anula en todo punto (“no q entonces no p”). El Teorema NO nos dice qué pasa si el Wronskiano se anula en todo punto. De hecho, el **Ejercicio 1.E** que veremos más adelante, es un ejemplo de un conjunto LI cuyo Wronskiano se anula en todo punto.

Veamos una aplicación del teorema anterior:

Ejercicio 1.11 b). Determinar cuáles de los siguientes conjuntos de funciones son L.I

- $\mathcal{G} = \{1 + 3\sin(x) - 2\cos(x), 3 + 5\sin(x) - 6\cos(x), -5\sin(x) + 6\cos(x)\}$.

Recordemos que, dadas $f, g : I \to \mathbb{R}$ (es decir dos funciones definidas en un intervalo $I \subseteq \mathbb{R}$ que toman valores en $\mathbb{R}$). Entonces

$$f = g \text{ si y sólo si } f(x) = g(x) \text{ para todo } x \in I.$$

Dem. Vamos a resolverlo de dos maneras. Usando el teorema anterior y por definición.

Para eso llamemos, $f(x) = 1 + 3\sin(x) - 2\cos(x)$, $g(x) = 3 + 5\sin(x) - 6\cos(x)$, $h(x) = -5\sin(x) + 6\cos(x)$.

Usando el Teorema anterior:

Claramente $f, g, h \in C^2(\mathbb{R})$ (de hecho son infinitamente derivables con derivadas continuas). Calculamos el wronskiano.

$$W(f,g,h)(x) =$$
$$= \det\left(\begin{bmatrix} 1 + 3\sin(x) - 2\cos(x) & 3 + 5\sin(x) - 6\cos(x) & -5\sin(x) + 6\cos(x) \\ 3\cos(x) + 2\sin(x) & 5\cos(x) + 6\sin(x) & -5\cos(x) - 6\sin(x) \\ -3\sin(x) + 2\cos(x) & -5\sin(x) + 6\cos(x) & 5\sin(x) - 6\cos(x) \end{bmatrix}\right).$$

Ahora deberíamos calcular el determinante de esa matriz de 3×3. Recordar que el Teorema 1.5.1 nos dice que basta con encontrar un punto donde el Wronskiano no se anule para probar que el conjunto es LI. Entonces podemos hacer eso, es decir, proponemos un valor (conveniente) y vemos qué pasa con el wronskiano en ese valor. De esa manera, nos ahorramos de calcular el Wronskiano de manera general para todo x. Por ejemplo, podemos tomar $x = 0$. Entonces,

$$W(f,g,h)(0) = \det\left(\begin{bmatrix} -1 & -3 & 6 \\ 3 & 5 & -5 \\ 2 & 6 & -6 \end{bmatrix}\right) = 24 \neq 0.$$

Como encontramos un punto (en este caso $x = 0$) donde W no se anula, por el Teorema 1.5.1, concluimos que el conjunto es LI.

Por definición: Sean $a, b, c \in \mathbb{R}$ tales que

$$af + bg + ch = \mathbf{0}. \tag{1.4}$$

Veamos que $a = b = c = 0$. Observar que en la ecuación anterior $\mathbf{0}$ representa a la función nula, es decir la función tal que $\mathbf{0}(x) = 0$ para todo $x \in \mathbb{R}$. Vimos que $\mathbf{0}$ es el elemento neutro del $\mathbb{R}$-espacio vectorial $C(\mathbb{R})$ (de manera similar se puede probar que $\mathbf{0}$ es el elemento de neutro de $C^n(\mathbb{R})$ para cualquier $n \in \mathbb{N}$). Observar que en (1.4) tenemos una igualdad de funciones. Entonces, usando lo que vimos arriba para igualdad de funciones, tenemos que vale que:

$$af(x) + bg(x) + ch(x) = \mathbf{0}(x) = 0 \text{ para todo } x \in \mathbb{R}.$$

Donde ahora el 0 obviamente es el número real 0. Escribiendo la expresión de las funciones, nos queda:

$$a(1 + 3\sin(x) - 2\cos(x)) + b(3 + 5\sin(x) - 6\cos(x)) + c(-5\sin(x) + 6\cos(x)) = 0 \text{ para todo } x \in \mathbb{R}.$$

Usando esa expresión, tenemos que probar que $a = b = c = 0$. Como la ecuación anterior vale para todo $x \in \mathbb{R}$, podemos proponer 3 valores de x y evaluar dicha ecuación en esos valores. Si hacemos eso, nos quedará un sistema de 3 ecuaciones con 3 incógnitas, resolvemos el sistema y vemos si su solución es la trivial.

Por ejemplo, si proponemos $x = 0$, nos queda: $-a - 3b + 6c = 0$. Si proponemos, $x = \frac{\pi}{2}$, nos queda $4a + 8b - 5c = 0$. Por último, si proponemos $x = -\frac{\pi}{2}$, nos queda $-2a - 2b + 5c = 0$.

Es decir, nos queda el sistema: $\begin{cases} -a - 3b + 6c = 0 \\ 4a + 8b - 5c = 0 \\ -2a - 2b + 5c = 0 \end{cases}$.

Resolviendo, nos queda que $a = b = c = 0$, es decir la única solución es la trivial. Por lo tanto el conjunto $\{f, g, h\}$ es LI. ∎

Ejercicio 1.D. Supongamos que con los 3 valores de x que propusimos en la resolución del ejercicio anterior (por definición) hubieramos obtenido un sistema de ecuaciones indeterminado. En ese caso, ¿se puede concluir entonces que el sistema es LD?

Spoiler alert: NO. Tratar de meditar esto.

Ejercicio 1.E. Sean $f, g \in C^1(\mathbb{R})$ las funciones definidas por $f(x) := x^3$ y $g(x) := |x|^3$. Comprobar que $\{f, g\}$ es un conjunto LI pero el Wronskiano de dicho conjunto es nulo para todo $x \in \mathbb{R}$.

Dem. Notar que $g = x^3$ si $x \geq 0$ y $g(x) = -x^3$ si $x < 0$. Además, es claro que $g'(x) = 3x^2$ si $x > 0$, $g'(x) = -3x^2$ si $x < 0$ y, por definición, podemos probar que $g'(0) = 0$. Entonces,

$$W(f, g)(x) = \det(\begin{bmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{bmatrix}).$$

Calculemos el Wronskiano para los distintos valores que puede tomar x. Si $x > 0$, tenemos que

$$W(f, g)(x) = \det(\begin{bmatrix} x^3 & x^3 \\ 3x^2 & 3x^2 \end{bmatrix}) = 0.$$

Si $x < 0$, tenemos que

$$W(f, g)(x) = \det(\begin{bmatrix} x^3 & -x^3 \\ 3x^2 & -3x^2 \end{bmatrix}) = 0.$$

Si $x = 0$,

$$W(f, g)(0) = \det(\begin{bmatrix} f(0) & g(0) \\ f'(0) & g'(0) \end{bmatrix}) = 0.$$

Por lo tanto, $W(f, g)(x) = 0$ para todo $x \in \mathbb{R}$. Sin embargo, el conjunto $\{f, g\}$ es LI. De hecho, sean $a, b \in \mathbb{R}$ tales que

$$af + bg = \mathbf{0},$$

donde **0** denota la función nula (el elemento neutro del espacio vectorial $C^1(\mathbb{R})$). Entonces

$$af(x) + bg(x) = ax^3 + b|x|^3 = 0 \text{ para todo } x \in \mathbb{R}.$$

En particular, si $x = 1$, tenemos que $a(1)^3 + b|1|^3 = a + b = 0$ y, si $x = -1$, tenemos que $a(-1)^3 + b|-1|^3 = -a + b = 0$. Entonces, $a + b = a - b = 0$, por lo tanto $a = b = 0$ y el conjunto $\{f, g\}$ resulta LI. ■

Ejercicio 1.F (De Parcial)**.** Observar que el siguiente conjunto de funciones está contenido en el $\mathbb{R}$-espacio vectorial $C^\infty(\mathbb{R})$ y comprobar que es linealmente independiente.

- $\{e^{\lambda_1 x}, e^{\lambda_2 x}, \cdots, e^{\lambda_n x}\}$, donde $\lambda_1 < \lambda_2 < \cdots < \lambda_n$.

Observar que si $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ como $\lambda_i - \lambda_n < 0$ para todo $i \in \{1, 2, \cdots, n-1\}$ vale que

$$\lim_{x\to\infty} e^{(\lambda_i - \lambda_n)x} = 0,$$

para todo $i \in \{1, 2, \cdots, n-1\}$. Este es un resultado visto en el CBC en Análisis Matemático (si no están convencidos, vuelvan a repasar eso).

Más aún, si $a_1, \cdots, a_{n-1} \in \mathbb{R}$, usando las propiedades de límite (de sucesiones convergentes) también vale que

$$\lim_{x\to\infty} \sum_{i=1}^{n-1} a_i e^{(\lambda_i - \lambda_n)x} = 0.$$

Dem. Claramente $\{e^{\lambda_1 x}, e^{\lambda_2 x}, \cdots, e^{\lambda_n x}\} \subseteq C^\infty(\mathbb{R})$, ya que todas las funciones del conjunto en cuestión son infinitamente derivables con derivadas continuas.

A continuación vamos a probar que el conjunto dado es LI por inducción en i.

Si $i = 1$, el conjunto $\{e^{\lambda_1 x}\}$ es LI, porque tiene un sólo elemento (que no es la función nula).

Si $i = 2$ (no es necesario probar este caso pero lo vamos a hacer para fijar ideas) el conjunto $\{e^{\lambda_1 x}, e^{\lambda_2 x}\}$ también es LI. Supongamos que no, es decir supongamos que dicho conjunto es LD. Entonces existe $a_1 \in \mathbb{R}$ tal que

$$e^{\lambda_2 x} = a_1 e^{\lambda_1 x} \text{ para todo } x \in \mathbb{R}.$$

Equivalentemente, como $e^{\lambda_2 x} > 0$ para todo x (es decir nunca se anula), tenemos que

$$1 = a_1 \frac{e^{\lambda_1 x}}{e^{\lambda_2 x}} = a_1 e^{(\lambda_1 - \lambda_2)x} \text{ para todo } x \in \mathbb{R}.$$

Ahora, tomando límite a ambos lados de la igualdad anterior, tenemos que

$$1 = \lim_{x\to\infty} 1 = \lim_{x\to\infty} a_1 e^{(\lambda_1 - \lambda_2)x} = 0,$$

donde usamos lo que probamos arriba. Por lo tanto llegamos a un absurdo y el conjunto $\{e^{\lambda_1 x}, e^{\lambda_2 x}\}$ no puede ser LD. Entonces es LI.

Hipótesis inductiva (HI): para $i = n-1$, el conjunto $\{e^{\lambda_1 x}, e^{\lambda_2 x}, \cdots, e^{\lambda_{n-1} x}\}$ es LI.

Con la HI, vamos a probar que para $i = n$ el conjunto $\{e^{\lambda_1 x}, e^{\lambda_2 x}, \cdots, e^{\lambda_{n-1} x}, e^{\lambda_n x}\}$ también es LI. Supongamos que no, es decir que dicho conjunto es LD. Entonces, por hipótesis inductiva, no queda otra que $e^{\lambda_n x}$ sea una CL de los elementos de $\{e^{\lambda_1 x}, e^{\lambda_2 x}, \cdots, e^{\lambda_{n-1} x}\}$ (antes de continuar, pensar por qué pasa esto). Es decir, existen $a_1, \cdots, a_{n-1} \in \mathbb{R}$, tales que

$$e^{\lambda_n x} = a_1 e^{\lambda_1 x} + a_2 e^{\lambda_2 x} + \cdots + a_{n-1} e^{\lambda_{n-1} x}.$$

Operando de la misma manera que para el caso $i = 2$, tenemos que

$$1 = \frac{a_1 e^{\lambda_1 x} + a_2 e^{\lambda_2 x} + \cdots + a_{n-1} e^{\lambda_{n-1} x}}{e^{\lambda_n x}} = \sum_{i=1}^{n-1} a_i e^{(\lambda_i - \lambda_n)x} \text{ para todo } x \in \mathbb{R}.$$

Ahora, tomando límite a ambos lados de la igualdad anterior (como en el caso $i = 2$), tenemos que

$$1 = \lim_{x\to\infty} 1 = \lim_{x\to\infty} \sum_{i=1}^{n-1} a_i e^{(\lambda_i-\lambda_n)x} = 0,$$

donde usamos lo que probamos arriba. Por lo tanto llegamos a un absurdo y el conjunto $\{e^{\lambda_1 x}, e^{\lambda_2 x}, \cdots, e^{\lambda_{n-1}x}, e^{\lambda_n x}\}$ no es LD. Por lo tanto es LI y probamos lo que queríamos. ■

1.6. Bases de espacios vectoriales

Definición. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial, $\mathcal{S} \subseteq \mathbb{V}$ un subespacio. El conjunto $\{v_1, v_2, \cdots, v_r\}$ es una *base* de $\mathcal{S}$ si

i) $\mathcal{S} = gen\{v_1, v_2, \cdots, v_r\}$,

ii) el conjunto $\{v_1, v_2, \cdots, v_r\}$ es LI.

En este caso, se define la dimensión de $\mathcal{S}$ como $\dim(\mathcal{S}) = \#\{v_1, v_2, \cdots, v_r\} = r$, es decir el cardinal (el número de elementos) de alguna base de $\mathcal{S}$.

Observar que de la definición de base se sigue que $v_1, v_2, \cdots, v_r \in \mathcal{S}$.

Ejercicio 1.5 a). Hallar una base y determinar la dimensión de cada uno de los siguientes subespacios:

- $\mathcal{S} = \{p \in \mathbb{R}_2[x] : p(1) = 0\}$.

Dem. Si $p \in \mathcal{S}$, entonces $p \in \mathbb{R}_2[x]$ es decir $p(x) = ax^2 + bx + c$, con $a, b, c \in \mathbb{R}$ y además, $0 = p(1) = a \cdot 1^2 + b \cdot 1 + c = a + b + c$. Despejando, nos queda (por ejemplo) $a = -b - c$. Por lo tanto,

$$p(x) = ax^2 + bx + c = (-b - c)x^2 + bx + c = b(-x^2 + x) + c(-x^2 + 1) \text{ con } b, c \in \mathbb{R}.$$

Por lo tanto, $\mathcal{S} = \{p \in \mathbb{R}_2[x] : p(x) = b(-x^2+x)+c(-x^2+1) \text{ con } b, c \in \mathbb{R}\} = gen\{-x^2+x, -x^2+1\}$. Una base de $\mathcal{S}$, podría ser

$$B_{\mathcal{S}} = \{-x^2 + x, -x^2 + 1\}.$$

Entonces, $\dim(\mathcal{S}) = 2$. ■

Ejercicio 1.G. Hallar una base y determinar la dimensión del siguiente subespacio:

- $\mathcal{S} = \{p \in \mathbb{R}[x] : p^{(n)} = 0\}$. Con n algún número natural.

Dem. Primero entendamos qué elementos de $\mathbb{R}[x]$ viven en $\mathcal{S}$. Observar que $p \in \mathbb{R}[x]$ pertenece a $\mathcal{S}$ si y sólo si $p^{(n)} = 0$. Donde $p^{(n)}$ denota la derivada n-ésima de p y 0 denota al polinomio nulo, es decir el polinomio que vale 0 en todo punto (en otros ejercicios lo hemos notado como $\mathbf{0}$). Como verán, tenemos en realidad una igualdad de polinomios (de funciones), es decir, $p \in \mathcal{S}$ si y sólo si $p^{(n)}(x) = 0$ para todo $x \in \mathbb{R}$, donde ahora 0 denota al número real 0.

Paréntesis para ver si se entendió lo de arriba: Es crucial entender cuándo el 0 denota una función y cuando denota un número real dado que la notación del ejercicio no diferencia entre ambos casos, pero por contexto lo podemos inferir. Es decir como $p^{(n)}$ es un polinomio, no tendría sentido la igualdad $p^{(n)} = 0$ con 0 un número real, porque un polinomio y un número no se comparan. En cambio, como $p^{(n)}(x)$ es un número real (es decir al evaluar un polinomio en un punto x obtenemos un número real) en la igualdad $p^{(n)}(x) = 0$, el símbolo 0 denota al número real 0. **Fin del paréntesis**

En conclusión, $\mathcal{S}$ contiene a todos los polinomios de grado menor o igual a $n-1$ y al polinomio nulo (que técnicamente no tiene grado). Esto es así porque si el grado de p es $n-1$ (o menor), su derivada n-ésima es nula en todo x; por el contrario, si el grado de p es n o mayor, su derivada n-ésima no se anula. Conclusión

$$\mathcal{S} = \mathbb{R}_{n-1}[x] = \{p \in \mathbb{R}[x] : grado(p) \leq n-1\} \cup \{0\}.$$

Una base de $\mathcal{S}$ podría ser $B_{\mathcal{S}} = \{1, x, \cdots, x^{n-1}\}$ y $\dim(\mathcal{S}) = n$. ■

Ejercicio 1.H (De Parcial)**.** Consideremos $\mathbb{V} = \mathbb{C}^n$ con cuerpo $\mathbb{K} = \mathbb{R}$ y sean $v_1, v_2, \cdots, v_m \in \mathbb{R}^n$ (m vectores con todas sus componentes reales). Demostrar que el conjunto $B = \{v_1, v_2, \cdots, v_m\}$ no puede ser base de $\mathbb{V}$.

Hay muchas maneras de demostrar este ejercicio. Propongo que piensen otras formas alternativas.

Dem. Vamos a probarlo por el absurdo. Supongamos que B es una base de $\mathbb{V}$. Entonces, dado $v \in \mathbb{C}^n$, existen $a_1, \cdots, a_m \in \mathbb{K} = \mathbb{R}$ tales que

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_m v_m.$$

Entonces $v \in \mathbb{R}^n$ puesto que es la combinación lineal real de m vectores de $\mathbb{R}^n$. Como v era cualquiera concluimos entonces que $\mathbb{C}^n \subseteq \mathbb{R}^n$ (básicamente porque cualquier vector v de $\mathbb{C}^n$ pertenece también a $\mathbb{R}^n$). Eso claramente es absurdo, por ejemplo el vector $v = [i\ 0\ \cdots\ 0]^T \in \mathbb{C}^n$ pero $v \notin \mathbb{R}^n$ (la primera componente es no real). El absurdo se produjo por creer que B es una base de $\mathbb{V}$. Por lo tanto B no puede ser base de $\mathbb{V}$. ■

1.7. Igualdad de subespacios

Recordemos que dados A, B dos conjuntos, entonces:

- $A \subseteq B,$ si para cada $x \in A$ vale que $x \in B$.
- $A \subsetneq B,$ si para cada $x \in A$ vale que $x \in B$ y existe $z \in B$ tal que $z \notin A$.
- $A = B,$ si $A \subseteq B$ y $B \subseteq A$.

Si ahora los conjuntos $\mathcal{S}, \mathcal{T}$ son además **subespacios** de un $\mathbb{K}$-espacio vectorial, vale la siguiente propiedad que usaremos ampliamente:

Teorema 1.7.1. *Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial, $\mathcal{S}, \mathcal{T} \subseteq \mathbb{V}$ dos subespacios. Entonces $\mathcal{S} = \mathcal{T}$ si y sólo si $\mathcal{S} \subseteq \mathcal{T}$ y $\dim(\mathcal{S}) = \dim(\mathcal{T})$.*

Dem. Una de las implicaciones es obvia. Porque si $\mathcal{S} = \mathcal{T}$ entonces claramente $\mathcal{S} \subseteq \mathcal{T}$ y $\dim(\mathcal{S}) = \dim(\mathcal{T})$.

Recíprocamente, si $\mathcal{S} \subseteq \mathcal{T}$ y $\dim(\mathcal{S}) = \dim(\mathcal{T})$. Veamos que $\mathcal{S} = \mathcal{T}$. Lo vamos a probar por el absurdo. Supongamos que $\mathcal{S} \neq \mathcal{T}$, esto significa que $\mathcal{S} \subsetneq \mathcal{T}$ (si no son iguales, como por hipótesis $\mathcal{S} \subseteq \mathcal{T}$ no queda otra que $\mathcal{S}$ esté totalmente contenido en $\mathcal{T}$). Es decir, existe un vector $z \in \mathcal{T}$ pero $z \notin \mathcal{S}$. Supongamos que $\dim(\mathcal{S}) = r$ y sea $\{s_1, \cdots, s_r\}$ una base de $\mathcal{S}$. Entonces el conjunto $\{s_1, \cdots, s_r, z\}$ es LI (eso es porque $z \notin \mathcal{S}$, meditar por qué pasa eso) y además $\{s_1, \cdots, s_r, z\} \subseteq \mathcal{T}$ puesto que $s_1, \cdots, s_r \in \mathcal{S} \subseteq \mathcal{T}$ y teníamos que $z \in \mathcal{T}$. Entonces $\mathcal{T}$ contiene al menos $r+1$ vectores LI, por lo tanto $\dim(\mathcal{T}) \geq r+1$, pero $\dim(\mathcal{T}) = \dim(\mathcal{S}) = r$, lo cual es absurdo ($r$ no es mayor o igual que $r+1$). Por lo tanto $\mathcal{S} = \mathcal{T}$. ■

A continuación veremos dos ejercicios donde aplicaremos el Teorema 1.7.1.

Ejercicio 1.I (De Parcial)**.** Sea $\mathbb{V}$ un $\mathbb{K}-$espacio vectorial tal que $\dim(\mathbb{V}) = n$. Demostrar que:

i) n vectores LI de $\mathbb{V}$ forman una base de $\mathbb{V}$.

ii) n vectores que generan $\mathbb{V}$ forman una base de $\mathbb{V}$.

Dem. Veamos el item i). El otro hacerlo de ejercicio.

Sea $\{v_1, \cdots, v_n\}$ algún conjunto de n vectores de $\mathbb{V}$ linealmente independientes. Entonces, $\dim(gen\{v_1, \cdots, v_n\}) = n = \dim(\mathbb{V})$. Como claramente $gen\{v_1, \cdots, v_n\}$ y $\mathbb{V}$ son subespacios y $gen\{v_1, \cdots, v_n\} \subseteq \mathbb{V}$, por el el Teorema 1.7.1, tenemos que $gen\{v_1, \cdots, v_n\} = \mathbb{V}$. Entonces el conjunto $\{v_1, \cdots, v_n\}$ es una base de $\mathbb{V}$. ■

Ejercicio 1.J (De Parcial)**.** Sean $\mathcal{S} = \{p \in \mathbb{R}_3[x] : p(1) = p'(1) = 0\}$ y $\mathcal{T} = gen\{x^2 + mx + 1, x^3 + x^2 - 5x + (5 + m), -x^3 + (m + 3)x^2 + x - 1\}$. Hallar $m \in \mathbb{R}$ (si existe) tal que $\mathcal{S} = \mathcal{T}$.

Dem. Como $\mathcal{S}$ y $\mathcal{T}$ son subespacios, podemos usar el Teorema 1.7.1 y probar para qué valores de m se cumple que $\mathcal{T} \subseteq \mathcal{S}$ y $\dim(\mathcal{S}) = \dim(\mathcal{T})$. Si pasa eso, entonces por el resultado anterior, $\mathcal{S} = \mathcal{T}$.

Llamemos $p(x) = x^2 + mx + 1,\ q(x) = x^3 + x^2 - 5x + (5 + m),\ r(x) = -x^3 + (m + 3)x^2 + x - 1$. Entonces $\mathcal{T} \subseteq \mathcal{S}$ si y sólo si $p, q, r \in \mathcal{S}$ (¿por qué?).

Verifiquemos si $p \in \mathcal{S}$. Evaluando $p(1) = m + 2 = 0$ entonces $m = -2$ y $p'(1) = 2(1) + m = 2 - 2 = 0$. Entonces $m = -2$ (si m toma otro valor, no hay manera de que $\mathcal{T} \subseteq \mathcal{S}$). Reemplazando $m = -2$, tenemos que $p(x) = x^2 - 2x + 1,\ q(x) = x^3 + x^2 - 5x + 3,\ r(x) = -x^3 + x^2 + x - 1$. Falta verificar que $q, r \in \mathcal{S}$ (si eso no ocurre entonces no existe tal m y nunca podría pasar que $\mathcal{S} = \mathcal{T}$). Verificando, se ve que $q(1) = 0$, $q'(1) = 0$ y $r(1) = 0$, $r'(1) = 0$. Entonces, si $m = -2$, $\mathcal{T} \subseteq \mathcal{S}$. Falta verificar que para ese valor de m, vale que $\dim(\mathcal{S}) = \dim(\mathcal{T})$. En ese caso, observar $\dim(\mathcal{S}) = 2$, pues $\mathcal{S} = gen\{x^3 - 3x + 2, x^2 - 2x + 1\}$. Además, como $\frac{q+r}{2} = p$, vale que $\mathcal{T} = gen\{p, q, r\} = gen\{q, r\}$. Como q y r no son multiplos entonces son LI y $\dim(\mathcal{T}) = 2 = \dim(\mathcal{S})$. Entonces, si $m = -2$ tenemos que $\mathcal{S} = \mathcal{T}$. ■

1.8. Ecuaciones diferenciales de segundo orden

A continuación comenzaremos a estudiar ecuaciones diferenciales lineales homogéneas de orden 2 que tienen la forma

$$y'' + a_1 y' + a_0 y = 0,$$

con $a_0, a_1 \in \mathbb{R}$. En el siguiente ejercicio analizaremos las soluciones de dicha ecuación diferencial.

Ejercicio 1.17. Sean $a_0, a_1 \in \mathbb{R}$. Sabiendo que el espacio solución de la ecuación diferencial lineal homogénea de orden 2

$$y'' + a_1 y' + a_0 y = 0, \tag{1.5}$$

tiene dimensión 2, comprobar que

a) cualesquiera sean $a, b \in \mathbb{R}$ con $b \neq a$, el conjunto de funciones $\{e^{ax}, e^{bx}\}$ es una base del espacio solución de la ecuación $y'' - (a + b)y' + aby = 0$;

(b) para cualquier $a \in \mathbb{R}$, el conjunto de funciones $\{e^{ax}, xe^{ax}\}$ es una base del espacio solución de la ecuación $y'' - 2ay' + a^2 y = 0$;

(c) cualesquiera sean $a, b \in \mathbb{R}$ con $b \neq 0$, el conjunto de funciones $\{e^{ax}\cos(bx), e^{ax}\sin(bx)\}$ es una base del espacio solución de la ecuación $y'' - 2ay' + (a^2 + b^2)y = 0$.

Dem. $a)$: Llamemos $\mathcal{S} := \{y \in C^\infty(\mathbb{R}) : y'' - (a + b)y' + aby = 0\}$. Es decir, $\mathcal{S}$ es el conjunto de soluciones de la ecuación diferencial $y'' - (a + b)y' + aby = 0$. Por un lado, es claro que $\mathcal{S}$ es un subespacio y por el otro lado, tomando $a_1 = -(a + b)$ y $a_0 = ab$ en (1.5), se sigue que $\mathcal{S}$ es un subespacio de dimensión 2.

Notar que las funciones e^{ax} y e^{bx} son soluciones de la ecuación $y'' - (a+b)y' + aby = 0$. De hecho,

$$(e^{ax})'' - (a+b)(e^{ax})' + ab(e^{ax}) = a^2e^{ax} - a(a+b)e^{ax} + ab(e^{ax}) = 0.$$

De la misma manera se prueba que e^{bx} es una solución de la ecuación $y'' - (a+b)y' + aby = 0$. Por lo tanto $e^{ax} \in \mathcal{S}$ y $e^{bx} \in \mathcal{S}$. Veamos que el conjunto $\{e^{ax}, e^{bx}\}$ es LI, podemos probarlo calculando el Wronskiano (por ejemplo):

$$W(e^{ax}, e^{bx})(x) = \det(\begin{bmatrix} e^{ax} & e^{bx} \\ ae^{ax} & be^{bx} \end{bmatrix}) = (b-a)e^{(a+b)x} \neq 0,$$

donde usamos que $e^{(a+b)x} > 0$ para todo $x \in \mathbb{R}$ y $b \neq a$. Por lo tanto, $\{e^{ax}, e^{bx}\}$ es un conjunto LI. Conclusión: tenemos 2 vectores LI de $\mathcal{S}$ (que es un subespacio de dimensión 2). Entonces, por el Ejercicio 1.I, el conjunto $\{e^{ax}, e^{bx}\}$ es una base de $\mathcal{S}$ o equivalentemente, $\{e^{ax}, e^{bx}\}$ es una base del espacio solución de la ecuación $y'' - (a+b)y' + aby = 0$.

b) : Se resuelve igual que el item *a*) viendo que los vectores e^{ax} y xe^{ax} son soluciones de la ecuación $y'' - 2ay' + a^2y = 0$ y observando que el conjunto $\{e^{ax}, xe^{ax}\}$ es un conjunto LI (pueden probarlo calculando el Wronskiano por ejemplo). Tomando, $a_1 = -2a$ y $a_0 = a^2$ en (1.5), tenemos que el subespacio que es el conjunto de soluciones de la ecuación $y'' - 2ay' + a^2y = 0$ tiene dimensión 2. Por lo tanto, por el Ejercicio 1.I, el conjunto $\{e^{ax}, xe^{ax}\}$ es una base del espacio solución de la ecuación $y'' - 2ay' + a^2y = 0$.

c) : Se resuelve igual que el item *a*) viendo que los vectores $e^{ax}\cos(bx)$ y $e^{ax}\sin(bx)$ son soluciones de la ecuación $y'' - 2ay' + (a^2 + b^2)y = 0$ y observando que, como $b \neq 0$, el conjunto $\{e^{ax}\cos(bx), e^{ax}\sin(bx)\}$ es un conjunto LI (pueden probarlo calculando el Wronskiano por ejemplo). Tomando $a_1 = -2a$ y $a_0 = a^2 + b^2$ en (1.5), tenemos que el subespacio que es el conjunto de soluciones de la ecuación $y'' - 2ay' + (a^2 + b^2)y = 0$ tiene dimensión 2. Por lo tanto, por el Ejercicio 1.I, el conjunto $\{e^{ax}\cos(bx), e^{ax}\sin(bx)\}$ es una base del espacio solución de la ecuación $y'' - 2ay' + (a^2 + b^2)y = 0$. ■

Veamos una aplicación del **Ejercicio 1.17**.

Ejercicio 1.18 b). Hallar la solución $y \in C^\infty(\mathbb{R})$ de la ecuación diferencial indicada que satisface las siguientes condiciones: $y(0) = 1$, $y'(0) = 1$.

- $y'' + 4y' + 4y = 0$.

Dem. Notar que la ecuación diferencial dada la podemos pensar como

$$y'' - 2(-2)y + (-2)(-2)y = 0.$$

Es decir, la ecuación diferencial es de la forma $y'' - 2ay' + a^2y = 0$ con $a = -2$. Entonces, por el **Ejercicio 1.17**, $\{e^{-2x}, xe^{-2x}\}$ es una base del espacio solución de la ecuación $y'' + 4y' + 4y = 0$. Es decir, todas las soluciones de la ecuación diferencial son de la forma

$$y(x) = \alpha e^{-2x} + \beta xe^{-2x} \text{ con } \alpha, \beta \in \mathbb{R}.$$

Buscamos de todas las soluciones, aquellas que cumplen que $y(0) = 1$, $y'(0) = 1$. Es decir, $1 = y(0) = \alpha e^{-2\cdot 0} + \beta 0 e^{-2\cdot 0} = \alpha$ y como $y'(x) = -2\alpha e^{-2x} + \beta(e^{-2x} - 2xe^{-2x})$, tenemos que $1 = y'(0) = -2\alpha + \beta = -2 \cdot 1 + \beta$. Entonces $\beta = 3$. Por lo tanto la solución de la ecuación diferencial que satisface $y(0) = 1$, $y'(0) = 1$ es

$$y(x) = e^{-2x} + 3xe^{-2x}.$$

■

1.9. Subespacios fundamentales de una matriz

Definición. Sea $A \in \mathbb{K}^{n\times m}$ (es decir una matriz de $n \times m$ cuyas componentes son elementos de $\mathbb{K}$). Supongamos que $A = [A_1 \ A_2 \ \cdots \ A_m]$, donde $A_1, A_2, \cdots, A_m \in \mathbb{K}^n$ son las columnas de A. Se define:

- El *espacio nulo de* A: $\operatorname{nul}(A) := \{x \in \mathbb{K}^m : Ax = 0_{\mathbb{K}^n}\} \subseteq \mathbb{K}^m$,
- El *espacio columna de* A: $\operatorname{col}(A) := gen\{A_1, A_2, \cdots, A_m\} \subseteq \mathbb{K}^n$,
- El *espacio fila de* A: $\operatorname{fil}(A) := \operatorname{col}(A^T) \subseteq \mathbb{K}^m$.

Dada $A \in \mathbb{K}^{n\times m}$, observar que el rango de A, es decir el número de columnas (o filas) de A linealmente independientes (lo notamos $\operatorname{rg}(A)$), coincide con la dimensión del subespacio $\operatorname{col}(A)$. Es decir $\operatorname{rg}(A) = \dim(\operatorname{col}(A))$.

Recordemos el Teorema de la dimensión para matrices que usaremos ampliamente en estos ejercicios.

Teorema 1.9.1. *Sea $A \in \mathbb{R}^{n\times m}$. Entonces,*

$$\dim(\operatorname{nul}(A)) + \operatorname{rg}(A) = m.$$

Sea $A \in \mathbb{C}^{n\times m}$ y supongamos que estamos trabajando en el cuerpo $\mathbb{K} = \mathbb{C}$. Entonces,

$$\dim(\operatorname{nul}(A)) + \operatorname{rg}(A) = m.$$

Sea $A \in \mathbb{C}^{n\times m}$ y supongamos que estamos trabajando en el cuerpo $\mathbb{K} = \mathbb{R}$. Entonces,

$$\dim(\operatorname{nul}(A)) + \operatorname{rg}(A) = 2m.$$

En el caso en que $A \in \mathbb{C}^{n\times m}$, para el teorema de la dimensión debemos discriminar en los casos en que estamos trabajando en el cuerpo $\mathbb{K} = \mathbb{C}$ ó $\mathbb{K} = \mathbb{R}$. Esto es así, porque en el primer caso, $\dim(\mathbb{C}^m_{\mathbb{C}}) = m$ y en el segundo caso $\dim(\mathbb{C}^m_{\mathbb{R}}) = 2m$ y eso se ve reflejado en dicho teorema.

La demostración del Teorema 1.9.1 se vio en el CBC, de todas maneras, más adelante enunciaremos y demostraremos una generalización de este teorema para transformaciones lineales.

Proposición 1.9.2. *Sea $A \in \mathbb{K}^{m\times n}$ y $b \in \mathbb{K}^n$. Entonces, si existe solución del sistema $Ax = b$, todas las soluciones x_s del sistema $Ax = b$ se pueden expresar como*

$$x_s = x_p + x_h,$$

donde x_p es una solución particular (es decir $Ax_p = b$) y $x_h \in \operatorname{nul}(A)$ (es decir x_h es solucion del sistema homogéneo $Ax = 0$).

Dem. Primero veamos que efectivamente x_s resuelve el sistema $Ax = b$. Esto es así pues, $Ax_s = A(x_p + x_h) = Ax_p + Ax_h = b + 0 = b$.

Por útlimo, veamos que toda solución del sistema se puede escribir como queremos. Sea x_p una solución particular del sistema $Ax = b$ (que existe por hipótesis), entonces $Ax_p = b$. Por otra parte, supongamos que x_s es cualquier solución del sistema $Ax = b$. Entonces $Ax_s = b$. Restando estas dos ecuaciones, nos queda que $0 = b - b = Ax_s - Ax_p = A(x_s - x_p)$, entonces si llamamos $x_h := x_s - x_p$, claramente $x_h \in \operatorname{nul}(A)$. Finalmente, $x_s = x_p + (x_s - x_p) = x_p + x_h$, con x_p una solución particular y $x_h \in \operatorname{nul}(A)$, y obtenemos lo que queríamos probar. ■

El siguiente ejercicio parece sencillo a primera vista, pero requiere bastante atención.

Ejercicio 1.20. Sea $A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$. Hallar una base de cada uno de los cuatro subespacios fundamentales de A.

Sea $b = \begin{bmatrix} 2-3i \\ 3+2i \end{bmatrix}$. ¿Existe $x \in \mathbb{C}^2$ tal que $Ax = b$? Si la respuesta es afirmativa, hallar todas las soluciones del sistema $Ax = b$.

Dem.

- Primero vamos a obtener el subespacio $\mathrm{col}(A)$. Fácilmente vemos que $\mathrm{col}(A) = gen\{\begin{bmatrix} 1 \\ i \end{bmatrix}, \begin{bmatrix} i \\ -1 \end{bmatrix}\}$. La dificultad del ejercicio está en que no se aclara el cuerpo donde se está trabajando. Veremos que las respuestas serán distintas si tomamos como cuerpo $\mathbb{R}$ ó $\mathbb{C}$. Si pensamos a $\mathbb{C}^2$ como $\mathbb{C}$-espacio vectorial, como $\begin{bmatrix} 1 \\ i \end{bmatrix} = -i \begin{bmatrix} i \\ -1 \end{bmatrix}$, entonces $\mathrm{col}(A) = gen\{\begin{bmatrix} 1 \\ i \end{bmatrix}\}$ y una base de $\mathrm{col}(A)$ puede ser $B_{\mathrm{col}(A)} = \{\begin{bmatrix} 1 \\ i \end{bmatrix}\}$. Por otra parte, si pensamos a $\mathbb{C}^2$ como $\mathbb{R}$-espacio vectorial, no existe $a \in \mathbb{R}$ tal que $\begin{bmatrix} 1 \\ i \end{bmatrix} = a \begin{bmatrix} i \\ -1 \end{bmatrix}$, por lo tanto el conjunto $\{\begin{bmatrix} 1 \\ i \end{bmatrix}, \begin{bmatrix} i \\ -1 \end{bmatrix}\}$ el LI y en este caso, una base de $\mathrm{col}(A)$ puede ser $B_{\mathrm{col}(A)} = \{\begin{bmatrix} 1 \\ i \end{bmatrix}, \begin{bmatrix} i \\ -1 \end{bmatrix}\}$.
- Para el calculo de $\mathrm{nul}(A)$ también tendremos que pensar en esos dos casos. Si pensamos a $\mathbb{C}^2$ como $\mathbb{C}$-espacio vectorial, como en este caso $\dim(\mathbb{C}^2_{\mathbb{C}}) = 2$, el Teorema de la Dimensión 1.9.1, nos dice que $\dim(\mathrm{nul}(A)) + \mathrm{rg}(A) = 2$. Entonces, $\dim(\mathrm{nul}(A)) = 2 - \mathrm{rg}(A) = 2 - 1 = 1$. Entonces, como (por ejemplo) el vector $\begin{bmatrix} 1 \\ i \end{bmatrix} \in \mathrm{nul}(A)$, tenemos que $\mathrm{nul}(A) = gen\{\begin{bmatrix} 1 \\ i \end{bmatrix}\}$.

  Otra forma de encontrar $\mathrm{nul}(A)$ (en este caso donde pensamos a $\mathbb{C}^2$ como $\mathbb{C}$-espacio vectorial) es resolver el sistema homogéneo $Ax = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Conclusión, una base de $\mathrm{nul}(A)$ podría ser $B_{\mathrm{nul}(A)} = \{\begin{bmatrix} 1 \\ i \end{bmatrix}\}$.

  Por otra parte, si pensamos a $\mathbb{C}^2$ como $\mathbb{R}$-espacio vectorial (recordemos que en este caso $\dim(\mathbb{C}^2_{\mathbb{R}}) = 4$), también por el Teorema de la Dimensión 1.9.1, tenemos que $\dim(\mathrm{nul}(A)) = 4 - \mathrm{rg}(A) = 4 - 2 = 2$. Entonces como $\begin{bmatrix} 1 \\ i \end{bmatrix}$ y $\begin{bmatrix} i \\ -1 \end{bmatrix} \in \mathrm{nul}(A)$ y además, como no existe $a \in \mathbb{R}$ tal que $\begin{bmatrix} i \\ -1 \end{bmatrix} = a \begin{bmatrix} 1 \\ i \end{bmatrix}$, el conjunto $\{\begin{bmatrix} 1 \\ i \end{bmatrix}, \begin{bmatrix} i \\ -1 \end{bmatrix}\}$ es LI. Por lo tanto, una base de $\mathrm{nul}(A)$ podría ser $B_{\mathrm{nul}(A)} = \{\begin{bmatrix} 1 \\ i \end{bmatrix}, \begin{bmatrix} i \\ -1 \end{bmatrix}\}$.

  Otra forma de hallar $\mathrm{nul}(A)$ cuando pensamos a $\mathbb{C}^2$ como $\mathbb{R}$-espacio vectorial es la siguiente: x pertenece a $\mathrm{nul}(A)$ si $x \in \mathbb{C}^2$ y además $Ax = 0$. En este caso, una base de $\mathbb{C}^2_{\mathbb{R}}$ puede ser $B_{\mathbb{C}^2_{\mathbb{R}}} = \{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} i \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ i \end{bmatrix}\}$ y se ve claramente que $\dim(\mathbb{C}^2_{\mathbb{R}}) = 4$. Como $x \in \mathbb{C}^2$, existen $a, b, c, d \in \mathbb{R}$ tales que

  $$x = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} i \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 1 \end{bmatrix} + d \begin{bmatrix} 0 \\ i \end{bmatrix}.$$

  Entonces, si $0 = Ax$ tenemos que $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = A(a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} i \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 1 \end{bmatrix} + d \begin{bmatrix} 0 \\ i \end{bmatrix}) = a \begin{bmatrix} 1 \\ i \end{bmatrix} + b \begin{bmatrix} i \\ -1 \end{bmatrix} + c \begin{bmatrix} i \\ -1 \end{bmatrix} + d \begin{bmatrix} -1 \\ -i \end{bmatrix}$. Operando, nos queda el sistema $0 = a + bi + ci - d$, $0 = ai - b - c - di$. Entonces,

  $$(b + c)i = d - a \text{ y } b + c = (a - d)i.$$

  Observar que $a, b, c, d \in \mathbb{R}$ por lo que $b + c \in \mathbb{R}$ y $d - a \in \mathbb{R}$. Pero las ecuaciones anteriores nos indican que $b + c = (a - d)i$, entonces $b + c$ también es un número complejo puro (pues es igual al número real $a-d$ multiplicado por i), pero el único número que es real y complejo puro a la vez es 0. Por lo tanto $b+c=0$, por lo que $c=-b$ y $d=a$.

Volviendo a la expresión de x, tenemos que $x = a\begin{bmatrix}1\\0\end{bmatrix} + b\begin{bmatrix}i\\0\end{bmatrix} + c\begin{bmatrix}0\\1\end{bmatrix} + d\begin{bmatrix}0\\i\end{bmatrix} = a\begin{bmatrix}1\\0\end{bmatrix} + b\begin{bmatrix}i\\0\end{bmatrix} + (-b)\begin{bmatrix}0\\1\end{bmatrix} + a\begin{bmatrix}0\\i\end{bmatrix} = a\begin{bmatrix}1\\i\end{bmatrix} + b\begin{bmatrix}i\\-1\end{bmatrix}$, con $a,b\in\mathbb{R}$. Por lo tanto $\text{nul}(A) = gen\{\begin{bmatrix}1\\i\end{bmatrix}, \begin{bmatrix}i\\-1\end{bmatrix}\}$ y como esos generadores son LI, una base de $\text{nul}(A)$ podría ser $B_{\text{nul}(A)} = \{\begin{bmatrix}1\\i\end{bmatrix}, \begin{bmatrix}i\\-1\end{bmatrix}\}$.

- Para obtener $\text{fil}(A)$ y $\text{nul}(A^T)$, observar que en este caso en particular (es decir en este ejercicio) $A^T = A$. Entonces $\text{nul}(A^T) = \text{nul}(A)$ y como $\text{fil}(A) = \text{col}(A^T)$, tenemos que $\text{fil}(A) = \text{col}(A^T) = \text{col}(A) = gen\{\begin{bmatrix}1\\i\end{bmatrix}, \begin{bmatrix}i\\-1\end{bmatrix}\}$. Entonces, si pensamos a $\mathbb{C}^2$ como $\mathbb{C}$-espacio vectorial, una base de $\text{fil}(A)$ puede ser $B_{\text{fil}(A)} = \{\begin{bmatrix}1\\i\end{bmatrix}\}$ y una base de $\text{nul}(A^T) = \text{nul}(A)$ podría ser $B_{\text{nul}(A^T)} = \{\begin{bmatrix}1\\i\end{bmatrix}\}$.

  Si pensamos a $\mathbb{C}^2$ como $\mathbb{R}$-espacio vectorial, una base de $\text{fil}(A)$ puede ser $B_{\text{fil}(A)} = \{\begin{bmatrix}1\\i\end{bmatrix}, \begin{bmatrix}i\\-1\end{bmatrix}\}$ y una base de $\text{nul}(A^T)$ podría ser $B_{\text{nul}(A^T)} = \{\begin{bmatrix}1\\i\end{bmatrix}, \begin{bmatrix}i\\-1\end{bmatrix}\}$.

- Finalmente, observar que $b = \begin{bmatrix}2-3i\\3+2i\end{bmatrix} = 2\begin{bmatrix}1\\i\end{bmatrix} + (-3)\begin{bmatrix}i\\-1\end{bmatrix} = A\begin{bmatrix}2\\-3\end{bmatrix}$. Entonces, si llamamos $x_p := \begin{bmatrix}2\\-3\end{bmatrix} \in \mathbb{C}^2$, tenemos que $Ax_p = b$, es decir el sistema $Ax=b$ tiene al menos una solución (es compatible).

  Cuando existe solución del sistema $Ax=b$, vimos en la Proposición 1.9.2, que todas las soluciones de dicho sistema se pueden expresar como $x_s = x_p + x_h$, donde x_p es una solución particular y $x_h \in \text{nul}(A)$. Entonces, si pensamos a $\mathbb{C}^2$ como $\mathbb{C}$-espacio vectorial, todas las soluciones del sistema son $x_s = \begin{bmatrix}2\\-3\end{bmatrix} + a\begin{bmatrix}1\\i\end{bmatrix}$, con $a\in\mathbb{C}$. Por otra parte, si pensamos a $\mathbb{C}^2$ como $\mathbb{R}$-espacio vectorial, todas las soluciones del sistema son $x_s = \begin{bmatrix}2\\-3\end{bmatrix} + a\begin{bmatrix}1\\i\end{bmatrix} + b\begin{bmatrix}i\\-1\end{bmatrix}$, con $a,b\in\mathbb{R}$.

■

El siguiente ejercicio resultará útil para resolver el **Ejercicio 1.19**.

Ejercicio 1.K. Sean $A\in\mathbb{K}^{n\times m}$ y $b\in\mathbb{K}^n$.

a) El sistema $Ax=b$ es compatible si y sólo si $b\in\text{col}(A)$.

b) Si $b\in\text{col}(A)$, el sistema $Ax=b$ tiene un única solución si y sólo si $\text{nul}(A)=\{0\}$.

Dem. *a*): Vamos a suponer que $A=[A_1\ A_2\ \cdots\ A_m]$, donde $A_1,A_2,\cdots,A_m\in\mathbb{K}^n$ son las columnas de A. Si el sistema $Ax=b$ es compatible, entonces existe $x_0 = \begin{bmatrix}x_{01}\\x_{02}\\\vdots\\x_{0m}\end{bmatrix} \in\mathbb{K}^m$ tal que $Ax_0=b$.

Entonces, tenemos que $b = Ax_0 = [A_1\ A_2\ \cdots\ A_m]\begin{bmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0m} \end{bmatrix} = x_{01}A_1 + x_{02}A_2 + \cdots + x_{0m}A_m$. Por lo tanto (como b es combinación lineal de las columnas de A) $b \in gen\{A_1, A_2, \cdots, A_m\} = \text{col}(A)$.

Recíprocamente, si $b \in \text{col}(A)$, entonces existen $a_1, a_2, \cdots, a_m \in \mathbb{K}$ tales que

$$b = a_1A_1 + a_2A_2 + \cdots + a_mA_m.$$

Operando, nos queda que $b = a_1A_1 + a_2A_2 + \cdots + a_mA_m = [A_1\ A_2\ \cdots\ A_m]\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$. Llamando $x_0 := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \in \mathbb{K}^m$, claramente $Ax_0 = b$ (es decir encontramos $x_0 \in \mathbb{K}^m$ tal que $b = Ax_0$). Entonces el sistema $Ax = b$ es compatible.

b)): Si $b \in \text{col}(A)$, entonces por a) vale que el sistema $Ax = b$ es compatible, es decir tiene al menos una solución $x_0 \in \mathbb{K}^m$ tal que $Ax_0 = b$. Supongamos que el sistema tiene otra solución $x_0' \in \mathbb{K}^m$, es decir $Ax_0' = b$. Restanto estas dos igualdades nos queda que $0_{\mathbb{K}^n} = b - b = Ax_0 - Ax_0' = A(x_0 - x_0')$. Por lo tanto $x_0 - x_0' \in \text{nul}(A)$. Entonces $\text{nul}(A) = \{0\}$ si y sólo si $x_0 - x_0' = 0_{\mathbb{K}^m}$ si y sólo si $x_0 = x_0'$ (y el sistema $Ax = b$ tiene una única solución). ■

A partir del **Ejercicio 1.K**, tenemos otra manera de caracterizar el espacio $\text{col}(A)$:

Proposición 1.9.3. *Sea $A \in \mathbb{K}^{n\times m}$. Entonces*

$$\text{col}(A) = \{y \in \mathbb{K}^n : \ \textit{existe } x \in \mathbb{K}^m : y = Ax\} = \{Ax : x \in \mathbb{K}^m\}.$$

Dem. Vamos probar la doble inclusión, sea $y \in \text{col}(A)$, entonces, por el **Ejercicio 1.K**, el sistema $Ax = y$ es compatible, es decir existe $x_0 \in \mathbb{K}^m$ tal que $y = Ax_0$, entonces $y \in \{y \in \mathbb{K}^n : \ \text{existe } x \in \mathbb{K}^m : y = Ax\}$.

Reciprocamente, si $y \in \{y \in \mathbb{K}^n : \ \text{existe } x \in \mathbb{K}^m : y = Ax\}$, entonces existe $x \in \mathbb{K}^m$ tal que $y = Ax$. Entonces el sistema es compatible y, por el **Ejercicio 1.K**, $y \in \text{col}(A)$. ■

Ejercicio 1.L (De Parcial)**.** Si $A \in \mathbb{R}^{5\times 4}$ es una matriz tal que $\text{rg}(A) = 4$ y $b \in \mathbb{R}^5$, entonces el sistema $Ax = b$ no tiene solución cuando la matriz ampliada del sistema $[A\ b] \in \mathbb{R}^{5\times 5}$ es inversible.

Dem. Vamos a probar el contrarrecíproco de lo que queremos ver. Es decir, vamos a probar que si la matriz ampliada del sistema $[A\ b] \in \mathbb{R}^{5\times 5}$ no es inversible, entonces el sistema $Ax = b$ tiene solución (es decir probaremos "no q entonces no p" por lo que también valdrá que "p entonces q").

Supongamos que $A = [A_1\ A_2\ A_3\ A_4]$, donde $A_1, A_2, A_3, A_4 \in \mathbb{R}^5$ son las 4 columnas de A. Si la matriz ampliada del sistema $[A\ b] \in \mathbb{R}^{5\times 5}$ no es inversible, eso significa que el conjunto $\{A_1, A_2, A_3, A_4, b\}$ (es decir, el conjunto de las columnas de la matriz ampliada $[A\ b]$) no es LI (ó lo que es lo mismo es LD). Pero como $4 = \text{rg}(A) = \dim(\text{col}(A))$, tenemos que el conjunto $\{A_1, A_2, A_3, A_4\}$ (las columnas de A) es LI. Por lo tanto, la única posibilidad que queda es que b sea una CL de A_1, A_2, A_3, A_4 (si no están convencidos de esto, meditarlo un poco). Es decir, existen $a_1, a_2, a_3, a_4 \in \mathbb{R}$, tales que

$$b = a_1A_1 + a_2A_2 + a_3A_3 + a_4A_4 = [A_1\ A_2\ A_3\ A_4]\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = A\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}.$$

Llamando $x_0 := \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \in \mathbb{R}^4$. Tenemos que $Ax_0 = b$, y el sistema es compatible. En conclusión probamos que si $[A\ b]$ es no inversible entonces el sistema $Ax = b$ tiene solución, por contrarrecíproco vale que si $Ax = b$ no tiene solución entonces $[A\ b]$ es inversible y probamos lo que queríamos. ■

Ejercicio 1.M (De Parcial)**.** Sean $A \in \mathbb{R}^{p\times m}$ y $B \in \mathbb{R}^{m\times n}$. Demostrar que:

a) $\text{col}(AB) \subseteq \text{col}(A)$. Además, si $\text{rg}(B) = m$ vale la igualdad, pero la recíproca no es cierta.

b) $\text{nul}(B) \subseteq \text{nul}(AB)$. Además, si $\text{rg}(A) = m$ vale la igualdad, pero la recíproca no es cierta.

c) $\text{rg}(AB) \leq \text{mín}\{\text{rg}(A), \text{rg}(B)\}$.

Dem. *a*) : Para probar $\text{col}(AB) \subseteq \text{col}(A)$, tomemos $y \in \text{col}(AB)$ y veamos que $y \in \text{col}(A)$. Sea $y \in \text{col}(AB)$, entonces por la Proposición 1.9.3), existe $x \in \mathbb{R}^n$ tal que $y = ABx = A(Bx)$; entonces si llamamos $x' := Bx \in \mathbb{R}^m$, claramente $y = Ax'$. Encontramos un vector x' tal que $y = Ax'$, entonces (por la misma proposición) $y \in \text{col}(A)$. Y probamos que $\text{col}(AB) \subseteq \text{col}(A)$.

Si $\text{rg}(B) = m$, entonces como $\text{col}(B) \subseteq \mathbb{R}^m$ y $\dim(\text{col}(B)) = \text{rg}(B) = m$, tenemos que $\text{col}(B) = \mathbb{R}^m$ (usamos la propiedad que dice que dos subespacios son iguales si y sólo si uno está contenido en el otro y tienen misma dimensión). Entendida esta observación, veamos la otra inclusión. Sea $y \in \text{col}(A)$ entonces, existe $x \in \mathbb{R}^m$ tal que $y = Ax$, pero como $x \in \mathbb{R}^m = \text{col}(B)$, existe $z \in \mathbb{R}^n$ tal que $x = Bz$, por lo tanto $y = Ax = A(Bz) = (AB)z$ y tenemos que $y \in \text{col}(AB)$. Entonces probamos que $\text{col}(A) \subseteq \text{col}(AB)$ y como siempre vale la otra inclusión, en este caso, tenemos que $\text{col}(AB) = \text{col}(A)$.

La recíproca no vale. Es decir $\text{col}(AB) = \text{col}(A)$, no implica que $\text{rg}(B) = m$. De hecho, si tomamos $A = 0_{\mathbb{R}^{p\times m}}$, $B = 0_{\mathbb{R}^{m\times n}}$ tenemos que $\text{col}(A) = \{0_{\mathbb{R}^p}\}$ y $\text{rg}(B) = 0 \neq m$. Sin embargo, $\text{col}(AB) = \text{col}(0_{\mathbb{R}^{p\times n}}) = \{0_{\mathbb{R}^p}\} = \text{col}(A)$.

b) : Si $x \in \text{nul}(B)$ entonces $Bx = 0$. Entonces, $ABx = A(Bx) = A0 = 0$, por lo tanto $x \in \text{nul}(AB)$ y tenemos la inclusión $\text{nul}(B) \subseteq \text{nul}(AB)$.

Si $\text{rg}(A) = m$, por el Teorema de la dimensión, tenemos que $\dim(\text{nul}(A)) = m - \text{rg}(A) = 0$. Entonces $\text{nul}(A) = \{0\}$. En este caso, veamos que tenemos la otra inclusión. Si $x \in \text{nul}(AB)$ entonces $0 = ABx = A(Bx)$. Entonces $Bx \in \text{nul}(A)$, pero como $\text{nul}(A) = \{0\}$, tenemos que $Bx = 0$ y por lo tanto $x \in \text{nul}(B)$. Entonces tenemos que $\text{nul}(AB) \subseteq \text{nul}(B)$ y como siempre vale la otra inclusión, concluimos que $\text{nul}(AB) = \text{nul}(B)$.

La recíproca no vale. Es decir $\text{nul}(AB) = \text{nul}(B)$, no implica que $\text{rg}(A) = m$. Observar que sirve el mismo ejemplo que usamos en *a*).

c) : Por el ítem *a*) tenemos que $\text{col}(AB) \subseteq \text{col}(A)$, entonces $\text{rg}(AB) = \dim(\text{col}(AB)) \leq \dim(\text{col}(A)) = \text{rg}(A)$. Por otra parte, por el ítem *b*), tenemos que $\dim(\text{nul}(B)) \leq \dim(\text{nul}(AB))$. Entonces, por el Teorema de la dimensión, tenemos que $\text{rg}(AB) = n - \dim(\text{nul}(AB)) \leq n - \dim(\text{nul}(B)) = \text{rg}(B)$. Esto implica que $\text{rg}(AB) \leq \text{mín}\{\text{rg}(A), \text{rg}(B)\}$.

Para enteder mejor por qué vale esto, observar que sólo tenemos dos opciones

$$\text{mín}\{\text{rg}(A), \text{rg}(B)\} = \text{rg}(A) \text{ ó } \text{mín}\{\text{rg}(A), \text{rg}(B)\} = \text{rg}(B).$$

Por ejemplo si $\text{rg}(A) < \text{rg}(B)$, tendremos la primera opción. Como vimos que $\text{rg}(AB) \leq \text{rg}(B)$ y $\text{rg}(AB) \leq \text{rg}(A)$, tenemos lo que queríamos probar. ■

Ejercicio 1.24. Sean $A \in \mathbb{R}^{3\times 3}$ y $B \in \mathbb{R}^{3\times 4}$ dos matrices tales que

$$AB = \begin{bmatrix} 10 & -10 & -5 & 5 \\ 11 & -11 & -4 & 7 \\ 11 & -11 & -5 & 6 \end{bmatrix},$$

donde $\text{rg}(A) = 3$, y B satisface que $B[1\ 1\ 1\ 1]^T = [0\ 3\ 1]^T, B[1\ 0\ 1\ 0]^T = [5\ 7\ 6]^T$. Hallar todas las soluciones del sistema $Bx = [5\ 1\ 4]^T$.

Dem. Observar que
$$[5\ 7\ 6]^T - 2[0\ 3\ 1]^T = [5\ 1\ 4]^T.$$

Por lo tanto

$$B([1\ 0\ 1\ 0]^T - 2[1\ 1\ 1\ 1]^T) = B[1\ 0\ 1\ 0]^T + (-2)B[1\ 1\ 1\ 1]^T = [5\ 7\ 6]^T - 2[0\ 3\ 1]^T = [5\ 1\ 4]^T.$$

Entonces, el vector $[1\ 0\ 1\ 0]^T - 2[1\ 1\ 1\ 1]^T = [-1\ \ -2\ \ -1\ \ -2]^T$ es una solución del sistema $Bx = [5\ 1\ 4]^T$. En la Proposición 1.9.2 vimos que como existe solución del sistema $Bx = [5\ 1\ 4]^T$, todas las soluciones x_s del sistema se pueden expresar como

$$x_s = x_p + x_h,$$

donde x_p es una solución particular y $x_h \in \mathrm{nul}(B)$. Ya encontramos una solución particular $x_p = [-1\ \ -2\ \ -1\ \ -2]^T$, sólo nos resta calcular $\mathrm{nul}(B)$.

Haciendo la cuenta podemos ver que

$$\mathrm{nul}(AB) = gen\{[1\ 1\ 0\ 0]^T, [-1\ 0\ \ -1\ 1]^T\},$$

pero por el **Ejercicio 1.M**, como $\mathrm{rg}(A) = 3$, vale que $\mathrm{nul}(AB) = \mathrm{nul}(B)$ (notar que en general sólo vale que $\mathrm{nul}(B) \subseteq \mathrm{nul}(AB)$, pero como vimos en el **Ejercicio 1.M**, en este caso como $\mathrm{rg}(A) = 3$ tenemos la igualdad).

Por lo tanto $\mathrm{nul}(B) = gen\{[1\ 1\ 0\ 0]^T, [-1\ 0\ \ -1\ 1]^T\}$ y todas las soluciones del sistema $Bx = [5\ 1\ 4]^T$ son

$$x_s = x_p + x_h = [-1\ \ -2\ \ -1\ \ -2]^T + \alpha[1\ 1\ 0\ 0]^T + \beta[-1\ 0\ \ -1\ 1]^T \text{ con } \alpha, \beta \in \mathbb{R}.$$

■

1.10. Algoritmos para producir una base a partir de un sistema de generadores

Para resolver el **Ejercicio 1.13** necesitamos recordar varias definiciones, notaciones y un resultado que será fundamental.

Sea $A \in \mathbb{K}^{m\times n}$, llamamos $A_{i*} \in \mathbb{K}^n$ a la i-ésima fila de A, denotamos por $A_{*j} \in \mathbb{K}^m$ la j-ésima columna de A y $A_{ij} \in \mathbb{K}$ es la entrada ij de A. Por ejemplo, si

$$A = \begin{bmatrix} 2 & -5 & 8 & 9 & 3 \\ 0 & 0 & -2 & 1 & 7 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \text{ Entonces, } A_{2*} = \begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \\ 7 \end{bmatrix}, \ A_{*3} = \begin{bmatrix} 8 \\ -2 \\ 0 \\ 0 \end{bmatrix} \text{ y } A_{25} = 7.$$

Sea $A \in \mathbb{K}^{m\times n}$, definimos por r al *número de filas no nulas de* A

$$r := \#\{i \in \{1, \cdots, m\} : A_{i*} \neq 0\}.$$

En cada fila no nula, denotemos por p_i al *índice de la primera entrada no nula de* A:

$$p_i := \min\{j \in \{1, \cdots, n\} : A_{ij} \neq 0\} \text{ para } i = 1, 2, \cdots, r.$$

Diremos que una matriz es *escalonada por filas* si vale que:

1. $A_{i*} \neq 0$ para todo $1 \leq i \leq r$ y $A_{i*} = 0$ para $r < i \leq m$.
2. $p_1 < p_2 < \cdots < p_r$.

Si la matriz A es escalonada, entonces las entradas con índices (i, p_i) con $1 \leq i \leq r$ se llaman *pivotes*. A las columnas correspondientes A_{*p_i} con $i \leq i \leq r$ las llamamos *columnas pivotales*.

Veamos ejemplos de matrices que son escalonadas por filas y matrices que no lo son:

Ejemplo 1.e.

- $A = \begin{bmatrix} 2 & -5 & 8 & 9 & 3 \\ 0 & 0 & -2 & 1 & 7 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

  Entonces, $r = 3, A_{1*} \neq 0, A_{2*} \neq 0, A_{3*} \neq 0$ y $A_{4*} = 0$. Por otra parte, $p_1 = 1, p_2 = 3$ y $p_3 = 4$ y tenemos que $p_1 < p_2 < p_3$. Por lo tanto, A es escalonada por filas.

- $A = \begin{bmatrix} 0 & -5 & 7 & 1 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

  Entonces, $r = 2, A_{1*} \neq 0, A_{2*} \neq 0$ y $A_{3*} = 0$. Por otra parte, $p_1 = 2$ y $p_2 = 3$ y tenemos que $p_1 < p_2$. Por lo tanto, A es escalonada por filas.

- $A = \begin{bmatrix} 2 & -5 & 8 & 9 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

  Entonces, $r = 2$ pero $A_{2*} = 0$. Entonces, A NO es escalonada por filas.

- $A = \begin{bmatrix} 2 & -5 & 8 & 9 & 3 \\ 0 & 1 & 3 & 4 & 0 \\ 0 & 5 & -2 & 1 & 7 \end{bmatrix}$.

  Entonces, $r = 3$, $p_2 = p_3 = 2$. Por lo tanto $p_2 \not< p_3$. Entonces, A NO es escalonada por filas.

Una matriz se llama *escalonada reducida (por filas)* si cumple con las propiedades 1. y 2. y además con las siguientes propiedades 3. y 4.:

3. $A_{ip_i} = 1$ para todo $1 \leq i \leq r$. Es decir, en cada fila no nula, el elemento delantero diferente de cero (pivote) es igual a uno.

4. $A_{k,p_i} = 0$ para todo $1 \leq k \leq i - 1$ y $2 \leq i \leq r$. Es decir, todos los elementos por encima de los pivotes son nulos.

Veamos ejemplos de una matriz que es escalonadas reducida (por filas) y otra que no:

Ejemplo 1.f.

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

  Entonces, $r = 3, A_{1*} \neq 0, A_{2*} \neq 0, A_{3*} \neq 0$ y $A_{4*} = 0$. Por otra parte, $p_1 = 2, p_2 = 4$ y $p_3 = 5$ y tenemos que $p_1 < p_2 < p_3$. Por lo tanto, A es escalonada por filas. Además, $A_{1p1} = A_{12} = 1$, $A_{2p_2} = A_{24} = 1$ y $A_{3p_3} = A_{35} = 1$. Finalmente, $A_{1p_2} = A_{14} = 0$, $A_{1p_3} = A_{15} = 0$, $A_{2p_3} = A_{25} = 0$. Por lo tanto, A es escalonada reducida por filas. En este caso A_{*2}, A_{*4} y A_{*5} son las columnas pivotales de A.

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

Entonces, $r = 2, A_{1*} \neq 0, A_{2*} \neq 0$ y $A_{3*} = 0$. Por otra parte, $p_1 = 1$ y $p_2 = 2$ y tenemos que $p_1 < p_2$. Por lo tanto, A es escalonada por filas. Además, $A_{1p_1} = A_{11} = 1$, y $A_{2p_2} = A_{22} = 1$. Sin embargo, $A_{1p_2} = A_{12} = 1 \neq 0$. Por lo tanto, A NO es escalonada reducida por filas.

El siguiente resultado será fundamental para resolver el **Ejercicio 1.13 b)**:

Proposición 1.10.1. *Sea $A \in \mathbb{K}^{m\times n}$. Entonces, las columnas de A correspondientes a las columnas pivotales de E (la forma reducida por filas de A) forman una base de* $\operatorname{col}(A)$.

Antes de demostrar la Proposición 1.10.1 veamos un ejemplo concreto:

Ejemplo 1.g. Sea $A = \begin{bmatrix} 0 & 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 3 & -6 & 6 & 4 & -5 \\ 0 & 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$. Mostrar que las columnas de A correspondientes a las columnas pivotales de E (la forma reducida por filas de A) forman una base de $\operatorname{col}(A)$.

En primer lugar, obtengamos E la forma escalonada reducida (por filas) de A. De hecho, triangulando A podemos ver que

$$E = \begin{bmatrix} 0 & 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Entonces, $r = 3$ y los índices de E son $p_1 = 2, p_2 = 3$ y $p_3 = 6$. Las columnas pivotales de E son $E_{*p_1} = E_{*2}$, $E_{*p_2} = E_{*3}$ y $E_{*p_3} = E_{*6}$. Claramente, el conjunto $\{E_{*p_1}, E_{*p_2}, E_{*p_3}\}$ es un conjunto LI, pues $E_{*p_1} = e_1, E_{*p_2} = e_2$ y $E_{*p_3} = e_3$, donde e_1, e_2, e_3 son los primeros 3 elementos de la base canónica de $\mathbb{R}^4$.

Recordemos que como E es la forma escalonada reducida (por filas) de A vale que

$$Ax = 0 \text{ si y sólo si } Ex = 0. \tag{1.6}$$

Es decir, las soluciones al sistema homogéneo asociado a A y a E coinciden.

Veamos que las columnas de A correspondientes a las columnas pivotales de E que son $\{A_{*p_1}, A_{*p_2}, A_{*p_3}\} = \{A_{*2}, A_{*3}, A_{*6}\}$ forman un conjunto LI. De hecho, sean $a_1, a_2, a_3 \in \mathbb{R}$ tales que

$$a_1 A_{*p_1} + a_2 A_{*p_2} + a_3 A_{*p_3} = 0.$$

Entonces $A \begin{bmatrix} 0 \\ a_1 \\ a_2 \\ 0 \\ 0 \\ a_3 \\ 0 \end{bmatrix} = 0$ y, por (1.6), tenemos que $E \begin{bmatrix} 0 \\ a_1 \\ a_2 \\ 0 \\ 0 \\ a_3 \\ 0 \end{bmatrix} = 0 = a_1 E_{*p_1} + a_2 E_{*p_2} + a_3 E_{*p_3}$. Pero, como $\{E_{*p_1}, E_{*p_2}, E_{*p_3}\}$ es un conjunto LI, tenemos que $a_1 = a_2 = a_3 = 0$ y por lo tanto, $\{A_{*p_1}, A_{*p_2}, A_{*p_3}\} = \{A_{*2}, A_{*3}, A_{*6}\}$ es un conjunto LI.

Por otra parte, las columnas no pivotales de E son $u_1 = E_{*1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $u_2 = E_{*4} = \begin{bmatrix} -2 \\ -2 \\ 0 \\ 0 \end{bmatrix}$, $u_3 = E_{*5} = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix}$ y $u_4 = E_{*7} = \begin{bmatrix} -24 \\ -7 \\ 4 \\ 0 \end{bmatrix}$. Como

$$u_{ik} = 0 \text{ para todo } 1 \leq i \leq 4, \quad 3 < k \leq 4,$$

se sigue las columnas u_i para $i = 1, 2, 3, 4$ son una C.L. de las 3 columnas pivotales de E. Es decir,

$$\begin{aligned}
E_{*1} &= u_1 = 0E_{*p_1} + 0E_{*p_2} + 0E_{*p_3} = 0e_1 + 0e_2 + 0e_3,\\
E_{*4} &= u_2 = -2E_{*p_1} + -2E_{*p_2} + 0E_{*p_3} = -2e_1 + -2e_2 + 0e_3,\\
E_{*5} &= u_3 = 3E_{*p_1} + 2E_{*p_2} + 0E_{*p_3} = 3e_1 + 2e_2 + 0e_3,\\
E_{*7} &= u_4 = -24E_{*p_1} - 7E_{*p_2} + 4E_{*p_3} = -24e_1 - 7e_2 + 4e_3.
\end{aligned}$$

Entonces (escribiendo las ecuaciones de arriba en forma matricial), nos queda:

$$E\begin{bmatrix}-1\\0\\0\\0\\0\\0\\0\end{bmatrix} = 0,\quad E\begin{bmatrix}0\\-2\\-2\\-1\\0\\0\\0\end{bmatrix} = 0,\quad E\begin{bmatrix}0\\3\\2\\0\\-1\\0\\0,\end{bmatrix} = 0\quad E\begin{bmatrix}0\\-24\\-7\\0\\0\\4\\-1\end{bmatrix} = 0.$$

Entonces, nuevamente por (1.6), vale que

$$A\begin{bmatrix}-1\\0\\0\\0\\0\\0\\0\end{bmatrix} = 0,\quad A\begin{bmatrix}0\\-2\\-2\\-1\\0\\0\\0\end{bmatrix} = 0,\quad A\begin{bmatrix}0\\3\\2\\0\\-1\\0\\0,\end{bmatrix} = 0\quad A\begin{bmatrix}0\\-24\\-7\\0\\0\\4\\-1\end{bmatrix} = 0.$$

Entonces

$$\begin{aligned}
A_{*1} &= 0A_{*p_1} + 0A_{*p_2} + 0A_{*p_3} = 0A_{*2} + 0A_{*3} + 0A_{*6},\\
A_{*4} &= -2A_{*p_1} - 2A_{*p_2} + 0A_{*p_3} = -2A_{*2} - 2A_{*3} + 0A_{*6},\\
A_{*5} &= 3A_{*p_1} + 2A_{*p_2} + 0A_{*p_3} = 3A_{*2} + 2A_{*3} + 0A_{*6},\\
A_{*7} &= -24A_{*p_1} - 7A_{*p_2} + 4A_{*p_3} = -24A_{*2} - 7A_{*3} + 4A_{*6}.
\end{aligned}$$

Por lo tanto las columnas de A correspondientes a las columnas no pivotales de E que son $A_{*1}, A_{*4}, A_{*5}, A_{*7}$, son C.L. de las columnas de A correspondientes a las columnas pivotales de E.

Por lo tanto, $\mathrm{col}(A) = gen\{A_{*1}, A_{*2}, A_{*3}, A_{*4}, A_{*5}, A_{*6}, A_{*7}\} = gen\{A_{*2}, A_{*3}, A_{*6}\}$. Entonces $\{A_{*2}, A_{*3}, A_{*6}\}$ es un conjunto LI que genera $\mathrm{col}(A)$ por lo tanto es una base de $\mathrm{col}(A)$ y probamos lo que queríamos.

Demostración. (de la Proposición 1.10.1)

Sea $E := [w_{p_1}\ w_{p_2} \cdots\ w_{p_r}\ u_1 \cdots\ u_{n-p_r}]$ la forma escalonada reducida (por filas) de A. Donde, $w_{p_1}, w_{p_2}, \cdots, w_{p_r}$ son las r columnas pivotales de E y $u_1, \cdots, u_{n-p_r}$ son las columnas no pivotales de E. **Nota:** para simplificar la demostración estamos suponiendo que las primeras r filas de E corresponden a las columnas pivotales de E. Eso no tiene por qué ocurrir en general, pero la prueba sería similar si las columnas pivotales de E tuvieran otro orden.

Entonces, si e_i denota el elemento i de la base canónica de $\mathbb{K}^m$, tenemos que

$$w_{p_1} = e_1,\quad w_{p_2} = e_2,\quad \cdots,\quad w_{p_r} = e_r.$$

Por lo tanto, como $p_1 < p_2 < \cdots < p_r$, es claro que $\{w_{p_1}, w_{p_2}, \cdots, w_{p_r}\}$ es un conjunto LI. Por otra parte, como $u_1, \cdots, u_{n-p_r}$ son las columnas no pivotales de E, por definición vale que

$$u_{ik} = 0 \text{ para todo } 1 \leq i \leq n - p_r,\quad r < k \leq m.$$

Entonces, es claro que las columnas u_i son una C.L. de las r columnas pivotales de E. Es decir

$$u_i = \sum_{k=1}^{r} u_{ik} w_{p_k} = \sum_{k=1}^{r} u_{ik} e_k. \tag{1.7}$$

Recordemos que, como E es la forma escalonada reducida (por filas) de A vale que

$$Ax = 0 \text{ si y sólo si } Ex = 0. \tag{1.8}$$

Es decir, las soluciones al sistema homogéneo asociado a A y a E coinciden.

Vamos a escribir $\hat{A} = [v_{p_1}\ v_{p_2} \cdots\ v_{p_r}\ z_1 \cdots\ z_{n-p_r}]$, donde $v_{p_1}\ v_{p_2} \cdots v_{p_r}$ son las r columnas de A correspondientes a las columnas pivotales de E y $z_1, \cdots, z_{n-p_r}$ son las columnas de A correspondientes a las columnas no pivotales de E. Como sólo reordenamos las columnas de A para obtener $\hat{A}$, si bien A y $\hat{A}$ podrían no coincidir, es claro que, $\operatorname{col}(A) = \operatorname{col}(\hat{A})$. Veamos que $\{v_{p_1}, v_{p_2}, \cdots, v_{p_r}\}$ es un conjunto LI de $\mathbb{K}^n$. De hecho, sean $a_1, \cdots, a_r \in \mathbb{K}$ tales que

$$a_1 v_{p_1} + a_2 v_{p_2} + \cdots + a_r v_{p_r} = 0.$$

Entonces, (escribiendo las ecuaciones de arriba en forma matricial), nos queda:

$$A \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_r \\ 0 \\ \cdots \\ 0 \end{bmatrix} = a_1 v_{p_1} + a_2 v_{p_2} + \cdots + a_r v_{p_r} = 0 \text{ y, por (1.8), } E \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_r \\ 0 \\ \cdots \\ 0 \end{bmatrix} = 0 = a_1 w_{p_1} + a_2 w_{p_2} + \cdots a_r w_{p_r}.$$

Pero, como $\{w_{p_1}, w_{p_2}, \cdots, w_{p_r}\}$ es un conjunto LI, tenemos que $a_1 = a_2 = \cdots = a_r = 0$ y por lo tanto, $\{v_{p_1},\ v_{p_2}, \cdots, v_{p_r}\}$ es un conjunto LI.

Por otra parte, observar que por (1.7), si $i = 1$, vale que

$$0 = -u_1 + \sum_{k=1}^{r} u_{1k} w_{p_k}.$$

Es decir, $E \begin{bmatrix} u_{11} \\ u_{12} \\ \cdots \\ u_{1r} \\ -1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} = -u_1 + \sum_{k=1}^{r} u_{1k} w_{p_k} = 0$ entonces, por (1.8), $A \begin{bmatrix} u_{11} \\ u_{12} \\ \cdots \\ u_{1r} \\ -1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} = -z_1 + \sum_{k=1}^{r} u_{1k} v_{p_k} =$ 0. Por lo tanto $z_1 = \sum_{k=1}^{r} u_{ik} v_{p_k}$ y z_1 es C.L. de las las columnas de A correspondientes a las columnas pivotales de E. De la misma manera, se prueba que las otras columnas de A correspondientes a las columnas no pivotales de E que son $z_2, \cdots, z_{n-p_r}$ son C.L. de las columnas de A correspondientes a las columnas pivotales de E. Por lo tanto,

$$\operatorname{col}(A) = \operatorname{col}(\hat{A}) = gen\{v_{p_1}, v_{p_2}, \cdots, v_{p_r}, z_1 \cdots, z_{n-p_r}\} = gen\{v_{p_1}, v_{p_2}, \cdots, v_{p_r}\}.$$

Entonces $\{v_{p_1}, v_{p_2}, \cdots, v_{p_r}\}$ es un conjunto LI que genera $\operatorname{col}(A)$ por lo tanto es una base de $\operatorname{col}(A)$ y probamos lo que queríamos. ∎

Ejercicio 1.13. Explicar por qué los siguientes algoritmos producen una base $\mathcal{B}$ de un subespacio $\mathcal{S}$ a partir de un sistema de generadores $\mathcal{G}$ del mismo.

(a) **Algorithm 1** espacio filas:

Requiere $\mathcal{G} = \{v_1, \cdots, v_m\} \subseteq \mathbb{K}^n$ sistema de generadores de $\mathcal{S}$. **Ensure:** $\mathcal{B}$ una base de $\mathcal{S}$.

1. Construir $A \in \mathbb{K}^{m\times n}$ cuyas filas son $v_1^T, \cdots, v_m^T$.
2. Construir $E \in \mathbb{K}^{m\times n}$ la matriz escalonada por filas reducidas de A.
3. Listar las filas no nulas de $E : \{E_{p_1 *}, \cdots, E_{p_r *}\}$
4. $\mathcal{B} \leftarrow \{E_{p_1 *}, \cdots, E_{p_r *}\}$.
5. **return** $\mathcal{B}$.

(b) **Algorithm 2** espacio columnas:

Requiere $\mathcal{G} = \{v_1, \cdots, v_n\} \subseteq \mathbb{K}^m$ sistema de generadores de $\mathcal{S}$. **Ensure:** $\mathcal{B}$ una base de $\mathcal{S}$.

1. Construir $A \in \mathbb{K}^{m\times n}$ cuyas columnas son $v_1, \cdots, v_n$.
2. Construir $E \in \mathbb{K}^{m\times n}$ la matriz escalonada por filas reducidas de A.
3. Listar las columnas de A que corresponden a las columnas pivotales de $E : \{v_{j_1}, \cdots, v_{j_p}\}$
4. $\mathcal{B} \leftarrow \{v_{j_1}, \cdots, v_{j_p}\}$
5. **return** $\mathcal{B}$.

Demostración. $a :)$ Sea $\mathcal{G} := \{v_1, \cdots, v_m\} \subseteq \mathbb{K}^n$ un sistema de generadores de $\mathcal{S}$. Construimos la matriz cuyas filas son $v_1^T, \cdots, v_m^T$, es decir tomamos

$$A := \begin{bmatrix} v_1^T \\ \cdots \\ v_m^T \end{bmatrix}. \text{ Notar que } \mathcal{S} = \text{fil}(A) = \text{col}(A^T).$$

Luego, triangulando A obtenemos E la matriz escalonada por filas de A y listamos las filas no nulas de E que llamaremos $\{E_{p_1 *}, \cdots, E_{p_r *}\}$. Entonces, por la Proposición 1.4.1, se sigue que $\mathcal{S} = \text{fil}(A) = \text{fil}(E) = gen\{E_{p_1 *}, \cdots, E_{p_r *}\}$ y además, como $\{E_{p_1 *}, \cdots, E_{p_r *}\}$ es un conjunto LI, tenemos que $\{E_{p_1 *}, \cdots, E_{p_r *}\}$ es una base de $\text{fil}(A) = \text{fil}(E) = \mathcal{S}$. Por lo tanto, el algoritmo produce una base $\mathcal{B}$ de $\mathcal{S}$ a partir del sistema de generadores $\mathcal{G}$.

$b :)$ Sea $\mathcal{G} := \{v_1, \cdots, v_n\} \subseteq \mathbb{K}^m$ un sistema de generadores de $\mathcal{S}$. Construimos la matriz cuyas columnas son $v_1, \cdots, v_n$, es decir tomamos

$$A := [v_1 \ \cdots \ v_n]. \text{ Notar que } \mathcal{S} = \text{col}(A).$$

Luego, construimos E la matriz escalonada por filas reducida de A y listamos las columnas de A que corresponden a las columnas pivotales de E. Entonces, por la Proposición 1.10.1, las columnas de A correspondientes a las columnas pivotales de E forman una base de $\text{col}(A) = \mathcal{S}$. Por lo tanto, el algoritmo produce una base $\mathcal{B}$ de $\mathcal{S}$ a partir del sistema de generadores $\mathcal{G}$. ■

1.11. Coordenadas de un vector en una base

Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial de dimensión finita y sea $B = \{v_1; \cdots; v_n\}$ una base (ordenada) de $\mathbb{V}$. Dado $x \in \mathbb{V}$ existen únicos $a_1, \cdots, a_n \in \mathbb{K}$ tales que $x = a_1 v_1 + \cdots + a_n v_n$. El vector $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{K}^n$ se llama el *vector de coordenadas de x en la base B* y será denotado por $[x]^B$, por $C_B(x)$ ó por $[x]_B$.

Vamos a usar la notación $[x]^B$, que es la que se usa en la guía. Para entender la definición (y la notación) que acabamos de ver, supongamos que estamos en $\mathbb{R}_2[x]$ y tomamos $B = \{1, x^2, x\}$ que es una base de $\mathbb{R}_2[x]$. Si $q(x) = 1-x$ entonces $[q]^B = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \in \mathbb{R}^3$, porque $q(x) = 1{\cdot}x{+}0{\cdot}x^2{+}(-1){\cdot}x$, para todo $x \in \mathbb{R}$.

Por otra parte, si $p \in \mathbb{R}_2[x]$ es tal que $[p]^B = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, entonces $p(x) = 1 \cdot 1 + 2 \cdot x^2 + 3 \cdot x = 1 + 2x^2 + 3x$ y de esa manera obtenemos el vector p sabiendo sus coordenadas en una base.

Notar que la expresión $[\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}]^B$ NO tiene sentido. En la notación que definimos: $[v]^B$, el vector v es un elemento de $\mathbb{V}$. En este caso en particular, $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \in \mathbb{R}^3$, claramente $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \notin \mathbb{R}_2[x]$ y no tiene sentido calcularle las coordenadas en una base de $\mathbb{R}_2[x]$ a un vector que no esté en dicho espacio.

Observar que si $\mathbb{V}$ es un $\mathbb{K}$ espacio vectorial tal que la $\dim(\mathbb{V}) = n$, dado $v \in \mathbb{V}$ el vector de coordenadas $[v]^B$ siempre es un vector de $\mathbb{K}^n$. Y si $B = \{v_1, \cdots, v_n\}$ es una base de $\mathbb{V}$ y $[v]^B = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{K}^n$ entonces $v = a_1v_1 + \cdots + a_nv_n$ y recuperamos al vector v.

Proposición 1.11.1. *Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial de dimensión finita y sea $B = \{v_1; \cdots; v_n\}$ una base (ordenada) de $\mathbb{V}$. Si $v, w \in \mathbb{V}$ y $\alpha, \beta \in \mathbb{K}$, entonces*

$$[\alpha v + \beta w]^B = \alpha[v]^B + \beta[w]^B.$$

Es decir, "tomar coordenadas" es lineal

Dem. Supongamos que $[v]^B = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ y $[w]^B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$. Entonces $v = a_1v_1 + \cdots + a_nv_n$ y $w = b_1v_1 + \cdots + b_nv_n$. Entonces, $\alpha v + \beta w = \alpha(a_1v_1 + \cdots + a_nv_n) + \beta(b_1v_1 + \cdots + b_nv_n) = (\alpha a_1 + \beta b_1)v_1 + \cdots + (\alpha a_n + \beta b_n)v_n$. Por lo tanto,

$$[\alpha v + \beta w]^B = \begin{bmatrix} \alpha a_1 + \beta b_1 \\ \vdots \\ \alpha a_1 + \beta b_1 \end{bmatrix} = \alpha[v]^B + \beta[w]^B.$$

■

Este tema suele traer muchas confusiones por lo que a continuación vamos resolver algunos ejercicios para fijar ideas.

Ejercicio 1.N. Sean $\mathbb{V} = gen\{1, e^x, e^{-x}\}$ un $\mathbb{R}$-espacio vectorial, $B = \{1, e^x, e^{-x}\}$ un conjunto de $\mathbb{V}$ y $g(x) = senh(x)$. Entonces:

a) Demostrar que B es una base de $\mathbb{V}$.

b) Obtener $[g]^B$.

c) Encontrar un vector $h \in \mathbb{V}$ tal que $[h]^B = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$. ¿Existe otro vector h' con esas coordenadas en la base B?

Antes de comenzar a resolver el ejercicio, es importante entender que cada elemento del conjunto B es una función (no un número real). En general se suele abusar un poco de la notación y eso puede llegar a confundir un poco. Estrictamente hablando, si definimos las funciones $f_1(x) = 1$ (la función que siempre vale 1), $f_2(x) = e^x$, $f_3(x) = e^{-x}$, entonces $B = \{f_1, f_2, f_3\}$ y es como se debe interpretar esa notación. Meditar un poco sobre esto porque es crucial para resolver el ejercicio.

Dem. a): Observar que cada elemento de B pertenece a $\mathbb{V}$. Por definición de $\mathbb{V}$, el conjunto $\{f_1, f_2, f_3\}$ es un sistema de generadores de $\mathbb{V}$. Sólo falta probar que el conjunto es LI. Para eso, por ejemplo, podemos calcular el Wronskiano.

$$W(f_1, f_2, f_3)(x) = \det(\begin{bmatrix} 1 & e^x & e^{-x} \\ 0 & e^x & -e^{-x} \\ 0 & e^x & e^{-x} \end{bmatrix}) = e^x e^{-x} + e^{-x} e^x = 1 + 1 = 2 \neq 0.$$

Por lo tanto el conjunto es LI. Entonces B es una base de $\mathbb{V}$ y $\dim(\mathbb{V}) = 3$.

b) : Recordar que $g(x) = senh(x) = \frac{e^x - e^{-x}}{2} \in \mathbb{V}$. Si $[g]^B = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, entonces $\frac{e^x - e^{-x}}{2} = a \cdot 1 + b \cdot e^x + c \cdot e^{-x}$ para todo $x \in \mathbb{R}$. Entonces $0 = a \cdot 1 + (b - \frac{1}{2}) \cdot e^x + (c + \frac{1}{2}) \cdot e^{-x}$, para todo $x \in \mathbb{R}$. Si $\mathbf{0}$ es la función nula (que es el elemento neutro de $\mathbb{V}$), la ecuación anterior es equivalente a $\mathbf{0} = af_1 + (b - \frac{1}{2})f_2 + (b + \frac{1}{2})f_3$. Como recién vimos que B es una base (en particular un conjunto LI) y tenemos una CL de los elementos de B igualadas al elemento neutro de $\mathbb{V}$, se sigue que $a = 0$, $b - \frac{1}{2} = 0$ y $c + \frac{1}{2} = 0$. Entonces $a = 0$, $b = \frac{1}{2}$, $c = -\frac{1}{2}$ y por lo tanto $[g]^B = \begin{bmatrix} 0 \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$.

c) : Si $[h]^B = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$, por definición de las coordenadas de un vector en una base, $h(x) = 1 + 2e^x - e^{-x}$. Por supuesto que no existe otra función con esas coordenadas en base B. Supongamos que sí, es decir, que existe $h' \in \mathbb{V}$ tal que $[h']^B = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$. Entonces $h'(x) = 1 + 2e^x - e^{-x} = h(x)$ para todo $x \in \mathbb{R}$. Entonces $h' = h$. ■

Ejercicio 1.26. Sean $p_1(x) = \frac{1}{2}(x-1)(x-2)$, $p_2(x) = -x(x-2)$ y $p_3(x) = \frac{1}{2}x(x-1)$.

a) Demostrar que $B = \{p_1, p_2, p_3\}$ es una base de $\mathbb{R}_2[x]$.

b) Si $p \in \mathbb{R}_2[x]$ entonces

$$[p]^B = \begin{bmatrix} p(0) \\ p(1) \\ p(2) \end{bmatrix}.$$

c) Hallar $[p]^B$ si $p(x) = x^2 - x + 1$.

Dem. a): Claramente, $p_1, p_2, p_3 \in \mathbb{R}_2[x]$. Recordar que en el **Ejercicio 1.I**, probamos que “n vectores LI de un subespacio de dimensión n forman un base de dicho subespacio”. Entonces, como $\dim(\mathbb{R}_2[x])$=3, sólo basta ver que B es un conjunto LI para probar que B es una base de $\mathbb{R}_2[x]$. Veamos que es LI, por ejemplo, calculando el Wronskiano:

$$W(p_1, p_2, p_3)(x) = det(\begin{bmatrix} \frac{1}{2}(x-1)(x-2) & -x(x-2) & \frac{1}{2}x(x-1) \\ x - \frac{3}{2} & -2x + 2 & x - \frac{1}{2} \\ 1 & -2 & 1 \end{bmatrix}).$$

Si evaluamos el Wronskiano en $x = 1$, nos queda $W(p_1, p_2, p_3)(1) = det(\begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ 1 & -2 & 1 \end{bmatrix}) =$ $(-1)(-\dfrac{1}{2} - \dfrac{1}{2}) = 1 \neq 0$. Encontramos un punto donde el Wronskiano no se anula, por lo tanto B es un conjunto LI y (por los argumentos que dimos arriba) es una base de $\mathbb{R}_2[x]$.

b): Sea $p \in \mathbb{R}_2[x]$, si $[p]^B = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, entonces $p(x) = ap_1(x) + bp_2(x) + cp_3(x)$, para todo $x \in \mathbb{R}$. Entonces,

$$p(0) = ap_1(0) + bp_2(0) + cp_3(0) = a \cdot 1 + b \cdot 0 + c \cdot 0 = a,$$
$$p(1) = ap_1(1) + bp_2(1) + cp_3(1) = a \cdot 0 + b \cdot 1 + c \cdot 0 = b,$$
$$p(2) = ap_1(2) + bp_2(2) + cp_3(2) = a \cdot 0 + b \cdot 0 + c \cdot 1 = c.$$

Por lo tanto $[p]^B = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} p(0) \\ p(1) \\ p(2) \end{bmatrix}$.

c): Usando *b*), tenemos que $[p]^B = \begin{bmatrix} p(0) \\ p(1) \\ p(2) \end{bmatrix} = \begin{bmatrix} 0^2 - 0 + 1 \\ 1^2 - 1 + 1 \\ 2^2 - 2 + 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$. ■

1.12. Matriz de cambio de base

En esta sección estudiaremos una herramienta que nos permitirá obtener las coordenadas de un cierto vector en una base determinada sabiendo las coordenadas de dicho vector en otra base. Dicha herramienta es la matriz de cambio de base:

Definición. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial de dimensión n y sean $B_1 = \{v_1, \cdots, v_n\}$ y $B_2 = \{w_1, \cdots, w_n\}$ dos bases de $\mathbb{V}$. Se llama la *matriz de cambio de base de* B_1 *a* B_2 y se nota $M_{B_1}^{B_2} \in \mathbb{K}^{n\times n}$, a la matriz

$$M_{B_1}^{B_2} := [[v_1]^{B_2}\ [v_2]^{B_2}\ \cdots\ [v_n]^{B_2}],$$

es decir a la matriz cuya columna i-ésima es el vector de coordenadas de v_i (de la base B_1) en la base B_2, para cada $i \in \{1, 2, \cdots, n\}$.

Para entender la definición anterior hagamos el siguiente ejercicio, similar al **Ejercicio 1.28**.

Ejercicio 1.Ñ. Supongamos que estamos en $\mathbb{R}_2[x]$ y tomamos como bases $B_1 = \{1, x, x^2\}$ y $B_2 = \{p_1, p_2, p_3\}$ (donde p_1, p_2, p_3 son los polinomios definidos en el **Ejercicio 1.26**), obtener $M_{B_1}^{B_2}$.

Dem. Por definición,

$$M_{B_1}^{B_2} = [[1]^{B_2}\ [x]^{B_2}\ [x^2]^{B_2}] = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix},$$

donde usamos $[p]^{B_2} = \begin{bmatrix} p(0) \\ p(1) \\ p(2) \end{bmatrix}$, que ya lo probamos en el **Ejercicio 1.26**. ■

Si $\mathbb{V}$ es un $\mathbb{K}$-espacio vectorial de dimensión n y $B_1 = \{v_1, \cdots, v_n\}$ y $B_2 = \{w_1, \cdots, w_n\}$ son dos bases de $\mathbb{V}$. Dado $v \in \mathbb{V}$, se cumple que

$$[v]^{B_2} = M_{B_1}^{B_2}\ [v]^{B_1} \tag{1.9}$$

Entonces, la matriz de cambio de base de B_1 a B_2 (como era de esperar) nos permite obtener las coordenadas de v en la base B_2 sabiendo las coordenadas de dicho v en base B_1.

La prueba de esta propiedad es muy simple. Supongamos que $[v]^{B_1} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, entonces

$$M_{B_1}^{B_2}\ [v]_{B_1} = [[v_1]^{B_2}\ [v_2]^{B_2}\ \cdots\ [v_n]^{B_2}] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = a_1[v_1]^{B_2} + a_2[v_2]^{B_2} + \cdots + a_n[v_n]^{B_2} =$$

$$= [a_1v_1 + a_2v_2 + \cdots + a_nv_n]^{B_2},$$

donde usamos que tomar coordenadas es lineal (Proposición 1.11.1). Por lo tanto, como $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$, tenemos que $M_{B_1}^{B_2}\ [v]^{B_1} = [a_1v_1 + a_2v_2 + \cdots + a_nv_n]^{B_2} = [v]^{B_2}$ y probamos lo que queríamos.

El siguiente ejercicio es similar al **Ejercicio 1.30**.

Ejercicio 1.O. Sea $E = \{1, x, x^2\}$ la base canónica de $\mathbb{R}_2[x]$.

a) Hallar una base B de $\mathbb{R}_2[x]$ tal que $M_E^B = \begin{bmatrix} 1 & -2 & 3 \\ 0 & -1 & 2 \\ 0 & 1 & -1 \end{bmatrix}$.

b) Para cada $p \in \mathbb{R}_2[x]$ determinar la expresión de las coordenadas de p en la base B.

c) Hallar todos los vectores $p \in \mathbb{R}_2[x]$ tales que $[p]^B = [p]^E$.

Dem. *a*): Por definición $M_E^B = [[1]^B\ [x]^B\ [x^2]^B]$. Entonces, si $B = \{q, r, s\}$ (con q, r, s polinomios a determinar), tenemos que

$$1 = 1 \cdot q(x) + 0 \cdot r(x) + 0 \cdot s(x), \text{ para todo } x \in \mathbb{R}, \tag{1.10}$$

entonces $q(x) = 1$. También

$$x = (-2) \cdot q(x) + (-1) \cdot r(x) + 1 \cdot s(x) = -2 - r(x) + s(x), \text{ para todo } x \in \mathbb{R}, \tag{1.11}$$

y finalmente,

$$x^2 = 3 \cdot q(x) + 2 \cdot r(x) + (-1) \cdot s(x) = 3 + 2r(x) - s(x), \text{ para todo } x \in \mathbb{R}. \tag{1.12}$$

Despejando en (1.11), nos queda que $s(x) = x + 2 + r(x)$, para todo $x \in \mathbb{R}$ y reemplazando en (1.12), tenemos que $x^2 = 3 + 2r(x) - s(x) = 3 + 2r(x) - (x + 2 + r(x)) = 1 - x + r(x)$, para todo $x \in \mathbb{R}$. Entonces $r(x) = x^2 + x - 1$ y $s(x) = r(x) + 2 + x = x^2 + x - 1 + x + 2 = x^2 + 2x + 1$. Por lo tanto, $B = \{q, r, s\} = \{1, x^2 + x - 1, x^2 + 2x + 1\}$.

Nota: En el próximo ejercicio, veremos que la matriz de cambio de base siempre es inversible y que $(M_E^B)^{-1} = M_B^E$. Entonces, otra manera de resolver este ejercicio sería calcular

$$(M_E^B)^{-1} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix} = M_B^E = [[q]^E\ [r]^E\ [s]^E].$$

Por lo tanto $q(x) = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 = 1,\ r(x) = -1 \cdot 1 + 1 \cdot x + 1 \cdot x^2 = -1 + x + x^2,\ s(x) = 1 \cdot 1 + 2 \cdot x + 1 \cdot x^2 = 1 + 2x + x^2$, y obtenemos lo mismo.

$b)$: Si $p \in \mathbb{R}_2[x]$ tiene como expresión $p(x) = a+bx+cx^2$ con $a, b, c \in \mathbb{R}$, entonces $[p]^E = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

Por lo tanto

$$[p]^B = M_E^B \ [p]^E = \begin{bmatrix} 1 & -2 & 3 \\ 0 & -1 & 2 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a-2b+3c \\ -b+2c \\ b-c \end{bmatrix}.$$

$c)$: Si $p \in \mathbb{R}_2[x]$ es tal que $[p]^B = [p]^E$, entonces, usando que $[p]^B = M_E^B \ [p]^E$, nos queda que

$$[p]^B = M_E^B \ [p]^E = [p]^E.$$

Entonces

$$M_E^B \ [p]^E - [p]^E = (M_E^B - I)[p]^E = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Entonces $[p]^E \in \text{nul}(M_E^B - I)$. Entonces, calculemos el espacio nulo de

$$M_E^B - I = \begin{bmatrix} 1 & -2 & 3 \\ 0 & -1 & 2 \\ 0 & 1 & -1 \end{bmatrix} - I = \begin{bmatrix} 0 & -2 & 3 \\ 0 & -2 & 2 \\ 0 & 1 & -2 \end{bmatrix}.$$

Resolviendo, tenemos que $\text{nul}(M_E^B - I) = gen\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\}$. Entonces, $[p]^B = [p]^E$ si y sólo si $[p]^E \in gen\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\}$. Es decir, $[p]^E = a\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}$, con $a \in \mathbb{R}$. Entonces, $p(x) = a \cdot 1 + 0 \cdot x + 0 \cdot x^2 = a$, con $a \in \mathbb{R}$. Escrito de otra manera $\{p \in \mathbb{R}_2[x] : [p]^E = [p]^B\} = gen\{1\}$. ■

Ejercicio 1.P (De Parcial)**.** Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial de dimensión n y sean B_1, B_2, B_3 tres bases de $\mathbb{V}$. Demostrar las siguientes afirmaciones:

a) $M_{B_1}^{B_2}$ es una matriz inversible y además vale que $M_{B_2}^{B_1} = (M_{B_1}^{B_2})^{-1}$.

b) $M_{B_1}^{B_2} = M_{B_3}^{B_2} M_{B_1}^{B_3}$.

Dem. $a)$: Supongamos que $B_1 = \{v_1, \cdots, v_n\}$ y $B_2 = \{w_1, \cdots, w_n\}$. Entonces, por definición

$$M_{B_1}^{B_2} := [[v_1]^{B_2} \ [v_2]^{B_2} \ \cdots \ [v_n]^{B_2}].$$

Veamos que las columnas de la matriz $M_{B_1}^{B_2}$ (que son vectores de $\mathbb{K}^n$) son LI. Sean $a_1, a_2, \cdots, a_n \in \mathbb{K}$ tales que

$$a_1[v_1]^{B_2} + a_2[v_2]^{B_2} \ \cdots \ a_n[v_n]^{B_2} = 0_{\mathbb{K}}{}^n.$$

Entonces, como tomar coordenadas es lineal (Proposición 1.11.1), tenemos que

$$0_{\mathbb{K}}{}^n = a_1[v_1]^{B_2} + a_2[v_2]^{B_2} \ \cdots \ a_n[v_n]^{B_2} = [a_1v_1 + a_2v_2 + \cdots + a_nv_n]^{B_2}.$$

Entonces,

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0w_1 + \cdots + 0w_n = 0_{\mathbb{V}}.$$

Finalmente, como $\{v_1, \cdots, v_n\}$ es un conjunto LI (forman una base), y tenemos una CL de dichos elementos igualada al elemento neutro de $\mathbb{V}$, se sigue que $a_1 = a_2 = \cdots = a_n = 0$. Es decir las columnas de $M_{B_1}^{B_2}$ son LI. Por lo tanto $\mathrm{rg}(M_{B_1}^{B_2}) = \dim(\mathrm{col}(M_{B_1}^{B_2})) = n$ y como $M_{B_1}^{B_2} \in \mathbb{K}^{n\times n}$, tenemos que $M_{B_1}^{B_2}$ es inversible.

Por otra parte, sabemos que para todo $v \in \mathbb{V}$ vale que $[v]^{B_2} = M_{B_1}^{B_2}[v]^{B_1}$. Entonces, multiplicando por $(M_{B_1}^{B_2})^{-1}$ a ambos lados de la igualdad anterior, tenemos que

$$(M_{B_1}^{B_2})^{-1}[v]^{B_2} = (M_{B_1}^{B_2})^{-1}M_{B_1}^{B_2}[v]^{B_1} = I[v]^{B_1} = [v]^{B_1}.$$

También sabemos que $[v]^{B_1} = M_{B_2}^{B_1}[v]^{B_2}$. Entonces, igualando ambas expresiones, nos queda que

$$(M_{B_1}^{B_2})^{-1}[v]^{B_2} = [v]^{B_1} = M_{B_2}^{B_1}[v]^{B_2} \text{ para todo } v \in \mathbb{V}.$$

Entonces $M_{B_2}^{B_1} = (M_{B_1}^{B_2})^{-1}$ como queríamos ver.

Justificación del último paso de la demostración anterior: Si no están convencidos de este último paso, observar que tenemos que $M_{B_2}^{B_1}[v]^{B_2} = (M_{B_1}^{B_2})^{-1}[v]^{B_2}$, para todo $v \in \mathbb{V}$. Recordemos que $B_2 = \{w_1, \cdots, w_n\}$. Entonces, por ejemplo, $[w_1]^{B_2} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ (pues $w_1 = 1.w_1 + 0.w_2 + \cdots + 0.w_n$) y entonces, volviendo a la igualdad anterior,

$$M_{B_2}^{B_1}\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = (M_{B_1}^{B_2})^{-1}\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Siempre que hacemos $A\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ obtenemos la primera columna de la matriz A (si no se ve, hacer la cuenta). Entonces, la igualdad anterior nos indica que la primera columna de $M_{B_2}^{B_1}$ y de $(M_{B_2}^{B_1})^{-1}$ coinciden. Si hacemos lo mismo para $v = w_2$, tendremos que la segunda columna de dichas matrices coinciden y así sucesivamente. En definitiva, todas la columnas de ambas matrices coinciden y por lo tanto son iguales.

b) : Sabemos que para todo $v \in \mathbb{V}$ vale que $[v]^{B_2} = M_{B_1}^{B_2}[v]^{B_1}$, $[v]^{B_2} = M_{B_3}^{B_2}[v]^{B_3}$ y $[v]^{B_3} = M_{B_1}^{B_3}[v]^{B_1}$. Entonces

$$M_{B_1}^{B_2}[v]^{B_1} = [v]^{B_2} = M_{B_3}^{B_2}[v]^{B_3} = M_{B_3}^{B_2}(M_{B_1}^{B_3}[v]^{B_1}) = M_{B_3}^{B_2}M_{B_1}^{B_3}[v]^{B_1} \text{ para todo } v \in \mathbb{V}.$$

Es decir, $M_{B_1}^{B_2}[v]^{B_1} = M_{B_3}^{B_2}M_{B_1}^{B_3}[v]^{B_1}$ para todo $v \in \mathbb{V}$. Por lo tanto, con la misma justificación que usamos en *a*), tenemos que $M_{B_1}^{B_2} = M_{B_3}^{B_2}M_{B_1}^{B_3}$ y probamos lo que queríamos. ■

1.13. Operaciones entre subespacios

A continuación vamos a estudiar dos tipos de operaciones entre subespacios que nos devuelven otro subespacio. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y $\mathcal{S}$ y $\mathcal{T}$ subespacios de $\mathbb{V}$:

- Se llama *intersección* de $\mathcal{S}$ y $\mathcal{T}$ a $\mathcal{S} \cap \mathcal{T} := \{v \in \mathbb{V} : v \in \mathcal{S} \text{ y } v \in \mathcal{T}\}$.
- Se llama *suma* de $\mathcal{S}$ y $\mathcal{T}$ a $\mathcal{S} + \mathcal{T} := \{v \in \mathbb{V} : \exists\, s \in \mathcal{S}, t \in \mathcal{T} : v = s + t\} = \{s + t : s \in \mathcal{S}, t \in \mathcal{T}\}$.

Proposición 1.13.1. *Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y $\mathcal{S}$ y $\mathcal{T}$ subespacios de $\mathbb{V}$. Entonces:*

- $\mathcal{S}_1 \cap \mathcal{S}_2$ *es el mayor subespacio contenido en* $\mathcal{S}_1$ *y* $\mathcal{S}_2$.
- $\mathcal{S}_1 + \mathcal{S}_2$ *es el menor subespacio que contiene a* $\mathcal{S}_1$ *y* $\mathcal{S}_2$.

Antes de probar este resultado, es importante destacar que en esta proposición, con "mayor" y "menor" nos referimos respecto a la inclusión de conjuntos.

La inclusión de conjuntos permite"ordenar" conjuntos. Pero notar que a diferencia del orden que conocemos de los números reales, este orden de conjuntos no es un orden total. Es decir, si $a, b \in \mathbb{R}$, sabemos que sólo tenemos tres opciones $a = b$ ó $a < b$ ó $b < a$. En el caso de conjuntos (y la inclusión) no necesariamente pasa lo mismo que en $\mathbb{R}$ con el orden usual. Por ejemplo, en $\mathbb{R}^2$, consideremos $\mathcal{S}_1 = gen\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \}$ y $\mathcal{S}_2 = gen\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \}$. Observar que, $\mathcal{S}_1 \neq \mathcal{S}_2$, $\mathcal{S}_1 \not\subseteq \mathcal{S}_2$ y $\mathcal{S}_2 \not\subseteq \mathcal{S}_1$, es decir, en este caso, $\mathcal{S}_1$ y $\mathcal{S}_2$ no se comparan (no podemos decir nada acerca de quien es "mayor, menor ó igual" respecto del orden que provee la inclusión de conjuntos). Sin embargo, hay casos (como el de esta proposición) donde los conjuntos se comparan y uno es "mayor" o "menor" que el otro.

Dem. Se deja como ejercicio probar que $\mathcal{S}_1 \cap \mathcal{S}_2$ es un subespacio. Veamos ahora que $\mathcal{S}_1 \cap \mathcal{S}_2$ es el mayor subespacio contenido en $\mathcal{S}_1$ y $\mathcal{S}_2$. Por un lado, claramente $\mathcal{S}_1 \cap \mathcal{S}_2 \subseteq \mathcal{S}_1$, pues si $x \in \mathcal{S}_1 \cap \mathcal{S}_2$ entonces $x \in \mathcal{S}_1$ (y también $x \in \mathcal{S}_2$) y tenemos la inclusión $\mathcal{S}_1 \cap \mathcal{S}_2 \subseteq \mathcal{S}_1$. De la misma manera, $\mathcal{S}_1 \cap \mathcal{S}_2 \subseteq \mathcal{S}_2$. Es decir, probamos que $\mathcal{S}_1 \cap \mathcal{S}_2$ es un subespacio que está contenido en $\mathcal{S}_1$ y $\mathcal{S}_2$. Falta ver que es el mayor (respecto de la inclusión). Para eso, supongamos que tenemos otro subespacio $\mathcal{T}$ tal que $\mathcal{T} \subseteq \mathcal{S}_1$ y $\mathcal{T} \subseteq \mathcal{S}_2$ (es decir contenido en $\mathcal{S}_1$ y $\mathcal{S}_2$) entonces $\mathcal{T} \subseteq \mathcal{S}_1 \cap \mathcal{S}_2$, pues si $x \in \mathcal{T}$ por hipótesis, tenemos que $x \in \mathcal{S}_1$ y $x \in \mathcal{S}_2$. Entonces $x \in \mathcal{S}_1 \cap \mathcal{S}_2$ y vale la inclusión que escribimos. Por lo tanto, como $\mathcal{T}$ era cualquiera, concluimos que $\mathcal{S}_1 \cap \mathcal{S}_2$ es el mayor subespacio (respecto de la inclusión) contenido en $\mathcal{S}_1$ y $\mathcal{S}_2$.

Se deja como ejercicio probar que $\mathcal{S}_1 + \mathcal{S}_2$ es un subespacio. Finalmente, veamos que $\mathcal{S}_1 + \mathcal{S}_2$ es el menor subespacio que contiene a $\mathcal{S}_1$ y $\mathcal{S}_2$. Primero veamos que $\mathcal{S}_1 \subseteq \mathcal{S}_1 + \mathcal{S}_2$. Si $x \in \mathcal{S}_1$ entonces como $x = x + 0_{\mathbb{V}}$ y $0_{\mathbb{V}} \in \mathcal{S}_2$ (porque $\mathcal{S}_2$ es un subespacio), por definición de la suma de subespacios, $x \in \mathcal{S}_1 + \mathcal{S}_2$ y probamos la inclusión. De la misma manera, se prueba que $\mathcal{S}_2 \subseteq \mathcal{S}_1 + \mathcal{S}_2$. Hasta acá vimos que $\mathcal{S}_1 + \mathcal{S}_2$ es un subespacio que contiene a $\mathcal{S}_1$ y a $\mathcal{S}_2$. Falta ver que es el menor (respecto de la inclusión). Para eso, supongamos que $\mathcal{T}$ es un subespacio tal que $\mathcal{S}_1 \subseteq \mathcal{T}$ y $\mathcal{S}_1 \subseteq \mathcal{T}$. Entonces $\mathcal{S}_1 + \mathcal{S}_2 \subseteq \mathcal{T}$, porque si $x \in \mathcal{S}_1 + \mathcal{S}_2$, existen $y \in \mathcal{S}_1$ y $z \in \mathcal{S}_2$ tales que $x = y + z$, pero por hipótesis $y \in \mathcal{T}$ y $z \in \mathcal{T}$. Además, como $\mathcal{T}$ es un subespacio $x + y \in \mathcal{T}$ (la suma de elementos de $\mathcal{T}$ pertenece a $\mathcal{T}$) entonces $x = y + z \in \mathcal{T}$ y tenemos que $\mathcal{S}_1 + \mathcal{S}_2 \subseteq \mathcal{T}$. Como $\mathcal{T}$ era cualquiera, concluimos que $\mathcal{S}_1 + \mathcal{S}_2$ es el menor subespacio (respecto de la inclusión) que contiene a $\mathcal{S}_1$ a $\mathcal{S}_2$. ∎

La siguiente propiedad es muy útil para calcular un sistema de generadores de $\mathcal{S} + \mathcal{T}$ conociendo algún sistema de generadores de $\mathcal{S}$ y $\mathcal{T}$. La idea vale para espacios vectoriales de cualquier dimensión, pero para simplificar la notación, lo vamos a ver para un espacio vectorial de dimensión finita.

Proposición 1.13.2. *Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial de dimensión finita y $\mathcal{S}$ y $\mathcal{T}$ subespacios de $\mathbb{V}$ tales que $\mathcal{S} = gen\{v_1, \cdots, v_s\}$ y $\mathcal{T} = gen\{w_1, \cdots, w_t\}$. Entonces*

$$\mathcal{S} + \mathcal{T} = gen\{v_1, \cdots, v_s, w_1, \cdots, w_t\}.$$

Dem. Vamos a probar la doble inclusión. Si $v \in \mathcal{S} + \mathcal{T}$, entonces existen $s \in \mathcal{S}$ y $t \in \mathcal{T}$ tales que $v = s + t$. Entonces, existen $a_1, \cdots, a_s \in \mathbb{K}$ y $b_1, \cdots, b_t \in \mathbb{K}$ tales que $s = a_1v_1 + \cdots + a_sv_s$ y $t = b_1w_1 + \cdots + b_tw_t$. Por lo tanto

$$v = s + t = a_1v_1 + \cdots + a_sv_s + b_1w_1 + \cdots + b_tw_t \in gen\{v_1, \cdots, v_s, w_1, \cdots, w_t\}.$$

Recíprocamente, si $v \in gen\{v_1, \cdots, v_s, w_1, \cdots, w_t\}$ entonces existen $a_1, \cdots, a_s \in \mathbb{K}$ y $b_1, \cdots, b_t \in \mathbb{K}$ tales que $v = a_1v_1 + \cdots + a_sv_s + b_1w_1 + \cdots + b_tw_t$. Sean $s := a_1v_1 + \cdots + a_sv_s \in \mathcal{S}$ y

$t := b_1w_1 + \cdots + b_tw_t \in \mathcal{T}$. Entonces $v = s + t$ y $v \in \mathcal{S} + \mathcal{T}$. Como probamos la doble inclusión, se sigue que $\mathcal{S} + \mathcal{T} = gen\{v_1, \cdots, v_s, w_1, \cdots, w_t\}$. ■

Otra propiedad que usaremos ampliamente es la del *Teorema de la dimensión* para la suma de subespacios:

Teorema 1.13.3. *Sea $\mathbb{V}$ es un $\mathbb{K}$-espacio vectorial de dimensión finita y $\mathcal{S}$ y $\mathcal{T}$ subespacios de $\mathbb{V}$, entonces*

$$\dim(\mathcal{S} + \mathcal{T}) = \dim(\mathcal{S}) + \dim(\mathcal{T}) - \dim(\mathcal{S} \cap \mathcal{T}).$$

Dem. Sean $s = \dim(\mathcal{S})$, $t = \dim(\mathcal{T})$ y $r = \dim(\mathcal{S} \cap \mathcal{T})$. Si $s = 0$ entonces $\mathcal{S} = \{0\}$, $\mathcal{S} \cap \mathcal{T} = \{0\} \cap \mathcal{T} = \{0\}$, entonces $r = 0$ y $\mathcal{S} + \mathcal{T} = \{0\} + \mathcal{T} = \mathcal{T}$. Por lo tanto, se sigue trivialmente que $\dim(\mathcal{S} + \mathcal{T}) = t = 0 + t - 0 = \dim(\mathcal{S}) + \dim(\mathcal{T}) - \dim(\mathcal{S} \cap \mathcal{T})$ y la igualdad vale. Lo mismo si $t = 0$.

Supongamos que $\mathcal{S} \cap \mathcal{T} = \{0\}$. Entonces $r = 0$ y sean $B_{\mathcal{S}} = \{v_1, \cdots, v_s\}$ una base de $\mathcal{S}$ y $B_{\mathcal{T}} = \{w_1, \cdots, w_t\}$ una base de $\mathcal{T}$. Entonces, $\{v_1, \cdots, v_s, w_1, \cdots, w_t\}$ es una base de $\mathcal{S} + \mathcal{T}$. De hecho, por la Proposición 1.13.2, dicho conjunto es un sistema de generadores de $\mathcal{S} + \mathcal{T}$. Veamos que dicho conjunto es LI. Sean $a_1, \cdots, a_s, b_1, \cdots, b_t \in \mathbb{K}$ tales que

$$a_1v_1 + \cdots + a_sv_s + b_1w_1 + \cdots + b_tw_t = 0_{\mathbb{V}}.$$

Entonces,

$$a_1v_1 + \cdots + a_sv_s = -b_1w_1 - \cdots - b_tw_t \in \mathcal{T}.$$

Por lo tanto, $a_1v_1 + \cdots + a_sv_s \in \mathcal{S} \cap \mathcal{T} = \{0\}$. Entonces $a_1v_1 + \cdots + a_sv_s = b_1w_1 + \cdots + b_tw_t = 0_{\mathbb{V}}$ y como $\{v_1, \cdots, v_s\}$ es una base de $\mathcal{S}$ y $\{w_1, \cdots, w_t\}$ una base de $\mathcal{T}$ se sigue que $a_1 = \cdots = a_s = b_1 = \cdots = b_t = 0$. Por lo tanto, efectivamente $\{v_1, \cdots, v_s, w_1, \cdots, w_t\}$ es una base de $\mathcal{S} + \mathcal{T}$. Entonces,

$$\dim(\mathcal{S} + \mathcal{T}) = s + t = s + t - 0 = \dim(\mathcal{S}) + \dim(\mathcal{T}) - \dim(\mathcal{S} \cap \mathcal{T})$$

y vale la igualdad.

Finalmente, supongamos que $\mathcal{S} \cap \mathcal{T} \neq \{0\}$ y $\dim(\mathcal{S} \cap \mathcal{T}) = r > 0$. Sea $B_{\mathcal{S} \cap \mathcal{T}} = \{u_1, \cdots, u_r\}$ una base de $\mathcal{S} \cap \mathcal{T}$. Completemos dichos vectores para obtener bases de $\mathcal{S}$ y $\mathcal{T}$ es decir, sean $v_{r+1}, \cdots, v_s \in \mathcal{S}$ tales que $\{u_1, \cdots, u_r, v_{r+1}, \cdots, v_s\}$ es una base de $\mathcal{S}$ y sean $w_{r+1}, \cdots, w_t \in \mathcal{T}$ tales que $\{u_1, \cdots, u_r, w_{r+1}, \cdots, w_t\}$ es una base de $\mathcal{T}$. Entonces, $\{u_1, \cdots, u_r, v_{r+1}, \cdots, v_s, w_{r+1}, \cdots, w_t\}$ es una base de $\mathcal{S} + \mathcal{T}$. Es claro que, por la Proposición 1.13.2, dicho conjunto es un sistema de generadores de $\mathcal{S}+\mathcal{T}$. Veamos que dicho conjunto es LI. Sean, $c_1, \cdots, c_r, a_{r+1}, \cdots, a_s, b_{r+1}, \cdots, b_t \in \mathbb{K}$ tales que

$$c_1u_1 + \cdots + c_ru_r + a_{r+1}v_{r+1} + \cdots + a_sv_s + b_{r+1}w_{r+1} + \cdots + b_tw_t = 0_{\mathbb{V}}.$$

Entonces,

$$c_1u_1 + \cdots + c_ru_r + a_{r+1}v_{r+1} + \cdots + a_sv_s = -b_{r+1}w_{r+1} - \cdots - b_tw_t \in \mathcal{T}$$

Por lo tanto, $-b_{r+1}w_{r+1} - \cdots - b_tw_t \in \mathcal{S} \cap \mathcal{T}$. Entonces, existen $d_1, \cdots, d_r \in \mathbb{K}$ tales que

$$-b_{r+1}w_{r+1} - \cdots - b_tw_t = d_1u_1 + \cdots + d_ru_r.$$

Entonces,

$$d_1u_1 + \cdots + d_ru_r + b_{r+1}w_{r+1} + \cdots + b_tw_t = 0_{\mathbb{V}}.$$

Pero, como $\{u_1, \cdots, u_r, w_{r+1}, \cdots, w_t\}$ es una base de $\mathcal{T}$, tenemos que $d_1 = \cdots = d_r = b_{r+1} = \cdots = b_t = 0$. Entonces, $c_1u_1 + \cdots + c_ru_r + a_{r+1}v_{r+1} + \cdots + a_sv_s = -b_{r+1}w_{r+1} - \cdots - b_tw_t = 0_{\mathbb{V}}$. De igual manera, como $\{u_1, \cdots, u_r, v_{r+1}, \cdots, v_s\}$ es una base de $\mathcal{S}$, se sigue que $c_1 = \cdots = c_r =$

$a_{r+1} = \cdots = a_s = 0$. Por lo tanto, $\{u_1, \cdots, u_r, v_{r+1}, \cdots, v_s, w_{r+1}, \cdots, w_t\}$ es un conjunto LI y por ende es una base de $\mathcal{S} + \mathcal{T}$. Entonces,

$$\dim(\mathcal{S} + \mathcal{T}) = r + (s - r) + (t - r) = s + t - r = \dim(\mathcal{S}) + \dim(\mathcal{T}) - \dim(\mathcal{S} \cap \mathcal{T})$$

y vale la igualdad. ■

En el siguiente ejercicio (muy similar al **Ejercicio 1.31**) veremos por qué la unión de subespacios en general no es un subespacio.

Ejercicio 1.Q. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y sean $\mathcal{S}_1$ y $\mathcal{S}_2$ dos subespacios tales que ninguno contiene al otro. Entonces $\mathcal{S}_1 \cup \mathcal{S}_2$ NO es un subespacio.

Dem. Vamos a probar el ejercicio por el absurdo. Por hipótesis, $\mathcal{S}_1 \not\subseteq \mathcal{S}_2$ y $\mathcal{S}_2 \not\subseteq \mathcal{S}_1$, entonces existen vectores (distintos) v y w tales que $v \in \mathcal{S}_2 \setminus \mathcal{S}_1$ y $w \in \mathcal{S}_1 \setminus \mathcal{S}_2$. Supongamos que $\mathcal{S}_1 \cup \mathcal{S}_2$ es un subespacio, entonces como $v \in \mathcal{S}_1 \cup \mathcal{S}_2$ y $w \in \mathcal{S}_1 \cup \mathcal{S}_2$, tendríamos que $v + w \in \mathcal{S}_1 \cup \mathcal{S}_2$ entonces, $v+w \in \mathcal{S}_1$ ó $v+w \in \mathcal{S}_2$. Como $\mathcal{S}_1$ es un subespacio, si $v+w \in \mathcal{S}_1$, tenemos que $v = (v+w) - w \in \mathcal{S}_1$, pues $w \in \mathcal{S}_1$. Entonces $v \in \mathcal{S}_1$ lo cual es absurdo (teníamos que $v \in \mathcal{S}_2 \setminus \mathcal{S}_1$). De la misma manera, como $\mathcal{S}_2$ es un subespacio, si $v + w \in \mathcal{S}_2$, tenemos que $w = (v + w) - v \in \mathcal{S}_2$, pues $v \in \mathcal{S}_2$. Entonces $w \in \mathcal{S}_2$ lo cual es absurdo (teníamos que $w \in \mathcal{S}_1 \setminus \mathcal{S}_2$). El absurdo se produjo por suponer que $\mathcal{S}_1 \cup \mathcal{S}_2$ era un subespacio. Por lo tanto, $\mathcal{S}_1 \cup \mathcal{S}_2$ NO es un subespacio. ■

Ejercicio 1.R. En $\mathbb{V} = gen\{1, e^x, e^{-x}\}$, consideremos $\mathcal{S} = gen\{3, 4e^x\}$ y $\mathcal{T} = \{g \in \mathbb{V} : g'(0) = 0\}$. Hallar bases de $\mathcal{S} \cap \mathcal{T}$ y $\mathcal{S} + \mathcal{T}$.

Dem. Primero obtengamos $\mathcal{S} \cap \mathcal{T}$. Si $f \in \mathcal{S} \cap \mathcal{T}$, entonces $f \in \mathcal{S}$, es decir $f(x) = a3 + b4e^x$, con $a, b \in \mathbb{R}$ y además, $f \in \mathcal{T}$, entonces como $f'(x) = (a3 + b4e^x)' = 4be^x$, tenemos que $f'(0) = 4b = 0$, entonces $b = 0$. Por lo tanto $f(x) = 3a$, con $a \in \mathbb{R}$. Entonces, $\mathcal{S} \cap \mathcal{T} = gen\{1\}$ y una base podría ser $B_{\mathcal{S}\cap\mathcal{T}} = \{1\}$.

Ahora obtengamos $\mathcal{S}+\mathcal{T}$. Primero busquemos un sistema de generadores de $\mathcal{T}$. Si $g \in \mathcal{T}$ entonces $g(x) = a1 + be^x + ce^{-x}$ con $a, b, c \in \mathbb{R}$ y además como $g'(x) = be^x - ce^{-x}$, tenemos que $g'(0) = b - c = 0$, es decir $b = c$. Entonces $g(x) = a1 + b(e^x + e^{-x})$, con $a, b \in \mathbb{R}$. Entonces $\mathcal{T} = gen\{1, e^x + e^{-x}\}$ y fácilmente vemos que $\dim(\mathcal{T}) = 2$. De la misma manera vemos que $\dim(\mathcal{S}) = 2$. Entonces, usando el teorema de la dimensión, $\dim(\mathcal{S} + \mathcal{T}) = \dim(\mathcal{S}) + \dim(\mathcal{T}) - \dim(\mathcal{S} \cap \mathcal{T}) = 2 + 2 - 1 = 3$. Por lo tanto como $\mathcal{S} + \mathcal{T} \subseteq \mathbb{V}$ y $\dim(\mathbb{V}) = \dim(\mathcal{S} + \mathcal{T}) = 3$, tenemos que $\mathcal{S} + \mathcal{T} = \mathbb{V}$. Una base de $\mathcal{S} + \mathcal{T}$ es cualquier base de $\mathbb{V}$, por ejemplo, como vimos arriba, $B_{\mathcal{S}+\mathcal{T}} = \{1, e^x, e^{-x}\}$ es una base posible. ■

Ejercicio 1.32. Para las siguientes elecciones de $\mathcal{S}_1$ y $\mathcal{S}_2$ del espacio vectorial $\mathbb{V}$, hallar una base del mayor subespacio contenido en ambos y otra del menor subespacio que los contiene.

a) $\mathcal{S}_1 := \{x = [x_1\ x_2\ x_2\ x_4]^T \in \mathbb{R}^4 : x_2 + x_3 + x_4 = 0\}$ y $\mathcal{S}_2 := \{x = [x_1\ x_2\ x_2\ x_4]^T \in \mathbb{R}^4 : x_1 + x_2 = x_3 - 2x_4 = 0\}$.

b) $\mathcal{S}_1 = \text{col}(A)$ y $\mathcal{S}_2 = \text{nul}(A)$ con $A = \begin{bmatrix} -1 & 1 & 1 & -2 & 1 \\ -1 & 0 & 3 & -4 & 2 \\ -1 & 0 & 3 & -5 & 3 \\ -1 & 0 & 3 & -6 & 4 \\ -1 & 0 & 3 & -6 & 4 \end{bmatrix}$.

c) $\mathcal{S}_1 = gen\{\begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}\}$ y $\mathcal{S}_2 = gen\{\begin{bmatrix} 4 \\ 2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \end{bmatrix}\}$.

Usando la Proposición 1.13.1 tenemos que $\mathcal{S}_1 + \mathcal{S}_2$ es el menor subespacio que contiene a $\mathcal{S}_1$ y $\mathcal{S}_2$ y $\mathcal{S}_1 \cap \mathcal{S}_2$ es el mayor subespacio contenido en $\mathcal{S}_1$ y $\mathcal{S}_2$.

Dem. $a)$: Calculemos $\mathcal{S}_1 \cap \mathcal{S}_2$. Si $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \in \mathcal{S}_1 \cap \mathcal{S}_2$ entonces, por un lado $x \in \mathcal{S}_1$ y por lo tanto $x_2 + x_3 + x_4 = 0$ y, por el otro lado $x \in \mathcal{S}_2$ y entonces $x_1 + x_2 = 0$ y $x_3 - 2x_4 = 0$. Entonces, x cumple el sistema

$$x_2 + x_3 + x_4 = x_1 + x_2 = x_3 - 2x_4 = 0.$$

Haciendo cuentas, se ve que las soluciones del sistema anterior son

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3x_4 \\ -3x_4 \\ 2x_4 \\ x_4 \end{bmatrix} \text{ con } x_4 \in \mathbb{R}.$$

Entonces $\mathcal{S}_1 \cap \mathcal{S}_2 = gen\{\begin{bmatrix} 3 \\ -3 \\ 2 \\ 1 \end{bmatrix}\}$ y una base de $\mathcal{S}_1 \cap \mathcal{S}_2$ puede ser $B_{\mathcal{S}_1 \cap \mathcal{S}_2} = \{\begin{bmatrix} 3 \\ -3 \\ 2 \\ 1 \end{bmatrix}\}$.

Por otra parte, es claro que $S_1 = gen\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}\}$ y $S_2 = gen\{\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}\}$. Por lo tanto, $\dim(S_1) = 3$ y $\dim(S_2) = 2$ y por el Teorema de la dimensión para la suma de subespacios tenemos que $\dim(\mathcal{S}_1 + \mathcal{S}_2) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) - \dim(\mathcal{S}_1 \cap \mathcal{S}_2) = 3 + 2 - 1 = 4$. Entonces $\mathcal{S}_1 + \mathcal{S}_2 = \mathbb{R}^4$ y una base de $\mathcal{S}_1 + \mathcal{S}_2$ es cualquier base de $\mathbb{R}^4$ (por ejemplo la canónica).

$b)$ Resolviendo el sistema homogéneo asociado a A, se puede ver que $\mathcal{S}_2 := \text{nul}(A) = gen\{\begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}\}$. Una base de $\mathcal{S}_2 = \text{nul}(A)$ puede ser $B_{\mathcal{S}_2} = \{\begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}\}$. Por otra parte, por el Teorema de la Dimensión, vale que $\dim(\text{col}(A)) = \#$ columnas de $A - \dim(\text{nul}(A)) = 5 - 2 = 3$. Además, es claro que $\mathcal{S}_1 := \text{col}(A) = gen\{\begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 3 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \\ -5 \\ -6 \\ -6 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 4 \end{bmatrix}\}$. Entonces, para obtener una base de $\mathcal{S}_1 = \text{col}(A)$ basta tomar 3 generadores LI de $\text{col}(A)$. Por ejemplo, $B_{\mathcal{S}_1} = \{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \\ -5 \\ -6 \\ -6 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 4 \end{bmatrix}\}$. Ahora sólo resta calcular $\mathcal{S}_1 \cap \mathcal{S}_2$ y $\mathcal{S}_1 + \mathcal{S}_2$.

Si $x \in \mathcal{S}_1 \cap \mathcal{S}_2$ entonces, por un lado $x \in \mathcal{S}_2$, por lo tanto $x = a \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ con $a, b \in \mathbb{R}$ y, por el otro lado $x \in \mathcal{S}_1$, entonces $x = c\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix} + d\begin{bmatrix}-2\\-4\\-5\\-6\\-6\end{bmatrix} + e\begin{bmatrix}1\\2\\3\\4\\4\end{bmatrix}$ con $c, d, e \in \mathbb{R}$. Por lo tanto

$$a\begin{bmatrix}-2\\-1\\0\\1\\1\end{bmatrix} + b\begin{bmatrix}3\\2\\1\\0\\0\end{bmatrix} - c\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix} - d\begin{bmatrix}-2\\-4\\-5\\-6\\-6\end{bmatrix} - e\begin{bmatrix}1\\2\\3\\4\\4\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix}.$$

Entonces, tenemos que resolver el sistema homogeneo

$$\begin{bmatrix}-2 & 3 & -1 & 2 & -1\\ -1 & 2 & 0 & 4 & -2\\ 0 & 1 & 0 & 5 & -3\\ 1 & 0 & 0 & 6 & -4\\ 1 & 0 & 0 & 6 & -4\end{bmatrix}\begin{bmatrix}a\\b\\c\\d\\e\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix}.$$

Resolviendo, nos queda que

$$a = 4e - 6d,\ b = 3e - 5d,\ c = -d.$$

Volviendo a cualquiera de las dos ecuaciones de x, tenemos que

$$x = a\begin{bmatrix}-2\\-1\\0\\1\\1\end{bmatrix} + b\begin{bmatrix}3\\2\\1\\0\\0\end{bmatrix} = (4e-6d)\begin{bmatrix}-2\\-1\\0\\1\\1\end{bmatrix} + (3e-5d)\begin{bmatrix}3\\2\\1\\0\\0\end{bmatrix} = d\begin{bmatrix}-3\\-4\\-5\\-6\\-6\end{bmatrix} + e\begin{bmatrix}1\\2\\3\\4\\4\end{bmatrix} \text{ con } d, e \in \mathbb{R}.$$

Por lo tanto $\mathcal{S}_1 \cap \mathcal{S}_2 = gen = \{\begin{bmatrix}-3\\-4\\-5\\-6\\-6\end{bmatrix}, \begin{bmatrix}1\\2\\3\\4\\4\end{bmatrix}\}$ y una base de $\mathcal{S}_1 \cap \mathcal{S}_2$ puede ser $B_{\mathcal{S}_1 \cap \mathcal{S}_2} = \{\begin{bmatrix}-3\\-4\\-5\\-6\\-6\end{bmatrix}, \begin{bmatrix}1\\2\\3\\4\\4\end{bmatrix}\}$.

Finalmente, es claro que $\mathcal{S}_1 + \mathcal{S}_2 = gen\{\begin{bmatrix}-2\\-1\\0\\1\\1\end{bmatrix}, \begin{bmatrix}3\\2\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}-2\\-4\\-5\\-6\\-6\end{bmatrix}, \begin{bmatrix}1\\2\\3\\4\\4\end{bmatrix}\}$. Además, $\dim(S_1 + \mathcal{S}_2) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) - \dim(\mathcal{S}_1 \cap \mathcal{S}_2) = 2 + 3 - 2 = 3$. Por lo tanto, debemos extraer 3 generadores LI de $\mathcal{S}_1 + \mathcal{S}_2$. Notar que

$$\begin{bmatrix}-2\\-4\\-5\\-6\\-6\end{bmatrix} = -6\begin{bmatrix}-2\\-1\\0\\1\\1\end{bmatrix} - 5\begin{bmatrix}3\\2\\1\\0\\0\end{bmatrix} + \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}1\\2\\3\\4\\4\end{bmatrix} = 4\begin{bmatrix}-2\\-1\\0\\1\\1\end{bmatrix} + 3\begin{bmatrix}3\\2\\1\\0\\0\end{bmatrix}.$$

Entonces, una base de $\mathcal{S}_1 + \mathcal{S}_2$ puede ser $B_{\mathcal{S}_1 + \mathcal{S}_2} = \{\begin{bmatrix}-2\\-1\\0\\1\\1\end{bmatrix}, \begin{bmatrix}3\\2\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix}\}$.

$c)$: Por la Proposición 1.13.2, tenemos que

$$\mathcal{S}_1 + \mathcal{S}_2 = gen\{\begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \end{bmatrix}\}.$$

Entonces sólo tenemos que extraer una base, viendo la independencia lineal de los generadores de $\mathcal{S}_1 + \mathcal{S}_2$. Lo podemos hacer por definición o colocando los vectores en la fila de una matriz y triangulando. De cualquier manera veremos que $\begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \end{bmatrix}$ es CL de los otros vectores. Entonces $B_{\mathcal{S}_1+\mathcal{S}_2} = \{\begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 2 \\ 0 \end{bmatrix}\}$. Finalmente, calculemos $\mathcal{S}_1 \cap \mathcal{S}_2$. Sea $x \in \mathcal{S}_1 \cap \mathcal{S}_2$ entonces, $x \in \mathcal{S}_1$ y $x \in \mathcal{S}_2$. Entonces $x = a\begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ con $a, b \in \mathbb{R}$ y $x = c\begin{bmatrix} 4 \\ 2 \\ 2 \\ 0 \end{bmatrix} + d\begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \end{bmatrix}$ con $c, d \in \mathbb{R}$. Entonces

$$a\begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = c\begin{bmatrix} 4 \\ 2 \\ 2 \\ 0 \end{bmatrix} + d\begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \end{bmatrix}.$$

Operando, nos queda el siguiente sistema de ecuaciones,

$$a + b - 4c - 2d = 0,\ b - 2c = 0,\ 2a + b - 2c - 2d = 0,\ a + b = 0.$$

Despejando nos queda, $a = -2c$, $b = 2c$, $d = -2c$. Volviendo a la expresión de x, nos queda $x = a\begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = c\begin{bmatrix} 0 \\ 2 \\ -2 \\ 0 \end{bmatrix}$, con $c \in \mathbb{R}$. Nos hubiera dado lo mismo si reemplazamos en la otra expresión de x. Es decir, $x = c\begin{bmatrix} 4 \\ 2 \\ 2 \\ 0 \end{bmatrix} + d\begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \end{bmatrix} = c\begin{bmatrix} 0 \\ 2 \\ -2 \\ 0 \end{bmatrix}$ con $c \in \mathbb{R}$. En conclusión $\mathcal{S}_1 \cap \mathcal{S}_2 = gen\{\begin{bmatrix} 0 \\ 2 \\ -2 \\ 0 \end{bmatrix}\}$ y una base podría ser $B_{\mathcal{S}_1\cap\mathcal{S}_2} = \{\begin{bmatrix} 0 \\ 2 \\ -2 \\ 0 \end{bmatrix}\}$. ■

Suma directa de subespacios

Definición. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y $\mathcal{S}$ y $\mathcal{T}$ subespacios de $\mathbb{V}$. Diremos que $\mathcal{S}$ y $\mathcal{T}$ están en suma directa (y la notamos $\mathcal{S} \oplus \mathcal{T}$) si para cada $v \in \mathcal{S} + \mathcal{T}$ existen únicos $s \in \mathcal{S}$ y $t \in \mathcal{T}$ tales que $v = s + t$.

Proposición 1.13.4. *Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y $\mathcal{S}$ y $\mathcal{T}$ subespacios de $\mathbb{V}$. Entonces $\mathcal{S}$ y $\mathcal{T}$ están en suma directa si y sólo si $\mathcal{S} \cap \mathcal{T} = \{0\}$.*

Dem. Supongamos que $\mathcal{S}$ y $\mathcal{T}$ están en suma directa. Sea $x \in \mathcal{S} \cap \mathcal{T}$, entonces $x \in \mathcal{S}$ y $x \in \mathcal{T}$. Entonces

$$x = x + 0_{\mathbb{V}} = 0_{\mathbb{V}} + x.$$

Como la escritura de x es única (estamos suponiendo que $\mathcal{S}$ y $\mathcal{T}$ están en suma directa), entonces $x = 0_{\mathbb{V}}$ y se sigue que $\mathcal{S} \cap \mathcal{T} = \{0\}$.

Recíprocamente, supongamos que $\mathcal{S} \cap \mathcal{T} = \{0\}$. Sea $v \in \mathcal{S} + \mathcal{T}$ y supongamos que $v = s_1 + t_1 = s_2 + t_2$ con $s_1, s_2 \in \mathcal{S}$ y $t_1, t_2 \in \mathcal{T}$. Entonces $(s_1 - s_2) = (t_2 - t_1)$. Por lo tanto, $(s_1 - s_2) \in \mathcal{S} \cap \mathcal{T} = \{0\}$. Entonces $s_1 = s_2$ y $t_1 = t_2$. Es decir, cada $v \in \mathcal{S} + \mathcal{T}$ se escribe de manera única como $v = s + t$ con $s \in \mathcal{S}$ y $t \in \mathcal{T}$. Por lo tanto, $\mathcal{S}$ y $\mathcal{T}$ están en suma directa. ∎

El siguiente ejercicio es muy similar al **Ejercicio 1.34**.

Ejercicio 1.S. En $\mathbb{R}_3[x]$ consideremos los subespacios $\mathcal{S}_1 = gen\{1 + x, 1 - x^2, 1 + x^3\}$ y $\mathcal{S}_2 = \{p \in \mathbb{R}_3[x] : p(x) = 0\}$. Hallar un subespacio $\mathcal{T}$ de $\mathbb{R}_3[x]$ tal que

$$\mathcal{S}_1 \oplus \mathcal{T} = \mathcal{S}_2 \oplus \mathcal{T} = \mathbb{R}_3[x].$$

¿Es único?

Dem. Primero busquemos un sistema de generadores de $\mathcal{S}_2$. Sea $p \in \mathcal{S}_2$, entonces $p(x) = ax^3 + bx^2 + cx + d$ con $a, b, c, d \in \mathbb{R}$ y además $p(0) = d = 0$. Entonces $\mathcal{S}_2 = gen\{x^3, x^2, x\}$. Observar que $\dim(\mathcal{S}_1) = \dim(\mathcal{S}_2) = 3$. Entonces, como $\dim(\mathbb{R}_3[x]) = 4$, por el Teorema de la dimensión (Teorema 1.13.3), $\dim(\mathcal{T}) = \dim(\mathcal{S}_1 \oplus \mathcal{T}) - \dim(\mathcal{S}_1) - \dim(\mathcal{S}_1 \cap \mathcal{T}) = \dim(\mathbb{R}_3[x]) - 3 - 0 = 4 - 3 = 1$ y $\mathcal{T}$ debe ser un subespacio de dimensión 1 que está en suma directa con $\mathcal{S}_1$ y $\mathcal{S}_2$, es decir, el generador de $\mathcal{T}$ debe ser LI con los 3 generadores de $\mathcal{S}_1$ y $\mathcal{S}_2$. Proponemos $\mathcal{T} = gen\{1\}$. Claramente $\dim(\mathcal{T}) = 1$, además como los conjuntos $\{1, 1 + x, 1 - x^2, 1 + x^3\}$ y $\{1, x, x^2, x^3\}$ son LI, se sigue que $\mathcal{S}_1 \cap \mathcal{T} = \mathcal{S}_2 \cap \mathcal{T} = \{0\}$ y el $\mathcal{T}$ propuesto cumple. El $\mathcal{T}$ posible no es único. Por ejemplo, el subespacio $\mathcal{T} := gen\{2 + x\}$ también sirve. ∎

El siguiente ejercicio es muy similar al **Ejercicio 1.35**.

Ejercicio 1.T (De Parcial)**.** En $\mathbb{R}_3[x]$ consideremos los subespacios $\mathcal{S}_3 = gen\{1 + x, 1 - x^2\}$ y $\mathcal{S}_4 = \{p \in \mathbb{R}_3[x] : p(-1) = 0\}$. Hallar un subespacio $\mathcal{T}$ de $\mathbb{R}_3[x]$ tal que $\mathcal{S}_3 \oplus \mathcal{T} = \mathcal{S}_4$.

Dem. Primero observar que $\mathcal{S}_3 \subseteq \mathcal{S}_4$ porque sino el ejercicio no tendría sentido. De hecho, si $p \in \mathcal{S}_3$ entonces $p(x) = a(1 + x) + b(1 - x^2)$ con $a, b \in \mathbb{R}$. Entonces $p(-1) = a \cdot 0 + b \cdot 0 = 0$ y $p \in \mathcal{S}_4$. Entonces vale $\mathcal{S}_3 \subseteq \mathcal{S}_4$.

Por otra parte, busquemos un sistema de generadores de $\mathcal{S}_4$. Si $p \in \mathcal{S}_4$ entonces $p(x) = ax^3 + bx^2 + cx + d$ con $a, b, c, d \in \mathbb{R}$ y además $p(-1) = -a + b - c + d = 0$. Entonces $a = b - c + d$ y volviendo a la expresión de p, nos queda $p(x) = b(x^3 + x^2) + c(x^3 - x) + d(x^3 + 1)$, entonces $\mathcal{S}_4 = gen\{x^3 + x^2, x^3 - x, x^3 + 1\}$ y $\dim(\mathcal{S}_4) = 3$. Entonces, el subespacio $\mathcal{T}$ que buscamos debe cumplir 3 cosas:

1. $\mathcal{T} \subseteq \mathcal{S}_4$.
2. $\dim(\mathcal{T}) = \dim(\mathcal{S}_3 \oplus \dim \mathcal{T}) - \dim(\mathcal{S}_3) - \dim(\mathcal{S}_3 \cap \mathcal{T}) = \dim(\mathcal{S}_4) - 2 - 0 = 3 - 2 = 1$.
3. El generador de $\mathcal{T}$ debe ser LI con los dos generadores de $\mathcal{S}_3$.

Proponemos $\mathcal{T} = gen\{x^3 + 1\}$. Claramente, $\mathcal{T} \subseteq \mathcal{S}_4$, $\dim(\mathcal{T}) = 1$ y como el conjunto $\{1 + x, 1 - x^2, x^3 + 1\}$ es LI entonces $\mathcal{S}_3 \cap \mathcal{T} = \{0\}$. Por lo tanto, el $\mathcal{T}$ propuesto cumple, ya que $\mathcal{T}$ y $\mathcal{S}_3$ están en suma directa, $\mathcal{T} \oplus \mathcal{S}_3 \subseteq \mathcal{S}_4$ y $\dim(\mathcal{T} \oplus \mathcal{S}_3) = 3 = \dim(\mathcal{S}_4)$, lo que nos da que $\mathcal{S}_3 \oplus \mathcal{T} = \mathcal{S}_4$. ∎

En el ejercicio anterior, observar que, por ejemplo, el subespacio $\mathcal{T}' = gen\{x^3\}$ está en suma directa con $\mathcal{S}_3$ y tiene dimensión 1, pero como $\mathcal{T}' \not\subseteq \mathcal{S}_4$ no sirve, es decir $\mathcal{S}_3 \oplus \mathcal{T} \neq \mathcal{S}_4$.

CAPÍTULO 2

Transformaciones Lineales

En este capítulo introduciremos el concepto de transformaciones lineales que son las funciones con las que trabajaremos en la materia. Veremos varias propieadades y resultados asociados a estas funciones.

Definición. Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales. Una función $T : \mathbb{V} \to \mathbb{W}$ se llama *transformación lineal* (TL) de $\mathbb{V}$ en $\mathbb{W}$ si cumple:

i) $T(v + v') = T(v) + T(v'),$ para todo $v,\ v' \in \mathbb{V}$,

ii) $T(\alpha v) = \alpha T(v),$ para todo $\alpha \in \mathbb{K},\ v \in \mathbb{V}$.

El conjunto de todas las transformaciones lineales de $\mathbb{V}$ a $\mathbb{W}$ se denota por $\mathcal{L}(\mathbb{V}, \mathbb{W})$. Si $\mathbb{V} = \mathbb{W}$ escribimos $\mathcal{L}(\mathbb{V}) := \mathcal{L}(\mathbb{V}, \mathbb{V})$. A las transformaciones lineales de $\mathbb{V}$ a $\mathbb{W} = \mathbb{K}$ (es decir aquellas cuya imagen cae en el cuerpo) se las llama *funcionales lineales*.

Ejercicio 2.1 c). Verificar que la siguiente aplicación es una transformación lineal: $T_3 : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$T_3([x_1\ x_2\ x_3]^T) := [-3x_2 + 2x_3\ 3x_1 - x_3\ \ -2x_1 + x_2]^T.$$

Dem. Observar que $T_3(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}) := \begin{bmatrix} -3x_2 + 2x_3 \\ 3x_1 - x_3 \\ -2x_1 + x_2 \end{bmatrix} = \begin{bmatrix} 0 & -3 & 2 \\ 3 & 0 & -1 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Si llamamos $A := \begin{bmatrix} 0 & -3 & 2 \\ 3 & 0 & -1 \\ -2 & 1 & 0 \end{bmatrix}$ entonces, si $v = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$, tenemos que $T_3(v) = Av$. Con esto en mente, verifiquemos $i)$ y $ii)$ de la definición de transformaciones lineales.

Sean $\alpha \in \mathbb{R}$ y $v, v' \in \mathbb{R}^3$, entonces

i) $T(v + v') = A(v + v') = Av + Av' = T(v) + T(v')$.

ii) $T(\alpha v) = A(\alpha v) = \alpha(Av) = \alpha T(v)$.

Por lo tanto, como T cumple con $i)$ y $ii)$ de la definición de transformaciones lineales, tenemos que T es una transformación lineal, es decir $T \in \mathcal{L}(\mathbb{R}^3)$. ∎

Operando tal como hicimos arriba podemos demostrar que si $T \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ entonces (siempre) existe $A \in \mathbb{K}^{m\times n}$ tal que $T(x) = Ax,$ para todo $x \in \mathbb{K}^n$. Ver el **Ejercicio 2.2**.

Veamos otros ejemplos de transformaciones lineales:

Ejemplo 2.a. Verificar las siguientes afirmaciones:

- Para cada $A \in \mathbb{K}^{m\times n}$, la aplicación $T_A : \mathbb{K}^{m\times n} \to \mathbb{K}$ definida por

  $$T_A(X) := \operatorname{tr}(A^* X)$$

  es una funcional lineal de $\mathbb{K}^{m\times n}$.

- Dados $n \in \mathbb{N}$ y $a_0, a_1, \cdots, a_{n-1} \in \mathbb{R}$, la aplicación $L : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ definida por

$$L[y] := \frac{d^n y}{dx^n} + a_{n-1}\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1\frac{dy}{dx} + a_0 y$$

es una trasformación lineal de $C^\infty(\mathbb{R})$ a $C^\infty(\mathbb{R})$.

Recordemos las propiedades de la traza de una matriz (se sugiere verificarlas haciendo la cuenta):

1. $\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$, para todo $A, B \in \mathbb{K}^{n\times n}$.
2. $\operatorname{tr}(\alpha A) = \alpha \operatorname{tr}(A)$, para todo $\alpha \in \mathbb{K}$, $A \in \mathbb{K}^{n\times n}$.
3. $\operatorname{tr}(A^*) = \overline{\operatorname{tr}(A)}$, para todo $A \in \mathbb{K}^{n\times n}$,
4. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, para todo $A \in \mathbb{K}^{n\times m}, B \in \mathbb{K}^{m\times n}$.

Dem. Verifiquemos $i)$ y $ii)$ de la definición de transformaciones lineales usando las propiedades que acabamos de ver. Sean $\alpha \in \mathbb{R}$ y $X, Y \in \mathbb{K}^{m\times n}$, entonces

i) $T_A(X+Y) = \operatorname{tr}(A^*(X+Y)) = \operatorname{tr}(A^*X + A^*Y) = \operatorname{tr}(A^*X) + \operatorname{tr}(A^*Y) = T_A(X) + T_A(Y)$, donde usamos la propiedad 1 de la traza.

ii) $T_A(\alpha X) = \operatorname{tr}(A^*(\alpha X)) = \operatorname{tr}(\alpha A^* X) = \alpha \operatorname{tr}(A^*X) = \alpha T_A(X)$, donde usamos la propiedad 2 de la traza.

Entonces como T_A cumple con $i)$ y $ii)$ de la definición de transformaciones lineales, tenemos que $T_A \in \mathcal{L}(\mathbb{K}^{m\times n}, \mathbb{K})$ es decir es una funcional lineal de $\mathbb{K}^{m\times n}$.

Verifiquemos $i)$ y $ii)$ de la definición de transformaciones lineales usando la linealidad de la derivada. Sean $\alpha \in \mathbb{R}$ e $y, z \in C^\infty(\mathbb{R})$, entonces

i) $L[y+z] = \frac{d^n(y+z)}{dx^n} + a_{n-1}\frac{d^{n-1}(y+z)}{dx^{n-1}} + \cdots + a_1\frac{d(y+z)}{dx} + a_0(y+z) = \frac{d^n y}{dx^n} + \frac{d^n z}{dx^n} + a_{n-1}\frac{d^{n-1}y}{dx^{n-1}} + a_{n-1}\frac{d^{n-1}z}{dx^{n-1}} + \cdots + a_1\frac{dy}{dx} + a_1\frac{dz}{dx} + a_0 y + a_0 z = L[y] + L[z]$.

ii) $L[\alpha y] = \frac{d^n(\alpha y)}{dx^n} + a_{n-1}\frac{d^{n-1}(\alpha y)}{dx^{n-1}} + \cdots + a_1\frac{d(\alpha y)}{dx} + a_0(\alpha y) = \alpha\frac{d^n y}{dx^n} + a_{n-1}\alpha\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1\alpha\frac{dy}{dx} + a_0\alpha y = \alpha(\frac{d^n y}{dx^n} + a_{n-1}\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1\frac{dy}{dx} + a_0 y) = \alpha L[y]$.

Entonces como L cumple con $i)$ y $ii)$ de la definición de transformaciones lineales, tenemos que $L \in \mathcal{L}(C^\infty(\mathbb{R}))$. ■

Ejercicio 2.A. Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales. Demostrar que

si $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ entonces $T(0_\mathbb{V}) = 0_\mathbb{W}$.

Dem. Por un lado, observar que como $0_\mathbb{V} = 0_\mathbb{V} + 0_\mathbb{V}$ y T es TL, tenemos que $T(0_\mathbb{V}) = T(0_\mathbb{V} + 0_\mathbb{V}) = T(0_\mathbb{V}) + T(0_\mathbb{V})$. Entonces, si $-T(0_\mathbb{V})$ es el inverso aditivo de $T(0_\mathbb{V})$ (existe porque $\mathbb{W}$ es un $\mathbb{K}$-espacio vectorial), tenemos que $0_\mathbb{W} = T(0_\mathbb{V}) + -T(0_\mathbb{V}) = (T(0_\mathbb{V}) + T(0_\mathbb{V})) + -T(0_\mathbb{V}) = T(0_\mathbb{V}) + (T(0_\mathbb{V}) + -T(0_\mathbb{V})) = T(0_\mathbb{V}) + 0_\mathbb{W} = T(0_\mathbb{V})$. ■

Ejercicio 2.B. Explicar porque las siguientes funciones NO son transformaciones lineales:

a) $T : \mathbb{K}^{2\times 2} \to \mathbb{K}$ definida por $T(X) := \det(X)$.

b $T : \mathbb{R}^3 \to \mathbb{R}^3$, definida por $T(x) := \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

c) $f : \mathbb{C} \to \mathbb{C}$, definida por $f(z) := \overline{z}$.

Dem. Para demostrar que una función NO es una transformación lineal, basta con encontrar algún contraejmplo donde $i)$ o $ii)$ de la definición de transformaciones lineales NO se cumpla.

$a)$: Sean $X = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ e $Y = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Entonces $T(X) + T(Y) = \det(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}) + \det(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}) = 0 + 0 = 0 \neq 1 = \det(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}) = T(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}) = T(X + Y)$. Entonces T NO es una transformación lineal.

$b)$: Basta tomar cualquier vector $x, y \in \mathbb{R}^3$ y ver que $T(x) + T(y) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = T(x + y)$. Entonces T NO es una transformación lineal.

$c)$: Sean $\alpha = i$ y $x = 1$. Entonces $f(\alpha x) = f(i \cdot 1) = f(i) = \overline{i} = -i \neq i \cdot 1 = i \cdot f(1) = \alpha f(x)$. Entonces f NO es una funcional lineal. ■

2.1. Imagen y núcleo de una transformación lineal

Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales y $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. El *núcleo* de T se denota por $\mathrm{Nu}(T)$ y se define como

$$\mathrm{Nu}(T) := \{v \in \mathbb{V} : T(v) = 0_{\mathbb{W}}\} \subseteq \mathbb{V}.$$

La *imagen* de T se denota por $\mathrm{Im}(T)$ y se define como

$$\mathrm{Im}(T) := \{w \in \mathbb{W} : \exists v \in \mathbb{V} : w = T(v)\} = \{T(v) : v \in \mathbb{V}\} \subseteq \mathbb{W}.$$

Por otro lado, sean $X \subseteq \mathbb{V}$ y $Z \subseteq \mathbb{W}$ dos subconjuntos de $\mathbb{V}$ y $\mathbb{W}$ respectivamente, entonces: el conjunto de todas las imágenes de los elementos de X se designa por $T(X)$ y se define como:

$$T(X) := \{w \in \mathbb{W} : \exists x \in X : w = T(x)\} = \{T(x) : x \in X\} \subseteq \mathbb{W}.$$

El conjunto de todos los elementos de $\mathbb{V}$ cuyas imágenes pertenecen a Y se designa con $T^{-1}(Y)$ y se define como:

$$T^{-1}(Y) := \{v \in \mathbb{V} : T(v) \in Y\} \subseteq \mathbb{V}.$$

No confundir la preimagen de un conjunto con la existencia de inversa de una transformación lineal. La preimagen de un conjunto siempre está definida (podría eventualmente dar el conjunto vacío) sin embargo no siempre existe la inversa de una transformación lineal.

> Observar que $\mathrm{Nu}(T) = T^{-1}(\{0_{\mathbb{W}}\})$ e $\mathrm{Im}(T) = T(\mathbb{V})$.

Ejercicio 2.C (De Parcial)**.** Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales y $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Demostrar las siguientes afirmaciones:

a) $\mathrm{Nu}(T)$ es un subespacio de $\mathbb{V}$.

b) $\mathrm{Im}(T)$ es un subespacio de $\mathbb{W}$.

c) Si $\mathcal{S}$ es un subespacio de $\mathbb{V}$ entonces $T(\mathcal{S})$ es un subespacio de $\mathbb{W}$.

d) Si $\mathcal{T}$ es un subespacio de $\mathbb{W}$ entonces $T^{-1}(\mathcal{T})$ es un subespacio de $\mathbb{V}$.

e) Supongamos que $\mathcal{S}$ es un subespacio de $\mathbb{V}$ tal que $\mathcal{S} = gen\{v_1, v_2, \cdots, v_r\}$. Entonces

$$T(\mathcal{S}) = gen\{T(v_1), T(v_2), \cdots, T(v_r)\}.$$

f) Supongamos que $\mathbb{V} = gen\{v_1, v_2, \cdots, v_n\}$. Entonces

$$\mathrm{Im}(T) = gen\{T(v_1), T(v_2), \cdots, T(v_n)\}.$$

Dem. a): Veamos que $\mathrm{Nu}(T)$ es un subespacio de $\mathbb{V}$.

i) $0_{\mathbb{V}} \in \mathrm{Nu}(T)$ pues como T es una TL, tenemos que $T(0_{\mathbb{V}}) = 0_{\mathbb{W}}$. Ver **Ejercicio 2.A**.

ii) si $u, v \in \mathrm{Nu}(T)$ entonces $T(u) = 0_{\mathbb{W}}$ y $T(v) = 0_{\mathbb{W}}$. Entonces, como T es TL, $T(u+v) = T(u) + T(v) = 0_{\mathbb{W}} + 0_{\mathbb{W}} = 0_{\mathbb{W}}$. Por lo tanto $u + v \in \mathrm{Nu}(T)$.

iii) si $\alpha \in \mathbb{K}$ y $v \in \mathrm{Nu}(T)$ entonces $T(v) = 0_{\mathbb{W}}$. Entonces, como T es TL, $T(\alpha v) = \alpha T(v) = \alpha 0_{\mathbb{W}} = 0_{\mathbb{W}}$. Por lo tanto $\alpha v \in \mathrm{Nu}(T)$.

b): Veamos que $\mathrm{Im}(T)$ es un subespacio de $\mathbb{V}$.

i) $0_{\mathbb{W}} \in \mathrm{Im}(T)$ pues como T es una TL, tenemos que $T(0_{\mathbb{V}}) = 0_{\mathbb{W}}$ (existe $x \in \mathbb{V}$ tal que $T(x) = 0_{\mathbb{W}}$, en este caso $x = 0_{\mathbb{V}}$). Ver **Ejercicio 2.A**.

ii) si $u, v \in \mathrm{Im}(T)$ entonces existen $x, y \in \mathbb{V}$ tales que $T(x) = u$ y $T(y) = v$. Entonces, como T es TL, $T(x+y) = T(x) + T(y) = u + v$. Encontramos un vector, que en este caso es $x + y$, tal que $T(x+y) = u + v$. Por lo tanto $u + v \in \mathrm{Im}(T)$.

iii) si $\alpha \in \mathbb{K}$ y $v \in \mathrm{Im}(T)$ entonces existe $x \in \mathbb{V}$ tal que $T(x) = v$. Entonces, como T es TL, $T(\alpha x) = \alpha T(x) = \alpha v$. Encontramos un vector, que en este caso es αx, tal que $T(\alpha x) = \alpha v$. Por lo tanto $\alpha v \in \mathrm{Im}(T)$.

c): Se deja como ejercicio (es muy parecido al ítem b).

d): Veamos que $T^{-1}(\mathcal{T})$ es un subespacio de $\mathbb{V}$.

i) Veamos que $0_{\mathbb{V}} \in T^{-1}(\mathcal{T})$. Como T es una TL, $T(0_{\mathbb{V}}) = 0_{\mathbb{W}}$ (ver **Ejercicio 2.A**) y como $\mathcal{T}$ es un subespacio $0_{\mathbb{W}} \in \mathcal{T}$. Conclusión, $T(0_{\mathbb{V}}) = 0_{\mathbb{W}} \in \mathcal{T}$ entonces $0_{\mathbb{V}} \in T^{-1}(\mathcal{T})$.

ii) si $u, v \in T^{-1}(\mathcal{T})$ entonces $T(u) \in \mathcal{T}$ y $T(v) \in \mathcal{T}$. Entonces, como $\mathcal{T}$ es un subespacio, $T(u) + T(v) \in \mathcal{T}$. Además como T es TL, tenemos que $T(u+v) = T(u) + T(v) \in \mathcal{T}$. Entonces $u + v \in T^{-1}(\mathcal{T})$.

iii) si $\alpha \in \mathbb{K}$ y $v \in T^{-1}(\mathcal{T})$ entonces $T(v) \in \mathcal{T}$. Entonces, como $\mathcal{T}$ es un subespacio, $\alpha T(v) \in \mathcal{T}$. Además como T es TL, tenemos que $T(\alpha v) = \alpha T(v) \in \mathcal{T}$. Entonces $\alpha v \in T^{-1}(\mathcal{T})$.

e): Tenemos que probar una igualdad de conjuntos, así que vamos a probar la doble inclusión.

Por un lado, como $v_1, v_2, \cdots, v_r \in \mathcal{S}$, tenemos que $T(v_i) \in T(\mathcal{S})$ para todo $i = 1, 2, \cdots, r$ (recordar la definición de $T(\mathcal{S})$) y además observar que en el ítem c) probamos que $T(\mathcal{S})$ es un subspacio (porque $\mathcal{S}$ es un subespacio). Entonces

$$gen\{T(v_1), T(v_2), \cdots, T(v_r)\} \subseteq T(\mathcal{S}).$$

Veamos la otra inclusión. Supongamos que $w \in T(\mathcal{S})$, entonces existe $s \in \mathcal{S}$ tal que $w = T(s)$. Entonces, como $s \in \mathcal{S} = gen\{v_1, v_2, \cdots, v_r\}$, existen $a_1, a_2, \cdots, a_r \in \mathbb{K}$ tales que

$$s = a_1 v_1 + a_2 v_2 + \cdots + a_r v_r.$$

Entonces como T es TL,

$$w = T(s) = T(a_1v_1 + a_2v_2 + \cdots + a_rv_r) = a_1T(v_1) + a_2T(v_2) + \cdots + a_rT(v_r).$$

Por lo tanto $w \in gen\{T(v_1), T(v_2), \cdots, T(v_r)\}$ y $T(\mathcal{S}) \subseteq gen\{T(v_1), T(v_2), \cdots, T(v_r)\}$. Como probamos la doble inclusión, vale la igualdad de conjuntos.

$f)$: Como $\mathrm{Im}(T) = T(\mathbb{V})$, el item $f)$ es un caso particular del item $e)$. ■

Ejercicio 2.D. Sea $\mathbb{K} = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$ y sea $A \in \mathbb{K}^{n\times m}$. Consideremos la transformación lineal $T_A : \mathbb{K}^m \to \mathbb{K}^n$ definida por

$$T_A(x) := Ax,$$

para cada $x \in \mathbb{K}^m$. Demostrar que $\mathrm{Nu}(T_A) = \mathrm{nul}(A)$ e $\mathrm{Im}(T_A) = \mathrm{col}(A)$.

Dem. Para demostrar lo que nos pide el ejercicio, basta con aplicar la definición de los subespacios involucrados.

Por un lado, $x \in \mathrm{Nu}(T_A)$ si y sólo si

$$T_A(x) = Ax = 0_{\mathbb{K}^n},$$

si y sólo si $x \in \mathrm{nul}(A)$. Finalmente, $y \in \mathrm{Im}(T_A)$ si y sólo si existe $x \in \mathbb{K}^m$ tal que

$$y = T_A(x) = Ax,$$

si y sólo si $y \in \mathrm{col}(A)$. ■

A continuación probaremos el teorema de la dimensión para transformaciones lineales:

Teorema 2.1.1. *Sean $\mathbb{V}, \mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales, $\mathbb{V}$ de dimensión finita y $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Entonces*

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

Dem. Sea $n = \dim(\mathbb{V})$ y $r = \dim(\mathrm{Nu}(T))$. Supongamos que $0 < r < n$ y sea $\{v_1, \cdots, v_r\}$ una base de $\mathrm{Nu}(T)$. Sean $v_{r+1}, \cdots, v_n \in \mathbb{V}$ de manera tal que $\{v_1, \cdots, v_r, v_{r+1}, \cdots, v_n\}$ es una base de $\mathbb{V}$.

Afirmamos que $\{T(v_{r+1}), \cdots, T(v_n)\}$ es una base de $\mathrm{Im}(T)$. De hecho, claramente $T(v_{r+1}), \cdots, T(v_n) \in \mathrm{Im}(T)$. Por lo que $gen\{T(v_{r+1}), \cdots, T(v_n)\} \subseteq \mathrm{Im}(T)$. Recíprocamente, sea $w \in \mathrm{Im}(T)$, entonces existe $v \in \mathbb{V}$ tal que $w = T(v)$. Como $\{v_1, \cdots, v_r, v_{r+1}, \cdots, v_n\}$ es una base de $\mathbb{V}$, existen $a_1, \cdots, a_r, a_{r+1}, \cdots, a_n \in \mathbb{K}$ tales que $v = a_1v_1 + \cdots + a_rv_r + a_{r+1}v_{r+1} + \cdots + a_nv_n$. Entonces, usando que T es TL, tenemos que

$$\begin{aligned} w = T(v) &= T(a_1v_1 + \cdots + a_rv_r + a_{r+1}v_{r+1} + \cdots + a_nv_n) \\ &= a_1T(v_1) + \cdots + a_rT(v_r) + a_{r+1}T(v_{r+1}) + \cdots + a_nT(v_n) \\ &= a_10_{\mathbb{W}} + \cdots + a_r0_{\mathbb{W}} + a_{r+1}T(v_{r+1}) + \cdots + a_nT(v_n) \\ &= a_{r+1}T(v_{r+1}) + \cdots + a_nT(v_n) \in gen\{T(v_{r+1}), \cdots, T(v_n)\}, \end{aligned}$$

donde usamos que $\{v_1, \cdots, v_r\}$ una base de $\mathrm{Nu}(T)$. Entonces, $\mathrm{Im}(T) = gen\{T(v_{r+1}), \cdots, T(v_n)\}$.

Veamos que $\{T(v_{r+1}), \cdots, T(v_n)\}$ es un conjunto LI. Sean $\alpha_{r+1}, \cdots, \alpha_n \in \mathbb{K}$ tales que

$$\alpha_{r+1}T(v_{r+1}) + \cdots + \alpha_nT(v_n) = 0_{\mathbb{W}}.$$

Entonces, como T es TL tenemos que

$$T(\alpha_{r+1}v_{r+1} + \cdots + \alpha_nv_n) = 0_{\mathbb{W}}.$$

Entonces, $\alpha_{r+1}v_{r+1} + \cdots + \alpha_nv_n \in \mathrm{Nu}(T)$. Como $\{v_1, \cdots, v_r\}$ una base de $\mathrm{Nu}(T)$, existen $\alpha_1, \cdots, \alpha_r \in \mathbb{K}$ tales que

$$\alpha_{r+1}v_{r+1} + \cdots + \alpha_nv_n = \alpha_1v_1 + \cdots + \alpha_rv_r.$$

Entonces

$$\alpha_{r+1}v_{r+1} + \cdots + \alpha_n v_n - \alpha_1 v_1 - \cdots - \alpha_r v_r = 0_{\mathbb{V}}.$$

Como $\{v_1, \cdots, v_r, v_{r+1}, \cdots, v_n\}$ es una base de $\mathbb{V}$, en particular es un conjunto LI y tenemos una CL de dichos vectores igualada a $0_{\mathbb{V}}$. Por lo tanto, $\alpha_1 = \cdots = \alpha_r = \cdots = \alpha_n = 0$. Entonces, $\{T(v_{r+1}), \cdots, T(v_n)\}$ es una base de $\mathrm{Im}(T)$ y se sigue que $\dim(\mathrm{Im}(T)) = n - r$. Entonces,

$$\dim(\mathbb{V}) = n = r + (n - r) = \dim(\mathrm{Nu}(T)) + \dim(\mathrm{Im}(T))$$

y probamos lo que queríamos.

Observar que si $r = n$, entonces T es la función nula. En ese caso $\dim(\mathrm{Im}(T)) = 0$ y el teorema vale.

Si $r = 0$, entonces de manera similar al caso $0 < r < n$ se puede ver que si $\{v_1, \cdots, v_n\}$ es una base de $\mathbb{V}$ entonces $\{T(v_1), \cdots, T(v_n)\}$ es una base de $\mathrm{Im}(T)$. Por lo tanto $\dim(\mathrm{Im}(T)) = n$ y el teorema vale. ■

A partir del Teorema de la dimensión para transformaciones lineales, se puede probar fácilmente el **Teorema 1.9.1** para matrices: Sea $A \in \mathbb{R}^{n\times m}$. Entonces,

$$\dim(\mathrm{nul}(A)) + \mathrm{rg}(A) = m.$$

Sea $A \in \mathbb{C}^{n\times m}$ y supongamos que estamos trabajando en el cuerpo $\mathbb{K} = \mathbb{C}$. Entonces,

$$\dim(\mathrm{nul}(A)) + \mathrm{rg}(A) = m.$$

Sea $A \in \mathbb{C}^{n\times m}$ y supongamos que estamos trabajando en el cuerpo $\mathbb{K} = \mathbb{R}$. Entonces,

$$\dim(\mathrm{nul}(A)) + \mathrm{rg}(A) = 2m.$$

Dem. Para demostrar este resultado, vamos a considerar la transformación lineal asociada a la matriz A, es decir, la transformación lineal $T_A : \mathbb{K}^m \to \mathbb{K}^n$ definida por

$$T_A(x) := Ax,$$

para cada $x \in \mathbb{K}^m$.

En el **Ejercicio 2.D** probamos que $\mathrm{Nu}(T_A) = \mathrm{nul}(A)$ e $\mathrm{Im}(T_A) = \mathrm{col}(A)$. Entonces, $\dim(\mathrm{Nu}(T_A)) = \dim(\mathrm{nul}(A))$ y $\dim(\mathrm{Im}(T_A)) = \dim(\mathrm{col}(A)) = \mathrm{rg}(A)$. Entonces, si $A \in \mathbb{R}^{n\times m}$, $T_A \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ y como $\dim(\mathbb{R}^m) = m$, por el Teorema 2.1.1, tenemos que $m = \dim(\mathrm{nul}(A)) + \mathrm{rg}(A)$.

Análogamente, si $A \in \mathbb{C}^{n\times m}$ y estamos trabajando en el cuerpo $\mathbb{K} = \mathbb{C}$, como $\dim(\mathbb{C}^m_{\mathbb{C}}) = m$, por el Teorema 2.1.1, tenemos que $m = \dim(\mathrm{nul}(A)) + \mathrm{rg}(A)$.

Finalmente, si $A \in \mathbb{C}^{n\times m}$ y estamos trabajando en el cuerpo $\mathbb{K} = \mathbb{R}$, como $\dim(\mathbb{C}^m_{\mathbb{R}}) = 2m$, por el Teorema 2.1.1, tenemos que $2m = \dim(\mathrm{nul}(A)) + \mathrm{rg}(A)$. ■

Ejercicio 2.5. Sea $T : \mathbb{R}_3[x] \to \mathbb{R}_3[x]$ definda por

$$T(p) := p + (1 - x)p'.$$

a) Explicar por qué T está bien definida y es una transformación lineal.

b) Hallar una base del núcleo de T.

c) Hallar una base de la imagen de T.

d) Comprobar que el polinomio $q(x) = 1 + x + x^2 - x^3$ pertenece a la imagen de T y resolver la ecuación $T(p) = q$.

Dem. $a)$: Para ver que T está bien definida tenemos que probar que a cada elemento del dominio de T (es decir a cada $p \in \mathbb{R}_3[x]$) se le asigna un único elemento del conjunto de llegada (también $\mathbb{R}_3[x]$).

Observar que si p pertenece a $\mathbb{R}_3[x]$ (el dominio de T) entonces p es derivable y la operación $p + (1-x)p'$ se puede hacer sin problemas. Entonces, a todo elemento del dominio se le puede aplicar T. Por otra parte, si p pertenece a $\mathbb{R}_3[x]$, en primer lugar su derivada es única (esto ahora suena obvio, pero en algún momento uno tuvo que probar, que la derivada de una función es única) y además la derivada de un elemento de $\mathbb{R}_3[x]$ es un elemento de $\mathbb{R}_2[x]$ (es decir es un polinomio de grado igual o menor a 2 o el polinomio nulo). Finalmente, es claro que, el producto de un polinomio de $\mathbb{R}_2[x]$ y el polinomio $1-x$ nos devuelve un polinomio de $\mathbb{R}_3[x]$ y la suma de elementos de $\mathbb{R}_3[x]$ también es un elemento de $\mathbb{R}_3[x]$. Por lo tanto, cuando hacemos la operación $p+(1-x)p'$ obtenemos un (único) elemento del espacio de llegada que es $\mathbb{R}_3[x]$. Conclusión, a cada elemento del dominio de T le corresponde un único elemento del espacio de llegada y entonces T está bien definida.

Ahora veamos que T es lineal, para eso, llamemos $r(x) := 1 - x$. Entonces, $T(p) = p + rp'$.

Sean $p_1, p_2 \in \mathbb{R}_3[x]$. Usando que la derivada es lineal,

$$\begin{aligned} T(p_1 + p_2) &= p_1 + p_2 + r(p_1 + p_2)' = p_1 + p_2 + r(p_1' + p_2') \\ &= p_1 + rp_1' + p_2 + rp_2' = T(p_1) + T(p_2). \end{aligned}$$

De la misma manera, sea $p \in \mathbb{R}_3[x]$ y $\alpha \in \mathbb{R}$. Usando que la derivada es lineal

$$T(\alpha p) = \alpha p + r(\alpha p)' = \alpha + r\alpha p' = \alpha(p + rp') = \alpha T(p).$$

$b)$: Vamos a hallar una base de $\mathrm{Nu}(T) = \{p \in \mathbb{R}_3[x] : T(p) = 0\}$. Si $p \in \mathrm{Nu}(T)$ entonces $p \in \mathbb{R}_3[x]$, es decir $p(x) = a + bx + cx^2 + dx^3$ con $a, b, c, d \in \mathbb{R}$ y, para todo $x \in \mathbb{R}$,

$$\begin{aligned} 0 = T(p)(x) &= a + bx + cx^2 + dx^3 + (1-x)(a + bx + cx^2 + dx^3)' \\ &= a + bx + cx^2 + dx^3 + (1-x)(b + 2cx + 3dx^2) \\ &= 1 \cdot (a + b) + x(b + 2c - b) + x^2(c + 3d - 2c) + x^3(d - 3d) \\ &= 1 \cdot (a + b) + x \cdot (2c) + x^2(-c + 3d) + x^3(-2d). \end{aligned}$$

Entonces, $-2d = 0,\ -c + 3d = 0, 2c = 0$ y $a + b = 0$. Por lo tanto, $a = -b$ y $c = d = 0$. Reemplazando en la expresión de p, nos queda $p(x) = a + bx + cx^2 + dx^3 = a - ax$, con $a \in \mathbb{R}$. Por lo tanto,

$$\mathrm{Nu}(T) = \{p \in \mathbb{R}_3[x] : p(x) = a(1-x), a \in \mathbb{R}\} = gen\{1 - x\}.$$

Una base de $\mathrm{Nu}(T)$ puede ser, $B_{\mathrm{Nu}(T)} = \{1 - x\}$.

$c)$: En el Ejercicio 2.C, vimos que como $\{1, x, x^2, x^3\}$ es una base de $\mathbb{R}_3[x]$ (en particular es un sistema de generadores), entonces

$$\mathrm{Im}(T) = gen\{T(1), T(x), T(x^2), T(x^3)\}.$$

Haciendo cuentas, nos queda que $T(1) = 1 + (1-x)(1)' = 1,\ T(x) = x + (1-x)(x)' = x + 1 - x = 1,\ T(x^2) = x^2 + (1-x)(x^2)' = x^2 + (1-x)2x = -x^2 + 2x,\ T(x^3) = x^3 + (1-x)(x^3)' = x^3 + (1-x)3x^2 = 3x^2 - 2x^3$. Por lo tanto,

$$\mathrm{Im}(T) = gen\{1, 1, -x^2 + 2x, 3x^2 - 2x^3\} - = gen\{1, -x^2 + 2x, 3x^2 - 2x^3\}.$$

Observar que los vectores $1, -x^2 + 2x, 3x^2 - 2x^3$ son LI, eso se puede ver (por ejemplo) observando que los polinomios del conjunto son de distinto grado. Entonces, una base de $\mathrm{Im}(T)$ puede ser

$$B_{\mathrm{Im}(T)} = \{1, -x^2 + 2x, 3x^2 - 2x^3\}.$$

$d)$: Notar que

$$q(x) = 1 + x + x^2 - x^3 = 1 \cdot 1 + \frac{1}{2} \cdot (-x^2 + 2x) + \frac{1}{2} \cdot (3x^2 - 2x^3).$$

Por lo tanto, $q \in \mathrm{Im}(T)$ y entonces, la ecuación $T(p) = q$ tiene solución. En este caso (al igual que con ecuaciones matriciales) tenemos que $T(p) = q$ si y sólo si

$$p = p_h + p_p,$$

donde $p_h \in \mathrm{Nu}(T)$(solución del sistema homogéneo) y p_p es una solución particular de $T(p) = q$.

Observar que

$$\begin{aligned} T(1 \cdot 1 + \frac{1}{2} \cdot x^2 + \frac{1}{2} \cdot x^3) &= 1T(1) + \frac{1}{2}T(x^2) + \frac{1}{2}T(x^3) \\ &= 1 \cdot 1 + \frac{1}{2} \cdot (-x^2 + 2x) + \frac{1}{2} \cdot (3x^2 - 2x^3) = q(x). \end{aligned}$$

Por lo tanto, si $p_p(t) := 1 \cdot 1 + \frac{1}{2} \cdot x^2 + \frac{1}{2} \cdot x^3$, tenemos que p_p es una solución particular del sistema $T(p) = q$. Entonces, usando que $\mathrm{Nu}(T) = gen\{1 - x\}$, tenemos que todas las soluciones de la ecuación $T(p) = q$ son

$$p(x) = 1 + \frac{x^2}{2} + \frac{x^3}{2} + a(1 - x) \text{ con } a \in \mathbb{R}.$$

■

Para resolver el siguiente ejercicio vamos a usar algunas de las propiedades que probamos en el **Ejercicio 2.C**.

Ejercicio 2.7. Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{R}$-espacios vectoriales, $T : \mathbb{V} \to \mathbb{W}$ una TL y $\mathcal{V} = \{v_1, \cdots, v_n\} \in \mathbb{V}$ un conjunto de n puntos de $\mathbb{V}$. Sea

$$C(\mathcal{V}) := \{\sum_{i=1}^{n} p_i v_i \ : p_1, \cdots, p_n \in \mathbb{R}^+ \text{ y } \sum_{i=1}^{n} p_i = 1\}$$

la *cápsula convexa* de $\mathcal{V}$. Comprobar que la imagen de la cápsula convexa de $\mathcal{V}$ por T es la cápsula convexa de la imagen de $\mathcal{V}$ por T. Es decir,

$$T(C(\mathcal{V})) = C(T(\mathcal{V})).$$

Dem. Tenemos que probar una igualdad de conjuntos así que probaremos la doble inclusión.

Sea $w \in T(C(\mathcal{V}))$, entonces existe $v \in C(\mathcal{V})$ tal que $w = T(v)$. Como $v \in C(\mathcal{V})$, existen $p_1, \cdots, p_n \in \mathbb{R}^+$ y $\sum_{i=1}^n p_i = 1$ tales que $v = p_1 v_1 + \cdots + p_n v_n$. Entonces, como T es TL,

$$w = T(v) = T(p_1 v_1 + \cdots + p_n v_n) = p_1 T(v_2) + \cdots + p_n T(v_n).$$

Por lo tanto $w \in C(T(\mathcal{V})) = \{\sum_{i=1}^n p_i T(v_i) \ : p_1, \cdots, p_n \in \mathbb{R}^+ \text{ y } \sum_{i=1}^n p_i = 1\}$ y tenemos que $T(C(\mathcal{V})) \subseteq C(T(\mathcal{V}))$.

Recíprocamente, si $w \in C(T(\mathcal{V}))$, existen $p_1, \cdots, p_n \in \mathbb{R}^+$ y $\sum_{i=1}^n p_i = 1$ tales que

$$w = p_1 T(v_1) + \cdots + p_n T(v_n).$$

A continuación, tenemos que probar que $w \in T(C(\mathcal{V}))$, es decir debemos proponer $v \in C(\mathcal{V})$ tal que $w = T(v)$. Proponemos

$$v := p_1 v_1 + \cdots + p_n v_n,$$

claramente $v \in C(\mathcal{V})$ y por otra parte, como T es TL, tenemos que

$$T(v) = T(p_1 v_1 + \cdots + p_n v_n) = p_1 T(v_1) + \cdots + p_n T(v_n) = w$$

y entonces $w \in T(C(\mathcal{V}))$. Por lo tanto $C(T(\mathcal{V})) \subseteq T(C(\mathcal{V}))$. Como probamos la doble inclusión, concluimos que $T(C(\mathcal{V})) = C(T(\mathcal{V}))$ que era lo que queríamos probar. ■

Ejercicio 2.E (De Parcial)**.** Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{R}$-espacios vectoriales, $B = \{w_1, w_2, w_3\}$ una base de $\mathbb{W}$. Dadas las funcionales lineales $f_i : \mathbb{V} \to \mathbb{R}$ con $i = 1, 2, 3$, se define $T : \mathbb{V} \to \mathbb{W}$ como

$$T(v) = f_1(v)w_1 + f_2(v)w_2 + f_3(v)w_3.$$

Demostrar que:

a) T es una TL.

b) $\text{Nu}(T) = \text{Nu}(f_1) \cap \text{Nu}(f_2) \cap \text{Nu}(f_3)$.

Dem. $a)$: Basta con probar que la función T cumple con $i)$ y $ii)$ de la definición de TL.

i) Sean $v, v' \in \mathbb{V}$, entonces $T(v + v') = f_1(v + v')w_1 + f_2(v + v')w_2 + f_3(v + v')$, usando que f_i son funcionales lineales para $i = 1, 2, 3$, tenemos entonces que

$$T(v + v') = (f_1(v) + f_1(v'))w_1 + (f_2(v) + f_2(v'))w_1 + (f_3(v) + f_3(v'))w_1,$$

como $f_i(v), f_i(v') \in \mathbb{R}$ (el cuerpo de $\mathbb{W}$) para $i = 1, 2, 3$, podemos aplicar la propiedad distributiva (recordar que era uno de los axiomas de los espacios vectoriales). Entonces

$$T(v + v') = f_1(v)w_1 + f_1(v')w_1 + f_2(v)w_2 + f_2(v')w_2 + f_3(v)w_3 + f_3(v')w_3.$$

Por último, asociando, nos queda

$$T(v + v') = f_1(v)w_1 + f_2(v)w_2 + f_3(v)w_3 + f_1(v')w_1 + f_2(v')w_2 + f_3(v')w_3 = T(v) + T(v').$$

ii) Sea $v \in \mathbb{V}$ y $\alpha \in \mathbb{R}$. Usando que f_i son funcionales lineales para $i = 1, 2, 3$, tenemos que

$$T(\alpha v) = f_1(\alpha v)w_1 + f_2(\alpha v)w_2 + f_3(\alpha v) = \alpha f_1(v)w_1 + \alpha f_2(v)w_2 + \alpha f_3(v)w_3 = \alpha T(v).$$

$b)$: Tenemos que probar una igualdad de conjuntos. Vamos a probar la doble inclusión.

Si $v \in \text{Nu}(f_1) \cap \text{Nu}(f_2) \cap \text{Nu}(f_3)$ entonces $v \in \text{Nu}(f_1)$, $v \in \text{Nu}(f_2)$ y $v \in \text{Nu}(f_3)$. Es decir $f_1(v) = f_2(v) = f_3(v) = 0$. Entonces

$$T(v) = f_1(v)w_1 + f_2(v)w_2 + f_3(v)w_3 = 0w_1 + 0w_2 + 0w_3 = 0_{\mathbb{W}} + 0_{\mathbb{W}} + 0_{\mathbb{W}} = 0_{\mathbb{W}},$$

entonces $v \in \text{Nu}(T)$ y tenemos que $\text{Nu}(f_1) \cap \text{Nu}(f_2) \cap \text{Nu}(f_3) \subseteq \text{Nu}(T)$.

Recíprocamente, si $v \in \text{Nu}(T)$ entonces,

$$0_{\mathbb{W}} = T(v) = f_1(v)w_1 + f_2(v)w_2 + f_3(v)w_3.$$

Observar que $f_1(v), f_2(v), f_3(v) \in \mathbb{R}$ (el cuerpo de $\mathbb{W}$), entonces la ecuación anterior es una combinación lineal de los vectores w_1, w_2, w_3 igualada al elemento neutro de $\mathbb{W}$. Como $\{w_1, w_2, w_3\}$ es una base de $\mathbb{W}$, los escalares $f_1(v), f_2(v), f_3(v)$ son todos nulos. Es decir $f_1(v) = f_2(v) = f_3(v) = 0$, entonces $v \in \text{Nu}(f_1) \cap \text{Nu}(f_2) \cap \text{Nu}(f_3)$ y tenemos la otra inclusión, $\text{Nu}(T) \subseteq \text{Nu}(f_1) \cap \text{Nu}(f_2) \cap \text{Nu}(f_3)$. Como probamos la doble inclusión, concluimos que $\text{Nu}(T) = \text{Nu}(f_1) \cap \text{Nu}(f_2) \cap \text{Nu}(f_3)$. ∎

2.2. Buena definición de transformaciones lineales

Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales, $\mathbb{V}$ de dimensión finita. La siguiente proposición prueba que si definimos una transformación lineal T sobre una base de $\mathbb{V}$ entonces T está bien definida y además es única.

Proposición 2.2.1. *Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales, $\mathbb{V}$ de dimensión finita. Sea $B = \{v_1, v_2, \cdots, v_n\}$ una base de $\mathbb{V}$ y sean $w_1, w_2, \cdots, w_n \in \mathbb{W}$ vectores arbitrarios. Entonces existe una única transformación lineal $T : \mathbb{V} \to \mathbb{W}$ tal que*

$$T(v_i) = w_i, \text{ para cada } i \in \{1, 2, \cdots, n\}.$$

Dem. **Buena definición.** Una transformación lineal está bien definida si: a cada elemento v del dominio de T (que en este caso es $\mathbb{V}$) le corresponde un único elemento de $\mathbb{W}$. A dicho elemento se lo denota $T(v)$.

Sea $v \in \mathbb{V}$ un elemento del dominio de T. Entonces, como B es una base de $\mathbb{V}$, existen **únicos** $a_1, a_2, \cdots, a_n \in \mathbb{K}$ tales que $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. Definimos

$$T(v) := a_1w_1 + \cdots + a_nw_n \in \mathbb{W}.$$

Por lo tanto T está bien definida, ya que a cada elemento del dominio de T le corresponde un único elemento de $\mathbb{W}$.

Veamos que T es una transformación lineal. Sean $v, v' \in \mathbb{V}$, como B es una base de $\mathbb{V}$, existen **únicos** $a_1, a_2, \cdots, a_n \in \mathbb{K}$ tales que $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$ y existen **únicos** $b_1, b_2, \cdots, b_n \in \mathbb{K}$ tales que $v' = b_1v_1 + b_2v_2 + \cdots + b_nv_n$. Entonces,

$$v + v' = a_1v_1 + a_2v_2 + \cdots + a_nv_n + b_1v_1 + b_2v_2 + \cdots + b_nv_n = (a_1 + b_1)v_1 + \cdots + (a_n + b_n)v_n.$$

Por lo tanto, por definición

$$T(v+v') = (a_1+b_1)w_1+\cdots+(a_n+b_n)w_n = a_1w_1+\cdots+a_nw_n+b_1w_1+\cdots+b_nw_n = T(v)+T(v').$$

De manera similar se prueba que si $\alpha \in \mathbb{K}$ y $v \in \mathbb{V}$ entonces $T(\alpha v) = \alpha T(v)$.

Unicidad. Supongamos que existe otra transformación lineal, digamos $S : \mathbb{V} \to \mathbb{W}$, tal que

$$S(v_i) = w_i, \text{ para cada } i \in \{1, 2, \cdots, n\}.$$

Sea $v \in \mathbb{V}$ (el dominio de S y de T), como B es una base de $\mathbb{V}$, existen únicos $a_1, a_2, \cdots, a_n \in \mathbb{K}$ tales que $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. Entonces

$$S(v) = S(a_1v_1 + \cdots + a_nv_n) = a_1S(v_1) + \cdots + a_nS(v_n) = a_1w_1 + \cdots + a_nw_n = T(v).$$

Como la igualdad anterior vale para todo $v \in \mathbb{V}$, concluimos que $S = T$ como queríamos ver. ■

El siguiente ejercicio que se relaciona con el **Ejercicio 2.9**, permite diferenciar entre los conceptos de existencia, unicidad y buena definición de una transformación lineal. Estos tres conceptos suelen generar confusión por lo que se recomienda meditar un poco sobre qué diferencia hay entre ellos.

Ejercicio 2.F. Consideremos:

1. $T_1 : \mathbb{R}_2[x] \to \mathbb{R}^3$, definida por
$$T_1(1) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \; T_1(x) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \; T_2(x^2) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

2. $T_2 : \mathbb{R}_2[x] \to \mathbb{R}^3$, definida por
$$T_2(1) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \; T_2(x) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \; T_2(x+1) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

3. $T_3 : \mathbb{R}_2[x] \to \mathbb{R}^3$, definida por
$$T_3(1) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \; T_3(x) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \; T_3(x+1) = \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix}.$$

a) Indicar si las transformaciones lineales de los items 1., 2. y 3. están bien definidas.

b) Indicar si existe una transformación lineal que satisfaga la condiciones de los items 1., 2. ó 3. y de existir dicha transformación lineal, estudiar si es única.

Dem.

1. Como $\{1, x, x^2\}$ es una base de $\mathbb{R}_2[x]$, por la Proposición 2.2.1, la transformación lineal T_1 está bien definida. Por la misma proposición, T_1 es la única transformación lineal que satisface las condiciones del item 1.

2. Como el conjunto $\{1, x, 1 + x\}$ NO es una base de $\mathbb{R}_2[x]$, la transformación lineal T_2 NO está bien definida. Observar que el vector x^2 pertenece al dominio de T_2, sin embargo, por como se definió T_2 no hay ningún elemento de $\mathbb{R}^3$ asignado a x^2, por lo tanto T_2 NO está bien definida.

   Observar por otra parte que,

$$T_2(1) + T_2(x) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = T_2(1 + x).$$

   Por lo tanto, NO puede existir ninguna transformación lineal que cumpla con las condiciones del item 2.

3. Como el conjunto $\{1, x, 1 + x\}$ NO es una base de $\mathbb{R}_2[x]$, la transformación lineal T_3 NO está bien definida. Hay elementos del dominio de T_3 a los cuáles no se le asignó ningún elemento de $\mathbb{R}^3$.

   En este caso, como

$$T_3(1) + T_3(x) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix} = T_3(1 + x),$$

   existen (infinitas) transformaciones lineales que cumplen las condiciones del item 3. Por ejemplo, consideremos $T_v : \mathbb{R}_3[x] \to \mathbb{R}^3$, definida por

$$T_v(1) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \ T_v(x) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ T_v(x^2) = v,$$

   donde v es cualquier vector de $\mathbb{R}^3$. Entonces, para cada v de $\mathbb{R}^3$ que elijamos, la transformación lineal T_v cumple con las condiciones del item 3., está bien definida y es única (porque se definió sobre una base). Pero claramente, cambiando v, tendremos infinitas transformaciones lineales T_v que cumplan con las condiciones del item 3.

■

El siguiente ejercicio es muy similar al **Ejercicio 2.10**.

Ejercicio 2.G. Sea $B = \{x^2 + x, x^2 - x, 1\}$ una base de $\mathbb{R}_2[x]$ y sea $T \in \mathcal{L}(\mathbb{R}_2[x], \mathbb{R}^3)$ una TL que actúa sobre la base B de la siguiente manera

$$T(x^2 + x) = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}, \ T(x^2 - x) = \begin{bmatrix} -6 \\ 9 \\ -12 \end{bmatrix}, \ T(1) = \begin{bmatrix} 4 \\ -6 \\ 8 \end{bmatrix}.$$

a) Hallar $T(2x^2+3x+5)$.

b) Hallar bases de $\text{Nu}(T)$ e $\text{Im}(T)$.

c) Hallar, (si existen) todos los $p \in \mathbb{R}_2[x]$ tales que $T(p) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Dem. a): Para hallar $T(2x^2+3x+5)$, primero escribimos al vector $2x^2+3x+5$ como CL de los elementos de la base B. Entonces, si

$$2x^2+3x+5 = a(x^2+x)+b(x^2-x)+c = x^2(a+b)+x(a-b)+1\cdot c, \text{ para todo } x \in \mathbb{R}.$$

Tenemos que $2 = a+b,\ 3 = a-b,\ 5 = c$. Entonces $a = \frac{5}{2}$, $b = -\frac{1}{2}$ y $c = 5$. Entonces,

$$T(2x^2+3x+5) = T(\frac{5}{2}(x^2+x)-\frac{1}{2}(x^2-x)+5) = \frac{5}{2}T(x^2+x)-\frac{1}{2}T(x^2-x)+5T(1) =$$

$$= \frac{5}{2}\begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -6 \\ 9 \\ -12 \end{bmatrix} + 5\begin{bmatrix} 4 \\ -6 \\ 8 \end{bmatrix} = \begin{bmatrix} 28 \\ -42 \\ 56 \end{bmatrix}.$$

b): Recordemos que $\text{Nu}(T) = \{p \in \mathbb{R}_2[x] : T(p) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\}$. Entonces, si $p \in \text{Nu}(T)$, por un lado $p \in \mathbb{R}_2[x]$ y, como B es una base de $\mathbb{R}_2[x]$, $p(x) = a(x^2+x)+b(x^2-x)+c$, con $a,b,c \in \mathbb{R}$ y, por el otro lado,

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = T(p) = T(a(x^2+x)+b(x^2-x)+c) = aT(x^2+x)+bT(x^2-x)+cT(1) =$$

$$= a\begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} + b\begin{bmatrix} -6 \\ 9 \\ -12 \end{bmatrix} + c\begin{bmatrix} 4 \\ -6 \\ 8 \end{bmatrix} = \begin{bmatrix} 2 & -6 & 4 \\ -3 & 9 & -6 \\ 4 & -12 & 8 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Entonces, basta con resolver el sistema homogéneo de arriba. Operando, nos queda que $a = 3b-2c$. Entonces, volviendo a la expresión de p, nos queda

$$p(x) = a(x^2+x)+b(x^2-x)+c = (3b-2c)(x^2+x)+b(x^2-x)+c = b(4x^2+2x)+c(-2x^2-2x+1),$$

con $b,c \in \mathbb{R}$. Es decir $\text{Nu}(T) = gen\{4x^2+2x, -2x^2-2x+1\}$ y una base de $\text{Nu}(T)$ puede ser $B_{\text{Nu}(T)} = \{4x^2+2x, -2x^2-2x+1\}$.

Recordemos que como $\{x^2+x, x^2-x, 1\}$ es una base de $\mathbb{R}_2[x]$, entonces

$$\text{Im}(T) = gen\{T(x^2+x), T(x^2-x), T(1)\} = gen\{\begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} -6 \\ 9 \\ -12 \end{bmatrix}, \begin{bmatrix} 4 \\ -6 \\ 8 \end{bmatrix}\}.$$

Basta con extraer una base de ese conjunto. Por el Teorema de la dimensión, 2.1.1, sabemos que

$$\dim(T) = \dim(\mathbb{R}_2[x]) - \dim(\text{Nu}(T)) = 3-2 = 1.$$

Entonces, podemos obtener una base de $\text{Im}(T)$ con cualquier generador de $\text{Im}(T)$, por ejemplo $B_{\text{Im}(T)} = \{\begin{bmatrix} -6 \\ 9 \\ -12 \end{bmatrix}\}$.

c): Recordar que existe $p \in \mathbb{R}_2[x]$ tal que $T(p) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, si y sólo si $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \in \mathrm{Im}(T)$ y eso vale por definición de $\mathrm{Im}(T) = \{v \in \mathbb{R}^3 : \exists p \in \mathbb{R}_2[x] : v = T(p)\}$.

Como $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \notin \mathrm{Im}(T) = gen\{ \begin{bmatrix} -6 \\ 9 \\ -12 \end{bmatrix} \}$ concluimos que no existe p tal que $T(p) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. ■

En el siguiente ejercicio que se relaciona con los **Ejercicio 2.11 y 2.12**, aplicaremos varios de los conceptos que vimos hasta acá.

Ejercicio 2.H (De Parcial)**.** Sea $T : \mathbb{R}_2[x] \to \mathbb{R}^2$ la TL definida por

$$T(p) = \begin{bmatrix} p(0) - p(1) \\ p(0) + p(-1) \end{bmatrix}.$$

a) Hallar bases de $\mathrm{Nu}(T)$ e $\mathrm{Im}(T)$.

b) Sean $\mathcal{S} := \{p \in \mathbb{R}_2[x] : p(1) = p'(1) = 0\}$ y $\mathcal{U} := gen\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \}$. Hallar $T(\mathcal{S})$ y $T^{-1}(\mathcal{U})$.

Dem. a): Recordemos que $\mathrm{Nu}(T) = \{p \in \mathbb{R}_2[x] : T(p) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\}$. Entonces $p \in \mathrm{Nu}(T)$ si $p(x) = ax^2 + bx + c$, con $a, b, c \in \mathbb{R}$ y $T(p) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Observar que $p(0) = c,\ p(1) = a + b + c$ y $p(-1) = a - b + c$. Entonces

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = T(p) = \begin{bmatrix} -a - b \\ a - b + 2c \end{bmatrix}.$$

Entonces, $-a-b = 0$ y $a-b+2c = 0$. Entonces $a = -b$ y $c = b$. Por lo tanto, $p(x) = -bx^2+bx+b = b(-x^2 + x + 1)$, con $b \in \mathbb{R}$. Entonces, $\mathrm{Nu}(T) = gen\{-x^2 + x + 1\}$ y una base de $\mathrm{Nu}(T)$ puede ser $B_{\mathrm{Nu}(T)} = gen\{-x^2 + x + 1\}$.

Por otra parte, por el **Ejercicio 2.C**, como $\{1, x, x^2\}$ es una base de $\mathbb{R}_2[x]$ (en particular es un sistema de generadores de $\mathbb{R}_2[x]$) se sigue que

$$\mathrm{Im}(T) = gen\{T(1), T(x), T(x^2)\} = gen\{ \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \} = \mathbb{R}^2.$$

Otra manera de resolver este item es usando el Teorema de la dimensión 2.1.1: como $\dim(\mathrm{Im}(T)) = \dim(\mathbb{R}_2[x]) - \dim(\mathrm{Nu}(T)) = 3 - 1 = 2$ e $\mathrm{Im}(T) \subseteq \mathbb{R}^2$, tenemos que $\mathrm{Im}(T) = \mathbb{R}^2$.

b): Busquemos un sistema de generadores de $\mathcal{S}$. Si $p \in \mathcal{S}$ entonces $p(x) = ax^2 + bx + c$, con $a, b, c \in \mathbb{R}$ y como $p(1) = a + b + c$ y $p'(1) = 2a + b$, tenemos que $a + b + c = 2a + b = 0$. Entonces $b = -2a$ y $c = -a - b = -a + 2a = a$. Entonces $p(x) = ax^2 - 2ax + a = a(x^2 - 2x + 1)$, con $a \in \mathbb{R}$. Entonces $\mathcal{S} = gen\{x^2 - 2x + 1\}$ y, por el **Ejercicio 2.C**,

$$T(\mathcal{S}) = gen\{T(x^2 - 2x + 1)\} = gen\{ \begin{bmatrix} 1 \\ 5 \end{bmatrix} \}.$$

Finalmente, recordar que $T^{-1}(\mathcal{U}) = \{p \in \mathbb{R}_2[x] : T(p) \in \mathcal{U}\}$. Entonces $p \in T^{-1}(\mathcal{U})$ si $p(x) = ax^2+bx+c$, con $a, b, c \in \mathbb{R}$ y $T(p) \in \mathcal{U}$. Como $p(0) = c,\ p(1) = a+b+c$ y $p(-1) = a-b+c$, entonces

$$T(p) = \begin{bmatrix} -a - b \\ a - b + 2c \end{bmatrix} \in \mathcal{U} = gen\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \}.$$

Entonces, $\begin{bmatrix} -a-b \\ a-b+2c \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ con $\lambda \in \mathbb{R}$. Entonces $-a-b=\lambda$ y $a-b+2c=\lambda$. Entonces $a=-b-\lambda$ y $2c=\lambda-a+b=\lambda+b+\lambda+b=2b+2\lambda$, entonces $c=b+\lambda$. Por lo tanto,

$$p(x) = (-b-\lambda)x^2 + bx + b + \lambda = b(-x^2+x+1) + \lambda(-x^2+1),$$

con $b, \lambda \in \mathbb{R}$. Entonces $T^{-1}(\mathcal{U}) = gen\{-x^2+x+1, -x^2+1\}$. ■

Observar que en el ejercicio anterior $\mathrm{Nu}(T) \subseteq T^{-1}(\mathcal{U})$, eso no es casualidad. En general, vale la siguiente propiedad:

> Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales y $T : \mathbb{V} \to \mathbb{W}$ una TL. Si $\mathcal{U} \subseteq \mathbb{W}$ es un subespacio entonces
> $$\mathrm{Nu}(T) \subseteq T^{-1}(\mathcal{U}).$$

De hecho, si $v \in \mathrm{Nu}(T)$, entonces $T(v) = 0_{\mathbb{W}}$, como $\mathcal{U}$ es un subespacio, $T(v) = 0_{\mathbb{W}} \in \mathcal{U}$, entonces $v \in T^{-1}(\mathcal{U})$ y tenemos la inclusión que queríamos probar.

2.3. Transformaciones lineales inversibles

Sean $\mathbb{V}, \mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales y $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Diremos que:

- T es *monomorfismo* si T es inyectiva, es decir, si $T(v_1) = T(v_2)$ entonces $v_1 = v_2$.
- T es *epimorfismo* si T es sobreyectiva, es decir, si $\mathrm{Im}(T) = \mathbb{W}$.
- T es *isomorfismo* si T es biyectiva, es decir, si es monomorfismo y epimorfismo.

Por otra parte, para lo que viene a continuación, vamos a definir la composición de transformaciones lineales.

Sean $\mathbb{V}, \mathbb{W}, \mathbb{Z}$ tres $\mathbb{K}$-espacios vectoriales, $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ y $S \in \mathcal{L}(\mathbb{W}, \mathbb{Z})$. La composición de S con T es la función $S \circ T : \mathbb{V} \to \mathbb{Z}$ definida por

$$S \circ T(v) := S(T(v)).$$

Se puede probar (se deja como ejercicio) que

> si $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ y $S \in \mathcal{L}(\mathbb{W}, \mathbb{Z})$ entonces $S \circ T \in \mathcal{L}(\mathbb{V}, \mathbb{Z})$, es decir, la composición de transformaciones lineales es una transformación lineal.

Proposición 2.3.1. *Sean $\mathbb{V}, \mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales y $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Entonces:*

1. *T es monomorfismo si y sólo si $\mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$.*
2. *Si T es isomorfismo, entonces $\dim(\mathbb{V}) = \dim(\mathbb{W})$. En este caso, existe T^{-1} y $T^{-1} \in \mathcal{L}(\mathbb{W}, \mathbb{V})$ (es decir la inversa de una transformación lineal, también es una transformación lineal).*

Dem. 1.: Vamos a probar las dos implicaciones. En primer lugar, supongamos que T es monomorfismo y veamos que eso implica que $\mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$.

Si T es monomorfismo, T es inyectiva y entonces, $T(v_1) = T(v_2)$ implica que $v_1 = v_2$. Supongamos que $v \in \mathrm{Nu}(T)$, entonces $T(v) = 0_{\mathbb{W}}$, por otra parte, como T es TL, siempre vale que $T(0_{\mathbb{V}}) = 0_{\mathbb{W}} = T(v)$, entonces, por hipótesis, $v = 0_{\mathbb{V}}$ y $\mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$.

Recíprocamente, supongamos que $\mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$ y veamos que eso implica que T es monomorfismo. Supongamos que $T(v_1) = T(v_2)$ para ciertos $v_1, v_2 \in \mathbb{V}$. Entonces, como T es TL, tenemos que $T(v_1 - v_2) = T(v_1) - T(v_2) = 0_{\mathbb{W}}$, entonces $v_1 - v_2 \in \mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$ por hipótesis. Por lo tanto $v_1 = v_2$ y T es monomorfismo.

2. Si T es isomorfismo, entonces T es monomorfismo y $\mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$ (por el item 1.) y además T es epimorfismo, es decir $\mathrm{Im}(T) = \mathbb{W}$. Entonces, por el Teorema de la dimensión 2.1.1, tenemos que

$$\dim(\mathbb{V}) = \dim(\mathrm{Nu}(T)) + \dim(\mathrm{Im}(T)) = \dim(\{0_{\mathbb{V}}\}) + \dim(\mathbb{W}) = \dim(\mathbb{W}).$$

En este caso, si T es isomorfismo entonces T es biyectiva y existe T^{-1}.

De hecho, como $\mathrm{Im}(T) = \mathbb{W}$ y $\mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$, dado $w \in \mathbb{W}$ existe un único $v \in \mathbb{V}$ tal que $T(v) = w$. Vamos a definir $T^{-1} : \mathbb{W} \to \mathbb{V}$ como

$$T^{-1}(w) := v.$$

La función T^{-1} está bien definida porque: T^{-1} está definida en todo elemento del dominio $\mathbb{W}$ (justamente porque $\mathrm{Im}(T) = \mathbb{W}$) y además a cada elemento del dominio le corresponde sólo un elemento de $\mathbb{V}$ (justamente porque $\mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$). Entonces, por como definimos T^{-1}, observar que

$$T \circ T^{-1}(w) = T(T^{-1}(w)) = T(v) = w = I_{\mathbb{W}}(w),$$

como esa igualdad vale para todo $w \in \mathbb{W}$ tenemos que $T \circ T^{-1} = I_{\mathbb{W}}$. Además,

$$T^{-1} \circ T(v) = T^{-1}(T(v)) = T^{-1}(w) = v = I_{\mathbb{V}}(v),$$

como esa igualdad vale para todo $v \in \mathbb{V}$ tenemos que $T^{-1} \circ T = I_{\mathbb{V}}$.

Sólo nos resta probar que T^{-1} es una TL.

Sean $w, w' \in \mathbb{W}$ entonces existen únicos $v, v' \in \mathbb{V}$ tales que $T(v) = w$ y $T(v') = w'$ y entonces $T^{-1}(w) = v$ y $T^{-1}(w') = v'$. Como T es una TL, tenemos que $T(v + v') = T(v) + T(v') = w + w'$ entonces, por definición de T^{-1}, tenemos que $T^{-1}(w + w') = v + v' = T^{-1}(w) + T^{-1}(w')$.

De manera similar, se prueba que si $\alpha \in \mathbb{K}$ y $w \in \mathbb{W}$ entonces $T^{-1}(\alpha w) = \alpha T^{-1}(w)$ (hacerlo como ejercicio) y concluimos que $T^{-1} \in \mathcal{L}(\mathbb{W}, \mathbb{V})$. ■

Ejercicio 2.I. Sea $\lambda \in \mathbb{C} \setminus \{0\}$. Comprobar que la aplicación $T : \mathbb{C}_n[x] \to \mathbb{C}_n[x]$ definida por

$$T(p) = p' - \lambda p$$

es un isomorfismo.

Dem. Tomemos a $\mathbb{C}_n[x]$ como $\mathbb{C}$-espacio vectorial (se deja como tarea pensar cómo se resuelve el ejercicio si tomamos como cuerpo $\mathbb{R}$) y probemos que T es monomorfismo, es decir veamos que $\mathrm{Nu}(T) = \{\mathbf{0}\}$. Sea $p \in \mathrm{Nu}(T)$, entonces, por un lado $p \in \mathbb{C}_n[x]$ es decir

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

con $a_n, a_{n-1}, \cdots, a_1, a_0 \in \mathbb{C}$ y, por otro lado,

$$\mathbf{0} = T(p) = p' - \lambda p.$$

Es decir, para cada $x \in \mathbb{C}$, tenemos que

$$\begin{aligned} 0 = \mathbf{0}(x) &= T(p)(x) = p'(x) - \lambda p(x) \\ &= a_n n x^{n-1} + a_{n-1}(n-1)x^{n-2} + \cdots + a_2 2x + a_1 - \lambda(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0) \\ &= x^n(-\lambda a_n) + x^{n-1}(a_n n - \lambda a_{n-1}) + \cdots + x(a_2 2 - \lambda a_1) + 1(a_1 - \lambda a_0). \end{aligned}$$

Es decir, tenemos que

$$0 = x^n(-\lambda a_n) + x^{n-1}(a_n n - \lambda a_{n-1}) + \cdots + x(a_2 2 - \lambda a_1) + 1(a_1 - \lambda a_0), \text{ para todo } x \in \mathbb{C}.$$

Entonces, usando por ejemplo que $\{x^n, x^{n-1}, \cdots, x, 1\}$ es una base de $\mathbb{C}_n[x]$, (y arriba nos quedó una CL de los elementos de la base igualada al elemento neutro de $\mathbb{C}_n[x]$), tenemos que

$$-\lambda a_n = a_n n - \lambda a_{n-1} = \cdots = a_2 2 - \lambda a_1 = a_1 - \lambda a_0 = 0.$$

Entonces, como $-\lambda a_n = 0$ y $\lambda \neq 0$, tenemos que $a_n = 0$. De la misma manera, como $\lambda a_{n-1} = a_n n = 0.n = 0$ y $\lambda \neq 0$, tenemos que $a_{n-1} = 0$. Si seguimos así, en el paso $n-1$ tendremos que $a_2 = 0$, entonces como $\lambda a_1 = a_2 2 = 0$ y $\lambda \neq 0$, tenemos que $a_1 = 0$. Finalmente, como $\lambda a_0 = a_1 = 0$ y $\lambda \neq 0$, tenemos que $a_0 = 0$. Entonces, probamos que $a_n = a_{n-1} = \cdots = a_2 = a_1 = a_0 = 0$ y volviendo a la expresión de p, nos queda que $p(x) = 0$, para todo $x \in \mathbb{C}$, es decir $p = \mathbf{0}$. Por lo tanto $\mathrm{Nu}(T) = \{\mathbf{0}\}$ y T es monomorfismo. Finalmente, por el teorema de la dimensión,

$$\dim(\mathrm{Im}(T)) = \dim(\mathbb{C}_n[x]) - \dim(\mathrm{Nu}(T)) = \dim(\mathbb{C}_n[x]) - 0 = \dim(\mathbb{C}_n[x])$$

y como siempre vale que $\mathrm{Im}(T) \subseteq \mathbb{C}_n[x]$, tenemos que $\mathrm{Im}(T) = \mathbb{C}_n[x]$ y T es epimorfismo. Como T es monomorfismo y epimorfismo, concluimos que T es isomorfismo. ■

El siguiente ejercicio es un ejemplo muy importante de isomorfismo, el llamado *isomorfismo de coordenadas.*

Ejercicio 2.J. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial, y sea $\{v_1, \cdots, v_n\}$ un conjunto de vectores de $\mathbb{V}$. Sea $\Lambda : \mathbb{K}^n \to \mathbb{V}$ la aplicación en $\mathbb{V}$ definida por

$$\Lambda([x_1 \ \cdots \ x_n]^T) := \sum_{j=1}^{n} x_j v_j.$$

a) Probar que $\Lambda \in \mathcal{L}(\mathbb{K}^n, \mathbb{V})$ y describir la imagen de Λ.

b) Probar que Λ es isomorfimo si, y sólo si, $\{v_1, v_2, \cdots, v_n\}$ es una base de $\mathbb{V}$.

c) Mostrar que si $B = \{v_1, v_2, \cdots, v_n\}$ es una base de $\mathbb{V}$ entonces

$$\Lambda \circ \Phi = I_{\mathbb{V}} \text{ y } \Phi \circ \Lambda = I_{\mathbb{K}^n}.$$

Donde $\Phi : \mathbb{V} \to \mathbb{K}^n$ es la TL tal que $\Phi(v) = [v]^B$.

Dem. *a)* : Basta con probar que la aplicación Λ cumple con *i)* y *ii)* de la definición de TL.
Sean $x = [x_1 \ \cdots \ x_n]^T, y = [y_1 \ \cdots \ y_n]^T \in \mathbb{K}^n$. Entonces
$\Lambda(x+y) = \Lambda([x_1 \ \cdots \ x_n]^T + [y_1 \ \cdots \ y_n]^T) = \Lambda([x_1 + y_1 \ \cdots \ x_n + y_n]^T) = \sum_{j=1}^n (x_j + y_j)v_j = \sum_{j=1}^n (x_j v_j + y_j v_j) = \sum_{j=1}^n x_j v_j + \sum_{j=1}^n y_j v_j = \Lambda([x_1 \ \cdots \ x_n]^T) + \Lambda([y_1 \ \cdots \ y_n]^T) = \Lambda(x) + \Lambda(y)$.
Por otra parte, sean $x = [x_1 \ \cdots \ x_n]^T$ y $\alpha \in \mathbb{K}$. Entonces $\Lambda(\alpha x) = \Lambda(\alpha[x_1 \ \cdots \ x_n]^T) = \Lambda([\alpha x_1 \ \cdots \ \alpha x_n]^T) = \sum_{j=1}^n (\alpha x_j)v_j = \alpha \sum_{j=1}^n x_j v_j = \alpha\Lambda([x_1 \ \cdots \ x_n]^T) = \alpha\Lambda(x)$.
Recordemos que vimos que si tenemos un sistema de generadores (en particular una base) de $\mathbb{K}^n$ entonces la imagen de Λ está generada por los transformados de dichos generadores de $\mathbb{K}^n$.
Sean e_i los vectores de $\mathbb{K}^n$ cuyas componentes son 0 en todos lados excepto en el lugar i donde vale 1, por ejemplo, $e_1 = [1 \ 0 \ 0 \ \cdots \ 0]^T$. Claramente $\{e_1, e_2, \cdots, e_n\}$ es una base de $\mathbb{K}^n$ (de hecho es la base canónica de $\mathbb{K}^n$). Por lo tanto

$$\mathrm{Im}(\Lambda) = gen\{\Lambda(e_1), \Lambda(e_2), \cdots, \Lambda(e_n)\}.$$

Observar que $\Lambda(e_i) = \sum_{j=1}^n (e_i)_j v_j = v_i$, (porque solo sobrevive el sumando $j = i$ que es cuando la componente de e_i es 1). Entonces

$$\mathrm{Im}(\Lambda) = gen\{\Lambda(e_1), \Lambda(e_2), \cdots, \Lambda(e_n)\} = gen\{v_1, v_2, \cdots, v_n\}.$$

b) : Primero veamos que Λ es monomorfismo si y sólo si el conjunto $\{v_1, v_2, \cdots, v_n\}$ es LI. Como $\Lambda \in \mathcal{L}(\mathbb{K}^n, \mathbb{V})$ entonces Λ es monomorfismo si y sólo si $\mathrm{Nu}(\Lambda) = \{0_{\mathbb{K}^n}\}$. Primero, supongamos que Λ es monomorfismo y probemos que $\{v_1, v_2, \cdots, v_n\}$ es LI: sean $\alpha_1, \cdots, \alpha_n \in \mathbb{K}$ tales que $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0_{\mathbb{V}}$, para probar que $\{v_1, v_2, \cdots, v_n\}$ es LI basta ver que $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

Observar que $[\alpha_1\ \alpha_2\ \cdots\ \alpha_n]^T \in \mathbb{K}^n$ y además

$$\Lambda([\alpha_1\ \alpha_2\ \cdots\ \alpha_n]^T) = \sum_{j=1}^{n} \alpha_j v_j = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0_{\mathbb{V}}.$$

Por lo tanto, $[\alpha_1\ \alpha_2\ \cdots\ \alpha_n]^T \in \mathrm{Nu}(\Lambda) = \{0_{\mathbb{K}^n}\}$ (por hipótesis). Entonces $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, y probamos $\{v_1, v_2, \cdots, v_n\}$ es LI.

Recíprocamente, supongamos que $\{v_1, v_2, \cdots, v_n\}$ es LI veamos que $\mathrm{Nu}(\Lambda) = \{0_{\mathbb{K}^n}\}$: si $x = [x_1\ \cdots\ x_n]^T \in \mathrm{Nu}(\Lambda)$, entonces

$$0_{\mathbb{V}} = \Lambda(x) = \sum_{j=1}^{n} x_j v_j = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n.$$

Tenemos una combinación lineal de los vectores $v_1, v_2, \cdots, v_n$ igualada al elemento neutro de $\mathbb{V}$. Como $\{v_1, v_2, \cdots, v_n\}$ es LI (por hipótesis), entonces $x_1 = x_2 = \cdots = x_n = 0$. Entonces $x = [x_1\ \cdots\ x_n]^T = 0_{\mathbb{K}^n}$. Por lo tanto, $\mathrm{Nu}(\Lambda) = \{0_{\mathbb{K}^n}\}$ y Λ es monomorfismo.

Usando el ítem a), tenemos que $\mathrm{Im}(T) = gen\{v_1, v_2, \cdots, v_n\}$. Además, recordemos que Λ es epimorfismo si y sólo si $\mathrm{Im}(T) = \mathbb{V}$. Por lo tanto, Λ es epimorfismo si y sólo si $\mathbb{V} = \mathrm{Im}(T) = gen\{v_1, v_2, \cdots, v_n\}$.

Como Λ es isomorfismo si y sólo si $\mathrm{Im}(T) = \mathbb{V}$ y $\mathrm{Nu}(T) = \{0_{\mathbb{K}^n}\}$, usando lo que acabamos de probar tenemos que $\mathrm{Im}(T) = \mathbb{V}$ y $\mathrm{Nu}(T) = \{0_{\mathbb{K}^n}\}$ si y sólo si $\{v_1, v_2, \cdots, v_n\}$ es LI y $\mathbb{V} = gen\{v_1, v_2, \cdots, v_n\}$. Pero $\{v_1, v_2, \cdots, v_n\}$ es LI y $\mathbb{V} = gen\{v_1, v_2, \cdots, v_n\}$ si y sólo si $\{v_1, v_2, \cdots, v_n\}$ es una base de $\mathbb{V}$ (por definición). Conclusión, Λ es isomorfimo si y sólo si $\{v_1, v_2, \cdots, v_n\}$ es una base de $\mathbb{V}$.

c) : Ya vimos que si $B = \{v_1, v_2, \cdots, v_n\}$ es una base de $\mathbb{V}$ entonces Λ es un isomorfismo y por lo tanto existe $\Lambda^{-1} \in \mathcal{L}(\mathbb{V}, \mathbb{K}^n)$, lo que nos pide el ejercicio es verificar que $\Phi = \Lambda^{-1}$.

Sea $v \in \mathbb{V}$, entonces como B es una base de $\mathbb{V}$, existen $x_1, x_2, \cdots, x_n \in \mathbb{K}$ tales que

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Entonces $\Phi(v) = [v]^B = [x_1\ \cdots\ x_n]^T$. Por lo tanto

$$\Lambda \circ \Phi(v) = \Lambda(\Phi(v)) = \Lambda([x_1\ \cdots\ x_n]^T) = \sum_{j=1}^{n} x_j v_j = x_1 v_1 + \cdots x_n v_n = v = I_{\mathbb{V}}(v).$$

Como la iguadad anterior vale para todo $v \in \mathbb{V}$, concluimos que $\Lambda \circ \Phi = I_{\mathbb{V}}$.

Por otra parte, sea $x = [x_1\ \cdots\ x_n]^T \in \mathbb{K}^n$. Observar que como $B = \{v_1, v_2, \cdots, v_n\}$ es una base de $\mathbb{V}$, si tomamos el vector $v := x_1 v_1 + \cdots + x_n v_n$, entonces claramente $[v]^B = [x_1\ \cdots\ x_n]^T$. Entonces

$$\Phi \circ \Lambda(x) = \Phi \circ \Lambda([x_1\ \cdots\ x_n]^T) = \Phi(\Lambda([x_1\ \cdots\ x_n]^T)) =$$
$$= \Phi(x_1 v_1 + \cdots + x_n v_n) = [x_1 v_1 + \cdots + x_n v_n]^B = [x_1\ \cdots\ x_n]^T = I_{\mathbb{K}^n}(x).$$

Como la iguadad anterior vale para todo $x \in \mathbb{K}^n$, concluimos que $\Phi \circ \Lambda = I_{\mathbb{K}^n}$. ■

La siguiente propiedad vale cuando los espacios de salida y de llegada de una transformación lineal tienen la misma dimensión:

Proposición 2.3.2. *Sean $\mathbb{V}, \mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales tales que* $\dim(\mathbb{V}) = \dim(\mathbb{W})$ *y sea $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Entonces son equivalentes:*

i) *T es isomorfismo,*

ii) *T es monomorfismo,*

iii) *T es epimorfismo.*

Dem. $i) \Rightarrow ii)$: Si T es isomorfismo entonces (por definición) T es monomorfismo y entonces vale $ii)$.

$ii) \Rightarrow iii)$: Si T es monomorfismo, entonces $\mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$. Por hipótesis $\dim(\mathbb{V}) = \dim(\mathbb{W})$ entonces, por el Teorema de la dimensión 2.1.1, tenemos que

$$\dim(\mathrm{Im}(T)) = \dim(\mathbb{V}) - \dim(\mathrm{Nu}(T)) = \dim(\mathbb{V}) = \dim(\mathbb{W}).$$

Como $\mathrm{Im}(T) \subseteq \mathbb{W}$ y en este caso $\dim(\mathrm{Im}(T)) = \dim(\mathbb{W})$, concluimos que $\mathrm{Im}(T) = \mathbb{W}$, entonces T es epimorfismo y vale $iii)$.

$iii) \Rightarrow i)$: Si T es epimorfismo, entonces $\mathrm{Im}(T) = \mathbb{W}$. Por hipótesis $\dim(\mathbb{V}) = \dim(\mathbb{W})$ entonces, por el Teorema de la dimensión, tenemos que

$$\dim(\mathrm{Nu}(T)) = \dim(\mathbb{V}) - \dim(\mathrm{Im}(T)) = \dim(\mathbb{V}) - \dim(\mathbb{W}) = 0,$$

entonces $\mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$ y T también es monomorfismo. Por lo tanto T es isomorfismo y vale $i)$.

Como probamos las implicaciones $i) \Rightarrow ii)$, $ii) \Rightarrow iii)$ y $iii) \Rightarrow i)$, concluimos que, cuando $\dim(\mathbb{V}) = \dim(\mathbb{W})$, los 3 items son equivalentes. Es decir, T es isomorfismo si y sólo si T es monomorfismo, si y sólo si T es epimorfismo. ∎

El siguiente ejemplo muestra que si $\dim(\mathbb{V}) \neq \dim(\mathbb{W})$, el resultado anterior es falso.

Ejemplo 2.b. : Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ la transformación lineal definida por

$$T(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Observar que

$$\mathrm{Im}(T) = gen\{T(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}), T(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}), T(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix})\} = gen\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}\} = \mathbb{R}^2.$$

Entonces T es epimorfismo. Pero, por el Teorema de la dimensión, $\dim(\mathrm{Nu}(T)) = \dim(\mathbb{R}^3) - \dim(\mathrm{Im}(T)) = 3 - 2 = 1$. Más aún,

$$\mathrm{Nu}(T) = gen\{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}.$$

Entonces T NO es monomorfismo y T NO es isomorfismo.

En el **Ejercicio 2.J** probamos que si $\mathbb{V}$ es un $\mathbb{K}$-espacio vectorial y $B = \{v_1, v_2, \cdots, v_n\}$ es una base de $\mathbb{V}$ entonces, la transformación lineal $\Lambda : \mathbb{K}^n \to \mathbb{V}$ definida por

$$\Lambda([x_1\ x_2 \cdots\ x_n]^T) := x_1 v_1 + x_2 v_2 + \cdots + x_n v_n$$

es un isomorfismo, es más, su inversa (que obviamente también es un isomorfismo), es la transformación lineal $\Phi : \mathbb{V} \to \mathbb{K}^n$, definida por

$$\Phi(v) = [v]^B.$$

Este hecho, nos permite enunciar una propiedad que usaremos durante todo lo que resta de la materia:

Proposición 2.3.3. *Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial, B una base de $\mathbb{V}$ y $\{u_1, u_2, \cdots, u_r\}$ un conjunto de vectores de $\mathbb{V}$. Entonces*

el conjunto $\{u_1, u_2, \cdots, u_r\}$ es LI en $\mathbb{V}$ si y sólo si $\{[u_1]^B, [u_2]^B, \cdots, [u_r]^B\}$ es LI en $\mathbb{K}^n$.

Es decir el conjunto $\{u_1, u_2, \cdots, u_r\}$ es LI en $\mathbb{V}$ si y sólo si el conjunto de los vectores de sus coordenadas en cualquier base de $\mathbb{V}$ son LI en $\mathbb{K}^n$.

Dem. Vamos a usar la misma notación del **Ejercico 2.J**.

Supongamos que $\{u_1, u_2, \cdots, u_r\}$ es LI y veamos que $\{[u_1]^B, [u_2]^B, \cdots, [u_r]^B\}$ es LI en $\mathbb{K}^n$. Entonces, sean $a_1, a_2, \cdots, a_r \in \mathbb{K}$ tales que

$$a_1[u_1]^B + a_2[u_2]^B + \cdots + a_r[u_r]^B = 0_{\mathbb{K}^n},$$

entonces

$$0_{\mathbb{K}^n} = a_1\Phi(u_1) + a_2\Phi(u_2) + \cdots + a_r\Phi(u_r) = \Phi(a_1u_1 + a_2u_2 + \cdots + a_ru_r),$$

donde usamos la definición de Φ y que Φ es una transformación lineal. Entonces,

$$a_1u_1 + a_2u_2 + \cdots + a_ru_r \in \mathrm{Nu}(\Phi) = \{0_{\mathbb{V}}\},$$

donde usamos que Φ es isomorfismo (y por ende monomorfismo). Entonces,

$$0_{\mathbb{V}} = a_1u_1 + a_2u_2 + \cdots + a_ru_r$$

y como $\{u_1, u_2, \cdots, u_r\}$ es LI (por hipótesis) se sigue que $a_1 = a_2 = \cdots = a_r = 0$ y entonces $\{[u_1]^B, [u_2]^B, \cdots, [u_r]^B\}$ es LI en $\mathbb{K}^n$.

La recíproca se prueba de manera similar y la dejamos como ejercicio, sería buena idea tratar de ver si les sale probarlo. ∎

El siguiente ejemplo es una aplicación de este resultado:

Ejemplo 2.c. Demostrar que el conjunto $\{1 + x + x^2, x + x^2, x^2\}$ es LI en $\mathbb{R}_2[x]$.

Consideremos $E' = \{1, x, x^2\}$ la base canónica de $\mathbb{R}_2[x]$, entonces $[1 + x + x^2]^{E'} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $[x+x^2]^{E'} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ y $[x^2]^{E'} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Entonces, como el conjunto $\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}$ es LI en $\mathbb{R}^3$ (por ejemplo, si ponen los vectores como filas de una matriz, la matriz queda automáticamente triangulada sin filas nulas), por la Proposición 2.3.3, concluimos que $\{1 + x + x^2, x + x^2, x^2\}$ es LI en $\mathbb{R}_2[x]$.

2.4. Matriz de una transformación lineal

Definición. Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales de dimensión finita. Sean $B_1 = \{v_1, v_2, \cdots, v_n\}$ una base de $\mathbb{V}$ y $B_2 = \{w_1, w_2, \cdots, w_m\}$ una base de $\mathbb{W}$. Sea $T : \mathbb{V} \to \mathbb{W}$ tal que $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Se llama *matriz de T en las bases B_1 y B_2*, a la matriz $[T]_{B_1}^{B_2} \in \mathbb{K}^{m\times n}$ definida por

$$[T]_{B_1}^{B_2} = [\ [T(v_1)]^{B_2}\ [T(v_2)]^{B_2}\ \cdots\ [T(v_n)]^{B_2}].$$

Es decir $[T]_{B_1}^{B_2}$ es la matriz que en sus columnas tiene los vectores de coordenadas en base B_2 de los transformados de los vectores de la base B_1.

A partir de la definición de matriz de una transformación lineal, se deduce la siguiente propiedad,

$$[T(v)]^{B_2} = [T]_{B_1}^{B_2}\ [v]^{B_1},\ \text{para cada } v \in \mathbb{V}. \tag{2.1}$$

De hecho, si $v \in \mathbb{V}$ entonces existen $a_1, a_2, \cdots, a_n \in \mathbb{K}$ tales que $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. Entonces, como T es TL, tenemos que $T(v) = a_1T(v_1) + a_2T(v_2) + \cdots + a_nT(v_n)$. Entonces, usando que tomar coordenadas es lineal, lo probamos en la Proposición 1.11.1, tenemos que

$$[T(v)]^{B_2} = [a_1T(v_1) + a_2T(v_2) + \cdots + a_nT(v_n)]^{B_2} = a_1[T(v_1)]^{B_2} + a_2[T(v_2)]^{B_2} + \cdots + a_n[T(v_n)]^{B_2} =$$

$$= [T]_{B_1}^{B_2} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_4 \end{bmatrix} = [T]_{B_1}^{B_2}\ [v]^{B_1}.$$

Ejercicio 2.14 a). Sea $A \in \mathbb{K}^{m\times n}$. Definimos $T : \mathbb{K}^n \to \mathbb{K}^m$, como

$$T(x) := Ax.$$

Entonces si E es la base canónica de $\mathbb{K}^n$ y F es la base canónica de $\mathbb{K}^m$. Tenemos que

$$[T]_E^F = A.$$

Dem. Recordemos que la base canónica de $\mathbb{K}^n$ es $E = \{e_1, e_2, \cdots, e_n\}$, donde e_i es el vector de $\mathbb{K}^n$ que tiene 0 en todas sus componentes salvo en el lugar i donde vale 1. Por otra parte, la base canónica de $\mathbb{K}^m$ es $F = \{f_1, f_2, \cdots, f_m\}$, donde f_i es el vector de $\mathbb{K}^m$ que tiene 0 en todas sus componentes salvo en el lugar i donde vale 1.

Recordemos también, que si $w \in \mathbb{K}^m$ entonces $[w]^F = w$, porque F es la base canónica de $\mathbb{K}^m$. Si no se convencen de esto último, hagan la cuenta y les va a salir.

Por último, recordemos que $Ae_i = A_i$, donde A_i denota la columna i-ésima de la matriz A. Es decir, si multiplicamos A por el vector i-ésimo de la base canónica de K^n obtenemos la columna i-ésima de la matriz A.

Entonces, por definición de $[T]_E^F$, tenemos que

$$[T]_E^F = [\ [T(e_1)]^F\ [T(e_2)]^F\ \cdots\ [T(e_n)]^F] =$$

$$= [\ [Ae_1]^F\ [Ae_2]^F\ \cdots\ [Ae_n]^F] = [\ Ae_1\ Ae_2\ \cdots\ Ae_n] = [A_1\ A_2\ \cdots\ A_n] = A$$

y probamos lo que queríamos. ∎

Ejercicio 2.14 c). Sea $T : \mathbb{R}_3[x] \to \mathbb{R}^4$, la transformación lineal definida por

$$T(p) := [p(0)\ p(1)\ p(10)\ p(100)]^T.$$

Hallar $[T]_{E'}^E$, donde E' es la base canónica de $\mathbb{R}_3[x]$ y E la base canónica de $\mathbb{R}^4$.

Dem. Recordar que $E' = \{1, x, x^2, x^3\}$ y $E = \{[1\ 0\ 0\ 0]^T, [0\ 1\ 0\ 0]^T, [0\ 0\ 1\ 0]^T, [0\ 0\ 0\ 1]^T\}$. Además, $T(1) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $T(x) = \begin{bmatrix} 0 \\ 1 \\ 10 \\ 100 \end{bmatrix}$, $T(x^2) = \begin{bmatrix} 0 \\ 1 \\ 10^2 \\ 10^4 \end{bmatrix}$ y $T(x^3) = \begin{bmatrix} 0 \\ 1 \\ 10^3 \\ 10^6 \end{bmatrix}$. Entonces, por definición de $[T]_{E'}^E$, tenemos que

$$[T]_{E'}^E = [\ [T(1)]^E\ [T(x)]^E\ [T(x^2)]^E\ [T(x^3)]^E] =$$

$$= [\ [\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}]^E\ [\begin{bmatrix} 0 \\ 1 \\ 10 \\ 100 \end{bmatrix}]^E\ [\begin{bmatrix} 0 \\ 1 \\ 10^2 \\ 10^4 \end{bmatrix}]^E\ [\begin{bmatrix} 0 \\ 1 \\ 10^3 \\ 10^6 \end{bmatrix}]^E] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 10 & 10^2 & 10^3 \\ 1 & 100 & 10^4 & 10^6 \end{bmatrix}.$$

∎

El siguiente resultado lo usaremos ampliamente cuando trabajemos con una matriz de una transformación lineal.

Proposición 2.4.1. *Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales de dimensión finita. Sean B_1 una base de $\mathbb{V}$ y B_2 una base de $\mathbb{W}$. Sea $T : \mathbb{V} \to \mathbb{W}$ tal que $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Entonces,*

- $v \in \mathrm{Nu}(T)$ *si y sólo si* $[v]^{B_1} \in \mathrm{nul}([T]_{B_1}^{B_2})$.
- $w \in \mathrm{Im}(T)$ *si y sólo si* $[w]^{B_2} \in \mathrm{col}([T]_{B_1}^{B_2})$.

Dem. Supongamos que $B_1 = \{v_1, v_2, \cdots, v_n\}$ es una base de $\mathbb{V}$ y $B_2 = \{w_1, w_2, \cdots, w_m\}$ es una base de $\mathbb{W}$.

Primero, veamos que $v \in \mathrm{Nu}(T)$ si y sólo si $[v]^{B_1} \in \mathrm{nul}([T]_{B_1}^{B_2})$. Supongamos que $v \in \mathrm{Nu}(T)$, entonces $T(v) = 0_{\mathbb{W}}$ y por la ecuación (2.1), tenemos que

$$[T]_{B_1}^{B_2}\ [v]^{B_1} = [T(v)]^{B_2} = [0_{\mathbb{W}}]^{B_2} = 0_{\mathbb{K}^m}$$

y entonces $[v]^{B_1} \in \mathrm{nul}([T]_{B_1}^{B_2})$. Recíprocamente, si $[v]^{B_1} \in \mathrm{nul}([T]_{B_1}^{B_2})$, de nuevo por la ecuación (2.1), tenemos que

$$0_{\mathbb{K}^m} = [T]_{B_1}^{B_2}\ [v]^{B_1} = [T(v)]^{B_2}.$$

Entonces $T(v) = 0_{\mathbb{W}}$ y $v \in \mathrm{Nu}(T)$.

Por último, veamos que $w \in \mathrm{Im}(T)$ si y sólo si $[w]^{B_2} \in \mathrm{col}([T]_{B_1}^{B_2})$. Supongamos que $w \in \mathrm{Im}(T)$, entonces existe $z \in \mathbb{V}$ tal que $w = T(z)$, entonces por la ecuación (2.1), tenemos que

$$[w]^{B_2} = [T(z)]^{B_2} = [T]_{B_1}^{B_2}\ [z]^{B_1}.$$

Por lo tanto $[w]^{B_2} \in \mathrm{col}([T]_{B_1}^{B_2})$.

Finalmente si $[w]^{B_2} \in \mathrm{col}([T]_{B_1}^{B_2})$, entonces existe $x = [x_1\ x_2\ \cdots\ x_n]^T \in \mathbb{K}^n$ tal que

$$[w]^{B_2} = [T]_{B_1}^{B_2} x.$$

Sea $z := x_1 v_1 + x_2 v_2 + \cdots + x_n v_n \in \mathbb{V}$, entonces $[z]^B = x$ y, nuevamente por la ecuación (2.1), tenemos que

$$[w]^{B_2} = [T]_{B_1}^{B_2} x = [T]_{B_1}^{B_2} [z]^{B_1} = [T(z)]^{B_2}.$$

Entonces, usando que tomar coordenadas es lineal, nos queda que $[T(z) - w]^{B_2} = 0_{\mathbb{K}^m}$. Por lo tanto $T(z) = w$ y $w \in \mathrm{Im}(T)$. ■

Corolario 2.4.2. *Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales de dimensión finita. Sean B_1 una base de $\mathbb{V}$ y B_2 una base de $\mathbb{W}$. Sea $T : \mathbb{V} \to \mathbb{W}$ tal que $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Entonces,*

- $\{[x_1]^{B_1}, [x_2]^{B_1}, \cdots, [x_r]^{B_1}\}$ *es una base de* $\mathrm{nul}([T]_{B_1}^{B_2})$ *si y sólo si* $\{x_1, x_2, \cdots, x_r\}$ *es una base de* $\mathrm{Nu}(T)$.
- $\dim(\mathrm{Nu}(T)) = \dim(\mathrm{nul}([T]_{B_1}^{B_2}))$.
- $\{[y_1]^{B_2}, [y_2]^{B_2}, \cdots, [y_t]^{B_2}\}$ *es una base de* $\mathrm{col}([T]_{B_1}^{B_2})$ *si y sólo si* $\{y_1, y_2, \cdots, y_t\}$ *es una base de* $\mathrm{Im}(T)$.
- $\dim(\mathrm{Im}(T)) = \dim(\mathrm{col}([T]_{B_1}^{B_2}))$.

Dem. Se puede demostrar usando la Proposición 2.4.1 y la Proposición 2.3.3. ■

Usando el resultado anterior, podemos demostrar el próximo ejercicio.

Ejercicio 2.13. Sean $\mathbb{V}$ y $\mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales de dimensión finita. Sean B una base de $\mathbb{V}$ y C una base de $\mathbb{W}$. Sea $T : \mathbb{V} \to \mathbb{W}$ tal que $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ y $[T]_B^C$ la matriz de T respecto de las bases B y C. Demostrar que

a) T es monomorfismo si y sólo si, $\mathrm{nul}([T]_B^C) = \{0\}$.

b) T es epimorfismo si y sólo si, $\mathrm{col}([T]_B^C) = \mathbb{K}^{\dim(\mathbb{W})}$.

c) T es isomorfismo si y sólo si, $\dim(\mathbb{V}) = \dim(\mathbb{W})$ y $[T]_B^C$ es inversible.

Dem. *a*): Por la Proposición 2.3.1, T es monomorfismo si y sólo si $\mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$, si y sólo si $0 = \dim(\mathrm{Nu}(T)) = \dim(\mathrm{nul}([T]_B^C))$ (ver Corolario 2.4.2), si y sólo si $\mathrm{nul}([T]_B^C) = \{0\}$.

b): Por definición, T es epimorfismo si y sólo si $\mathrm{Im}(T) = \mathbb{W}$, si y sólo si $\dim(\mathbb{W}) = \dim(\mathrm{Im}(T))$, donde usamos que siempre vale que $\mathrm{Im}(T) \subseteq \mathbb{W}$ y la propiedad de que dos subespacios son iguales si tienen misma dimensión y uno está contenido en el otro. Por el Corolario 2.4.2), tenemos que $\dim(\mathrm{col}([T]_B^C) = \dim(\mathrm{Im}(T)) = \dim(\mathbb{W}) = \dim(\mathbb{K}^{\dim(\mathbb{W})})$, si y sólo si $\mathrm{col}([T]_B^C) = \mathbb{K}^{\dim(\mathbb{W})}$, donde usamos nuevamente que siempre vale que $\mathrm{col}([T]_B^C) \subseteq \mathbb{K}^{\dim(\mathbb{W})}$ y la propiedad de que dos subespacios son iguales si tienen misma dimensión y uno está contenido en el otro.

c) : **Nota:** La frase "$\dim(\mathbb{V}) = \dim(\mathbb{W})$" se agregó explícitamente porque sino la vuelta del ítem *c*) podría no ser cierta. Si $\dim(\mathbb{V}) \neq \dim(\mathbb{W})$ entonces la frase "$[T]_B^C$ es inversible" no tiene sentido ya que para que esté definida la inversa de una matriz, esa matriz en principio, tiene que ser cuadrada (misma cantidad de filas que de columnas).

Supongamos que T es isomorfismo veamos que $[T]_B^C$ es inversible.

Recordemos que si T es isomorfismo, entonces $\dim(\mathbb{V}) = \dim(\mathbb{W}) =: n$ (lo probamos en la Proposición 2.3.1). Por lo tanto $[T]_B^C \in \mathbb{K}^{n\times n}$. Por otra parte, si T es isomorfismo entonces T es epimorfismo entonces, por el item *b*), tenemos que $\dim(\mathrm{col}([T]_B^C)) = \mathrm{rg}([T]_B^C) = \dim(\mathbb{W}) = n$. Entonces, $[T]_B^C$ es una matriz de $n \times n$ y de rango completo, por lo tanto es inversible.

Recíprocamente, supongamos que $\dim(\mathbb{V}) = \dim(\mathbb{W}) = n$ y que $[T]_B^C \in \mathbb{K}^{n\times n}$ es inversible. Entonces $\dim(\mathrm{col}([T]_B^C)) = \mathrm{rg}([T]_B^C) = n$, por lo tanto $\mathrm{col}([T]_B^C) = \mathbb{K}^n$ y por el item b) tenemos que T es epimorfismo. Por otra parte, como $\dim(\mathbb{V}) = \dim(\mathbb{W})$ ya vimos en la Proposición 2.3.1 que si T es epimorfismo entonces T es isomorfismo y probamos lo que queríamos. ■

Ejercicio 2.15. Sea $T \in \mathcal{L}(\mathbb{R}_2[x], \mathbb{R}^3)$ la transformación lineal definida por

$$[T]_B^C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Donde $B = \{1 + x^2, 1 + x, x + x^2\}$ y $C = \{[1\ 1\ 0]^T, [1\ 0\ 1]^T, [0\ 1\ 1]^T\}$. Hallar $T^{-1}([1\ 0\ 1]^T)$.

Dem. Notar que $p \in T^{-1}([1\ 0\ 1]^T)$ si y sólo si $T(p) = [1\ 0\ 1]^T$. Es decir, buscamos todas las soluciones de esa ecuación. Tenemos como dato la matriz $[T]_B^C$ y las bases B y C correspondientes. Entonces, recordemos que

$$[T(p)]^C = [[1\ 0\ 1]^T]^C = [T]_B^C [p]^B.$$

Haciendo cuentas, se ve que $[[1\ 0\ 1]^T]^C = [0\ 1\ 0]^T$. Llamemos $z = [z_1\ z_2\ z_3]^T := [p]^B$ y resolvamos el sistema que nos queda:

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = [[1\ 0\ 1]^T]^C = [T]_B^C [p]^B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} z.$$

Todas las soluciones de dicho sistema son

$$z = \begin{bmatrix} -z_3 \\ 1 - z_3 \\ z_3 \end{bmatrix} \text{ con } z_3 \in \mathbb{R}.$$

Para no confundir, vamos a llamar $\alpha := z_3 \in \mathbb{R}$. Por lo tanto, las soluciones de dicho sistema nos quedan

$$z = \begin{bmatrix} -\alpha \\ 1 - \alpha \\ \alpha \end{bmatrix} \text{ con } \alpha \in \mathbb{R}.$$

Pero $z = [p]^B$. Entonces

$$p(x) = -\alpha(1+x^2) + (1-\alpha)(1+x) + \alpha(x+x^2) = (1+x) + \alpha(-2), \text{ con } \alpha \in \mathbb{R}.$$

Por lo tanto, $p \in T^{-1}([1\ 0\ 1]^T)$ si y sólo si $p(x) = (1+x) + \alpha(-2)$, con $\alpha \in \mathbb{R}$. ■

El siguiente resultado lo usaremos para resolver el **Ejercicio 2.18**.

Proposición 2.4.3. *Sean $\mathbb{V}, \mathbb{W}, \mathbb{Z}$ tres $\mathbb{K}$-espacios vectoriales de dimensión finita y B, C y D bases de los espacios $\mathbb{V}$, $\mathbb{W}$ y $\mathbb{Z}$ respectivamente. Sean $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ y $S \in \mathcal{L}(\mathbb{W}, \mathbb{Z})$. Entonces*

i) $[S \circ T]_B^D = [S]_C^D\ [T]_B^C$.

ii) *Si* $\dim(V) = \dim(\mathbb{W})$ *y T es un isomorfimo, entonces*
$$[T^{-1}]_C^B = ([T]_B^C)^{-1}.$$

Dem. $i)$: En primer lugar, observar que $S \circ T \in \mathcal{L}(\mathbb{V}, \mathbb{Z})$. Sea $v \in \mathbb{V}$, entonces aplicando la ecuación (2.1), tenemos por un lado que
$$[S(T(v))]^D = [S]_C^D\ [T(v)]^C$$
y, por otro lado,
$$[S \circ T(v)]^D = [S \circ T]_B^D\ [v]^B.$$
Como $S \circ T(v) = S(T(v))$, juntando todo lo anterior se sigue que
$$[S \circ T]_B^D\ [v]^B = [S \circ T(v)]^D = [S(T(v))]^D = [S]_C^D\ [T(v)]^C.$$
Aplicando nuevamente la ecuación (2.1), tenemos que $[T(v)]^C = [T]_B^C\ [v]^B$ y volviendo a la ecuación anterior,
$$[S \circ T]_B^D\ [v]^B = [S]_C^D\ [T(v)]^C = [S]_C^D\ [T]_B^C\ [v]^B.$$
Es decir, tenemos que $[S \circ T]_B^D\ [v]^B = [S]_C^D\ [T]_B^C\ [v]^B$, para todo $v \in \mathbb{V}$.

Recordemos que si dos matrices A y B son tales que $A[v]^B = B[v]^B$ para todo $v \in \mathbb{V}$ entonces $A = B$. Por lo tanto, usando eso, concluimos que
$$[S \circ T]_B^D = [S]_C^D\ [T]_B^C$$
y demostramos lo que queríamos.

$ii)$: Supongamos que $\dim(\mathbb{V}) = \dim(\mathbb{W}) =: n$ y T es un isomorfismo; por un lado, sabemos que eso implica que T es inversible, es decir existe $T^{-1} : \mathbb{W} \to \mathbb{V}$ y además probamos que $T^{-1} \in \mathcal{L}(\mathbb{W}, \mathbb{V})$. Por otra parte, en el **Ejercicio 2.13 c)**, probamos que si $\dim(\mathbb{V}) = \dim(\mathbb{W}) = n$ y T es un isomorfismo entonces $[T]_B^C$ es inversible. Veamos que $[T^{-1}]_C^B = ([T]_B^C)^{-1}$.

Como T^{-1} es la inversa de T, tenemos que $T \circ T^{-1} = I_{\mathbb{W}}$ y $T^{-1} \circ T = I_{\mathbb{V}}$.

Observar que $[I_{\mathbb{V}}]_B^B = I_{n\times n}$ y $[I_{\mathbb{W}}]_C^C = I_{n\times n}$, donde $I_{n\times n}$ denota la matriz identidad de $n \times n$. Si no están convencidos de esa igualdad, observar que si $B = \{v_1, v_2, \cdots, v_n\}$ entonces, por definición, tenemos que
$$[I_{\mathbb{V}}]_B^B = [[I_{\mathbb{V}}(v_1)]^B\ [I_{\mathbb{V}}(v_2)]^B\ \cdots\ [I_{\mathbb{V}}(v_n)]^B] = [[v_1]^B\ [v_2]^B\ \cdots\ [v_n]^B] = I_{n\times n}.$$
De manera similar, se prueba que $[I_{\mathbb{W}}]_C^C = I_{n\times n}$.

Entonces, usando el item $i)$, tenemos que
$$I_{n\times n} = [I_{\mathbb{W}}]_C^C = [T \circ T^{-1}]_C^C = [T]_B^C\ [T^{-1}]_C^B \text{ y}$$
$$I_{n\times n} = [I_{\mathbb{V}}]_B^B = [T^{-1} \circ T]_B^B = [T^{-1}]_C^B\ [T]_B^C.$$
Entonces, probamos que $[T]_B^C\ [T^{-1}]_C^B = [T^{-1}]_C^B\ [T]_B^C = I_{n\times n}$. Por lo tanto, la inversa de $[T]_B^C$ es $[T^{-1}]_C^B$, es decir
$$[T^{-1}]_C^B = ([T]_B^C)^{-1}.$$
■

Ejercicio 2.18. Sea $T_1 \in \mathcal{L}(\mathbb{R}^3)$ la TL definida en el **Ejercicio 2.10** y $T_2 \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}_2)$ la TL definida por

$$T_2(\begin{bmatrix} a \\ b \\ c \end{bmatrix}) := (a+b) + (a+c)x + (b+c)x^2.$$

a) Hallar las matrices de T_1, T_2 y T_2^{-1} en las bases canónicas que correspondan.

b) Hallar la matriz de $T_1 \circ T_2^{-1}$ en las bases canónicas que correspondan y hallar base de $\mathrm{Nu}(T_1 \circ T_2^{-1})$ (usando dicha matriz).

Dem. Sean $B = \{[1\ 1\ 1]^T, [1\ -1\ 0]^T, [0\ 0\ 1]^T\}$ y $E = \{[1\ 0\ 0]^T, [0\ 1\ 0]^T, [0\ 0\ 1]^T\}$ la base canónica de $\mathbb{R}^3$. Es inmediato ver que

$$[T_1]_B^E = [\ [T_1([1\ 1\ 1]^T)]^E\ [T_1([1\ -1\ 0]^T)]^E\ [T_1([0\ 0\ 1]^T)]^E\ ] = \begin{bmatrix} 1 & -3 & 2 \\ -\frac{3}{2} & \frac{9}{2} & -3 \\ 2 & -6 & 4 \end{bmatrix}.$$

El ejercicio, nos pide $[T_1]_E^E$. Observar que, por la ecuación (2.1), por un lado vale que

$$[T_1(v)]^E = [T_1]_B^E\ [v]^B$$

y por el otro lado

$$[T_1(v)]^E = [T_1]_E^E\ [v]^E.$$

Por otra parte, recordemos que

$$[v]^B = [M]_E^B\ [v]^E,$$

donde $[M]_E^B$ es la matriz de cambio de base de E a B. Entonces, volviendo a la ecuación anterior, tenemos que

$$[T_1]_E^E\ [v]^E = [T_1(v)]^E = [T_1]_B^E\ [v]^B = [T_1]_B^E\ [M]_E^B\ [v]^E.$$

Es decir, $[T_1]_E^E\ [v]^E = [T_1]_B^E\ [M]_E^B\ [v]^E$, para todo $v \in \mathbb{V}$. Entonces (como la igualdad anterior vale para todo $v \in \mathbb{V}$) concluimos que

$$[T_1]_E^E = [T_1]_B^E\ [M]_E^B.$$

Observar que $[M]_B^E = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Entonces, tal como vimos en el **Ejercicio 1.P** tenemos $[M]_E^B = ([M]_B^E)^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Entonces,

$$[T_1]_E^E = [T_1]_B^E\ [M]_E^B = \begin{bmatrix} 1 & -3 & 2 \\ -\frac{3}{2} & \frac{9}{2} & -3 \\ 2 & -6 & 4 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 2 & 2 \\ \frac{3}{2} & -3 & -3 \\ -2 & 4 & 4 \end{bmatrix}.$$

Por último, si $E' = \{1, x, x^2\}$ es la base canónica de $\mathbb{R}_2[x]$ entonces

$$[T_2]_E^{E'} = [\ [T_2([1\ 0\ 0]^T)]^{E'}\ [T_2([0\ 1\ 0]^T)]^{E'}\ [T_2([0\ 0\ 1]^T)]^{E'}\ ] =$$
$$= [\ [1+x]^{E'}\ [1+x^2]^{E'}\ [x+x^2]^{E'}] = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Aplicando lo que probamos en la Proposición 2.4.3, tenemos que

$$[T_1 \circ T_2^{-1}]_{E'}^E = [T_1]_E^E\ [T_2^{-1}]_{E'}^E = [T_1]_E^E\ ([T_2]_E^{E'})^{-1}.$$

Como $([T_2]_E^{E'})^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}$. Concluimos que

$$[T_1 \circ T_2^{-1}]_{E'}^{E} = [T_1]_E^E \ ([T_2]_E^{E'})^{-1} =$$

$$= \frac{1}{2}\begin{bmatrix} -1 & 2 & 2 \\ \frac{3}{2} & -3 & -3 \\ -2 & 4 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{5}{2} \\ \frac{3}{4} & \frac{3}{4} & -\frac{15}{4} \\ -1 & -1 & 5 \end{bmatrix}.$$

Finalmente, vamos a hallar una base de $\mathrm{Nu}(T_1 \circ T_2^{-1})$. En la Proposición 2.4.1 probamos que $v \in \mathrm{Nu}(T_1 \circ T_2^{-1})$ si y sólo si $[v]^{E'} \in \mathrm{nul}([T_1 \circ T_2^{-1}]_{E'}^{E})$.

Resolviendo el sistema homogéneo asociado a la matriz $[T_1 \circ T_2^{-1}]_{E'}^{E}$, es fácil ver que

$$\mathrm{nul}([T_1 \circ T_2^{-1}]_{E'}^{E}) = gen\{[-1\ 1\ 0]^T, [5\ 0\ 1]^T\}.$$

Por la misma proposición, tenemos que $p(x) := -1 \cdot 1 + 1 \cdot x = -1 + x \in \mathrm{Nu}(T_1 \circ T_2^{-1})$ y $q(x) := 5 \cdot 1 + 1 \cdot x^2 = 5 + x^2 \in \mathrm{Nu}(T_1 \circ T_2^{-1})$. Por el Corolario 2.4.2, tenemos que $\dim(\mathrm{Nu}(T_1 \circ T_2^{-1})) = \dim(\mathrm{nul}([T_1 \circ T_2^{-1}]_{E'}^{E})) = 2$.

Por lo tanto,

$$\mathrm{Nu}(T_1 \circ T_2^{-1}) = gen\{-1 + x, 5 + x^2\}$$

y una base de $\mathrm{Nu}(T_1 \circ T_2^{-1})$ puede ser $B_{\mathrm{Nu}(T_1 \circ T_2^{-1})} = \{-1 + x, 5 + x^2\}$. ■

2.5. Proyectores y simetrías

Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y $\mathcal{S}_1, \mathcal{S}_2$ dos subespacios complementarios de $\mathbb{V}$, es decir

$$\mathbb{V} = \mathcal{S}_1 \oplus \mathcal{S}_2.$$

Entonces, a cada $v \in \mathbb{V}$ le corresponden únicos $v_1 \in \mathcal{S}_1$, $v_2 \in \mathcal{S}_2$ tales que

$$v = v_1 + v_2.$$

La *proyección de* $\mathbb{V}$ *sobre* $\mathcal{S}_1$ *en la dirección de* $\mathcal{S}_2$, denotada por $\Pi_{\mathcal{S}_1, \mathcal{S}_2}$, es la transformación lineal de $\mathbb{V}$ en $\mathbb{V}$ definida por

$$\Pi_{\mathcal{S}_1, \mathcal{S}_2}(v) := v_1.$$

Antes de resolver el **Ejercicio 2.19**, veamos un ejemplo.

Ejemplo 2.d. En $\mathbb{R}^3$, sean $\mathcal{S}_1 := gen\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\}$ y $\mathcal{S}_2 := gen\{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}$. Claramente,

$$\mathcal{S}_1 \oplus \mathcal{S}_2 = \mathbb{R}^3$$

y, $B = \{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}$ es una base de $\mathbb{R}^3$. Consideremos la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$, definida en la base B como

$$T(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ T(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \ T(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Entonces, $T = \Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$ es la proyección de $\mathbb{R}^3$ sobre $\mathcal{S}_1$ en la dirección de $\mathcal{S}_2$.

Observar que $\mathrm{Im}(T) = \mathcal{S}_1 = gen\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\}$ y $\mathrm{Nu}(T) = \mathcal{S}_2 = gen\{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}$.

Ejercicio 2.19.

a) Explicar por qué $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$ es la única TL de $\mathbb{V}$ en $\mathbb{V}$ tal que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v) = v$ si $v \in \mathcal{S}_1$ y $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v) = 0_{\mathbb{V}}$, si $v \in \mathcal{S}_2$. Comprobar que

$$\mathrm{Im}(\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}) \oplus \mathrm{Nu}(\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}) = \mathbb{V}.$$

b) Demostrar que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$ es *idempotente*, es decir, $\Pi^2_{\mathcal{S}_1,\,\mathcal{S}_2} = \Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$.

c) Observar que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2} + \Pi_{\mathcal{S}_2,\,\mathcal{S}_1} = I_{\mathbb{V}}$, o lo que es lo mismo, $\Pi_{\mathcal{S}_2,\,\mathcal{S}_1} = I_{\mathbb{V}} - \Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$.

Dem. *a*) : Primero veamos que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$ está bien definida y es una TL.

Para ver qué está bien definida, basta ver que a todo elemento del dominio de $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$ (que es $\mathbb{V}$) le corresponde un único elemento del espacio de llegada (que también es $\mathbb{V}$). De hecho, como

$$\mathbb{V} = \mathcal{S}_1 \oplus \mathcal{S}_2$$

a cada $v \in \mathbb{V}$ le corresponden únicos $v_1 \in \mathcal{S}_1$, $v_2 \in \mathcal{S}_2$ tales que $v = v_1 + v_2$. Como $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v) = v_1$, efectivamente $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$ está bien definida.

Veamos que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$ es una TL. Sean $v, w \in \mathbb{V}$ y $\alpha \in \mathbb{K}$, entonces existen únicos $v_1, w_1 \in \mathcal{S}_1$ y $v_2, w_2 \in \mathcal{S}_2$ tales que $v = v_1 + v_2$ y $w = w_1 + w_2$. Entonces, por un lado, por definición de $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$, tenemos que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v) = v_1$ y $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(w) = w_1$. Por otro lado, observar que

$$v + w = v_1 + v_2 + w_1 + w_2 = (v_1 + w_1) + (w_1 + w_2), \tag{2.2}$$

además, $v_1 + w_1 \in \mathcal{S}_1$ (porque $\mathcal{S}_1$ es un subespacio) y $v_2 + w_2 \in \mathcal{S}_2$ (porque $\mathcal{S}_2$ es un subespacio). Entonces (2.2) es la (única) descomposición del vector $v + w$ como suma de elementos de $\mathcal{S}_1$ y $\mathcal{S}_2$. Por lo tanto

$$\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v + w) = v_1 + w_1 = \Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v) + \Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(w).$$

De manera similar se prueba que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(\alpha v) = \alpha \Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v)$ y entonces, probamos que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$ es una transformación lineal.

Veamos que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$ es la única TL que cumple con lo que dice el item *a*).

Primero, observar que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$ efectivamente cumple lo que dice el item *a*). Es decir que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v) = v$ si $v \in \mathcal{S}_1$ y $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v) = 0_{\mathbb{V}}$, si $v \in \mathcal{S}_2$.

De hecho, si $v \in \mathcal{S}_1$, como $v = v + 0_{\mathbb{V}}$ y $0_{\mathbb{V}} \in \mathcal{S}_2$ (pues $\mathcal{S}_2$ es un subespacio), esa es la única descomposición de v como como suma de elementos de $\mathcal{S}_1$ y $\mathcal{S}_2$. Entonces, por definición de $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$, tenemos que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v) = v$. De esa manera, si $v \in \mathcal{S}_2$, entonces $v = 0_{\mathbb{V}} + v$, $0_{\mathbb{V}} \in \mathcal{S}_1$ (pues $\mathcal{S}_1$ es un subespacio) y esa descomposición es única. Entonces, por definición de $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$, tenemos que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v) = 0_{\mathbb{V}}$.

Finalmente, supongamos que existe $T \in \mathcal{L}(\mathbb{V})$ tal que $T(v) = v$, si $v \in \mathcal{S}_1$ y $T(v) = 0_{\mathbb{V}}$, si $v \in \mathcal{S}_2$. Veamos que $T = \Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$.

De hecho, dado $v \in \mathbb{V}$, existen únicos $v_1 \in \mathcal{S}_1$, $v_2 \in \mathcal{S}_2$ tales que

$$v = v_1 + v_2.$$

Entonces, como $v_1 \in \mathcal{S}_1$, tenemos que $T(v_1) = v_1$ y como $v_2 \in \mathcal{S}_2$, tenemos que $T(v_2) = 0_{\mathbb{V}}$. Además, por definición de $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$, tenemos que $\Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v) = v_1$. Entonces

$$T(v) = T(v_1 + v_2) = T(v_1) + T(v_2) = v_1 + 0_{\mathbb{V}} = v_1 = \Pi_{\mathcal{S}_1,\,\mathcal{S}_2}(v).$$

Como esa igualdad vale para cualquier $v \in \mathbb{V}$, concluimos que $T = \Pi_{\mathcal{S}_1,\,\mathcal{S}_2}$.

Por último, veamos que
$$\mathrm{Im}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) \oplus \mathrm{Nu}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) = \mathbb{V}.$$

En primer lugar, veamos que $\mathrm{Im}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) = \mathcal{S}_1$ y que $\mathrm{Nu}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) = \mathcal{S}_2$, probando la doble inclusión.

Veamos que $\mathrm{Im}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) = \mathcal{S}_1$. Si $w \in \mathrm{Im}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2})$, entonces existe $v \in \mathbb{V}$ tal que $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = w$. Como $v \in \mathbb{V}$ existen únicos $v_1 \in \mathcal{S}_1$, $v_2 \in \mathcal{S}_2$ tales que $v = v_1 + v_2$ y $w = \Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = v_1$, entonces $w = v_1 \in \mathcal{S}_1$ y vale que $\mathrm{Im}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) \subseteq \mathcal{S}_1$. Recíprocamente, si $v \in \mathcal{S}_1$, acabamos de ver que $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = v$, entonces $v \in \mathrm{Im}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2})$ y probamos que $\mathcal{S}_1 \subseteq \mathrm{Im}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2})$. Como probamos la doble inclusión, se sigue que $\mathrm{Im}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) = \mathcal{S}_1$.

Ahora veamos que $\mathrm{Nu}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) = \mathcal{S}_2$. Si $v \in \mathrm{Nu}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2})$, entonces $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = 0_{\mathbb{V}}$. Como $v \in \mathbb{V}$ existen únicos $v_1 \in \mathcal{S}_1$, $v_2 \in \mathcal{S}_2$ tales que $v = v_1 + v_2$ y $0_{\mathbb{V}} = \Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = v_1$, entonces $v = v_1 + v_2 = 0_{\mathbb{V}} + v_2 = v_2 \in \mathcal{S}_2$ y vale que $\mathrm{Nu}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) \subseteq \mathcal{S}_2$. Recíprocamente, si $v \in \mathcal{S}_2$, acabamos de ver que $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = 0_{\mathbb{V}}$, entonces $v \in \mathrm{Nu}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2})$ y probamos que $\mathcal{S}_2 \subseteq \mathrm{Nu}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2})$. Como probamos la doble inclusión, se sigue que $\mathrm{Nu}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) = \mathcal{S}_2$.

Usando que $\mathbb{V} = \mathcal{S}_1 \oplus \mathcal{S}_2$, concluimos que $\mathrm{Im}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) \oplus \mathrm{Nu}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}) = \mathbb{V}$.

b) : Recordar que $\Pi^2_{\mathcal{S}_1,\ \mathcal{S}_2} := \Pi_{\mathcal{S}_1,\ \mathcal{S}_2} \circ \Pi_{\mathcal{S}_1,\ \mathcal{S}_2}$. Como tenemos que probar una igualdad de TL, basta ver que la igualdad vale para cada $v \in \mathbb{V}$.

Sea $v \in \mathbb{V}$, entonces existen únicos $v_1 \in \mathcal{S}_1$, $v_2 \in \mathcal{S}_2$ tales que $v = v_1 + v_2$ y por definición de $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}$, tenemos que $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = v_1$. Ahora, como $v_1 \in \mathcal{S}_1$, vimos en el item a) que $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v_1) = v_1$. Juntando todo esto, nos queda que

$$\Pi^2_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = \Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v)) = \Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v_1) = v_1 = \Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v).$$

Como la igualdad anterior vale para todo $v \in \mathbb{V}$, concluimos que $\Pi^2_{\mathcal{S}_1,\ \mathcal{S}_2} = \Pi_{\mathcal{S}_1,\ \mathcal{S}_2}$.

c) : Como tenemos que probar una igualdad de TL, basta ver que la igualdad vale para cada $v \in \mathbb{V}$. Sea $v \in \mathbb{V}$, entonces existen únicos $v_1 \in \mathcal{S}_1$, $v_2 \in \mathcal{S}_2$ tales que $v = v_1 + v_2$ y por definición de $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}$ y $\Pi_{\mathcal{S}_2,\ \mathcal{S}_1}$, tenemos que $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = v_1$ y $\Pi_{\mathcal{S}_2,\ \mathcal{S}_1}(v) = v_2$. Entonces

$$[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2} + \Pi_{\mathcal{S}_2,\ \mathcal{S}_1}](v) = \Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v) + \Pi_{\mathcal{S}_2,\ \mathcal{S}_1}(v) = v_1 + v_2 = v = I_{\mathbb{V}}(v).$$

Como la igualdad anterior vale para todo $v \in \mathbb{V}$, concluimos que $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2} + \Pi_{\mathcal{S}_2,\ \mathcal{S}_1} = I_{\mathbb{V}}$. ■

Antes de seguir con el **Ejercicio 2.19 d), e)** veamos una definición que usaremos a continuación.

Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y $\mathcal{S}_1, \mathcal{S}_2$ dos subespacios complementarios de $\mathbb{V}$, es decir

$$\mathbb{V} = \mathcal{S}_1 \oplus \mathcal{S}_2.$$

La *simetría de* $\mathbb{V}$ *respecto de* $\mathcal{S}_1$ *en la dirección* $\mathcal{S}_2$, denotada por $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}$, es la transformación lineal de $\mathbb{V}$ en $\mathbb{V}$ definida por

$$\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2} := I_{\mathbb{V}} - 2\Pi_{\mathcal{S}_2,\ \mathcal{S}_1}.$$

Ejercicio 2.19 d), e).

d) Mostrar que $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}$ es la única transformación lineal de $\mathbb{V}$ en $\mathbb{V}$ tal que $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = v$, si $v \in \mathcal{S}_1$ y $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = -v$, si $v \in \mathcal{S}_2$.

e) Probar que $\Sigma^2_{\mathcal{S}_1,\ \mathcal{S}_2} = I_{\mathbb{V}}$.

Dem. *d*) : Se deja como ejercicio ver que $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}$ está bien definida y que es una TL.

Veamos que efectivamente $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = v$, si $v \in \mathcal{S}_1$, y $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = -v$, si $v \in \mathcal{S}_2$.

Ya vimos que si $v \in \mathcal{S}_1$ entonces $\Pi_{\mathcal{S}_2,\ \mathcal{S}_1}(v) = 0_{\mathbb{V}}$ y si $v \in \mathcal{S}_2$ entonces $\Pi_{\mathcal{S}_2,\ \mathcal{S}_1}(v) = v$. Por lo tanto, si $v \in \mathcal{S}_1$,
$$\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = I_{\mathbb{V}}(v) - 2\ \Pi_{\mathcal{S}_2,\ \mathcal{S}_1}(v) = v - 2\ 0_{\mathbb{V}} = v.$$

Finalmente, si $v \in \mathcal{S}_2$,

$$\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = I_{\mathbb{V}}(v) - 2\Pi_{\mathcal{S}_2,\ \mathcal{S}_1}(v) = v - 2v = -v.$$

Y probamos lo que queríamos.

Por último, supongamos que existe $S \in \mathcal{L}(\mathbb{V})$ tal que $S(v) = v$, si $v \in \mathcal{S}_1$ y $S(v) = -v$, si $v \in \mathcal{S}_2$. Veamos que $T = \Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}$.

De hecho, dado $v \in \mathbb{V}$, existen únicos $v_1 \in \mathcal{S}_1$, $v_2 \in \mathcal{S}_2$ tales que $v = v_1 + v_2$. Entonces, como $v_1 \in \mathcal{S}_1$, tenemos que $S(v_1) = v_1$ y como $v_2 \in \mathcal{S}_2$, tenemos que $S(v_2) = -v_2$. Además, por definición de $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}$, tenemos que

$$\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = v - 2\Pi_{\mathcal{S}_2,\ \mathcal{S}_1}(v) = v_1 + v_2 - 2v_2 = v_1 - v_2.$$

Entonces

$$S(v) = S(v_1 + v_2) = S(v_1) + S(v_2) = v_1 - v_2 = \Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v).$$

Como esa igualdad vale para cualquier $v \in \mathbb{V}$, concluimos que $S = \Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}$.

e) : Sea $v \in \mathbb{V}$, entonces existen únicos $v_1 \in \mathcal{S}_1$, $v_2 \in \mathcal{S}_2$ tales que $v = v_1 + v_2$ y por definición de $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}$, tenemos que

$$\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = v - 2\Pi_{\mathcal{S}_2,\ \mathcal{S}_1}(v) = v_1 + v_2 - 2v_2 = v_1 - v_2.$$

Ahora, como $v_1 \in \mathcal{S}_1$, vimos en el item d) que $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_1) = v_1$ y como $v_2 \in \mathcal{S}_2$, vimos en el item d) que $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_2) = -v_2$. Juntando todo esto, nos queda que

$$\Sigma^2_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = \Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v)) = \Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_1 - v_2) = \Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_1) - \Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_2) =$$

$$= v_1 - (-v_2) = v_1 + v_2 = v = I_{\mathbb{V}}(v).$$

Como la igualdad anterior vale para todo $v \in \mathbb{V}$, concluimos que $\Sigma^2_{\mathcal{S}_1,\ \mathcal{S}_2} = I_{\mathbb{V}}$. ■

La siguiente propiedad la usaremos para probar el **Ejercicio 2.20**.

Proposición 2.5.1. *Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial de dimensión n y $\mathcal{S}_1, \mathcal{S}_2$ dos subespacios complementarios de $\mathbb{V}$, es decir*

$$\mathbb{V} = \mathcal{S}_1 \oplus \mathcal{S}_2.$$

Supongamos que $\dim(\mathcal{S}_1) = r$ y $B = \{v_1, \cdots, v_r, v_{r+1}, \cdots, v_n\}$ es una base de $\mathbb{V}$ donde $\{v_1, v_2, \cdots, v_r\}$ es una base de $\mathcal{S}_1$ y $\{v_{r+1}, \cdots, v_n\}$ es una base de $\mathcal{S}_2$. Entonces

$$[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B = \begin{bmatrix} I_{r\times r} & 0_{r\times n-r} \\ 0_{n-r\times r} & 0_{n-r\times n-r} \end{bmatrix}.$$

$$[\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B = \begin{bmatrix} I_{r\times r} & 0_{r\times n-r} \\ 0_{n-r\times r} & -I_{n-r\times n-r} \end{bmatrix}.$$

Donde $I_{r\times r}$ denota la matriz identidad de $r \times r$, $0_{r\times n-r}$ la matriz nula de $r \times n - r$, etc.

Dem. En el ejercicio anterior, vimos que $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = v$ si $v \in \mathcal{S}_1$ y $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = 0_{\mathbb{V}}$, si $v \in \mathcal{S}_2$. Entonces

$$\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v_i) = v_i, \text{ para todo } i = 1, 2, \cdots, r \text{ y } \Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v_i) = 0_{\mathbb{V}} \text{ para todo } i = r+1, r+2, \cdots, n.$$

Recordemos además que $[v_i]^B = e_i$, donde e_i es el vector i-ésimo de la base canónica de $\mathbb{K}^n$, es decir el vector de $\mathbb{K}^n$ con todas componentes nulas excepto en el lugar i donde vale 1. Por ejemplo,

$$[v_1]^B = [1\ 0 \cdots\ 0]^T = e_1.$$

Por lo tanto, por definición de $[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B$, tenemos que

$$\begin{aligned}
&[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B \\
&= [\ [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v_1)]^B\ [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v_2)]^B \cdots\ [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v_r)]^B\ [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v_{r+1})]^B\ \cdots\ [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v_n)]^B\ ] \\
&= [\ [v_1]^B\ [v_2]^B\ \cdots\ [v_r]\ [0_{\mathbb{V}}]^B\ \cdots\ [0_{\mathbb{V}}]^B\ ] = [\ e_1\ e_2\ \cdots\ e_r\ 0_{\mathbb{K}^n}\ \cdots\ 0_{\mathbb{K}^n}] \\
&= \begin{bmatrix} I_{r\times r} & 0_{r\times n-r} \\ 0_{n-r\times r} & 0_{n-r\times n-r} \end{bmatrix}.
\end{aligned}$$

De la misma manera, en el ejercicio anterior, vimos que $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = v$ si $v \in \mathcal{S}_1$ y $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v) = -v$, si $v \in \mathcal{S}_2$.

Entonces $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_i) = v_i$, para todo $i = 1, 2, \cdots, r$ y $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_i) = -v_i$ para todo $i = r+1, r+2, \cdots, n$. Por lo tanto, por definición de $[\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B$, tenemos que

$$\begin{aligned}
&[\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B \\
&= [\ [\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_1)]^B\ [\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_2)]^B \cdots\ [\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_r)]^B\ [\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_{r+1})]^B\ \cdots\ [\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v_n)]^B\ ] \\
&= [\ [v_1]^B\ [v_2]^B\ \cdots\ [v_r]\ [-v_{r+1}]^B\ \cdots\ [-v_n]^B\ ] = [\ e_1\ e_2\ \cdots\ e_r\ -e_{r+1}\ \cdots\ -e_n] \\
&= \begin{bmatrix} I_{r\times r} & 0_{r\times n-r} \\ 0_{n-r\times r} & -I_{n-r\times n-r} \end{bmatrix}.
\end{aligned}$$

■

Ejercicio 2.20. Sean $\mathbb{V}$ un $\mathbb{R}$-espacio vectorial de dimensión 3, $B = \{v_1, v_2, v_3\}$ una base de $\mathbb{V}$, $\mathcal{S}_1$ y $\mathcal{S}_2$ los subespacios de $\mathbb{V}$ definidos por $\mathcal{S}_1 = gen\{v_1 - 2v_2, v_1 + v_3\}$ y $\mathcal{S}_2 = gen\{v_2 - v_3\}$.

a) Comprobar que $\mathbb{V} = \mathcal{S}_1 \oplus \mathcal{S}_2$.

b) Hallar las matrices con respecto a la base B de las proyecciones y simetrías inducidas por la partición $V = \mathcal{S}_1 \oplus \mathcal{S}_2$.

Dem. *a)* : Observar que $\mathcal{S}_1 \cap \mathcal{S}_2 = \{0_{\mathbb{V}}\}$. De hecho, si $w \in \mathcal{S}_1 \cap \mathcal{S}_2$ entonces $w \in \mathcal{S}_1$ y existen $\alpha, \beta \in \mathbb{R}$ tales que $w = \alpha(v_1 - 2v_2) + \beta(v_1 + v_3)$ y $w \in \mathcal{S}_2$ y entonces existe $\gamma \in \mathbb{R}$ tal que $w = \gamma(v_2 - v_3)$. Entonces

$$\alpha(v_1 - 2v_2) + \beta(v_1 + v_3) - \gamma(v_2 - v_3) = w - w = 0_{\mathbb{V}}.$$

Operando y sacando factor común, nos queda que

$$v_1(\alpha + \beta) + v_2(-2\alpha - \gamma) + v_3(\beta + \gamma) = 0_{\mathbb{V}}.$$

Como $\{v_1, v_2, v_3\}$ una base de $\mathbb{V}$, y tenemos una CL de dichos vectores igualada al elemento neutro de $\mathbb{V}$, se sigue que

$$\alpha + \beta = -2\alpha - \gamma = \beta + \gamma = 0.$$

Entonces, resolviendo el sistema, es claro que $\alpha = \beta = \gamma = 0$, por lo tanto $w = 0_{\mathbb{V}}$ y $\mathcal{S}_1 \cap \mathcal{S}_2 = \{0_{\mathbb{V}}\}$.

Notar que $\dim(\mathcal{S}_2) = 1$ y $\dim(\mathcal{S}_1) = 2$. De hecho, $[v_1 - 2v_2]^B = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$ y $[v_2 - v_3]^B = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$, entonces como los vectores $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ son LI en $\mathbb{R}^3$, por la Proposición 2.3.3, $\{v_1 - 2v_2, v_1 + v_3\}$ es un conjunto LI de $\mathbb{V}$ y entonces $\dim(\mathcal{S}_1) = 2$.

Por otra parte, es claro que $\mathcal{S}_1 \oplus \mathcal{S}_2 \subseteq \mathbb{V}$ y, por el Teorema de la dimensión para la suma de subespacios, se sigue que

$$\dim(\mathcal{S}_1 \oplus \mathcal{S}_2) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) - \dim(\mathcal{S}_1 \cap \mathcal{S}_2) = 2 + 1 - 0 = 3 = \dim(\mathbb{V}).$$

Por lo tanto, $\mathcal{S}_1 \oplus \mathcal{S}_2 = \mathbb{V}$.

b) : Llamemos $B' := \{v_1 - 2v_2, v_1 + v_3, v_2 - v_3\}$. Por *a*), B' es una base de $\mathbb{V}$. Además, por la Proposición 2.5.1,

$$[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B'}^{B'} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Entonces, aplicando los cambios de bases correspondientes, se sigue que

$$[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B}^{B} = [M]_{B'}^{B}[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B'}^{B'}[M]_{B}^{B'} = [M]_{B'}^{B}[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B'}^{B'}([M]_{B'}^{B})^{-1},$$

donde

$$[M]_{B'}^{B} = \begin{bmatrix} 1 & 1 & 0 \\ -2 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

es la matriz de cambio de base de B' a B. Por lo tanto

$$\begin{aligned} [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B}^{B} &= [M]_{B'}^{B}[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B'}^{B'}([M]_{B'}^{B})^{-1} \\ &= \begin{bmatrix} 1 & 1 & 0 \\ -2 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 1 & -1 & -1 \\ 2 & 1 & 1 \\ 2 & 1 & -2 \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 3 & 0 & 0 \\ -2 & 2 & 2 \\ 2 & 1 & 1 \end{bmatrix}. \end{aligned}$$

Asímismo, por la Proposición 2.5.1,

$$[\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B'}^{B'} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Entonces, aplicando los cambios de bases correspondientes, se sigue que

$$[\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B}^{B} = [M]_{B'}^{B}[\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B'}^{B'}[M]_{B}^{B'} = [M]_{B'}^{B}[\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B'}^{B'}([M]_{B'}^{B})^{-1}.$$

Por lo tanto

$$\begin{aligned} [\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B}^{B} &= [M]_{B'}^{B}[\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B'}^{B'}([M]_{B'}^{B})^{-1} \\ &= \begin{bmatrix} 1 & 1 & 0 \\ -2 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 1 & -1 & -1 \\ 2 & 1 & 1 \\ 2 & 1 & -2 \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 3 & 0 & 0 \\ -4 & 1 & 4 \\ 4 & 2 & -1 \end{bmatrix}. \end{aligned}$$

■

Ejercicio 2.K. En cada uno de los siguientes casos, hallar la matriz respecto de la base canónica de las transformaciones lineales indicadas:

- La proyección de $\mathbb{R}^3$ sobre el plano $gen\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\}$ en la dirección, $gen\{\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}\}$.
- La simetría de $\mathbb{R}_2[x]$ con respecto a $gen\{1, x\}$ en la dirección $gen\{1 + x + x^2\}$.

Dem. Con la notación de la definición de $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}$, tenemos que $\mathcal{S}_1 = gen\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\}$ y $\mathcal{S}_2 = gen\{\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}\}$. Observar que como $\mathbb{R}^3 = \mathcal{S}_1 \oplus \mathcal{S}_2$, tenemos que $B := \{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}\}$ es una base de $\mathbb{R}^3$. Entonces, por la Proposición 2.5.1, se sigue que

$$[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B = \begin{bmatrix} I_{2\times 2} & 0_{2\times 1} \\ 0_{1\times 2} & 0_{1\times 1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Si $E = \{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}$ es la base canónica de $\mathbb{R}^3$, el ejercicio nos pide $[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_E^E$. Entonces, como tenemos $[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B$, sólo nos resta hacer un cambio de base. Si

$$[M]_B^E = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix},$$

denota la matriz de cambio de base de B a E, tenemos que

$$[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_E^E = [M]_B^E\ [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B\ [M]_E^B = [M]_B^E\ [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B\ ([M]_B^E)^{-1}.$$

Finalmente, como $([M]_B^E)^{-1} = \begin{bmatrix} 2 & -2 & 1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{bmatrix}$, tenemos que

$$[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_E^E = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & -2 & 1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & -2 & 1 \end{bmatrix}.$$

Observar que $(\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & -2 & 1 \end{bmatrix})^2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & -2 & 1 \end{bmatrix}$. Este hecho no es casualidad, después veremos que si $T \in \mathcal{L}(\mathbb{V})$ es un proyector, siempre vale que

$$([T]_B^B)^2 = [T]_B^B, \text{ para cualquier base } B \text{ de } \mathbb{V}.$$

Con la notación de la definición de $\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}$, tenemos que $\mathcal{S}_1 = gen\{1, x\}$ y $\mathcal{S}_2 = gen\{1+x+x^2\}$. Observar que como $\mathbb{R}_2[x] = \mathcal{S}_1 \oplus \mathcal{S}_2$, tenemos que $B' := \{1, x, 1+x+x^2\}$ es una base de $\mathbb{R}_2[x]$. Entonces, por la Proposición 2.5.1, se sigue que

$$[\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B'}^{B'} = \begin{bmatrix} I_{2\times 2} & 0_{2\times 1} \\ 0_{1\times 2} & -I_{1\times 1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Si $E' = \{1, x, x^2\}$ es la base canónica de $\mathbb{R}_2[x]$, el ejercicio nos pide $[\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_{E'}^{E'}$. Entonces, como tenemos $[\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_{B'}^{B'}$, sólo nos resta hacer un cambio de base. Si

$$[M]_{B'}^{E'} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

denota la matriz de cambio de base de B' a E', tenemos que

$$[\Sigma_{\mathcal{S}_1,\,\mathcal{S}_2}]_{E'}^{E'} = [M]_{B'}^{E'}\,[\Sigma_{\mathcal{S}_1,\,\mathcal{S}_2}]_{B'}^{B'}\,[M]_{E'}^{B'} = [M]_{B'}^{E'}\,[\Sigma_{\mathcal{S}_1,\,\mathcal{S}_2}]_{B'}^{B'}\,([M]_{B'}^{E'})^{-1}.$$

Finalmente, como $([M]_{B'}^{E'})^{-1} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$, tenemos que

$$[\Sigma_{\mathcal{S}_1,\,\mathcal{S}_2}]_{E'}^{E'} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \end{bmatrix}.$$

Observar que $(\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \end{bmatrix})^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_{3\times 3}$. Este hecho no es casualidad, después veremos que si $S \in \mathcal{L}(\mathbb{V})$ es una simetría, siempre vale que

$$([S]_B^B)^2 = I_{\dim(\mathbb{V})\times\dim(\mathbb{V})},\ \text{para cualquier base } B \text{ de } \mathbb{V}.$$

■

El siguiente ejercicio recolecta varias propiedades de proyectores y simetrías.

Ejercicio 2.23. Verificar las siguientes afirmaciones:

a) Si $T \in \mathcal{L}(\mathbb{V})$ es tal que $T^2 = T$, entonces T es la proyección de $\mathbb{V}$ sobre $\mathrm{Im}(T)$ en la dirección de $\mathrm{Nu}(T)$ es decir
$$T = \Pi_{\mathrm{Im}(T),\ \mathrm{Nu}(T)}.$$

b) Si $T \in \mathcal{L}(\mathbb{V})$ es tal que $T^2 = T$, entonces $S := I_{\mathbb{V}} - 2T$ es tal que $S^2 = I_{\mathbb{V}}$.

c) Si $S \in \mathcal{L}(\mathbb{V})$ es tal que $S^2 = I_{\mathbb{V}}$, entonces $T := \frac{1}{2}(I_{\mathbb{V}} - S)$ es tal que $T^2 = T$.

d) Si $S \in \mathcal{L}(\mathbb{V})$ es tal que $S^2 = I_{\mathbb{V}}$, entonces S es la simetría de $\mathbb{V}$ con respecto a $\mathrm{Nu}(\frac{1}{2}(I_{\mathbb{V}} - S))$ en la dirección de $\mathrm{Im}(\frac{1}{2}(I_{\mathbb{V}} - S))$.

e) Si $S \in \mathcal{L}(\mathbb{V})$ es tal que $S^2 = I_{\mathbb{V}}$, entonces
$$\mathbb{V} = \mathrm{Nu}(S - I_{\mathbb{V}}) \oplus \mathrm{Nu}(S + I_{\mathbb{V}}).$$

Dem. *a*) Este ejercicio nos está pidiendo indirectamente que probemos que

$$\mathbb{V} = \mathrm{Im}(T) \oplus \mathrm{Nu}(T),$$

que es cuando está definido el proyector $\Pi_{\mathrm{Im}(T),\ \mathrm{Nu}(T)}$ (tal como vimos más arriba). Si probamos eso, luego veremos que $T = \Pi_{\mathrm{Im}(T),\ \mathrm{Nu}(T)}$.

Vamos a ver entonces que $\mathbb{V} = \mathrm{Im}(T) \oplus \mathrm{Nu}(T)$, usando como hipótesis que $T^2 = T \circ T = T$.

Primero veamos que $\mathrm{Im}(T) \cap \mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$. Supongamos que $v \in \mathrm{Im}(T) \cap \mathrm{Nu}(T)$, entonces por un lado $T(v) = 0_{\mathbb{V}}$ y por el otro $v \in \mathrm{Im}(T)$. Como $v \in \mathrm{Im}(T)$ existe $u \in \mathbb{V}$ tal que $v = T(u)$ entonces, usando que $T \circ T = T$, se sigue que $v = T(u) = T(T(u)) = T(v)$. Acabamos de demostrar un hecho importante (usando como hipótesis $T^2 = T$):

$$\text{si } v \in \mathrm{Im}(T) \text{ entonces } T(v) = v.$$

Siguiendo con la demostración, teníamos que $T(v) = 0_{\mathbb{V}}$ (porque $v \in \mathrm{Nu}(T)$), entonces $v = T(v) = 0_{\mathbb{V}}$, por lo tanto $\mathrm{Im}(T) \cap \mathrm{Nu}(T) = \{0_{\mathbb{V}}\}$ e $\mathrm{Im}(T)$ y $\mathrm{Nu}(T)$ están en suma directa.

Sólo resta ver que $\mathbb{V} = \mathrm{Im}(T) \oplus \mathrm{Nu}(T)$, es decir queremos ver que, dado $v \in \mathbb{V}$ existen $v_1 \in \mathrm{Im}(T)$ y $v_2 \in \mathrm{Nu}(T)$ tales que $v = v_1 + v_2$.

Observar que siempre vale que $v = T(v) + (v - T(v))$, donde $T(v)$ claramente pertenece a $\mathrm{Im}(T)$. Veamos que $v - T(v) \in \mathrm{Nu}(T)$, de hecho, $T(v - T(v)) = T(v) - T(T(v)) = T(v) - T^2(v) = T(v) - T(v) = 0_{\mathbb{V}}$. Por lo tanto, concluimos que $v - T(v) \in \mathrm{Nu}(T)$ y como a v lo escribimos como $v = v_1 + v_2$, con $v_1 := T(v) \in \mathrm{Im}(T)$ y $v_2 := v - T(v) \in \mathrm{Nu}(T)$, se sigue que

$$\mathbb{V} = \mathrm{Im}(T) \oplus \mathrm{Nu}(T).$$

Por lo tanto, está definido el proyector $\Pi_{Im(T),\ \mathrm{Nu}(T)}$.

Por otra parte, si $v \in \mathrm{Im}(T)$ vimos arriba que $T(v) = v$ y (obviamente) $T(v) = 0_{\mathbb{V}}$ si $v \in \mathrm{Nu}(T)$. Entonces, como $\Pi_{\mathrm{Im}(T),\ \mathrm{Nu}(T)}$ es la única transformación lineal tal que $\Pi_{\mathrm{Im}(T),\ \mathrm{Nu}(T)}(v) = v$, si $v \in \mathrm{Im}(T)$ y $\Pi_{\mathrm{Im}(T),\ \mathrm{Nu}(T)}(v) = 0_{\mathbb{V}}$, si $v \in \mathrm{Nu}(T)$ (lo probamos en el **Ejercicio 2.19 a**), tenemos que $T = \Pi_{\mathrm{Im}(T),\ \mathrm{Nu}(T)}$.

b) : Como $T^2 = T$, por el item a), tenemos que $T = \Pi_{\mathrm{Im}(T),\ \mathrm{Nu}(T)}$. Entonces

$$S = I_{\mathbb{V}} - 2T = I_{\mathbb{V}} - 2\Pi_{Im(T),\ \mathrm{Nu}(T)} = \Sigma_{\mathrm{Nu}(T),\ \mathrm{Im}(T)}$$

y por el **Ejercicio 2.19 e)** se sigue que $S^2 = I_{\mathbb{V}}$.

c) : De hecho, sea $T := \frac{1}{2}(I_{\mathbb{V}} - S)$ entonces

$$\begin{aligned} T^2 = T \circ T = \frac{1}{2}(I_{\mathbb{V}} - S) \circ \frac{1}{2}(I_{\mathbb{V}} - S) = \frac{1}{4}[(I_{\mathbb{V}} - S) \circ (I_{\mathbb{V}} - S)] = \\ = \frac{1}{4}(I_{\mathbb{V}} \circ I_{\mathbb{V}} - I_{\mathbb{V}} \circ S - S \circ I_{\mathbb{V}} + S \circ S) = \frac{1}{4}(I_{\mathbb{V}} - 2S + S^2) = \\ = \frac{1}{4}(I_{\mathbb{V}} - 2S + I_{\mathbb{V}}) = \frac{1}{4}(2I_{\mathbb{V}} - 2S) = \frac{1}{2}(I_{\mathbb{V}} - S) = T, \end{aligned}$$

donde usamos la hipótesis $S^2 = I_{\mathbb{V}}$.

Otra forma de probar este hecho es verificando que $T^2(v) = T(v)$, para todo $v \in \mathbb{V}$.

d) : Si $S \in \mathcal{L}(\mathbb{V})$ es tal que $S^2 = I_{\mathbb{V}}$, usando el ítem c), vimos que si definimos $T := \frac{1}{2}(I_{\mathbb{V}} - S)$, entonces $T^2 = T$. Por otra parte, por el ítem a), como $T^2 = T$, tenemos que $T = \Pi_{\mathrm{Im}(T),\ \mathrm{Nu}(T)}$. Entonces, por definición de simetría, se sigue que

$$S = I_{\mathbb{V}} - 2T = I_{\mathbb{V}} - 2\Pi_{\mathrm{Im}(T),\ \mathrm{Nu}(T)} = \Sigma_{\mathrm{Nu}(T),\ \mathrm{Im}(T)}.$$

Como $\mathrm{Nu}(T) = \mathrm{Nu}(\frac{1}{2}(I_{\mathbb{V}} - S))$ e $\mathrm{Im}(T) = \mathrm{Im}(\frac{1}{2}(I_{\mathbb{V}} - S))$. Entonces S es la simetría de $\mathbb{V}$ respecto a $\mathrm{Nu}(\frac{1}{2}(I_{\mathbb{V}} - S))$ en la dirección $\mathrm{Im}(\frac{1}{2}(I_{\mathbb{V}} - S))$.

e) : Si $S \in \mathcal{L}(\mathbb{V})$ es tal que $S^2 = I_{\mathbb{V}}$, usando el ítem c), vimos que si definimos $T := \frac{1}{2}(I_{\mathbb{V}} - S)$ entonces $T^2 = T$. Como $T^2 = T$, por el ítem a) tenemos, que

$$\mathbb{V} = \mathrm{Nu}(T) \oplus \mathrm{Im}(T) = \mathrm{Nu}(\frac{1}{2}(I_{\mathbb{V}} - S)) \oplus \mathrm{Im}(\frac{1}{2}(I_{\mathbb{V}} - S)).$$

Observar que si L es una transformación lineal y $a \in \mathbb{K} \setminus \{0\}$ (un escalar no nulo), siempre vale que

$$\mathrm{Nu}(aL) = \mathrm{Nu}(L).$$

La prueba es muy simple, pero la vamos a hacer. Si $v \in \mathrm{Nu}(aL)$, entonces $aL(v) = 0_{\mathbb{V}}$, como $a \neq 0$, se sigue que $L(v) = \frac{1}{a}\ aL(v) = \frac{1}{a}\ 0_{\mathbb{V}} = 0_{\mathbb{V}}$ y entonces $v \in \mathrm{Nu}(L)$. Recíprocamente, si $v \in \mathrm{Nu}(L)$ entonces, $L(v) = 0_{\mathbb{V}}$ y claramente $aL(v) = a0_{\mathbb{V}} = 0_{\mathbb{V}}$, entonces $v \in Nu(aL)$ y concluimos que $\mathrm{Nu}(aL) = \mathrm{Nu}(L)$.

Usando la idea anterior, se sigue que $\mathrm{Nu}(\frac{1}{2}(I_{\mathbb{V}} - S)) = \mathrm{Nu}(T) = \mathrm{Nu}(I_{\mathbb{V}} - S)$.

Ahora, veamos que

$$\mathrm{Im}(\frac{1}{2}(I_{\mathbb{V}} - S)) = \mathrm{Im}(T) = \mathrm{Nu}(I_{\mathbb{V}} + S).$$

De hecho, si $v \in \mathrm{Im}(T)$, como $T^2 = T$, en el item a) probamos que eso implica que $T(v) = v$. Entonces, reemplazado por la definición de T, nos queda

$$v = T(v) = \frac{1}{2}(I_{\mathbb{V}} - S)v = \frac{1}{2}(v - Sv).$$

Entonces

$$(I_{\mathbb{V}} + S)(v) = (I_{\mathbb{V}} + S)(\frac{1}{2}(v - Sv)) = \frac{1}{2}[v - Sv + Sv - S^2v] = \frac{1}{2}[v - I_{\mathbb{V}}(v)] = \frac{1}{2}[v - v] = 0_{\mathbb{V}},$$

donde usamos que $S^2 = I_{\mathbb{V}}$. Por lo tanto $\mathrm{Im}(\frac{1}{2}(I_{\mathbb{V}} - S)) \subseteq \mathrm{Nu}(I_{\mathbb{V}} + S)$.

Finalmente, si $v \in \mathrm{Nu}(I_{\mathbb{V}} + S)$, entonces $(I_{\mathbb{V}} + S)(v) = v + S(v) = 0_{\mathbb{V}}$, es decir $S(v) = -v$. Como $T = \frac{1}{2}(I_{\mathbb{V}} - S)$, entonces $S = I_{\mathbb{V}} - 2T$. Por lo tanto

$$-v = S(v) = (I_{\mathbb{V}} - 2T)v = v - 2T(v),$$

entonces, $-2v = -2T(v)$, o lo que es lo mismo, $v = T(v)$, entonces $v \in \mathrm{Im}(T) = \mathrm{Im}(\frac{1}{2}(I_{\mathbb{V}} - S))$ es decir probamos que $\mathrm{Nu}(\mathbb{V} + S) \subseteq \mathrm{Im}(\frac{1}{2}(I_{\mathbb{V}} - S))$. Por lo tanto, vale que $\mathrm{Im}(\frac{1}{2}(I_{\mathbb{V}} - S)) = \mathrm{Im}(T) = \mathrm{Nu}(I_{\mathbb{V}} + S)$ y entonces

$$\mathbb{V} = \mathrm{Nu}(T) \oplus \mathrm{Im}(T) = \mathrm{Nu}(I_{\mathbb{V}} - S) \oplus \mathrm{Nu}(I_{\mathbb{V}} + S).$$

■

Las siguientes son algunas conclusiones que podemos obtener del **Ejercicio 2.23**.

Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial, entonces

$$T \text{ es un proyector si y sólo si } T^2 = T \circ T = T.$$

En este caso, se cumplen las siguientes propiedades:

- $T = \Pi_{\mathrm{Im}(T),\ \mathrm{Nu}(T)}$, es decir, T es la proyección de $\mathbb{V}$ sobre $\mathrm{Im}(T)$ en la dirección de $\mathrm{Nu}(T)$.
- $T(v) = v$ si y sólo si $v \in \mathrm{Im}(T)$.
- $\mathbb{V} = \mathrm{Im}(T) \oplus \mathrm{Nu}(T)$.

$$S \text{ es una simetría si y sólo si } S^2 = I.$$

En este caso, se cumplen las siguientes propiedades:

- $S = \Sigma_{\mathrm{Nu}(I_{\mathbb{V}}-S),\ \mathrm{Im}(I_{\mathbb{V}}-S)}$, es decir, S es la simetría de $\mathbb{V}$ sobre $\mathrm{Nu}(I_{\mathbb{V}} - S)$ en la dirección de $\mathrm{Im}(I_{\mathbb{V}} - S)$.
- $\mathbb{V} = \mathrm{Nu}(S - I_{\mathbb{V}}) \oplus \mathrm{Nu}(S + I_{\mathbb{V}})$.

Ejercicio 2.L. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial de dimensión n (finita) y $\mathcal{S}_1, \mathcal{S}_2$ dos subespacios complementarios de $\mathbb{V}$, es decir

$$\mathbb{V} = \mathcal{S}_1 \oplus \mathcal{S}_2.$$

Entonces, para cualquier base B de $\mathbb{V}$ se sigue que

$$([\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B)^2 = [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B \quad \text{y} \quad ([\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B)^2 = I_{n\times n}.$$

Dem. Recordar que si $v \in \mathbb{V}$, tenemos que

$$[\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v)]^B = [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B\ [v]^B.$$

En el **Ejercicio 2.19 item b)**, vimos que $\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}$ es idempotente, es decir $\Pi^2_{\mathcal{S}_1,\ \mathcal{S}_2} = \Pi_{\mathcal{S}_1,\ \mathcal{S}_2}$. Entonces, por un lado, tenemos que

$$[\Pi^2_{\mathcal{S}_1,\ \mathcal{S}_2}(v)]^B = [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v)]^B = [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]_B^B\ [v]^B$$

y, por otro lado,

$$[\Pi^2_{\mathcal{S}_1,\ \mathcal{S}_2}(v)]^B = [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v))]^B = [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B\ [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}(v)]^B_B =$$

$$= [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B\ [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B\ [v]^B = ([\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B)^2[v]^B.$$

Igualando las dos expresiones anteriores, nos queda que

$$([\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B)^2\ [v]^B = [\Pi^2_{\mathcal{S}_1,\ \mathcal{S}_2}(v)]^B = [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B\ [v]^B.$$

Es decir, tenemos que

$$([\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B)^2\ [v]^B = [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B)[v]^B,$$

para todo $v \in \mathbb{V}$. Recordar que si dos matrices A y B son tales que $A[v]^B = B[v]^B$ para todo $v \in \mathbb{V}$, entonces $A = B$. Usando, este hecho, concluimos que

$$([\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B)^2 = [\Pi_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B.$$

Finalmente, en el **Ejercicio 2.19 d), e)**, vimos que $\Sigma^2_{\mathcal{S}_1,\ \mathcal{S}_2} = I_{\mathbb{V}}$. Entonces, por un lado, tenemos que

$$[\Sigma^2_{\mathcal{S}_1,\ \mathcal{S}_2}(v)]^B = [I_{\mathbb{V}}(v)]^B = [v]^B$$

y, por otro lado,

$$[\Sigma^2_{\mathcal{S}_1,\ \mathcal{S}_2}(v)]^B = [\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v))]^B = [\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B\ [\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}(v)]^B_B =$$

$$= [\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B\ [\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B\ [v]^B = ([\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B)^2[v]^B.$$

Igualando las dos expresiones anteriores, nos queda que

$$([\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B)^2\ [v]^B = [\Sigma^2_{\mathcal{S}_1,\ \mathcal{S}_2}(v)]^B = [v]^B = I_{n\times n}\ [v]^B.$$

Es decir tenemos que $([\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B)^2\ [v]^B = I_{n\times n}\ [v]^B$, para todo $v \in \mathbb{V}$. Entonces, se sigue que

$$([\Sigma_{\mathcal{S}_1,\ \mathcal{S}_2}]^B_B)^2 = I_{n\times n}.$$

∎

Ejercicio 2.M (De Parcial)**.** Sea $\mathbb{V}$ un K-espacio vectorial y sean $f, g : \mathbb{V} \to \mathbb{V}$ dos proyectores. Mostrar que:

a) $f = g$ si y sólo si $\operatorname{Im}(f) = \operatorname{Im}(g)$ y $\operatorname{Nu}(f) = \operatorname{Nu}(g)$. Es decir un proyector queda unívocamente definido conociendo su núcleo e imagen.

b) En general el ítem a) no vale si f y g son transformaciones lineales pero f ó g no son proyectores.

c) Si f y g conmutan (es decir $f \circ g = g \circ f$) entonces $f \circ g$ es un proyector tal que $\operatorname{Im}(f \circ g) = \operatorname{Im}(f) \cap \operatorname{Im}(g)$ y $\operatorname{Nu}(f \circ g) = \operatorname{Nu}(f) + \operatorname{Nu}(g)$.

Dem. Como f y g son proyectores entonces f y g son transformaciones lineales de $\mathbb{V}$ en $\mathbb{V}$ tales que $f \circ f = f$ y $g \circ g = g$.

a) : Si $f = g$ entonces obviamente tenemos que $\operatorname{Im}(f) = \operatorname{Im}(g)$ y $\operatorname{Nu}(f) = \operatorname{Nu}(g)$.

Recíprocamente, acabamos de ver que si $f \circ f = f^2 = f$, entonces $f = \Pi_{\operatorname{Im}(f),\ \operatorname{Nu}(f)}$, y además si $g \circ g = g^2 = g$, entonces $g = \Pi_{\operatorname{Im}(g),\ \operatorname{Nu}(g)}$. Usando que $\operatorname{Im}(f) = \operatorname{Im}(g)$ y $\operatorname{Nu}(f) = \operatorname{Nu}(g)$, se sigue que $f = \Pi_{\operatorname{Im}(f),\ \operatorname{Nu}(f)} = g$.

b): Por ejemplo, tomar $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f = id_{\mathbb{R}^2}$ (la función identidad de $\mathbb{R}^2$) y tomar $g : \mathbb{R}^2 \to \mathbb{R}^2$, definida en una base de $\mathbb{R}^2$ por:

$$g(\begin{bmatrix} 1 \\ 1 \end{bmatrix}) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad g(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Entonces g no es proyector, pues $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathrm{Im}(g)$ pero $g(\begin{bmatrix} 1 \\ 1 \end{bmatrix}) \neq \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Además $\mathrm{Im}(f) = \mathrm{Im}(g) = \mathbb{R}^2$ y $\mathrm{Nu}(f) = \mathrm{Nu}(g) = \{0\}$. Pero $f \neq g$. Pues, por ejemplo, $f(\begin{bmatrix} 1 \\ 1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \neq g(\begin{bmatrix} 1 \\ 1 \end{bmatrix})$.

c): Primero observar que $f \circ g$ es una transformación lineal y

$$(f \circ g) \circ (f \circ g) = f \circ (g \circ f) \circ g = f \circ (f \circ g) \circ g = (f \circ f) \circ (g \circ g) = f \circ g.$$

Entonces $f \circ g$ es un proyector.

Siempre vale que $\mathrm{Im}(f \circ g) \subseteq \mathrm{Im}(f)$ y $\mathrm{Im}(g \circ f) \subseteq \mathrm{Im}(g)$. Entonces $\mathrm{Im}(f \circ g) = \mathrm{Im}(g \circ f) \subseteq \mathrm{Im}(f) \cap \mathrm{Im}(g)$. Si hay dudas sobre este hecho, se recomienda hacer la demostración como ejercicio.

Por otra parte, si $x \in \mathrm{Im}(f) \cap \mathrm{Im}(g)$, entonces $x \in \mathrm{Im}(f)$ y $x \in \mathrm{Im}(g)$. Entonces, como f y g son proyectores, se sigue que $x = f(x)$ y $x = g(x)$. Entonces,

$$x = f(x) = f(g(x)) = f \circ g(x).$$

Por lo tanto $x \in \mathrm{Im}(f \circ g)$ y se sigue que $\mathrm{Im}(f) \cap \mathrm{Im}(g) \subseteq \mathrm{Im}(f \circ g)$. Entonces, como probamos la doble inclusión, tenemos que $\mathrm{Im}(f \circ g) = \mathrm{Im}(f) \cap \mathrm{Im}(g)$.

Finalmente, siempre vale que $\mathrm{Nu}(f) \subseteq \mathrm{Nu}(g \circ f) = \mathrm{Nu}(f \circ g)$ y $\mathrm{Nu}(g) \subseteq \mathrm{Nu}(f \circ g)$. Entonces, tenemos que $\mathrm{Nu}(f) + \mathrm{Nu}(g) \subseteq \mathrm{Nu}(f \circ g)$. De hecho, si $z \in \mathrm{Nu}(f) + \mathrm{Nu}(g)$, entonces $z = z_1 + z_2$, con $z_1 \in \mathrm{Nu}(f)$ y $z_2 \in \mathrm{Nu}(g)$, entonces,

$$f \circ g(z) = f \circ g(z_1 + z_2) = f \circ g(z_1) + f \circ g(z_2) = g \circ f(z_1) + f(0_{\mathbb{V}}) = g(0_{\mathbb{V}}) + 0_{\mathbb{V}} = 0_{\mathbb{V}}$$

y $z \in \mathrm{Nu}(f \circ g)$.

Por otra parte, si $x \in \mathrm{Nu}(f \circ g)$. Entonces, $f(g(x)) = 0$ y $g(x) \in \mathrm{Nu}(f)$. Observar que

$$x = x - g(x) + g(x),$$

con $x - g(x) \in \mathrm{Nu}(g)$ pues g es proyector. De hecho, $g(x - g(x))) = g(x) - g \circ g(x) = g(x) - g(x) = 0_{\mathbb{V}}$. Entonces, como escribimos a $x = x_1 + x_2$ con $x_1 := x - g(x) \in \mathrm{Nu}(g)$ y $x_2 := g(x) \in \mathrm{Nu}(f)$, se sigue que $x \in \mathrm{Nu}(f) + \mathrm{Nu}(g)$, entonces $\mathrm{Nu}(f \circ g) \subseteq \mathrm{Nu}(f) + \mathrm{Nu}(g)$ y, como probamos la doble inclusión, tenemos que $\mathrm{Nu}(f \circ g) = \mathrm{Nu}(f) + \mathrm{Nu}(g)$. ■

2.6. Rotaciones

Para resolver lose ejercicios de esta sección, vamos usamos las siguientes propiedades trigonométricas: Sean $a, b \in \mathbb{R}$, entonces

- $\cos(a)^2 + \sin(b)^2 = 1$.
- $\cos(a \pm b) = \cos(a)\cos(b) \mp \sin(a)\sin(b)$.
- $\sin(a \pm b) = \sin(a)\cos(b) \pm \cos(a)\sin(b)$.

El siguiente ejercicio nos presenta a las transformaciones lineales llamadas *rotaciones*.

Ejercicio 2.25. Sea $O(2, R) := \{R_\theta, S_\theta : \theta \in \mathbb{R}\}$ el conjunto de todas las TL de $\mathbb{R}^2$ en $\mathbb{R}^2$ definidas por

$$R_\theta(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) := \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

$$S_\theta(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) := \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

a) Hallar la imagen de la base canónica de $\mathbb{R}^2$ por $R_{\frac{\pi}{3}}$ y explicar el significado geométrico de la acción de $R_{\frac{\pi}{3}}$ sobre los vectores de $\mathbb{R}^2$.

Sea $E = \{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \}$ la base canónica de $\mathbb{R}^2$ y $\theta = \frac{\pi}{3}$. Entonces,

$$R_{\frac{\pi}{3}}(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix},$$

$$R_{\frac{\pi}{3}}(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix},$$

que es la rotación de un ángulo $\frac{\pi}{3}$ en sentido antihorario de ambos vectores.

b) Sea la base $\{ \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix} \}$, hallar la imagen de dicha base por $S_{\frac{\pi}{3}}$ y explicar el significado geométrico de la acción de $S_{\frac{\pi}{3}}$ sobre los vectores de $\mathbb{R}^2$.

Entonces,

$$S_{\frac{\pi}{3}}(\begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}) = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$$S_{\frac{\pi}{3}}(\begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}) = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

En el ítem e) veremos el significado geométrico de S_θ.

c) Hallar la imagen de la base canónica de $\mathbb{R}^2$ por R_θ y explicar el significado geométrico de la acción de R_θ sobre los vectores de $\mathbb{R}^2$.

Sea $E = \{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \}$ la base canónica de $\mathbb{R}^2$. Entonces,

$$R_\theta(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix},$$

$$R_\theta(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix},$$

que es la rotación de un ángulo θ en sentido antihorario (si $\theta \geq 0$) o en sentido horario (si $\theta < 0$) de ambos vectores.

d) Comprobar que $\{ \begin{bmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{bmatrix}, \begin{bmatrix} -\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \end{bmatrix} \}$ es una base de $\mathbb{R}^2$ y hallar su imagen por S_θ.

Veamos que $\{ \begin{bmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{bmatrix}, \begin{bmatrix} -\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \end{bmatrix} \}$ es una base de $\mathbb{R}^2$. Para esto basta ver que el conjunto es LI. Sean $a, b \in \mathbb{R}$ tales que

$$a \begin{bmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{bmatrix} + b \begin{bmatrix} -\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Entonces, nos queda el siguiente sistema matricial

$$\begin{bmatrix} \cos(\frac{\theta}{2}) & -\sin(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Como
$$\det\left(\begin{bmatrix} \cos(\frac{\theta}{2}) & -\sin(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{bmatrix}\right) = \cos(\frac{\theta}{2})^2 + \sin(\frac{\theta}{2})^2 = 1 \neq 0,$$
la única solución del sistema homogéneo anterior es la trivial, es decir, $a = b = 0$.

Por lo tanto, $\{\begin{bmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{bmatrix}, \begin{bmatrix} -\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \end{bmatrix}\}$ es una base de $\mathbb{R}^2$.

Por otra parte, veamos la imagen de esos vectores por S_θ.

$$\begin{aligned} S_\theta\left(\begin{bmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{bmatrix}\right) &= \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix}\begin{bmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{bmatrix} \\ &= \begin{bmatrix} \cos(\theta)\cos(\frac{\theta}{2}) + \sin(\theta)\sin(\frac{\theta}{2}) \\ \sin(\theta)\cos(\frac{\theta}{2}) - \cos(\theta)\sin(\frac{\theta}{2}) \end{bmatrix} \\ &= \begin{bmatrix} \cos(\theta - \frac{\theta}{2}) \\ \sin(\theta - \frac{\theta}{2}) \end{bmatrix} = \begin{bmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{bmatrix} \end{aligned}$$

$$\begin{aligned} S_\theta\left(\begin{bmatrix} -\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \end{bmatrix}\right) &= \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix}\begin{bmatrix} -\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \end{bmatrix} \\ &= \begin{bmatrix} -\cos(\theta)\sin(\frac{\theta}{2}) + \sin(\theta)\cos(\frac{\theta}{2}) \\ -\sin(\theta)\sin(\frac{\theta}{2}) - \cos(\theta)\cos(\frac{\theta}{2}) \end{bmatrix} \\ &= \begin{bmatrix} \sin(\theta - \frac{\theta}{2}) \\ -\cos(\theta - \frac{\theta}{2}) \end{bmatrix} = \begin{bmatrix} \sin(\frac{\theta}{2}) \\ -\cos(\frac{\theta}{2}) \end{bmatrix} = -\begin{bmatrix} -\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \end{bmatrix} \end{aligned}$$

e) ¿Cuál es el signficado geométrico de la acción de S_θ sobre los vectores de $\mathbb{R}^2$?

Sean $\mathcal{S}_1 := gen\{\begin{bmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{bmatrix}\}$ y $\mathcal{S}_2 := gen\{\begin{bmatrix} -\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \end{bmatrix}\}$. En el item $d)$ probamos que $\{\begin{bmatrix} \cos(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) \end{bmatrix}, \begin{bmatrix} -\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2}) \end{bmatrix}\}$ es una base de $\mathbb{R}^2$. Por lo tanto
$$\mathbb{R}^2 = \mathcal{S}_1 \oplus \mathcal{S}_2.$$
En el item $d)$ también probamos que:
$$S_\theta(v) = v \text{ para todo } v \in \mathcal{S}_1 \text{ y } S_\theta(v) = -v \text{ para todo } v \in \mathcal{S}_2.$$
Por lo tanto, S_θ es la simetría respecto a $\mathcal{S}_1$ en la dirección $\mathcal{S}_2$.

f) Dados $\alpha, \beta \in \mathbb{R}$, hallar las matrices respecto a la base canónica de las siguientes transformaciones lineales:
$$R_\alpha \circ R_\beta,\ S_\alpha \circ S_\beta,\ S_\alpha \circ R_\beta,\ R_\beta \circ S_\alpha.$$
$[R_\alpha]_E^E$, $[S_\alpha]_E^E$, $[R_\beta]_E^E$ y $[S_\beta]_E^E$ denotan las matrices respecto de la base canónica E de $\mathbb{R}^2$ de las transformaciones lineales $R_\alpha, S_\alpha, R_\beta$ y S_β, respectivamente. Entonces, por la Proposición 2.4.3,
$$\begin{aligned} [R_\alpha \circ R_\beta]_E^E &= [R_\alpha]_E^E [R_\beta]_E^E \\ &= \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}\begin{bmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{bmatrix} \\ &= \begin{bmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & -\cos(\alpha)\sin(\beta) - \sin(\alpha)\cos(\beta) \\ \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) & -\sin(\alpha)\sin(\beta) + \cos(\alpha)\cos(\beta) \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix} = [R_{\alpha+\beta}]_E^E.$$

$$\begin{aligned}
[S_\alpha \circ S_\beta]_E^E &= [S_\alpha]_E^E [S_\beta]_E^E \\
&= \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{bmatrix} \begin{bmatrix} \cos(\beta) & \sin(\beta) \\ \sin(\beta) & -\cos(\beta) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) & \cos(\alpha)\sin(\beta) - \sin(\alpha)\cos(\beta) \\ \sin(\alpha)\cos(\beta) - \cos(\alpha)\sin(\beta) & \sin(\alpha)\sin(\beta) + \cos(\alpha)\cos(\beta) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\alpha-\beta) & \sin(\beta-\alpha) \\ \sin(\alpha-\beta) & \cos(\alpha-\beta) \end{bmatrix} = \begin{bmatrix} \cos(\alpha-\beta) & -\sin(\alpha-\beta) \\ \sin(\alpha-\beta) & \cos(\alpha-\beta) \end{bmatrix} = [R_{\alpha-\beta}]_E^E.
\end{aligned}$$

$$\begin{aligned}
[S_\alpha \circ R_\beta]_E^E &= [S_\alpha]_E^E [R_\beta]_E^E \\
&= \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{bmatrix} \begin{bmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) & -\cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta) \\ \sin(\alpha)\cos(\beta) - \cos(\alpha)\sin(\beta) & -\sin(\alpha)\sin(\beta) - \cos(\alpha)\cos(\beta) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\alpha-\beta) & \sin(\alpha-\beta) \\ \sin(\alpha-\beta) & -\cos(\alpha-\beta) \end{bmatrix} = [S_{\alpha-\beta}]_E^E.
\end{aligned}$$

$$\begin{aligned}
[R_\beta \circ S_\alpha]_E^E &= [R_\beta]_E^E [S_\alpha]_E^E \\
&= \begin{bmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{bmatrix} \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\beta)\cos(\alpha) - \sin(\beta)\sin(\alpha) & \cos(\beta)\sin(\alpha) + \sin(\beta)\cos(\alpha) \\ \sin(\beta)\cos(\alpha) + \cos(\beta)\sin(\alpha) & \sin(\beta)\sin(\alpha) - \cos(\beta)\cos(\alpha) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\alpha+\beta) & \sin(\alpha+\beta) \\ \sin(\alpha+\beta) & -\cos(\alpha+\beta) \end{bmatrix} = [S_{\alpha+\beta}]_E^E.
\end{aligned}$$

(g) Concluir que el conjunto $O(2,\mathbb{R})$ es cerrado por composiciones. Sean $R_\alpha, S_\alpha, R_\beta, S_\beta \in O(2,\mathbb{R})$. Entonces, por el item f), tenemos que

$$\begin{aligned}
R_\alpha \circ R_\beta\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) &= \begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= R_{\alpha+\beta}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right).
\end{aligned}$$

Por lo tanto, $R_\alpha \circ R_\beta = R_{\alpha+\beta} \in O(2,\mathbb{R})$.

De la misma manera, podemos ver que $S_\alpha \circ S_\beta = R_{\alpha-\beta} \in O(2,\mathbb{R})$, $S_\alpha \circ R_\beta = S_{\alpha-\beta} \in O(2,\mathbb{R})$ y $R_\beta \circ S_\alpha = S_{\alpha+\beta} \in O(2,\mathbb{R})$. Por lo tanto, el conjunto $O(2,\mathbb{R})$ es cerrado por composiciones.

h) Si $\theta = 0$, entonces,

$$R_0 = \begin{bmatrix} \cos(0) & -\sin(0) \\ \sin(0) & \cos(0) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_{\mathbb{R}^2}.$$

i) Si $[R_\theta]_E^E$ es la matriz respecto de la base canónica E de la transformación lineal R_θ, entonces $[R_\theta]_E^E = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$ y

$$\det([R_\theta]_E^E) = \cos(\theta)^2 + \sin(\theta)^2 = 1 \neq 0.$$

Entonces $[R_\theta]_E^E$ es inversible y, por lo tanto R_θ es un isomorfismo. Entonces, existe R_θ^{-1}. Observar que

$$([R_\theta]_E^E)^{-1} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Entonces $R_\theta^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ es la transformación lineal

$$(R_\theta)^{-1}(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

De manera similar, si $[S_\theta]_E^E$ es la matriz respecto de la base canónica E de la transformación lineal S_θ, entonces $[S_\theta]_E^E = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix}$ y

$$\det([S_\theta]_E^E) = -\cos(\theta)^2 - \sin(\theta)^2 = -1 \neq 0.$$

Entonces $[S_\theta]_E^E$ es inversible y, por lo tanto, S_θ es un isomorfismo.

Entonces, existe S_θ^{-1}. Observar que

$$([S_\theta]_E^E)^{-1} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix} = [S_\theta]_E^E.$$

Entonces $S_\theta^{-1} = S_\theta$.

Sigamos con otro ejercicio de rotaciones en $\mathbb{R}^3$. Observar que la transformación lineal $R : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$\begin{aligned} R([1\ 0\ 0]^T) &:= [\cos(\theta)\ \sin(\theta)\ 0]^T \\ R([0\ 1\ 0]^T) &:= [-\sin(\theta)\ \cos(\theta)\ 0]^T \\ R([1\ 0\ 0]^T) &:= [0\ 0\ 0]^T \end{aligned}$$

es la rotación del ángulo θ en sentido antihorario del plano xy alrededor del eje z.

Ejercicio 2.26.

a) Hallar la imagen de los vectores $v_1 = [1\ 0\ 0]^T, v_2 = [1\ 1\ 0]^T, v_3 = [1\ 0\ 1]^T$ por la rotación de un ángulo $\frac{\pi}{4}$ en sentido antihorario del plano xy alrededor del eje z.

Observar que la expresión general de la transformación lineal $R_z^\theta : \mathbb{R}^3 \to \mathbb{R}^3$ que es la rotación de un ángulo θ en sentido antihorario del plano xy alrededor del eje z es:

$$R_z^\theta(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \text{ con } \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \in \mathbb{R}^3.$$

Tomando $\theta = \frac{\pi}{4}$, tenemos que

$$R_z^{\frac{\pi}{4}}(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}) = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \text{ con } \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \in \mathbb{R}^3.$$

Entonces

$$R_z^{\frac{\pi}{4}}(v_1) = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}.$$

$$R_z^{\frac{\pi}{4}}(v_2) = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{2} \\ 0 \end{bmatrix}.$$

$$R_z^{\frac{\pi}{4}}(v_3) = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 1 \end{bmatrix}.$$

b) Hallar la matriz respecto de la base canónica de la rotación de un ángulo θ en sentido antihorario del plano yz alrededor del eje x.

La transformación lineal $R_x^\theta : \mathbb{R}^3 \to \mathbb{R}^3$ que es la rotación de un ángulo θ en sentido antihorario del plano yz alrededor del eje x tiene la expresión:

$$R_x^\theta(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \text{ con } \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \in \mathbb{R}^3.$$

Por lo tanto, la matriz respecto de la base canónica de R_x^θ es

$$[R_x^\theta]_E^E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix},$$

donde E denota la base canónica de $\mathbb{R}^3$.

c) Hallar la matriz respecto de la base canónica de la rotación de un ángulo θ en sentido antihorario del plano zx alrededor del eje y.

La transformación lineal $R_y^\theta : \mathbb{R}^3 \to \mathbb{R}^3$ que es la rotación de un ángulo θ en sentido antihorario del plano xz alrededor del eje y, tiene la expresión:

$$R_y^\theta(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}) = \begin{bmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \text{ con } \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \in \mathbb{R}^3.$$

Por lo tanto, la matriz respecto de la base canónica de R_y^θ es

$$[R_y^\theta]_E^E = \begin{bmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix},$$

donde E denota la base canónica de $\mathbb{R}^3$.

2.7. Ecuaciones diferenciales. Operador derivación

Sea $\mathbb{K}$ el cuerpo $\mathbb{R}$ o $\mathbb{C}$. El operador *derivación* $D : C^\infty(\mathbb{R}, \mathbb{K}) \to C^\infty(\mathbb{R}, \mathbb{K})$, está definido por

$$D[y] := \frac{dy}{dx}.$$

Dados $a_0, a_1, \cdots, a_{n-1} \in \mathbb{K}$, se define el *operador diferencial de orden n con coeficientes constantes* $L : C^\infty(\mathbb{R}, \mathbb{K}) \to C^\infty(\mathbb{R}, \mathbb{K})$, como

$$L = D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0 I,$$

donde $D^n = D \circ D \circ \cdots \circ D$ (componer n veces el operador D).

Llamaremos *polinomio característico del operador L* al polinomio $p \in \mathbb{K}_n[x]$ que se obtiene de L intercambiando papeles entre D y x, es decir

$$p(x) := x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Recordar que probamos que $L \in \mathcal{L}(C^\infty(\mathbb{R}, \mathbb{K}))$.

Definición. Sean $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial, $D \in \mathcal{L}(\mathbb{V})$ y $q(x) = a_n x^n + a_{n-1}X^{n-1} + \cdots + a_1 x + a_0 x$ un polinomio de $\mathbb{C}_n[x]$ se define $q(D) : \mathbb{V} \to \mathbb{V}$ como

$$q(D) := a_n D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0 I.$$

Es fácil ver que como $D \in \mathcal{L}(\mathbb{V})$ también tenemos que $q(D) \in \mathcal{L}(\mathbb{V})$. Por otra parte, si $q_1 \in \mathbb{C}_n[x], q_2 \in \mathbb{C}_m[x]$ son dos polinomios, $\alpha, \beta \in \mathbb{C}$ y $D \in \mathcal{L}(\mathbb{V})$, es fácil ver que

$$(\alpha q_1 + \beta q_2)(D) = \alpha q_1(D) + \beta q_2(D) \text{ y } q_1 q_2(D) = q_1(D) \circ q_2(D) \tag{2.3}$$

la prueba de esto es muy simple y se sugiere hacerla.

A partir de la definición anterior, observar que:

Si p es el polinomio característico del operador diferencial de orden n que llamamos L. Entonces

$$p(D) = L.$$

La siguiente propiedad la vamos a usar para resolver varios ejercicios.

Proposición 2.7.1. *Sea $\mathbb{K} = \mathbb{C}$ y supongamos que*

$$p(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n),$$

donde $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{C}$ son las n raíces de p. Entonces

$$L = (D - \lambda_1 I) \circ (D - \lambda_2 I) \circ \cdots \circ (D - \lambda_n I).$$

Dem. Vamos a usar el principio de inducción para demostrar esta propiedad. Supongamos que $k = 1$, entonces $p(x) = x + a_0 = (x - \lambda_1)$, donde $\lambda_1 := -a_0 \in \mathbb{C}$ es la única raíz de p, Entonces

$$L = p(D) = (D - \lambda_1 I).$$

Por lo tanto para $k = 1$ se cumple la propiedad.

Supongamos que para $k = n$ vale la propiedad (HI, hipótesis inductiva). Usando la HI vamos a probar que la propiedad también vale para $k = n + 1$. En este caso $p \in \mathbb{C}_{n+1}[x]$.

Sea $\lambda_{n+1} \in \mathbb{C}$ una raíz de p (existe por el Teorema Fundamental del Álgebra). Entonces, usando la división de polinomios (por ejemplo), tenemos que

$$p(x) = (x - \lambda_{n+1})q(x),$$

donde $q \in \mathbb{C}_n[x]$.

Entonces, usando la propiedad (2.3), tenemos que

$$L = p(D) = (D - \lambda_{n+1}I)q(D).$$

Sean $\lambda_n, \lambda_{n-1}, \cdots, \lambda_1 \in \mathbb{C}$, las n raíces del polinomio q, entonces por HI, tenemos que

$$q(D) = (D - \lambda_1 I) \circ (D - \lambda_2 I) \circ \cdots \circ (D - \lambda_n I).$$

Por otra parte,

$$p(x) = (x - \lambda_{n+1})q(x) = (x - \lambda_{n+1})[(x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)].$$

Entonces,

$$L = p(D) = (D - \lambda_{n+1}I) \circ q(D) = (D - \lambda_{n+1}I) \circ (D - \lambda_1 I) \circ (D - \lambda_2 I) \circ \cdots \circ (D - \lambda_n I) =$$

$$= (D - \lambda_1 I) \circ (D - \lambda_2 I) \circ \cdots \circ (D - \lambda_n I) \circ (D - \lambda_{n+1} I),$$

donde para la última igualdad usamos fuertemente que si $\alpha, \beta \in \mathbb{C}$, los operadores $(D - \alpha I)$ y $(D - \beta I)$ conmutan, es decir $(D - \alpha I) \circ (D - \beta I) = D^2 - \alpha D - \beta D + \alpha\beta I = (D - \beta I) \circ (D - \alpha I)$.

Finalmente, como a partir de la validez de la proposición para n deducimos la validez para $n + 1$, se sigue por inducción, que la propiedad vale para todo $n \in \mathbb{N}$. ■

Para resolver el **Ejercicio 2.27** vamos a usar la siguiente propiedad (su demostración es muy simple y se sugiere hacerla): sea $\lambda \in \mathbb{C}$ y D el operador derivación. Observar que, como $(D - \lambda I)^{k+1}$ es componer $k + 1$ veces el operador $D - \lambda I$, se sigue que para todo $k \geq 1$,

$$(D - \lambda I)^{k+1} = (D - \lambda I) \circ (D - \lambda I)^k = (D - \lambda I)^k \circ (D - \lambda I). \tag{2.4}$$

Ejercicio 2.27. Sea $D : C^\infty(\mathbb{R}, \mathbb{C}) \to C^\infty(\mathbb{R}, \mathbb{C})$ el operador derivación.

a) Sea $\lambda \in \mathbb{C}$. Verificar que para todo $k \in \mathbb{N}$ vale que

$$(D - \lambda I)^k[f(x)e^{\lambda x}] = f^k(x)e^{\lambda x},$$

para toda $f \in C^\infty(\mathbb{R}, \mathbb{C})$.

b) Comprobar que

$$\mathrm{Nu}(D - \lambda I) = gen\{e^{\lambda x}\}.$$

c) Para cada $k \in \mathbb{N}$ verificar que si

$$\mathrm{Nu}((D - \lambda I)^k) = \{p(x)e^{\lambda x} : p \in \mathbb{C}_{k-1}[x]\},$$

entonces

$$\mathrm{Nu}((D - \lambda I)^{k+1}) = \{p(x)e^{\lambda x} : p \in \mathbb{C}_k[x]\}.$$

d) Utilizar los incisos b) y c) junto al principio de inducción para demostrar que para todo $k \in \mathbb{N}$, el conjunto $\{x^i e^{\lambda x}\}$ con $i = \{0, 1, 2 \cdots, k - 1\}$, es una base de $\mathrm{Nu}((D - \lambda I)^k)$.

e) Sea $g \in C^\infty(\mathbb{R}, \mathbb{C})$. Comprobar que la ecuación

$$(D - \lambda I)^k[y] = g,$$

admite una solución particular de la forma $y_p(x) = f(x)e^{\lambda x}$, donde $f^k(x) = g(x)e^{-\lambda x}$.

Dem. $a)$: Vamos a probarlo por inducción. Primero, el caso $n = 1$. En este caso, dada $f \in C^\infty(\mathbb{R}, \mathbb{C})$, por definición de D tenemos:

$$(D - \lambda I)[f(x)e^{\lambda x}] = D[f(x)e^{\lambda x}] - \lambda f(x)e^{\lambda x} = \frac{d(f(x)e^{\lambda x})}{dx} - \lambda f(x)e^{\lambda x} =$$

$$= f'(x)e^{\lambda x} + \lambda f(x)e^{\lambda x} - \lambda f(x)e^{\lambda x} = f'(x)e^{\lambda x},$$

por lo que la igualdad se cumple para $n = 1$.

Supongamos que la igualdad vale para $n = k$ (HI), es decir,

$$(D - \lambda I)^k[f(x)e^{\lambda x}] = f^k(x)e^{\lambda x},$$

veamos que la igualdad vale para $n = k + 1$. Entonces, dada $f \in C^\infty(\mathbb{R}, \mathbb{C})$, por definición de D, usando la propiedad (2.4) y usando la HI, tenemos que

$$(D - \lambda I)^{k+1}[f(x)e^{\lambda x}] = ((D - \lambda I) \circ (D - \lambda I)^k)\ [f(x)e^{\lambda x}] = (D - \lambda I)[(D - \lambda I)^k\ [f(x)e^{\lambda x}]] =$$

$$= (D - \lambda I)[f^k(x)e^{\lambda x}] = \frac{d(f^k(x)e^{\lambda x})}{dx} - \lambda f^k(x)e^{\lambda x} = f^{k+1}(x)e^{\lambda x} + \lambda f^k(x)e^{\lambda x} - \lambda f^k(x)e^{\lambda x} =$$

$$= f^{k+1}(x)e^{\lambda x}.$$

Como a partir de la validez de la igualdad para $n = k$, deducimos la validez para $n = k + 1$, se sigue que por inducción, la igualdad vale para todo $k \in \mathbb{N}$.

$b)$: Tenemos que probar una igualdad de conjuntos así que veremos la doble inclusión.

Sea $f \in \mathrm{Nu}(D - \lambda I)$, entonces $(D - \lambda I)[f(x)] = 0$, para todo $x \in \mathbb{R}$. Entonces $D[f(x)] - \lambda f(x) = \frac{df}{dx} - \lambda f(x) = 0$, para todo $x \in \mathbb{R}$. Es decir,

$$f' - \lambda f = 0.$$

Para resolver esta ecuación diferencial, vamos a multiplicar ambos lados la última igualdad por $e^{-\lambda x}$. Entonces, para todo $x \in \mathbb{R}$,

$$0 = e^{-\lambda x} f'(x) - \lambda e^{-\lambda x} f(x) = (e^{-\lambda x} f(x))'.$$

Integrando, nos queda, para todo $x \in \mathbb{R}$,

$$e^{-\lambda x} f(x) = \int (e^{-\lambda x} f(x))' = \int 0 = c,$$

donde $c \in \mathbb{C}$ es una constante. Por lo tanto, despejando, nos queda

$$f(x) = c(e^{-\lambda x})^{-1} = ce^{\lambda x}.$$

Entonces, $f \in gen\{e^{\lambda x}\}$ y tenemos que $\mathrm{Nu}(D - \lambda I) \subseteq gen\{e^{\lambda x}\}$.

Recíprocamente, sea $f \in gen\{e^{\lambda x}\}$, entonces $f(x) = \alpha e^{\lambda x}$, para cierto $\alpha \in \mathbb{C}$. Entonces, para cada $x \in \mathbb{R}$,

$$(D - \lambda I)[f(x)] = f'(x) - \lambda f(x) = \alpha\lambda e^{\lambda x} - \alpha\lambda e^{\lambda x} = 0.$$

Entonces $f \in \mathrm{Nu}(D - \lambda I)$ y se sigue que $gen\{e^{\lambda x}\} \subseteq \mathrm{Nu}(D - \lambda I)$. Por lo tanto, como probamos la doble inclusión, tenemos que $\mathrm{Nu}(D - \lambda I) = gen\{e^{\lambda x}\}$.

$c)$: Usando la propiedad (2.4), tenemos que

$$\mathrm{Nu}((D - \lambda I)^{k+1}) = \mathrm{Nu}((D - \lambda I)^k \circ (D - \lambda I)).$$

Con esto, vamos a demostrar la igualdad de conjuntos que nos pide el item c), para eso vamos a probar la doble inclusión.

Sea $f \in \mathrm{Nu}((D-\lambda I)^{k+1}) = \mathrm{Nu}((D-\lambda I)^k \circ (D-\lambda I))$, entonces para cada $x \in \mathbb{R}$,

$$0 = ((D-\lambda I)^k \circ (D-\lambda I))[f(x)] = (D-\lambda I)^k[(D-\lambda I)[f(x)]].$$

Entonces $(D-\lambda I)[f] \in \mathrm{Nu}(D-\lambda I)^k$. Por hipótesis, $\mathrm{Nu}((D-\lambda I)^k) = \{p(x)e^{\lambda x} : p \in \mathbb{C}_{k-1}[x]\}$. Entonces, existe $p \in \mathbb{C}_{k-1}[x]$, tal que

$$(D-\lambda I)[f(x)] = f'(x) - \lambda f(x) = p(x)e^{\lambda x}.$$

Para resolver esta ecuación diferencial, vamos a multiplicar ambos lados la última igualdad por $e^{-\lambda x}$. Entonces, para todo $x \in \mathbb{R}$,

$$e^{-\lambda x}f'(x) - \lambda e^{-\lambda x}f(x) = e^{-\lambda x}p(x)e^{\lambda x} = p(x)e^{-\lambda x}e^{\lambda x} = p(x).$$

Entonces

$$(e^{-\lambda x}f(x))' = e^{-\lambda x}f'(x) - \lambda e^{-\lambda x}f(x) = p(x).$$

Entonces, integrando la ecuación anterior a cada lado, obtenemos

$$e^{-\lambda x}f(x) = \int (e^{-\lambda x}f(x))' = \int p(x).$$

Sea $q(x) := \int p(x)$, entonces como $p \in \mathbb{C}_{k-1}[x]$ (y estamos integrando p) tenemos que $q \in \mathbb{C}_k[x]$. Entonces, despejando f de la última igualdad, nos queda, para todo $x \in \mathbb{R}$,

$$f(x) = (e^{-\lambda x})^{-1}q(x) = e^{\lambda x}q(x).$$

Entonces $f \in \{p(x)e^{\lambda x} : p \in \mathbb{C}_k[x]\}$ y tenemos la primera inclusión: $\mathrm{Nu}((D-\lambda I)^{k+1}) \subseteq \{p(x)e^{\lambda x} : p \in \mathbb{C}_k[x]\}$.

Recíprocamente, sea $f \in \{p(x)e^{\lambda x} : p \in \mathbb{C}_k[x]\}$, entonces $f(x) = p(x)e^{\lambda x}$, para cierto $p \in \mathbb{C}_k[x]$. Entonces,

$$(D-\lambda)[f(x)] = \frac{d(p(x)e^{\lambda x})}{dx} - \lambda f(x) = p'(x)e^{\lambda x} + p(x)\lambda e^{\lambda x} - \lambda p(x)e^{\lambda x} = p'(x)e^{\lambda x}.$$

Entonces, como $p \in \mathbb{C}_k[x]$ su derivada $p' \in \mathbb{C}_{k-1}[x]$, entonces

$$(D-\lambda)[f] \in \{p(x)e^{\lambda x} : p \in \mathbb{C}_{k-1}[x]\} = \mathrm{Nu}((D-\lambda I)^k).$$

Entonces, para todo $x \in \mathbb{R}$,

$$0 = (D-\lambda I)^k[(D-\lambda)[f(x)]] = ((D-\lambda I)^k \circ (D-\lambda I))[f(x)] = (D-\lambda I)^{k+1}[f(x)].$$

Por lo tanto, $f \in \mathrm{Nu}((D-\lambda I)^{k+1})$ y se sigue que $\{p(x)e^{\lambda x} : p \in \mathbb{C}_k[x]\} \subseteq \mathrm{Nu}((D-\lambda I)^{k+1})$. Entonces, como probamos la doble inclusión, tenemos que

$$\mathrm{Nu}((D-\lambda I)^{k+1}) = \{p(x)e^{\lambda x} : p \in \mathbb{C}_k[x]\}.$$

d) : Vamos a demsostrar el inciso d) por inducción en k. Primero, el caso $k = 1$. En este caso, queremos ver, que el conjunto $\{e^{\lambda x}\}$ es una base de $\mathrm{Nu}(D-\lambda I)$. En el ítem b), vimos que $\mathrm{Nu}(D-\lambda I) = gen\{e^{\lambda x}\}$ y como $e^{\lambda x}$ no es la función nula, el conjunto $\{e^{\lambda x}\}$ es LI, por lo tanto $\{e^{\lambda x}\}$ es una base de $\mathrm{Nu}(D-\lambda I)$.

Ahora, supongamos que $\{e^{\lambda x}, xe^{\lambda x}, \cdots, x^{k-1}e^{\lambda x}\}$ es una base de $\mathrm{Nu}((D-\lambda I)^k)$ (HI).

Veamos que esto también vale para $k+1$, es decir, veamos que $\{e^{\lambda x}, xe^{\lambda x}, \cdots, x^{k-1}e^{\lambda x}, x^k e^{\lambda x}\}$ es una base de $\mathrm{Nu}((D-\lambda I)^{k+1})$.

Primero probemos que el conjunto $\{e^{\lambda x}, xe^{\lambda x}, \cdots, x^{k-1}e^{\lambda x}, x^k e^{\lambda x}\}$ es LI. De hecho, sean $a_0, a_1, \cdots, a_k \in \mathbb{R}$ tales que

$$
\begin{aligned}
&0 = a_0 e^{\lambda x} + a_1 x e^{\lambda x} + \cdots + a_{k-1}x^{k-1}e^{\lambda x} + a_k x^k e^{\lambda x}, \text{ para todo } x \in \mathbb{R}. \text{ Entonces}\\
&0 = e^{\lambda x}(a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} + a_k x^k), \text{ para todo } x \in \mathbb{R}. \text{ Entonces como } e^{\lambda x} > 0,\\
&0 = a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} + a_k x^k, \text{ para todo } x \in \mathbb{R}
\end{aligned}
$$

Entonces, como $\{1, x, \cdots, x^{k-1}, x^k\}$ es un conjunto LI, se sigue que $a_0 = a_1 = \cdots = a_{k-1} = a_k = 0$. Por lo tanto el conjunto $\{e^{\lambda x}, xe^{\lambda x}, \cdots, x^{k-1}e^{\lambda x}, x^k e^{\lambda x}\}$ es LI.

Ahora sólo nos resta probar que

$$gen\{e^{\lambda x}, xe^{\lambda x}, \cdots, x^{k-1}e^{\lambda x}, x^k e^{\lambda x}\} = \mathrm{Nu}((D - \lambda I)^{k+1}),$$

sabiendo que por (HI) vale

$$gen\{e^{\lambda x}, xe^{\lambda x}, \cdots, x^{k-1}e^{\lambda x}\} = \mathrm{Nu}((D - \lambda I)^{k}).$$

Pero eso, es exactamente lo que hicimos en el item c). Si no se convencen, notar que

$$gen\{e^{\lambda x}, xe^{\lambda x}, \cdots, x^{k-1}e^{\lambda x}, x^k e^{\lambda x}\} = \{p(x)e^{\lambda x} : p \in \mathbb{C}_k[x]\}.$$

De hecho, si $f \in gen\{e^{\lambda x}, xe^{\lambda x}, \cdots, x^{k-1}e^{\lambda x}, x^k e^{\lambda x}\}$, entonces existen $\alpha_0, \alpha_1, \cdots, \alpha_k \in \mathbb{C}$ tales que

$$f(x) = \alpha_0 e^{\lambda x} + \alpha_1 x e^{\lambda x} + \cdots + \alpha_k x^k e^{\lambda x} = (\alpha_0 + \alpha_1 x + \cdots + \alpha_k x^k)e^{\lambda x}.$$

Si llamamos $p(x) := \alpha_0 + \alpha_1 x + \cdots + \alpha_k x^k$ entonces $p \in \mathbb{C}_k[x]$ y $f(x) = p(x)e^{\lambda x}$, por lo tanto $f \in \{p(x)e^{\lambda x} : p \in \mathbb{C}_k[x]\}$ y tenemos que $gen\{e^{\lambda x}, xe^{\lambda x}, \cdots, x^k e^{\lambda x}\} \subseteq \{p(x)e^{\lambda x} : p \in \mathbb{C}_k[x]\}$.

Recíprocamente, si $f \in \{p(x)e^{\lambda x} : p \in \mathbb{C}_k[x]\}$, entonces $f(x) = p(x)e^{\lambda x}$, para cierto $p \in \mathbb{C}_k[x]$. Como $p \in \mathbb{C}_k[x]$, existen $a_0, a_1, \cdots, a_k \in \mathbb{C}$ tales que

$$p(x) = a_0 + a_1 x + \cdots + \alpha_k x^k.$$

Por lo tanto, $f(x) = (a_0 + a_1 x + \cdots + \alpha_k x^k)e^{\lambda x} = a_0 e^{\lambda x} + a_1 x e^{\lambda x} + \cdots + a_k x^k e^{\lambda x}$, entonces $f \in gen\{e^{\lambda x}, xe^{\lambda x}, \cdots, x^k e^{\lambda x}\}$ y tenemos la otra inclusión.

Conclusión, con la HI y el item c) probamos que $\{e^{\lambda x}, xe^{\lambda x}, \cdots, x^{k-1}e^{\lambda x}, x^k e^{\lambda x}\}$ es una base de $\mathrm{Nu}((D - \lambda I)^{k+1})$. Entonces por inducción, se sigue que para todo $k \in \mathbb{N}$, el conjunto $\{x^i e^{\lambda x}\}$ con $i = \{0, 1, 2 \cdots, k-1\}$, es una base de $\mathrm{Nu}((D - \lambda I)^k)$.

e) : Para probar que y_p es una solución particular del sistema $(D - \lambda I)^k[y] = g$, basta ver que $(D - \lambda I)^k[y_p] = g$.

Observar que como $f^k(x) \in C^\infty(\mathbb{R}, \mathbb{C})$ entonces $f \in C^\infty(\mathbb{R}, \mathbb{C})$ y por lo tanto podemos usar el item a) de la siguiente manera: sea $x \in \mathbb{R}$ entonces,

$$(D - \lambda I)^k[y_p(x)] = (D - \lambda I)^k[f(x)e^{\lambda x}] = f^k(x)e^{\lambda x} = g(x)e^{-\lambda x}e^{\lambda x} = g(x),$$

donde reemplazamos $f^k(x) = g(x)e^{-\lambda x}$.

Por lo tanto, probamos que $(D - \lambda I)^k[y_p] = g$ y entonces y_p es una solución particular del sistema $(D - \lambda I)^k[y] = g$. ■

En el próximo ejercicio, vamos a aplicar la Proposición 2.7.1 y lo que acabamos de demostrar en el **Ejercicio 2.27**.

Ejercicio 2.28 e), f). Resolver las siguientes ecuaciones diferenciales:

- $y'' - 2y' + y = (3 + 5x)e^{2x}$,
- $(D - I)^3[y] = (3 + 5x)e^{2x}$.

Dem. $e)$: Observar que si D es el operador derivación, entonces

$$L[y] := y'' - 2y' + y = (D^2 - 2D + I)[y].$$

Por lo tanto, si p es el polinomio característico de L tenemos que $p(x) = x^2 - 2x + 1$. Para aplicar la Proposición 2.7.1, observar que las raíces de p son $\lambda_1 = \lambda_2 = 1$. Entonces $p(x) = (x-1)(x-1)$ y por esa misma proposición, tenemos que

$$L = (D - \lambda_1 I)(D - \lambda_2 I) = (D - I)(D - I) = (D - I)^2.$$

Por lo tanto, vamos resolver el sistema

$$(D - I)^2[y] = g,$$

donde $g(x) = (3 + 5x)e^{2x}$. Como ya vimos, todas las soluciones del sistema a resolver y_s tienen la forma

$$y_s = y_h + y_p,$$

donde $y_h \in \mathrm{Nu}((D-I)^2)$ e y_p es una solución particular.

Por el **Ejercicio 2.27 c)** tenemos que una base de $\mathrm{Nu}((D-I)^2)$, es

$$B_{\mathrm{Nu}((D-I)^2)} = \{e^x, xe^x\}$$

y, por el **Ejercicio 2.27 d)**, tenemos que una posible solución particular es $y_p = f(x)e^x$, donde $f^2(x) = f''(x) = g(x)e^{-x} = (3 + 5x)e^{2x}e^{-x} = (3 + 5x)e^x$. Entonces, integrando, nos queda que

$$f'(x) = \int (3 + 5x)e^x = 3e^x + 5e^x(x-1) = (-2 + 5x)e^x,$$

observar que no agregamos la constante de integración porque basta encontrar una solución particular. Integrando nuevamente, tenemos que

$$f(x) = \int (-2 + 5x)e^x = -2e^x + 5e^x(x-1) = (-7 + 5x)e^x.$$

Por lo tanto

$$y_p = f(x)e^x = (-7 + 5x)e^x e^x = (-7 + 5x)e^{2x}$$

y las soluciones de la ecuación diferencial son:

$$y_s = \alpha e^x + \beta x e^x + (-7 + 5x)e^{2x},$$

con $\alpha, \beta \in \mathbb{C}$.

$f)$: Vamos resolver el sistema

$$(D - I)^3[y] = g,$$

donde $g(x) = (3 + 5x)e^{2x}$. Como ya vimos, todas las soluciones del sistema a resolver y_s tienen la forma

$$y_s = y_h + y_p,$$

donde $y_h \in \mathrm{Nu}((D-I)^3)$ e y_p es una solución particular.

Por el **Ejercicio 2.27 c)** tenemos que una base de $\mathrm{Nu}((D-I)^3)$, es

$$B_{\mathrm{Nu}((D-I)^3)} = \{e^x, xe^x, x^2e^x\}$$

y, por el **Ejercicio 2.27 d)**, tenemos que una posible solución particular es $y_p = f(x)e^x$, donde

$$f^3(x) = f'''(x) = g(x)e^{-x} = (3 + 5x)e^{2x}e^{-x} = (3 + 5x)e^x.$$

Entonces, usando lo que hicimos en el item anterior, tenemos que $f'(x) = (-7 + 5x)e^x$. Integrando nuevamente, nos queda que $f(x) = (-12 + 5x)e^x$. Por lo tanto $y_p = f(x)e^x = (-12 + 5x)e^x e^x = (-12 + 5x)e^{2x}$ y las soluciones de la ecuación diferencial son:

$$y_s = \alpha e^x + \beta x e^x + \gamma x^2 e^x + (-12 + 5x)e^{2x},$$

con $\alpha, \beta, \gamma \in \mathbb{C}$. ■

Ejercicio 2.29. Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y sean L y A dos transformaciones lineales de $\mathbb{V}$ en $\mathbb{V}$ que satisfacen las siguientes propiedades

i) $L \circ A = A \circ L$, es decir L y A conmutan.

ii) $\mathrm{Nu}(A \circ L)$ es de dimensión finita.

Verificar que:

a) $\mathrm{Nu}(L) + \mathrm{Nu}(A) \subseteq \mathrm{Nu}(A \circ L)$.

b) Si $w \in \mathrm{Nu}(A)$, entonces toda solución de la ecuación $L(v) = w$, pertence a $\mathrm{Nu}(A \circ L)$.

c) Si $w \in \mathrm{Nu}(A) \cap \mathrm{Im}(L)$ y $\mathcal{S}$ es un subespacio tal que $\mathrm{Nu}(L) \oplus \mathcal{S} = \mathrm{Nu}(A \circ L)$, entonces existe un único $v \in \mathcal{S}$ tal que $L(v) = w$.

d) Si además $\mathrm{Nu}(L) \cap \mathrm{Nu}(A) = \{0_{\mathbb{V}}\}$, entonces

- $\mathrm{Nu}(A \circ L) = \mathrm{Nu}(A) \oplus \mathrm{Nu}(L)$,
- para cada $w \in \mathrm{Nu}(A) \cap \mathrm{Im}(L)$ existe un único $v \in \mathrm{Nu}(A)$ tal que $L(v) = w$.

Dem. $a)$: Sea $x \in \mathrm{Nu}(L) + \mathrm{Nu}(A)$, entonces

$$x = x_1 + x_2,$$

con $x_1 \in \mathrm{Nu}(L)$ y $x_2 \in \mathrm{Nu}(A)$. Entonces

$$(A \circ L)(x) = A(L(x)) = A(L(x_1 + x_2)) = A(L(x_1) + L(x_2)) = A(L(x_1)) + A(L(x_2))$$

$$= A(L(x_1)) + L(A(x_2)) = A(0_{\mathbb{V}}) + L(0_{\mathbb{V}}) = 0_{\mathbb{V}} + 0_{\mathbb{V}} = 0_{\mathbb{V}},$$

donde usamos que $x_1 \in \mathrm{Nu}(L)$, $x_2 \in \mathrm{Nu}(A)$, que A y L conmutan y que $A, L \in \mathcal{L}(\mathbb{V})$. Por lo tanto, $x \in \mathrm{Nu}(A \circ L)$ y se sigue que $\mathrm{Nu}(L) + \mathrm{Nu}(A) \subseteq \mathrm{Nu}(A \circ L)$.

$b)$: Sea $v \in \mathbb{V}$, tal que $L(v) = w$, es decir v es una solución de la ecuación $L(v) = w$. Entonces

$$(A \circ L)(v) = A(L(v)) = A(w) = 0_{\mathbb{V}},$$

donde usamos que $w \in \mathrm{Nu}(A)$. Entonces $v \in \mathrm{Nu}(A \circ L)$.

$c)$: Como $w \in \mathrm{Nu}(A) \cap \mathrm{Im}(L)$, entonces $w \in \mathrm{Im}(L)$ y entonces existe un vector $x \in \mathbb{V}$ tal que $L(x) = w$.

En el item b) vimos que, como $w \in \mathrm{Nu}(A)$ y vale que $L(x) = w$, entonces $x \in \mathrm{Nu}(A \circ L)$. Como $\mathrm{Nu}(L) \oplus \mathcal{S} = \mathrm{Nu}(A \circ L)$ y $x \in \mathrm{Nu}(A \circ L)$, existen únicos $z \in \mathrm{Nu}(L)$ y $v \in \mathcal{S}$, tales que $x = z + v$. Entonces

$$w = L(x) = L(z + v) = L(z) + L(v) = 0_{\mathbb{V}} + L(v) = L(v),$$

porque $z \in \mathrm{Nu}(L)$.

Observar que v es único, porque si existe $v' \in \mathcal{S}$ tal que $L(v') = w$. Entonces, como

$$L(v) = w = L(v'),$$

tenemos que

$$0_{\mathbb{V}} = L(v) - L(v') = L(v - v'),$$

entonces $v - v' \in \mathrm{Nu}(L)$ y como $\mathcal{S}$ es un subespacio y $v, v' \in \mathcal{S}$, también tenemos que $v - v' \in \mathcal{S}$. Es decir $v - v' \in \mathrm{Nu}(L) \cap \mathcal{S} = \{0\}$. Entonces $v = v'$ y v es único.

$d)$: Este item no es nada fácil. La guía nos recomienda leer la demostración del teorema de la dimensión de transfomaciones lineales definidas en dominios de dimensión finita como ayuda, ver Teorema 2.1.1.

Dicho esto, vamos a demostrar

$$\mathrm{Nu}(A \circ L) = \mathrm{Nu}(A) \oplus \mathrm{Nu}(L).$$

Usando el item a), ya sabemos que $\mathrm{Nu}(A) \oplus \mathrm{Nu}(L) \subseteq \mathrm{Nu}(A \circ L)$. Sólo nos resta probar la otra inclusión. Para eso vamos a usar fuertemente que $\mathrm{Nu}(A \circ L)$ es de dimensión finita.

Observemos que como siempre vale que $\mathrm{Nu}(A) \subseteq \mathrm{Nu}(A) \oplus \mathrm{Nu}(L)$ (¿por qué?) y ya probamos que $\mathrm{Nu}(A) \oplus \mathrm{Nu}(L) \subseteq \mathrm{Nu}(A \circ L)$, tenemos que $\mathrm{Nu}(A) \subseteq \mathrm{Nu}(A \circ L)$. Entonces

$$\dim(\mathrm{Nu}(A)) \leq \dim(\mathrm{Nu}(A \circ L)),$$

por lo tanto $\mathrm{Nu}(A)$ también es de dimensión finita. Supongamos que $\dim(\mathrm{Nu}(A)) = r$, donde r es 0 ó un número natural. Vamos a dividir la prueba en dos casos, si $r = 0$ ó $r > 0$.

Caso $r = 0$

Si $r = 0$, es decir A es monomorfismo, en este caso $\mathrm{Nu}(L) = \mathrm{Nu}(A \circ L)$. Es fácil ver que $\mathrm{Nu}(L) \subseteq \mathrm{Nu}(A \circ L)$. Recíprocamente, si $x \in \mathrm{Nu}(A \circ L)$ entonces $A(L(x)) = 0_{\mathbb{V}}$, entonces $L(x) \in \mathrm{Nu}(A) = \{0_{\mathbb{V}}\}$, entonces $L(x) = 0_{\mathbb{V}}$ y $x \in \mathrm{Nu}(L)$. Por lo tanto, $\mathrm{Nu}(L) = \mathrm{Nu}(A \circ L)$ y tenemos que

$$\mathrm{Nu}(A \circ L) = \mathrm{Nu}(L) = \{0_{\mathbb{V}}\} \oplus \mathrm{Nu}(L) = \mathrm{Nu}(A) \oplus \mathrm{Nu}(L).$$

Entonces, para el caso $r = 0$, lo que queremos demostrar se cumple de manera trivial.

Caso $r > 0$

Supongamos que $r > 0$ (y finito). Sea $\{v_1, v_2, \cdots, v_r\}$ una base de $\mathrm{Nu}(A)$. La clave del ejercicio es, usando un poco de imaginación y la prueba del teorema de la dimensión, probar que

$$\{L(v_1), L(v_2), \cdots, L(v_r)\}$$

también es una base de $\mathrm{Nu}(A)$.

Veamos eso. Por un lado tenemos que, para $i = 1, 2, \cdots, r$

$$L(v_i) \in \mathrm{Nu}(A),$$

pues $A(L(v_i)) = (A \circ L)(v_i) = (L \circ A)(v_i) = L(A(v_i)) = L(0_{\mathbb{V}}) = 0_{\mathbb{V}}$, donde usamos que, como $\{v_1, v_2, \cdots, v_r\}$ es una base de $\mathrm{Nu}(A)$, entonces $v_1, v_2, \cdots, v_r \in \mathrm{Nu}(A)$.

Veamos que el conjunto $\{L(v_1), L(v_2), \cdots, L(v_r)\}$ es LI. Sean $a_1, a_2, \cdots, a_r \in \mathbb{K}$ tales que

$$a_1 L(v_1) + a_2 L(v_2) + \cdots + a_r L(v_r) = 0_{\mathbb{V}}.$$

Entonces, usando que L es TL, tenemos que

$$L(a_1 v_1 + a_2 v_2 + \cdots + a_r v_r) = 0_{\mathbb{V}},$$

entonces $a_1 v_1 + a_2 v_2 + \cdots + a_r v_r \in \mathrm{Nu}(L)$. Además, como $a_1 v_1 + a_2 v_2 + \cdots + a_r v_r$ es una CL de elementos de $\mathrm{Nu}(A)$, tenemos que también $a_1 v_1 + a_2 v_2 + \cdots + a_r v_r \in \mathrm{Nu}(A)$. Por lo tanto

$$a_1 v_1 + a_2 v_2 + \cdots + a_r v_r \in \mathrm{Nu}(L) \cap \mathrm{Nu}(A) = \{0\},$$

donde usamos la hipótesis. Entonces

$$a_1 v_1 + a_2 v_2 + \cdots + a_r v_r = 0_{\mathbb{V}},$$

y tenemos un CL de los vectores $v_1, v_2, \cdots, v_r$ igualada al elemento neutro de $\mathbb{V}$. Entonces como $\{v_1, v_2, \cdots, v_r\}$ es LI (porque es una base) se sigue que $a_1 = a_2 = \cdots = a_r = 0$. Entonces el conjunto $\{L(v_1), L(v_2), \cdots, L(v_r)\}$ es LI.

Probamos que el conjunto $\{L(v_1), L(v_2), \cdots, L(v_r)\}$ tiene r vectores LI que pertenecen a $\mathrm{Nu}(A)$, entonces como $\dim(\mathrm{Nu}(A)) = r$, tenemos que $\{L(v_1), L(v_2), \cdots, L(v_r)\}$ es una base de $\mathrm{Nu}(A)$.

Ahora sí, estamos listos para probar la otra inclusión que nos quedaba, es decir, veamos que $\mathrm{Nu}(A \circ L) \subseteq \mathrm{Nu}(A) \oplus \mathrm{Nu}(L)$.

Sea $x \in \text{Nu}(A \circ L)$ entonces $(A \circ L)(x) = A(L(x)) = 0_{\mathbb{V}}$. Entonces, $L(x) \in \text{Nu}(A)$. Como ya vimos que $\{L(v_1), L(v_2), \cdots, L(v_r)\}$ es una base de $\text{Nu}(A)$, existen $b_1, b_2, \cdots, b_r \in \mathbb{K}$ tales que

$$L(x) = b_1 L(v_1) + b_2 L(v_2) + \cdots + b_r L(v_r).$$

Entonces, como L es una TL, se sigue que

$$L(x) = L(b_1 v_1 + b_2 v_2 + \cdots + b_r v_r).$$

Pasando de miembro y usando nuevamente que L es una TL, nos queda que

$$0_{\mathbb{V}} = L(x) - L(b_1 v_1 + b_2 v_2 + \cdots + b_r v_r) = L(x - b_1 v_1 - b_2 v_2 - \cdots - b_r v_r),$$

entonces, $x - b_1 v_1 - b_2 v_2 - \cdots - b_r v_r \in \text{Nu}(L)$.

Observar que

$$x = [b_1 v_1 + b_2 v_2 + \cdots + b_r v_r] + [x - b_1 v_1 - b_2 v_2 - \cdots - b_r v_r].$$

Donde $x_1 := b_1 v_1 + b_2 v_2 + \cdots + b_r v_r \in \text{Nu}(A)$, pues x_1 es una CL de vectores de $\text{Nu}(A)$ y (acabamos de ver que) $x_2 := x - b_1 v_1 - b_2 v_2 - \cdots - b_r v_r \in \text{Nu}(L)$. Entonces tenemos que

$$x = x_1 + x_2,$$

con $x_1 \in \text{Nu}(A)$ y $x_2 \in \text{Nu}(L)$. Por lo tanto $x \in \text{Nu}(A) \oplus \text{Nu}(L)$, y tenemos que $\text{Nu}(A \circ L) \subseteq \text{Nu}(A) \oplus \text{Nu}(L)$. Como ya habíamos probado la otra inclusión, concluimos que $\text{Nu}(A \circ L) = \text{Nu}(A) \oplus \text{Nu}(L)$, como queríamos ver.

Finalmente, como probamos que $\text{Nu}(A \circ L) = \text{Nu}(L) \oplus \text{Nu}(A)$, usando el item c) es inmediato ver que para cada $w \in \text{Nu}(A) \cap Im(L)$ existe un único $v \in \text{Nu}(A)$ tal que $L(v) = w$. La prueba es idéntica al item c) tomando $\mathcal{S} = \text{Nu}(A)$. ■

Nota: Usando la hipótesis $\text{Nu}(A) \cap \text{Nu}(L) = \{0\}$, observar que para demostrar el ítem d) sólo usamos que $\dim(\text{Nu}(A))$ es finita y bastó con eso. Es decir si $\dim(\text{Nu}(L))$ no es finita pero $\dim(\text{Nu}(A))$ sí es finita (por lo que $\dim(\text{Nu}(A \circ L))$ tampoco es finita) sigue valiendo que $\text{Nu}(A \circ L) = \text{Nu}(A) \oplus \text{Nu}(L)$. Con argumentos muy parecidos a los que usamos para demostrar el ítem d), si pedimos que $\dim(\text{Nu}(L))$ es finita pero $\dim(\text{Nu}(A))$ es no finita también sigue valiendo que $\text{Nu}(A \circ L) = \text{Nu}(A) \oplus \text{Nu}(L)$. El problema ocurre cuando $\dim(\text{Nu}(A))$ y $\dim(\text{Nu}(L))$ son ambas no finitas (y por ende también $\dim(\text{Nu}(A \circ L))$ es no finita), en ese caso no tiene porque ser cierto que $\text{Nu}(A \circ L) = \text{Nu}(A) \oplus \text{Nu}(L)$.

Antes de pasar al próximo ejercicio, repasemos una propiedad que vimos en la prueba de la Proposición 2.7.1, si $\alpha, \beta \in \mathbb{C}$ y D es el operador diferencial, entonces

$$(D - \alpha I)^k \circ (D - \beta I)^n = (D - \beta I)^n \circ (D - \alpha I)^k, \tag{2.5}$$

para cualquier $k, n \in \mathbb{N}$. Es decir, para cualquier $k, n \in \mathbb{N}$, $(D - \alpha I)^k$ y $(D - \beta I)^n$ conmutan. La prueba de esta propiedad es muy simple y usa el princio de inducción, se sugiere hacerla de ejercicio.

Vamos a usar el **Ejercicio 2.29** para resolver el próximo ejercicio.

Ejercicio 2.30. Se considera el operador diferencial $L : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ definido por

$$L := (D - 2I)(D - 4I)(D + 3I)^2,$$

y la ecuación diferencial $L[y] = p$, donde $p(x) = 5x^3 e^{-3x}$.

a) Hallar una base $\mathcal{B}_L$ de $\text{Nu}(L)$.

b) Comprobar que el operador $A = (D + 3I)^4$, es un *aniquilador* de p: $A[p] = 0$.

c) Hallar una base $\mathcal{B}_{AL}$ de $\text{Nu}(A \circ L)$ que contenga a la base $\mathcal{B}_L$.

d) Comprobar que existe una solución particular y_p de la ecuación $L[y] = p$ perteneciente al subespacio $gen(\mathcal{B}_{AL} \setminus \mathcal{B}_L)$.

e) Hallar la solución general de la ecuación diferencial $L[y] = p$.

Dem. *a)* : Para resolver este ejercicio vamos a usar el **Ejercicio 2.27 b), c)** y el **Ejercicio 2.29**. Llamemos $R := (D - 2I)$, $S := (D - 4I)$ y $T := (D + 3I)^2$. Entonces,

$$L = R \circ S \circ T.$$

Por el **Ejercicio 2.27 b), c)** tenemos que $\mathcal{B}_R = \{e^{2x}\}$, $\mathcal{B}_S = \{e^{4x}\}$ y $\mathcal{B}_T = \{e^{-3x}, xe^{-3x}\}$ son bases de $\text{Nu}(R), \text{Nu}(S)$ y $\text{Nu}(T)$, respectivamente.

Para encontrar una base de $\text{Nu}(L)$, vamos a usar el **Ejercicio 2.29 c)** dos veces.

Primero, usando la propiedad (2.5), es fácil ver que $S \circ T = T \circ S$, por otra parte, es claro que

$$\text{Nu}(S) \cap \text{Nu}(T) = gen\{e^{4x}\} \cap gen\{e^{-3x}, xe^{-3x}\} = \{0\}.$$

Además, $\text{Nu}(S)$ y $\text{Nu}(T)$ son de dimensión finita. Entonces podemos usar el **Ejercicio 2.29 d)** y por lo tanto, tenemos que

$$\text{Nu}(S \circ T) = \text{Nu}(S) \oplus \text{Nu}(T) = gen\{e^{4x}, e^{-3x}, xe^{-3x}\}.$$

Segundo, usando nuevamente la propiedad (2.5), es fácil ver que $R \circ (S \circ T) = (S \circ T) \circ R$, por otra parte, es claro que

$$\text{Nu}(R) \cap \text{Nu}(S \circ T) = gen\{e^{2x}\} \cap gen\{e^{4x}, e^{-3x}, xe^{-3x}\} = \{0\}.$$

Además, $\text{Nu}(R)$ y $\text{Nu}(S \circ T)$ son de dimensión finita. Entonces podemos usar nuevamente el **Ejercicio 2.29d)** y por lo tanto, tenemos que

$$\text{Nu}(L) = \text{Nu}(R \circ S \circ T) = \text{Nu}(R) \oplus \text{Nu}(S \circ T) = gen\{e^{2x}\} \oplus gen\{e^{4x}, e^{-3x}, xe^{-3x}\} =$$

$$= gen\{e^{2x}, e^{4x}, e^{-3x}, xe^{-3x}\}.$$

Por lo tanto,

$$\mathcal{B}_L = \{e^{2x}, e^{4x}, e^{-3x}, xe^{-3x}\}.$$

b) : Llamemos $f(x) := 5x^3$, entonces claramente $f \in C^\infty(\mathbb{R})$ y $p(x) = f(x)e^{-3x}$. Sea $x \in \mathbb{R}$, entonces,

$$A[p(x)] = (D + 3I)^4[p(x)] = (D - (-3)I)^4[5x^3 e^{-3x}] = (D - (-3)I)^4[f(x)e^{-3x}] = f^4(x)e^{-3x} = 0,$$

donde usamos el **Ejercicio 2.27 a)** y el hecho de que $f^4(x) = (5x^3)'''' = (15x^2)''' = (30x)'' = (30)' = 0$. Por lo tanto $A[p] = \mathbf{0}$, donde $\mathbf{0}$ denota la transformación lineal nula.

c) : Observar que

$$A \circ L = (D + 3I)^4(D - 2I)(D - 4I)(D + 3I)^2 = (D - 2I)(D - 4I)(D + 3I)^6,$$

donde usamos que por la propiedad (2.5), todos los operadores involucrados conmutan.

Si llamamos $\tilde{T} := (D + 3I)^6$, con la misma notación que el item a), tenemos que

$$A \circ L = R \circ S \circ \tilde{T}$$

y, por el **Ejercicio 2.27 c)** tenemos que

$$\text{Nu}(\tilde{T}) = gen\{e^{-3x}, xe^{-3x}, x^2e^{-3x}, x^3e^{-3x}, x^4e^{-3x}, x^5e^{-3x}\}.$$

Es claro que $\text{Nu}(S) \circ \text{Nu}(\tilde{T}) = \{0\}$ (verificarlo) y ambos subespacios son de dimensión finita.

Por lo tanto, de la misma manera que lo hicimos en el item a), aplicando el **Ejercicio 2.29 d)**, se sigue que

$$\mathrm{Nu}(A \circ L) = \mathrm{Nu}(R \circ S \circ \tilde{T}) = \mathrm{Nu}(R) \oplus \mathrm{Nu}(S \circ \tilde{T}) = \mathrm{Nu}(R) \oplus [\mathrm{Nu}(S) \oplus \mathrm{Nu}(\tilde{T})] =$$

$$= gen\{e^{2x}\} \oplus gen\{e^{4x}, e^{-3x}, xe^{-3x}, x^2e^{-3x}, x^3e^{-3x}, x^4e^{-3x}, x^5e^{-3x}\} =$$

$$= gen\{e^{2x}, e^{4x}, e^{-3x}, xe^{-3x}, x^2e^{-3x}, x^3e^{-3x}, x^4e^{-3x}, x^5e^{-3x}\}.$$

La base $\mathcal{B}_{AL} = \{e^{2x}, e^{4x}, e^{-3x}, xe^{-3x}, x^2e^{-3x}, x^3e^{-3x}, x^4e^{-3x}, x^5e^{-3x}\}$ de $\mathrm{Nu}(A \circ L)$, contiene a la base $\mathcal{B}_L$.

d) : Primero, observemos que

$$\mathcal{B}_{AL} \setminus \mathcal{B}_L = \{y \in \mathcal{B}_{AL} : y \notin \mathcal{B}_L\} = \{x^2e^{-3x}, x^3e^{-3x}, x^4e^{-3x}, x^5e^{-3x}\}.$$

Entonces,

$$gen(\mathcal{B}_{AL} \setminus \mathcal{B}_L) = gen\{x^2e^{-3x}, x^3e^{-3x}, x^4e^{-3x}, x^5e^{-3x}\}.$$

Vamos a comprobar que existe una solución particular y_p perteneciente al subespacio $gen(\mathcal{B}_{AL} \setminus \mathcal{B}_L)$. Para eso, sea $y \in gen(\mathcal{B}_{AL} \setminus \mathcal{B}_L)$, entonces

$$y(x) = ax^2e^{-3x} + bx^3e^{-3x} + cx^4e^{-3x} + dx^5e^{-3x} = (ax^2 + bx^3 + cx^4 + dx^5)e^{-3x},$$

para ciertos $a, b, c, d \in \mathbb{C}$. Entonces, con la notación del item a) y usando el **Ejercicio 2.27 a)** con $f(x) = ax^2 + bx^3 + cx^4 + dx^5$, tenemos que $y(x) = f(x)e^{-3x}$ y

$$T[y(x)] = (D + 3I)^2[y(x)] = (D - (-3I))^2[f(x)e^{-3x}] = f''(x)e^{-3x} =$$

$$= (ax^2 + bx^3 + cx^4 + dx^5)''e^{-3x} = (2ax + 3bx^2 + 4cx^3 + 5dx^4)'e^{-3x} =$$

$$= (2a + 6bx + 12cx^2 + 20dx^3)e^{-3x}.$$

Ahora, llamemos $g(x) := 2a + 6bx + 12cx^2 + 20dx^3$, entonces $T[y(x)] = g(x)e^{-3x}$. Entonces

$$(S \circ T)[y(x)] = S(T[y(x)]) = S[g(x)e^{-3x}] =$$

$$= \frac{d[g(x)e^{-3x}]}{dx} - 4g(x)e^{-3x} = g'(x)e^{-3x} - 3g(x)e^{-3x} - 4g(x)e^{-3x} =$$

$$= (-7g(x) + g'(x))e^{-3x} = -7g(x)e^{-3x} + (6b + 24cx + 60dx^2)e^{-3x} =$$

$$= [-14a + 6b + (-42b + 24c)x + (-84c + 60d)x^2 - 140dx^3]e^{-3x}.$$

Finalmente, llamemos $h(x) := [-14a + 6b + (-42b + 24c)x + (-84c + 60d)x^2 - 140dx^3]$, entonces $(S \circ T)[y(x)] = h(x)e^{-3x}$. Entonces

$$L[y(x)] = (R \circ S \circ T)[y(x)] = R[h(x)e^{-3x}] =$$

$$= \frac{d(h(x)e^{-3x})}{dx} - 2h(x)e^{-3x} = h'(x)e^{-3x} - 3h(x)e^{-3x} - 2h(x)e^{-3x} =$$

$$= [-5h(x) + h'(x)]e^{-3x} =$$

$$= [70a - 72b + 24c + (210b - 288c + 120d)x + (420c - 720d)x^2 + 700x^3]e^{-3x}.$$

Como, para todo $x \in \mathbb{R}$, tenemos que

$$L[y(x)] = [70a - 72b + 24c + (210b - 288c + 120d)x + (420c - 720d)x^2 + 700dx^3]e^{-3x} = 5x^3e^{-3x},$$

simplificando el término e^{-3x} (que siempre es positivo), se sigue que

$$70a - 72b + 24c + (210b - 288c + 120d)x + (420c - 720d)x^2 + 700dx^3 = 5x^3.$$

Entonces

$$700d = 5,\ 420c - 720d = 0,\ 210b - 288c + 120d = 0,\ 70a - 72b + 24c = 0.$$

Entonces, $d = \frac{5}{700} = \frac{1}{140}$; $c = \frac{72}{42}d = \frac{72}{42\times 140} = \frac{3}{245}$; $b = \frac{288c-120d}{210} = \frac{107}{1225}$; $a = \frac{72b-24c}{70} = \frac{2872}{42875}$. Por lo tanto

$$y_p = (\frac{2872}{42875}x^2 + \frac{107}{1225}x^3 + \frac{3}{245}x^4 + \frac{1}{140}x^5)e^{-3x},$$

es una solución particular perteneciente al subespacio $gen(\mathcal{B}_{AL} \setminus \mathcal{B}_L)$. Se recomienda revisar bien las cuentas porque es altamente probable que alguna cuenta este mal.

$e)$: La solución general y_s tiene la forma $y_s = y_h + y_p$, donde $y_h \in \mathrm{Nu}(L)$ e y_p es una solución particular. Usando los item a) y d) tenemos que,

$$y_s = \alpha e^{2x} + \beta e^{4x} + \gamma e^{-3x} + \delta x e^{-3x} + (\frac{2872}{42875}x^2 + \frac{107}{1225}x^3 + \frac{3}{245}x^4 + \frac{1}{140}x^5)e^{-3x},$$

con $\alpha, \beta, \gamma, \delta \in \mathbb{C}$. ■

Observaciones y repaso de algunas propiedades
Antes de resolver el siguiente ejercicio, algunas observaciones:

- Con lo que vimos en el **Ejercicio 2.30**, es fácil ver que si tenemos el operador

  $$L := (D - \lambda_1)^{k_1}(D - \lambda_2)^{k_2} \cdots (D - \lambda_n)^{k_n}$$

  donde $k_1, \cdots, k_n \in \mathbb{N}$ y $\lambda_1, \cdots, \lambda_n \in \mathbb{C}$ todos distintos entre sí, entonces:

  $$\begin{aligned}\mathrm{Nu}(L) &= \{y \in C^\infty(\mathbb{R}, \mathbb{C}) : L[y] = 0\}\\ &= \{y \in C^\infty(\mathbb{R}, \mathbb{C}) : (D-\lambda_1)^{k_1}[y] = 0\} \oplus \cdots \oplus \{y \in C^\infty(\mathbb{R}, \mathbb{C}) : (D-\lambda_n)^{k_n}[y] = 0\}\\ &= \mathrm{Nu}((D-\lambda_1)^{k_1}) \oplus \cdots \oplus \mathrm{Nu}((D-\lambda_n)^{k_n}).\end{aligned} \tag{2.6}$$

  Es decir, la solución del sistema homogéneo $L[y] = 0$ se puede obtener factorizando el operador L como $L = (D-\lambda_1)^{k_1}(D-\lambda_2)^{k_2} \cdots (D-\lambda_n)^{k_n}$ donde $k_1, \cdots, k_n \in \mathbb{N}$ y $\lambda_1, \cdots, \lambda_n \in \mathbb{C}$ todas distintas entre sí (es decir las raíces del polinomio característico asociado a L se cuentan una vez). Luego, la solución del sistema homogéneo $L[y] = 0$, se obtiene sumando las soluciones de cada sistema homogéneo $(D - \lambda_i)^{k_i}[y] = 0$.

- Recordemos que si $z = a + ib \in \mathbb{C}$ (donde $a, b \in \mathbb{R}$) entonces

  $$e^z = e^{a+ib} = e^a(\cos(b) + i\sin(b)). \tag{2.7}$$

- Recordemos que en el **Ejercicio 1.3** se probó que, si estamos trabajando en un cuerpo complejo, dados $a, b \in \mathbb{R}$, se sigue que

  $$gen\{e^{ax}\cos(bx), e^{ax}\sin(bx)\} = gen\{e^{(a+ib)x}, e^{(a-ib)x}\}. \tag{2.8}$$

Vamos a resolver un ejemplo para aplicar estas propiedades.

Ejemplo 2.e. Obtener una base de soluciones del sistema

$$L[y] = y^{(6)} - 9y^{(5)} + 25y^{(4)} - 17y^{(3)} = 0.$$

El polinomio característico asociado a L es

$$p(x) = x^6 - 9x^5 + 25x^4 - 17x^3.$$

Las raíces de p son $\lambda_1 = 0$ (triple), $\lambda_2 = 4 + i$, $\lambda_3 = 4 - i$, $\lambda_4 = 1$. Entonces,

$$p(x) = x^3(x - (4 + i))(x - (4 - i))(x + 1).$$

Por la Proposición 2.7.1, L se factoriza como

$$L = D^3(D - (4 + i)I)(D - (4 - i)I)(D - I).$$

Por lo tanto, usando el **Ejercicio 2.27** tenemos que:

$$\mathrm{Nu}(D^3) = \mathrm{Nu}(D - 0I)^3 = gen\{1, x, x^2\},\ \mathrm{Nu}((D - I)) = gen\{e^x\},$$

$$\mathrm{Nu}((D - (4 + i)I)) = gen\{e^{(4+i)x}\},\ \mathrm{Nu}((D - (4 - i)I)) = gen\{e^{(4-i)x}\}.$$

Por la propiedad (2.8), tenemos que

$$gen\{e^{(4+i)x}, e^{(4-i)x}\} = gen\{e^{4x}\cos(x), e^{4x}\sin(x))\}.$$

Entonces, por la propiedad (2.6), se sigue que

$$\mathrm{Nu}(L) = gen\{1, x, x^2\} \oplus gen\{e^x\} \oplus gen\{e^{4x}\cos(x), e^{4x}\sin(x)\} =$$

$$= gen\{1, x, x^2, e^x, e^{4x}\cos(x), e^{4x}\sin(x)\}.$$

Por lo tanto, una base de soluciones del sistema es

$$\mathcal{B}_L = \{1, x, x^2, e^x, e^{4x}\cos(x), e^{4x}\sin(x)\}.$$

Ahora sí, con todas las propiedades que vimos vamos a poder resolver los ejercicios que restan.

Ejercicio 2.33 e). Construir una ecuación diferencial

$$L[y] = y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0,$$

con $a_0, a_1, \cdots, a_{n-1} \in \mathbb{R}$, del menor orden posible que tenga como soluciones:

- $y_1(t) = t,\ y_2(t) = \cos(3t),\ y_3(t) = e^{-t}$.

Dem. $e)$: Para resolver este item vamos a usar el **Ejercicio 2.27** y las propiedades (2.6), (2.7) y (2.8).

Observar que, si $y_1(t) = t = te^{0t}$ es una solución de $L[y] = 0$, es porque $\lambda_1 = 0$ es una raíz del polinomio característico asociado a L. Además, como t multiplica a e^{0t}, esa raíz es doble.

Además, si $y_2(t) = \cos(3t) = e^{0t}\cos(3t)$, es una solución de $L[y] = 0$, es porque $\lambda_1 = 3i$ es una raíz del polinomio característico asociado a L. Como $a_0, a_1, \cdots, a_{n-1} \in \mathbb{R}$, el polinomio característico asociado a L tendrá coeficientes reales, entonces si $\lambda_3 = 3i$ es una raíz, su conjugado $\lambda_4 = -3i$ también es una raíz.

Finalmente, si $y_3(t) = e^{-t}$ es una solución de $L[y] = 0$, es porque $\lambda_5 = -1$ es una raíz del polinomio característico asociado a L.

De esta manera, el polinomio mónico p de menor grado cuyas raíces son $\{0 \text{ (doble) }, 3i, -3i, -1\}$ es

$$p(x) = x^2(x - 3i)(x + 3i)(x + 1).$$

Entonces, por la Proposición 2.7.1, podemos factorizar al operador L como

$$L = (D - (0)I)^2(D - 3iI)(D + 3iI)(D + I) = D^5 + D^4 + 9D^3 + 9D^2.$$

Entonces y_1, y_2, y_3 son soluciones del sistema

$$L[y] = (D^5 + D^4 + 9D^3 + 9D^2)[y] = y^{(5)} + y^{(4)} + 9y^{(3)} + 9y'' = 0.$$

Es más, una base de soluciones de ese sistema es

$$\mathcal{B}_L = \{1, t, \cos(3t), \sin(3t), e^{-t}\}.$$

■

CAPÍTULO 3

Espacios Euclídeos

En esta unidad, comenzaremos con el estudio de $\mathbb{K}$-espacios vectoriales con producto interno. También, vamos a generalizar los conceptos geométricos de distancia, ortogonalidad y ángulo conocidas en $\mathbb{R}^2$ y en $\mathbb{R}^3$, a otros espacios vectoriales.

3.1. Producto interno y propiedades

Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial, $\mathbb{K} = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$. Un *producto interno* en $\mathbb{V}$ es una función $\langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{K}$, tal que:

i) Para cada $\lambda \in \mathbb{K}$ y para cada $x, y, z \in \mathbb{V}$

   1. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$,
   2. $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$.

ii) $\langle x, y \rangle = \overline{\langle y, x \rangle}$, para todo $x, y \in \mathbb{V}$.

iii) $\langle x, x \rangle > 0$ si $x \neq 0$.

Un $\mathbb{K}$-espacio vectorial con producto interno $\langle \cdot, \cdot \rangle$ se llama *espacio euclídeo* y lo denotaremos por $(\mathbb{V}, \langle \cdot, \cdot \rangle)$.

Observar que si $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ es un espacio euclídeo entonces:

- $\langle x, \lambda y \rangle = \overline{\langle \lambda y, x \rangle} = \overline{\lambda \langle y, x \rangle} = \overline{\lambda}\ \overline{\langle y, x \rangle} = \overline{\lambda}\ \langle x, y \rangle$,

  donde usamos i.2) y ii) de la definición de producto interno y que si $z, w \in \mathbb{C}$ entonces $\overline{zw} = \overline{z}\ \overline{w}$.

- $\langle x, x \rangle \geq 0$ para todo $x \in \mathbb{V}$ y $\langle x, x \rangle = 0$ si y sólo si $x = 0_{\mathbb{V}}$.

  En el **Ejercicio 1.I** vimos que $0_{\mathbb{V}} = 0 \cdot 0_{\mathbb{V}}$. Entonces $\langle 0_{\mathbb{V}}, 0_{\mathbb{V}} \rangle = \langle 0 \cdot\ 0_{\mathbb{V}}, 0_{\mathbb{V}} \rangle = 0 \langle 0_{\mathbb{V}}, 0_{\mathbb{V}} \rangle = 0$, donde usamos i.2) de la definición de producto interno. Entonces, por iii), se sigue que $\langle x, x \rangle \geq 0$ para todo $x \in \mathbb{V}$. Recíprocamente, como $\langle x, x \rangle \geq 0$ para todo $x \in \mathbb{V}$, si $\langle x, x \rangle = 0$ entonces $x = 0_{\mathbb{V}}$, donde usamos nuevamente iii).

Veamos algunos ejemplos de funciones que definen productos internos en los $\mathbb{K}$-espacios vectoriales correspondientes.

Ejercicio 3.A. Demostrar que las siguientes funciones definen productos internos en los $\mathbb{K}$-espacios vectoriales indicados:

a) $\langle\cdot,\cdot\rangle : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ definida por
$$\langle x, y\rangle = y^* A x,$$
donde $A = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix}$, define un producto interno en $\mathbb{C}^2$ (visto como $\mathbb{C}$-espacio vectorial).

b) $\langle\cdot,\cdot\rangle : C([0,1],\mathbb{R}) \times C([0,1],\mathbb{R}) \to \mathbb{R}$ definida por
$$\langle f, g\rangle = \int_0^1 f(t)\ g(t)dt,$$
define un producto interno en $C([0,1],\mathbb{R})$.

c) $\langle\cdot,\cdot\rangle : \mathbb{R}_2[x] \times \mathbb{R}_2[x] \to \mathbb{R}$ definida por
$$\langle p, q\rangle = p(0)q(0) + p(1)q(1) + p(2)q(2),$$
define un producto interno en $\mathbb{R}_2[x]$.

d) $\langle\cdot,\cdot\rangle : \mathbb{C}^2_{\mathbb{R}} \times \mathbb{C}^2_{\mathbb{R}} \to \mathbb{R}$ definida por
$$\langle x, y\rangle = \mathrm{Re}(y^* x)$$
define un producto interno en $\mathbb{C}^2$ como $\mathbb{R}$-espacio vectorial.

Dem. *a*): Veamos que la función $\langle\cdot,\cdot\rangle$ cumple con los axiomas del producto interno.

i) Para cada $\lambda \in \mathbb{C}$ y para cada $x, y, z \in \mathbb{C}^2$, tenemos que

1. $\langle x + y, z\rangle = z^* A(x + y) = z^* A x + z^* A y = \langle x, z\rangle + \langle y, z\rangle$.
2. $\langle \lambda x, y\rangle = y^* A(\lambda x) = \lambda y^* A x = \lambda \langle x, y\rangle$.

Observar que los items i.1) y i.2) valen para cualquier matriz A.

ii) Observar que $A^* = \overline{A^T} = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix} = A$. Entonces, para todo $x, y \in \mathbb{C}^2$,
$$\langle x, y\rangle = y^* A x = y^* A^* x = (x^* A y)^* = \overline{x^* A y} = \overline{\langle y, x\rangle}.$$
Donde usamos que como $x^* A y \in \mathbb{C}$ entonces $(x^* A y)^T = (x^* A y)$ y entonces $(x^* A y)^* = \overline{(x^* A y)^T} = \overline{(x^* A y)}$.

iii) Sea $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{C}^2$ entonces
$$\langle x, x\rangle = x^* A x = [\overline{x_1}\ \overline{x_2}] \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\overline{x_1}\ \overline{x_2}] \begin{bmatrix} 2x_1 + ix_2 \\ -ix_1 + 2x_2 \end{bmatrix} =$$
$$= 2\overline{x_1}x_1 + i\overline{x_1}x_2 - i\overline{x_2}x_1 + 2\overline{x_2}x_2 = [\overline{x_1}x_1 + i\overline{x_1}x_2 - i\overline{x_2}x_1 + \overline{x_2}x_2] + \overline{x_1}x_1 + \overline{x_2}x_2 =$$
$$(\overline{x_1 + ix_2})(x_1 + ix_2) + \overline{x_1}x_1 + \overline{x_2}x_2 = |x_1 + ix_2|^2 + |x_1|^2 + |x_2|^2.$$
Entonces, como $|x_1 + ix_2|^2 \geq 0, |x_1|^2 \geq 0$ y $|x_2|^2 \geq 0$, tenemos que $\langle x, x\rangle \geq 0$, para todo $x \in \mathbb{C}^2$. Además, si $\langle x, x\rangle = 0$ entonces $|x_1 + ix_2|^2 = 0, |x_1|^2 = 0$ y $|x_2|^2 = 0$, entonces $x_1 = x_2 = 0$. Por lo tanto, $\langle x, x\rangle > 0$ si y sólo si $x \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ y se cumple iii).

b): Veamos que la función $\langle\cdot,\cdot\rangle$ cumple con los axiomas del producto interno.

i) Para cada $\lambda \in \mathbb{R}$ y para cada $f, g, h \in C([0,1], \mathbb{R})$, tenemos que

1. $\langle f+g, h \rangle = \int_0^1 (f+g)(t)\ h(t)dt = \int_0^1 (f(t)+g(t))\ h(t)dt = \int_0^1 (f(t)h(t)+g(t)h(t))dt = \int_0^1 f(t)\ h(t)dt + \int_0^1 g(t)\ h(t)dt = \langle f, h \rangle + \langle g, h \rangle$.
2. $\langle \lambda f, g \rangle = \int_0^1 (\lambda f)(t)\ g(t)dt = \int_0^1 \lambda f(t)\ g(t)dt = \lambda \int_0^1 f(t)\ g(t)dt = \lambda \langle f, g \rangle$.

ii) Para cada $f, g \in C([0,1], \mathbb{R})$, tenemos que

$$\langle f, g \rangle = \int_0^1 f(t)\ g(t)dt = \int_0^1 g(t)\ f(t)dt = \langle g, f \rangle .$$

iii) Sea $f \in C([0,1], \mathbb{R})$, entonces, como $f(t)^2 \geq 0$ para todo $t \in [0,1]$, se sigue que

$$\langle f, f \rangle = \int_0^1 f(t)^2 dt \geq 0.$$

Ahora, supongamos que $\langle f, f \rangle = \int_0^1 f(t)^2 dt = 0$. Si $f \neq 0$, entonces existe $t_0 \in [0,1]$ tal que $f^2(t_0) > 0$. Como f es continua, f^2 también es continua. Por definición de continuidad, como $f^2(t_0) > 0$, existe un entorno alrededor de t_0 tal que f^2 es positiva en los puntos de ese entorno. Es decir, si llamamos $\epsilon := \frac{f(t_0)}{2}$, entonces existe $\delta > 0$ tal que si $|t - t_0| < \delta$ entonces $|f^2(t) - f^2(t_0)| < \epsilon$. Por lo tanto, si $|t - t_0| < \delta$, entonces

$$f^2(t) > f^2(t_0) - \epsilon = \frac{f^2(t_0)}{2}.$$

Entonces

$$\langle f, f \rangle = \int_0^1 f^2(t)dt = \int_0^{t_0-\delta} f^2(t)dt + \int_{t_0-\delta}^{t_0+\delta} f^2(t)dt + \int_{t_0+\delta}^1 f^2(t)dt \geq$$

$$\int_{t_0-\delta}^{t_0+\delta} f^2(t)dt \geq \int_{t_0-\delta}^{t_0+\delta} \frac{f^2(t_0)}{2} dt = \frac{f^2(t_0)}{2}(t_0 + \delta - (t_0 - \delta)) = f^2(t_0)\delta > 0.$$

Lo cual es absurdo, porque $\langle f, f \rangle = 0$. Entonces $\langle f, f \rangle = 0$ si y sólo si $f = 0$ o equivalentemente, $\langle f, f \rangle > 0$ si y sólo si $f \neq 0$ y se cumple iii).

c): Veamos que la función $\langle \cdot, \cdot \rangle$ cumple con los axiomas del producto interno.

i) Para cada $\lambda \in \mathbb{R}$ y para cada $p, q, r \in \mathbb{R}_2[x]$ tenemos que

1. $\langle p+q, r \rangle = (p+q)(0)r(0) + (p+q)(1)r(1) + (p+q)(2)r(2) = (p(0)+q(0))r(0) + (p(1)+q(1))r(1) + (p(2)+q(2))r(2) = [p(0)r(0) + p(1)r(1) + p(2)r(2)] + [q(0)r(0) + q(1)r(1) + q(2)r(2)] = \langle p, r \rangle + \langle q, r \rangle$.
2. $\langle \lambda p, q \rangle = (\lambda p)(0)r(0) + (\lambda p)(1)r(1) + (\lambda p)(2)r(2) = \lambda p(0)r(0) + \lambda p(1)r(1) + \lambda p(2)r(2) = \lambda \langle p, q \rangle$.

ii) Para cada $p, q \in \mathbb{R}_2[x]$ tenemos que

$$\langle p, q \rangle = p(0)q(0) + p(1)q(1) + p(2)q(2) = q(0)p(0) + q(1)p(1) + q(2)p(2) = \langle q, p \rangle .$$

iii) Sea $p \in \mathbb{R}_2[x]$ entonces, $\langle p, p \rangle = p(0)^2 + p(1)^2 + p(2)^2 \geq 0$. Por otra parte, si

$$\langle p, p \rangle = p(0)^2 + p(1)^2 + p(2)^2 = 0,$$

como $p(0)^2 \geq 0$, $p(1)^2 \geq 0$ y $p(2)^2 \geq 0$, tenemos que $p(0) = p(1) = p(2) = 0$. Es decir, el polinomio p tiene 3 raíces. Como $p \in \mathbb{R}_2[x]$, por el Teorema Fundamental del Álgebra, p tiene 2 raíces ó $p = 0$. Como acabamos de ver que p tiene 3 raíces, se sigue que $p = 0$.

Por lo tanto $\langle p, p \rangle > 0$ si y sólo si $p \neq 0$ y se cumple iii).

d): Veamos que la función $\langle \cdot, \cdot \rangle$ cumple con los axiomas del producto interno.

i) Para cada $\lambda \in \mathbb{R}$ y para cada $x, y \in \mathbb{C}^2$, tenemos que

1. $\langle x + y, z \rangle = \mathrm{Re}(z^*(x+y)) = \mathrm{Re}(z^*x + z^*y) = \mathrm{Re}(z^*x) + \mathrm{Re}(z^*y) = \langle x, z \rangle + \langle y, z \rangle$.
2. $\langle \lambda x, y \rangle = \mathrm{Re}(y^*(\lambda x)) = \mathrm{Re}(\lambda y^* x) = \lambda \, \mathrm{Re}(y^*x) = \lambda \langle x, y \rangle$.

ii) Para todo $x, y \in \mathbb{C}^2$,

$$\langle x, y \rangle = \mathrm{Re}(y^*x) = \mathrm{Re}((y^*x)^T) = \mathrm{Re}(x^T\overline{y}) = \mathrm{Re}(\overline{x^T\overline{y}}) = \mathrm{Re}(\overline{x^T}\overline{\overline{y}}) = \mathrm{Re}(x^*y) = \overline{\langle y, x \rangle}.$$

Donde usamos que como $y^*x \in \mathbb{C}$ entonces $(y^*x)^T = y^*x$ y además, $\mathrm{Re}(\overline{z}) = \mathrm{Re}(z)$, para todo $z \in \mathbb{C}$.

iii) Sea $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{C}^2$ entonces

$$\langle x, x \rangle = \mathrm{Re}(x^*x) = \mathrm{Re}(|x_1|^2 + |x_2|^2) = |x_1|^2 + |x_2|^2.$$

Entonces, como $|x_1|^2 \geq 0$ y $|x_2|^2 \geq 0$ tenemos que $\langle x, x \rangle \geq 0$ para todo $x \in \mathbb{C}^2$. Además, si $\langle x, x \rangle = 0$. Entonces $|x_1| = |x_2| = 0$, entonces $x_1 = x_2 = 0$. Por lo tanto, $\langle x, x \rangle > 0$ si y sólo si $x \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ y se cumple iii).

■

Ejercicio 3.B. Justificar por qué las siguientes funciones NO definen productos internos.

a) La función $\phi : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ definida por

$$\phi(x, y) = x^*Ay,$$

donde $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

b) La función $\phi : C([0,1], \mathbb{R}) \times C([0,1], \mathbb{R}) \to \mathbb{R}$ definida por

$$\phi(f, g) = \int_0^1 (t-2) \; f(t) \; g(t) dt.$$

c) La función $\phi : \mathbb{R}_2[x] \times \mathbb{R}_2[x] \to \mathbb{R}$ definida por

$$\phi(p, q) = p(0)q(0) + p(1)q(1).$$

Dem. Observar que las funciones ϕ de los items $a), b)$ y $c)$ cumplen los ítems i.1), i.2) y ii) de la definición de producto interno. El problema va a estar en el hecho de que ninguna cumple el item iii).

$a)$: Sea $x = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \in \mathbb{C}^2$, entonces

$$\phi(x, x) = [1 \; -1]^T \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0,$$

pero $x \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Por lo tanto ϕ no cumple el axioma iii) de la definición de producto interno.

b) : Sea $f(t) = 1 \in C([0,1], \mathbb{R})$. Entonces

$$\phi(f,f) = \int_0^1 (t-2)\ 1dt = (\frac{t^2}{2} - 2t)|_0^1 = -\frac{3}{2} < 0.$$

Por lo tanto ϕ no cumple el axioma iii) de la definición de producto interno.

c) Sea $p(x) = x(x-1) \in \mathbb{R}_2[x]$, entonces $p(0) = p(1) = 0$. Entonces

$$\phi(p,p) = p(0)^2 + p(1)^2 = 0,$$

pero $p \neq 0$ (el polinomio nulo). Por lo tanto ϕ no cumple el axioma iii) de la definición de producto interno. ■

Ejercicio 3.1. Sea $(\mathbb{V}, \langle\, \cdot, \cdot \,\rangle)$ un $\mathbb{C}$-espacio euclídeo finito dimensional y sea $v_0 \in \mathbb{V} \setminus \{0_{\mathbb{V}}\}$ un vector arbitrario pero fijo.

a) Comprobar que la aplicación $\phi : \mathbb{V} \to \mathbb{C}$ definida por

$$\phi(v) := \langle\, v, v_0 \,\rangle$$

es una funcional lineal de $\mathbb{V}$. Describir su núcleo e indicar la dimensión del mismo. Observar que

$$\mathbb{V} = \mathrm{Nu}(\phi) \oplus gen\{v_0\}.$$

b) Explicar por qué la aplicación $\psi : \mathbb{V} \to \mathbb{C}$ definida por $\psi(v) := \langle\, v_0, v \,\rangle$ no es una funcional lineal.

Dem. *a*) : Usando los items i.1) y i.2) de la definición de producto interno, tenemos que, dados $u, v \in \mathbb{V}$ y $\lambda \in \mathbb{C}$,

$$\phi(u+v) = \langle\, u+v, v_0 \,\rangle = \langle\, u, v_0 \,\rangle + \langle\, v, v_0 \,\rangle = \phi(u) + \phi(v),$$

$$\phi(\lambda u) = \langle\, \lambda u, v_0 \,\rangle = \lambda \langle\, u, v_0 \,\rangle = \lambda\phi(u).$$

Por lo tanto ϕ es una funcional lineal de $\mathbb{V}$.

Veamos que $\mathbb{V} = \mathrm{Nu}(\phi) \oplus gen\{v_0\}$. Primero, veamos que $\mathrm{Nu}(\phi) \cap gen\{v_0\} = \{0_{\mathbb{V}}\}$. Sea $v \in \mathrm{Nu}(\phi) \cap gen\{v_0\}$ entonces, $v \in \mathrm{Nu}(\phi)$ y $v \in gen\{v_0\}$. Como $v \in \mathrm{Nu}(\phi)$, tenemos que $\phi(v) = 0$ y como $v \in gen\{v_0\}$ tenemos que $v = \alpha v_0$ para cierto $\alpha \in \mathbb{C}$. Entonces

$$0 = \phi(v) = \langle\, v, v_0 \,\rangle = \langle\, \alpha v_0, v_0 \,\rangle = \alpha \langle\, v_0, v_0 \,\rangle .$$

Usando iii) de la definición de producto interno, tenemos que $\langle\, v_0, v_0 \,\rangle > 0$, pues $v_0 \neq 0$. Entonces se sigue que $\alpha = 0$, entonces $v = \alpha v_0 = 0 v_0 = 0_{\mathbb{V}}$, y probamos que $\mathrm{Nu}(\phi) \cap gen\{v_0\} = \{0_{\mathbb{V}}\}$.

Veamos ahora, que $\mathbb{V} = \mathrm{Nu}(\phi) \oplus gen\{v_0\}$. Sea $v \in \mathbb{V}$, entonces, como $\phi(v) \in \mathbb{C}$ y $\phi(v_0) \neq 0$ (recién vimos que $v_0 \notin \mathrm{Nu}(\phi)$), podemos definir el siguiente escalar complejo: $\alpha := \frac{\phi(v)}{\phi(v_0)} \in \mathbb{C}$. Entonces, es fácil ver que

$$v = [v - \alpha v_0] + \alpha v_0.$$

Entonces, usando que ϕ es una funcional lineal y la definición de α, tenemos que

$$\phi(v - \alpha v_0) = \phi(v) - \alpha\phi(v_0) = \phi(v) - \frac{\phi(v)}{\phi(v_0)}\phi(v_0) = 0.$$

Por lo tanto, $v - \alpha v_0 \in \mathrm{Nu}(\phi)$. Entonces $v = v_1 + v_2$ con $v_1 := v - \alpha v_0 \in \mathrm{Nu}(\phi)$ y $v_2 := \alpha v_0 \in gen\{v_0\}$. Por lo tanto $v \in \mathrm{Nu}(\phi) \oplus gen\{v_0\}$ y se sigue que

$$\mathbb{V} = \mathrm{Nu}(\phi) \oplus gen\{v_0\}.$$

Observar que para hacer estas cuentas no usamos que $\mathbb{V}$ es de dimensión finita.

Finalmente, ahora sí, como $\mathbb{V}$ es de dimensión finita, podemos aplicar el Teorema de la dimensión de la suma de subespacios y entonces,

$$\dim(\mathrm{Nu}(\phi)) = \dim(\mathbb{V}) - \dim(gen\{v_0\}) = \dim(\mathbb{V}) - 1.$$

Notar que

$$\mathrm{Nu}(\phi) = \{v \in \mathbb{V} : \phi(v) = 0\} = \{v \in \mathbb{V} : \langle v, v_0 \rangle = 0\} = gen\{v_0\}^{\perp},$$

para la última igualdad usamos la definición de subespacio ortogonal del conjunto $gen\{v_0\}$ que veremos con detalle en la Sección 3.4.

b) Tomemos $i \in \mathbb{C}$, entonces usando los items i.1) y ii) de la definición de producto interno, tenemos que

$$\psi(iv_0) = \langle v_0, iv_0 \rangle = \bar{i} \langle v_0, v_0 \rangle = -i\psi(v_0) \neq i\psi(v_0),$$

donde usamos que $\psi(v_0) = \langle v_0, v_0 \rangle > 0$ para afirmar que $-i\psi(v_0) \neq i\psi(v_0)$. Por lo tanto, ψ no es una funcional lineal, ya que existe un escalar $\alpha \in \mathbb{C}$ tal que $\psi(\alpha v_0) \neq \alpha\psi(v_0)$. ■

3.2. Espacios normados

Una *norma* en un $\mathbb{K}$-espacio vectorial $\mathbb{V}$ es una función $\|\cdot\| : \mathbb{V} \to \mathbb{R}^+$, tal que:

i) $\|x\| = 0$ si y sólo si $x = 0_{\mathbb{V}}$.

ii) $\|\lambda x\| = |\lambda|\ \|x\|$, para todo $x \in \mathbb{V}$, $\lambda \in \mathbb{C}$.

iii) $\|x + y\| \leq \|x\| + \|y\|$, para todo $x, y \in \mathbb{V}$ (desigualdad triangular).

El par $(\mathbb{V}, \|\cdot\|)$ se llama *espacio normado*. Si $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ es un $\mathbb{K}$-espacio euclídeo, la función

$$\|x\| := \langle x, x \rangle^{1/2}$$

define una norma en $\mathbb{V}$. Dicha norma se llama la *norma inducida por el producto interno*.

Se deja como ejercicio verificar que si $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ es un $\mathbb{K}$-espacio euclídeo, la función

$$\|x\| := \langle x, x \rangle^{1/2}$$

define una norma en $\mathbb{V}$. Ayuda: para el item iii) usar la desigualdad de Cauchy-Schwarz.

La siguiente es una propiedad que se deduce de la definición de norma y que usaremos más adelante:

Sean $u, v \in \mathbb{V}$. Entonces

$$|\ \|u\| - \|v\|\ | \leq \|u - v\|. \tag{3.1}$$

De hecho, por la desigualdad triangular, tenemos que $\|u\| = \|u - v + v\| \leq \|u - v\| + \|v\|$. Entonces, restando a ambos lados de la desigualdad por $\|v\|$ se sigue que

$$\|u\| - \|v\| \leq \|u - v\|.$$

De manera similar, $\|v\| = \|v - u + u\| \leq \|v - u\| + \|u\|$. Entonces

$$\|v\| - \|u\| \leq \|v - u\| = \|(-1)(u - v)\| = |-1|\|u - v\| = \|u - v\|.$$

Por lo tanto, tenemos que $-\|u-v\| \leq \|u\|-\|v\| \leq \|u-v\|$. Equivalentemente $|\ \|u\|-\|v\|\ | \leq \|u-v\|$, y probamos los que queríamos.

Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo, $\| \cdot \|$ la norma inducida por el producto interno y sean $u, v \in \mathbb{V}$. Las siguientes propiedades resultan útiles para la resolución de varios ejercicios.

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2 + 2\,\mathrm{Re}(\langle u, v \rangle). \tag{3.2}$$

De hecho, usando los axiomas de la definición de producto interno, tenemos

$$\begin{aligned} \|u+v\|^2 &= \langle u+v, u+v \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\ &= \|u\|^2 + \langle u, v \rangle + \overline{\langle u, v \rangle} + \|v\|^2 \\ &= \|u\|^2 + \|v\|^2 + 2\,\mathrm{Re}(\langle u, v \rangle). \end{aligned}$$

Teorema de Pitagoras: Si $\langle u, v \rangle = 0$ entonces

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2. \tag{3.3}$$

La prueba de (3.3) es inmediata usando que $\langle u, v \rangle = 0$ y la ecuación (3.2). Finalmente, si ahora $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ es un $\mathbb{R}$-espacio euclídeo, la ecuación (3.2), nos queda

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2 + 2\langle u, v \rangle. \tag{3.4}$$

Entonces,

$$\|u-v\|^2 = \|u+(-v)\|^2 = \|u\|^2 + \|-v\|^2 + 2\langle u, -v \rangle = \|u\|^2 + \|v\|^2 - 2\langle u, v \rangle. \tag{3.5}$$

Si hacemos la resta de las ecuaciones (3.4) y (3.5), se sigue que:

si $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ es un $\mathbb{R}$-espacio euclídeo entonces,

$$4\langle u, v \rangle = \|u+v\|^2 - \|u-v\|^2. \tag{3.6}$$

Si hacemos la suma de las ecuaciones (3.4) y (3.5), se sigue que:

si $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ es un $\mathbb{R}$-espacio euclídeo entonces,

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2). \tag{3.7}$$

3.3. Matriz de Gram, Gramiano y triángulos

Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo. Y sea $\mathcal{X} = \{x_1, x_2, \cdots, x_r\}$ un conjunto de vectores de $\mathbb{V}$, se llama la *matriz de Gram* y se denota por $G_{\mathcal{X}}$ a

$$G_{\mathcal{X}} := \begin{bmatrix} \langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle & \cdots & \langle x_1, x_r \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle & \cdots & \langle x_2, x_r \rangle \\ \vdots & \vdots & & \vdots \\ \langle x_r, x_1 \rangle & \langle x_r, x_2 \rangle & \cdots & \langle x_r, x_r \rangle \end{bmatrix}.$$

Observar que $G_{\mathcal{X}} \in \mathbb{C}^{r \times r}$. Llamaremos *Gramiano* de $\mathcal{X}$ al determinante de $G_{\mathcal{X}}$.

Antes de resolver el **Ejercicio 3.C**, recordemos algunas propiedades de la operación *. Sea $A \in \mathbb{C}^{m \times n}$ (cuando $n = 1$ tenemos un vector columna), se define

$$A^* = \overline{A^T}.$$

Entonces:

- $(A+B)^* = A^* + B^*,\ A, B \in \mathbb{C}^{m\times n}$.
- $(\alpha A)^* = \overline{\alpha} A^*,\ A \in \mathbb{C}^{m\times n},\ \alpha \in \mathbb{C}$.
- $(AB)^* = B^* A^*,\ A \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{n\times m}$.
- $(A^*)^* = A,\ A \in \mathbb{C}^{m\times n}$.
- Si $A \in \mathbb{C}^{1\times 1} = \mathbb{C}$ (es decir A es un escalar) como claramente $A^T = A$, tenemos que $A^* = \overline{A}$.

Si hay alguna duda con alguna de las propiedades se recomienda probarlas.

El siguiente ejercicio es similar al **Ejercicio 3.2** para $\dim(\mathbb{V}) = n$.

Ejercicio 3.C. Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{C}$-espacio euclídeo de dimensión n y sean $\mathcal{B} = \{v_1, v_2, \cdots, v_n\}$ una base de $\mathbb{V}$.

a) Comprobar que $G_{\mathcal{B}} := \begin{bmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle & \cdots & \langle v_1, v_n \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \cdots & \langle v_2, v_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle v_n, v_1 \rangle & \langle v_n, v_2 \rangle & \cdots & \langle v_n, v_n \rangle \end{bmatrix}$

es la única matriz de $\mathbb{C}^{n\times n}$ tal que

$$\langle x, y \rangle = ([x]^{\mathcal{B}})^T\ G_{\mathcal{B}}\ \overline{[y]^{\mathcal{B}}} = ([y]^{\mathcal{B}})^*\ G_{\mathcal{B}}^T\ [x]^{\mathcal{B}},$$

para todo $x, y \in \mathbb{V}$.

b) Comprobar que si $\mathcal{B}' = \{v_1', v_2', \cdots, v_n'\}$ es otra base de $\mathbb{V}$ vale que

$$G_{\mathcal{B}'} = (M_{\mathcal{B}'}^{\mathcal{B}})^T\ G_{\mathcal{B}}\ \overline{M_{\mathcal{B}'}^{\mathcal{B}}},$$

donde $M_{\mathcal{B}'}^{\mathcal{B}}$ es la matriz de cambio de coordenadas de las bases $\mathcal{B}'$ a $\mathcal{B}$.

Dem. $a)$: Sean $x, y \in \mathbb{V}$. Como $\mathcal{B}$ es una base $\mathbb{V}$, entonces $x = a_1v_1 + a_2v_2 + \cdots + a_nv_n$, para ciertos $a_1, a_2, \cdots, a_n \in \mathbb{C}$ y $y = b_1v_1 + b_2v_2 + \cdots + b_nv_n$, para ciertos $b_1, b_2, \cdots, b_n \in \mathbb{C}$. Entonces, es claro que

$$[x]^{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},\ [y]^{\mathcal{B}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Usando los items i.1), i.2) y ii) de la definición de producto interno, tenemos que

$$\begin{aligned} \langle x, y \rangle &= \langle a_1v_1 + \cdots + a_nv_n, b_1v_1 + \cdots + b_nv_n \rangle \\ &= a_1\overline{b_1}\langle v_1, v_1 \rangle + \cdots + a_1\overline{b_n}\langle v_1, v_n \rangle + \cdots + a_n\overline{b_1}\langle v_n, v_1 \rangle + \cdots + a_n\overline{b_n}\langle v_n, v_n \rangle \\ &= a_1[\langle v_1, v_1 \rangle\ \cdots\ \langle v_1, v_n \rangle] \begin{bmatrix} \overline{b_1} \\ \overline{b_2} \\ \vdots \\ \overline{b_n} \end{bmatrix} + \cdots + a_n[\langle v_n, v_1 \rangle\ \cdots\ \langle v_n, v_n \rangle] \begin{bmatrix} \overline{b_1} \\ \overline{b_2} \\ \vdots \\ \overline{b_n} \end{bmatrix} \\ &= [a_1\ a_2\ \cdots\ a_n]\ G_{\mathcal{B}} \begin{bmatrix} \overline{b_1} \\ \overline{b_2} \\ \vdots \\ \overline{b_n} \end{bmatrix} = ([x]^{\mathcal{B}})^T\ G_{\mathcal{B}}\ \overline{[y]^{\mathcal{B}}}. \end{aligned}$$

Por último, como $\langle x, y \rangle = [([x]^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[y]^{\mathcal{B}}}] \in \mathbb{C}$ arriba vimos que $[([x]^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[y]^{\mathcal{B}}}]^T = [([x]^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[y]^{\mathcal{B}}}]$, es decir la transpuesta de un número es el mismo número. Entonces, tenemos que

$$\langle x, y \rangle = ([x]^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[y]^{\mathcal{B}}} = [([x]^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[y]^{\mathcal{B}}}]^T = \overline{[y]^{\mathcal{B}}}^T G_{\mathcal{B}}^T \ (([x]^{\mathcal{B}})^T)^T = ([y]^{\mathcal{B}})^* \ G_{\mathcal{B}}^T \ [x]^{\mathcal{B}}.$$

Veamos que $G_{\mathcal{B}}$ es única. Supongamos que existe otra matriz $A \in \mathbb{C}^{n\times n}$ tal que

$$\langle x, y \rangle = ([x]^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[y]^{\mathcal{B}}} = ([x]^{\mathcal{B}})^T \ A \ \overline{[y]^{\mathcal{B}}}.$$

Entonces tomemos $x = y = v_1$, entonces, es claro que $[x]^{\mathcal{B}} = [y]^{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = e_1$ (el primer vector de la base canónica de $\mathbb{C}^n$). Entonces,

$$e_1^T \ G_{\mathcal{B}} \ e_1 = (G_{\mathcal{B}})_{11} = e_1^T \ A \ e_1 = A_{11}.$$

Es decir el lugar $(1, 1)$ de A y $G_{\mathcal{B}}$ coinciden.

Siguiendo con esa idea, tomemos $x = v_i$ e $y = v_j$ con $i \in \{1, 2, \cdots, n\}$ y $j \in \{1, 2, \cdots, n\}$. Entonces $[x]^{\mathcal{B}} = e_i$ y $[y]^{\mathcal{B}} = e_j$ (los vectores i-ésimos y j-ésimos de la base canónica de $\mathbb{C}^n$). Entonces

$$e_i^T \ G_{\mathcal{B}} \ e_j = (G_{\mathcal{B}})_{ij} = e_i^T \ A \ e_j = A_{ij}.$$

Por lo tanto, el lugar (i, j) de A y $G_{\mathcal{B}}$ coinciden para todo $i \in \{1, 2, \cdots, n\}$ y $j \in \{1, 2, \cdots, n\}$. Entonces $A = G_{\mathcal{B}}$ como queríamos ver.

b) : Recordar que $M_{\mathcal{B}'}^{\mathcal{B}}$ es la matriz cuyas columnas son las coordenadas de los vectores $v_1', v_2', \cdots, v_n'$ en la base $\mathcal{B}$, es decir

$$M_{\mathcal{B}'}^{\mathcal{B}} = [\ [v_1']^{\mathcal{B}} \ [v_2']^{\mathcal{B}} \cdots \ [v_n']^{\mathcal{B}}\ ].$$

Es fácil verificar que $(M_{\mathcal{B}'}^{\mathcal{B}})^T = \begin{bmatrix} ([v_1']^{\mathcal{B}})^T \\ ([v_2']^{\mathcal{B}})^T \\ \vdots \\ ([v_n']^{\mathcal{B}})^T \end{bmatrix}$. Entonces

$$(M_{\mathcal{B}'}^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{M_{\mathcal{B}'}^{\mathcal{B}}} = \begin{bmatrix} ([v_1']^{\mathcal{B}})^T \\ ([v_2']^{\mathcal{B}})^T \\ \vdots \\ ([v_n']^{\mathcal{B}})^T \end{bmatrix} G_{\mathcal{B}} \ [\ \overline{[v_1']^{\mathcal{B}}} \ \overline{[v_2']^{\mathcal{B}}} \cdots \ \overline{[v_n']^{\mathcal{B}}}\ ] =$$

$$= \begin{bmatrix} ([v_1']^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[v_1']^{\mathcal{B}}} & ([v_1']^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[v_2']^{\mathcal{B}}} & \cdots & ([v_1']^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[v_n']^{\mathcal{B}}} \\ ([v_2']^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[v_1']^{\mathcal{B}}} & ([v_2']^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[v_2']^{\mathcal{B}}} & \cdots & ([v_2']^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[v_n']^{\mathcal{B}}} \\ \vdots & \vdots & & \vdots \\ ([v_n']^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[v_1']^{\mathcal{B}}} & ([v_n']^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[v_2']^{\mathcal{B}}} & \cdots & ([v_n']^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[v_n']^{\mathcal{B}}} \end{bmatrix} =$$

$$= \begin{bmatrix} \langle v_1', v_1' \rangle & \langle v_1', v_2' \rangle & \cdots & \langle v_1', v_n' \rangle \\ \langle v_2', v_1' \rangle & \langle v_2', v_2' \rangle & \cdots & \langle v_2', v_n' \rangle \\ \vdots & \vdots & & \vdots \\ \langle v_n', v_1' \rangle & \left\langle v_, ' v_2' \right\rangle & \cdots & \langle v_n', v_n' \rangle \end{bmatrix} = G_{\mathcal{B}'},$$

donde, usamos que por el item a), vale que

$$\langle v_i', v_j' \rangle = ([v_i']^{\mathcal{B}})^T \ G_{\mathcal{B}} \ \overline{[v_j']^{\mathcal{B}}},$$

para cada $i \in \{1, 2, \cdots, n\}$ y $j \in \{1, 2, \cdots, n\}$. ■

A partir del ejercicio anterior podemos afirmar que

> La matriz de Gram de una base $\mathcal{B}$ determina unívocamente al producto interno $\langle \cdot, \cdot \rangle$. En este caso, llamamos a $G_{\mathcal{B}}$ la *matriz del producto interno* $\langle \cdot, \cdot \rangle$ *en la base* $\mathcal{B}$.

Ejercicio 3.3. En cada uno de los siguientes casos, verificar que la fórmula

$$\langle x, y \rangle := y^T G x,$$

define un producto interno en $\mathbb{R}^2$.

a) $G \in \mathcal{G}_1 := \{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}, \begin{bmatrix} 1 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 1 \end{bmatrix}, \begin{bmatrix} 1 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 1 \end{bmatrix} \}$.

b) $G \in \mathcal{G}_2 := \{ \begin{bmatrix} 1 & \cos(\theta) \\ \cos(\theta) & 1 \end{bmatrix} : \theta \in (0, \pi) \}$.

c) $G \in \mathcal{G}_3 := \{ \begin{bmatrix} l_1^2 & l_1 l_2 \cos(\theta) \\ l_1 l_2 \cos(\theta) & l_2^2 \end{bmatrix} : \theta \in (0, \pi),\ l_1, l_2 > 0 \}$.

d) $G \in \mathcal{G}_4 := \{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} : a > 0,\ \det \begin{bmatrix} a & b \\ b & c \end{bmatrix} > 0 \}$.

¿Se podrán definir otros productos internos en $\mathbb{R}^2$?

Dem. Observar que $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{G}_3 = \mathcal{G}_4$.

De hecho, tomando $\theta = \frac{\pi}{2}$, $\theta = \frac{\pi}{3}$, $\theta = \frac{\pi}{4}$ y $\theta = \frac{\pi}{6}$ en la matriz $G = \begin{bmatrix} 1 & \cos(\theta) \\ \cos(\theta) & 1 \end{bmatrix}$, obtenemos las 4 matrices de $\mathcal{G}_1$. Por lo tanto $\mathcal{G}_1 \subset \mathcal{G}_2$ y la inclusión es estricta porque por ejemplo si $\theta := \arccos(\frac{1}{3})$, tenemos que $\begin{bmatrix} 1 & \frac{1}{3} \\ \frac{1}{3} & 1 \end{bmatrix} \in \mathcal{G}_2 \setminus \mathcal{G}_1$.

Por otra parte, tomando $l_1 = l_2 = 1$, en la matriz $G = \begin{bmatrix} l_1^2 & l_1 l_2 \cos(\theta) \\ l_1 l_2 \cos(\theta) & l_2^2 \end{bmatrix}$ obtenemos todas las matrices de $\mathcal{G}_3$ y la inclusión es estricta porque por ejemplo si $l_1 = l_2 = 2$ y $\theta = \frac{\pi}{2}$, tenemos que $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \in \mathcal{G}_3 \setminus \mathcal{G}_2$.

Finalmente, sea $G = \begin{bmatrix} l_1^2 & l_1 l_2 \cos(\theta) \\ l_1 l_2 \cos(\theta) & l_2^2 \end{bmatrix}$. Entonces, para todo $l_1, l_2 > 0$ y $\theta \in (0, \pi)$, el lugar G_{12} coincide con el lugar G_{21}, el lugar $G_{11} = l_1^2 > 0$ y $\det(G) = l_1^2 l_2^2 - l_1^2 l_2^2 \cos(\theta)^2 = l_1^2 l_2^2 (1 - \cos(\theta)^2) = l_1^2 l_2^2 \sin(\theta)^2 > 0$. Entonces $\mathcal{G}_3 \subseteq \mathcal{G}_4$.

Por último, si $G = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ con $a > 0$ y $\det \begin{bmatrix} a & b \\ b & c \end{bmatrix} = ac - b^2 > 0$. Entonces $c > \frac{b^2}{a} > 0$. Tomemos $l_1 := \sqrt{a} > 0$ y $l_2 := \sqrt{c} > 0$. Por otra parte, notar que como $b^2 < ac$ y $ac > 0$, tomando raíz (que es una función creciente) se sigue que $|b| < \sqrt{a}\sqrt{c} = l_1 l_2$. Por lo tanto,

$$\frac{|b|}{l_1 l_2} < 1.$$

Entonces, podemos definir $\theta := \arccos(\frac{b}{l_1 l_2}) \in (0, \pi)$ y entonces $b = l_1 l_2 \cos(\theta)$. Por lo tanto $\mathcal{G}_4 \subseteq \mathcal{G}_3$ y, como ya probamos la otra inclusión, concluimos que $\mathcal{G}_3 = \mathcal{G}_4$.

Vamos a demostrar que $\langle x, y \rangle := y^T G x$, define un producto interno en $\mathbb{R}^2$ con $G \in \mathcal{G}_4$ si hacemos eso, como $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{G}_3 = \mathcal{G}_4$, lo habremos demostrado para los demás casos.

Sea $G = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ con $a > 0$ y $\det \begin{bmatrix} a & b \\ b & c \end{bmatrix} = ac - b^2 > 0$. Entonces

i) Para cada $\alpha \in \mathbb{R}$ y para cada $x, y, z \in \mathbb{R}^2$ tenemos que

1. $\langle x+z, y \rangle = y^T G(x+z) = y^T(Gx+Gz) = y^T Gx + y^T Gx = \langle x, y \rangle + \langle z, y \rangle$.
2. $\langle \alpha x, y \rangle = y^T G(\alpha x) = \alpha y^T Gx = \alpha \langle x, y \rangle$.

Observar que i.1 y i.2 valen para cualquier $G \in \mathbb{R}^{2\times 2}$.

ii) Para cada $x, y \in \mathbb{R}^2$ tenemos que

$$\langle x, y \rangle = y^T Gx = (y^T Gx)^T = x^T G^T (y^T)^T = x^T Gy = \langle y, x \rangle,$$

donde usamos que $G = \begin{bmatrix} a & b \\ b & c \end{bmatrix} = G^T$ y que $y^T Gx \in \mathbb{R}$.

iii) Sea $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$ entonces,

$$\langle x, x \rangle = x^T Gx = [x_1\ x_2] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = ax_1^2 + 2bx_1x_2 + cx_2^2.$$

Usando que $\sqrt{a} > 0$ podemos completar cuadrados y concluir que

$$ax_1^2 + 2bx_1x_2 + cx_2^2 = (\sqrt{a}x_1 + \frac{b}{\sqrt{a}}x_2)^2 + (c - \frac{b^2}{a})x_2^2.$$

Entonces, $\langle x, x \rangle = (\sqrt{2}x_1 + \frac{b}{\sqrt{a}}x_2)^2 + (c - \frac{b^2}{a})x_2^2 \geq 0$, porque $ac - b^2 > 0$ y, como $a > 0$, se sigue que $c - \frac{b^2}{a} > 0$. Por otra parte, si

$$\langle x, x \rangle = (\sqrt{a}x_1 + \frac{b}{\sqrt{a}}x_2)^2 + (c - \frac{b^2}{a})x_2^2 = 0,$$

como cada sumando es mayor o igual a cero, tenemos que $(\sqrt{a}x_1 + \frac{b}{\sqrt{a}}x_2)^2 = (c - \frac{b^2}{a})x_2^2 = 0$. Entonces, como $(c - \frac{b^2}{a}) > 0$ y $(c - \frac{b^2}{a})x_2^2 = 0$, se sige que $x_2 = 0$. Finalmente, como $0 = (\sqrt{a}x_1 + \frac{b}{\sqrt{a}}x_2)^2 = ax_1^2$ y $a > 0$, se sigue que $x_1 = 0$. Por lo tanto $x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Probamos que $\langle x, x \rangle > 0$ si y sólo si $x \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ y se cumple iii).

Finalmente, veamos que NO se pueden definir otros productos internos en $\mathbb{R}^2$. De hecho, supongamos que la fórmula $\langle x, y \rangle := y^T Gx$, con $G = \begin{bmatrix} a & b \\ b' & c \end{bmatrix} \in \mathbb{R}^{2\times 2}$ define un producto interno en $\mathbb{R}^2$. Veamos que $G \in \mathcal{G}_4$. De hecho, si $\langle x, y \rangle = y^T Gx$, define un producto interno entonces $\langle x, y \rangle = y^T Gx = x^T Gy = \langle y, x \rangle$, para todo $x, y \in \mathbb{R}^2$. En particular, podemos tomar $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ e $y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Entonces

$$y^T Gx = [0\ 1] \begin{bmatrix} a & b \\ b' & c \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = b' = [1\ 0] \begin{bmatrix} a & b \\ b' & c \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = b.$$

Por lo tanto $G = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. Por último, tenemos que

$$\langle x, x \rangle = x^T Gx > 0,$$

para todo $x \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. En particular, si tomamos $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, entonces

$$\left\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\rangle = [1\ 0] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = a > 0.$$

Sólo nos resta probar que $\det(G) > 0$. Para eso, consideremos la función $f : [0, 1] \to \mathbb{R}$, definida por

$$f(t) := \det(tI + (1 - t)G),$$

donde $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ es la matriz identidad de 2×2. Claramente f es una función continua (básicamente por que f en definitiva es un polinomio en t). Además, $f(0) = \det(G)$ y $f(1) = \det(I) = 1$. Supongamos que $x \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, entonces

$$x^T(tI + (1 - t)G)x = tx^T x + (1 - t)x^T Gx > 0,$$

pues $x^T x > 0$, $t \in [0, 1]$ y estamos usando que $x^T Gx > 0$. Entonces $f(t) = \det(tI + (1 - t)G) \neq 0$, porque si $tI + (1 - t)G$ fuera una matriz singular (no inversible o con determinante nulo), existiría algún vector no nulo $v \in \mathbb{R}^2$ tal que $(tI + (1 - t)G)v = 0$, con lo que $v^T[tI + (1 - t)G]v = 0$, que sería una contradicción. Por lo tanto, $f(t) \neq 0$, para todo $t \in [0, 1]$.

Como $f(1) = 1 > 0$ y f es continua, si $f(t) = \det(G) < 0$, por el Teorema de Bolzano, existiría algún $t \in [0, 1]$ tal que $f(t) = 0$. Como probamos que $f(t) \neq 0$ para todo $t \in [0, 1]$, se sigue que $f(0) = \det(G) > 0$. ■

La siguiente propiedad la vamos a usar para resolver el **Ejercicio 3.6** y tiene interés en sí misma.

Proposición 3.3.1. *Sea* $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ *un* $\mathbb{K}$*-espacio euclídeo de dimensión finita y* $\{x_1, x_2, \cdots, x_r\}$ *un conjunto de vectores de* $\mathbb{V}$*. Entonces*

$$\{x_1, x_2, \cdots, x_r\} \text{ es LD si y sólo si } \det(G_{\{x_1,x_2,\cdots,x_r\}}) = 0.$$

Es decir el Gramiano es nulo.

Antes de comenzar con la demostración, veamos una manera conveniente de escribir la matriz $G_{\{x_1,x_2,\cdots,x_r\}}$.

Sea $\mathcal{B} = \{v_1, \cdots, v_n\}$ una base ortonormal de $\mathbb{V}$. Es decir $\|v_i\| = 1$ para todo $i = 1, 2 \cdots, n$ y $\langle v_i, v_j \rangle = 0$ para todo $i \neq j$. Entonces, tenemos que $G_{\mathcal{B}} = I_{n\times n}$ (comprobarlo) y, por el **Ejercicio 3.C a)**,

$$\langle x_i, x_j \rangle = ([x_j]^{\mathcal{B}})^* G_{\mathcal{B}}^T\ [x_i]^{\mathcal{B}} = ([x_j]^{\mathcal{B}})^*\ [x_i]^{\mathcal{B}},$$

para todo $i = 1, 2, \cdots, r$ y $j = 1, 2, \cdots, r$. Sea

$$A := [[\ x_1]^{\mathcal{B}}\ [x_2]^{\mathcal{B}}\ \cdots [x_r]^{\mathcal{B}}\ ] \in \mathbb{C}^{n\times r},$$

la matriz que resulta de poner en sus columnas las coordenadas en base $\mathcal{B}$ de los vectores $x_1, x_2, \cdots, x_r$. Entonces, observar que

$$G^T_{\{x_1,x_2,\cdots,x_r\}} = \begin{bmatrix} \langle x_1, x_1 \rangle & \langle x_2, x_1 \rangle & \cdots & \langle x_r, x_1 \rangle \\ \langle x_1, x_2 \rangle & \langle x_2, x_2 \rangle & \cdots & \langle x_r, x_2 \rangle \\ \vdots & \vdots & & \vdots \\ \langle x_1, x_r \rangle & \langle x_2, x_r \rangle & \cdots & \langle x_r, x_r \rangle \end{bmatrix} =$$

$$= \begin{bmatrix} ([x_1]^{\mathcal{B}})^* \ [x_1]^{\mathcal{B}} & ([x_1]^{\mathcal{B}})^* \ [x_2]^{\mathcal{B}} & \cdots & ([x_1]^{\mathcal{B}})^* \ [x_r]^{\mathcal{B}} \\ ([x_2]^{\mathcal{B}})^* \ [x_1]^{\mathcal{B}} & ([x_2]^{\mathcal{B}})^* \ [x_2]^{\mathcal{B}} & \cdots & ([x_2]^{\mathcal{B}})^* \ [x_r]^{\mathcal{B}} \\ \vdots & \vdots & & \vdots \\ ([x_r]^{\mathcal{B}})^* \ [x_1]^{\mathcal{B}} & ([x_r]^{\mathcal{B}})^* \ [x_2]^{\mathcal{B}} & \cdots & ([x_r^{\mathcal{B}})^* \ [x_r]^{\mathcal{B}} \end{bmatrix} = A^*A.$$

Ahora sí vamos a demostar la proposición.

Dem. Supongamos que $\{x_1, x_2, \cdots, x_r\}$ es un conjunto LD de $\mathbb{V}$. Entonces, por definición de independencia lineal, existen escalares $a_1, a_2, \cdots, a_r \in \mathbb{K}$ no todos nulos, tales que

$$a_1x_1 + a_2x_2 + \cdots + a_rx_r = 0_{\mathbb{V}}.$$

Sea $z := [a_1 \ a_2 \ \cdots \ a_r]^T \neq 0_{\mathbb{K}^r}$. Entonces, tenemos que

$$Az = [[\ x_1]^{\mathcal{B}} \ [x_2]^{\mathcal{B}} \ \cdots [x_r]^{\mathcal{B}} \ ] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_4 \end{bmatrix} = a_1[x_1]^{\mathcal{B}} + a_2[x_2]^{\mathcal{B}} + \cdots + a_r[x_r]^{\mathcal{B}} =$$

$$= [a_1x_1 + a_2x_2 + \cdots + a_rx_r]^{\mathcal{B}} = [0_{\mathbb{V}}]^{\mathcal{B}} = 0_{\mathbb{K}^n}.$$

Entonces

$$G^T_{\{x_1,x_2,\cdots,x_r\}}z = A^*Az = A^*0_{\mathbb{K}^n} = 0_{\mathbb{K}^r}.$$

Por lo tanto $z \in \text{nul}(G^T_{\{x_1,x_2,\cdots,x_r\}})$. Entonces como $z \neq 0_{\mathbb{K}^r}$, se sigue que $\dim(\text{nul}(G^T_{\{x_1,x_2,\cdots,x_r\}})) > 0$ ó, usando el Teorema de la Dimensión, se sigue que $\text{rg}(G^T_{\{x_1,x_2,\cdots,x_r\}}) < r$. Por lo tanto, $G^T_{\{x_1,x_2,\cdots,x_r\}}$ es una matriz de $r \times r$ cuyo rango no es completo, entonces $\det(G^T_{\{x_1,x_2,\cdots,x_r\}}) = 0$. Finalmente como el determinante de una matriz y el determinante de su transpuesta coinciden, tenemos $\det(G_{\{x_1,x_2,\cdots,x_r\}}) = \det(G^T_{\{x_1,x_2,\cdots,x_r\}}) = 0$.

Recíprocamente, si $\det(G_{\{x_1,x_2,\cdots,x_r\}}) = 0$, entonces $\det(G^T_{\{x_1,x_2,\cdots,x_r\}}) = \det(G_{\{x_1,x_2,\cdots,x_r\}}) = 0$. Entonces, existe un vector no nulo $z = [z_1 \ z_2 \cdots \ z_r]^T \in \mathbb{K}^r$ tal que

$$G^T_{\{x_1,x_2,\cdots,x_r\}}z = A^*Az = 0_{\mathbb{K}^r}.$$

Entonces,

$$z^*A^*Az = z^*0_{\mathbb{K}^r} = 0.$$

Observar que $z^*A^*Az = (Az)^*(Az) = \|Az\|_{pic}$, donde $\|\cdot\|_{pic}$ denota el producto interno canónico de $\mathbb{K}^n$. Entonces,

$$\|Az\|_{pic} = 0.$$

Por lo tanto, por definición de norma, $Az = 0_{\mathbb{K}^n}$, entonces

$$0_{\mathbb{K}^n} = Az = z_1[x_1]^{\mathcal{B}} + z_2[x_2]^{\mathcal{B}} + \cdots + z_r[x_r]^{\mathcal{B}} = [z_1x_1 + z_2x_2 + \cdots + z_rx_r]^{\mathcal{B}}.$$

Entonces,

$$z_1x_1 + z_2x_2 + \cdots + z_rx_r = 0_{\mathbb{V}},$$

y como $z_1, z_2, \cdots, z_r \in \mathbb{K}$ son escalares no todos nulos, se sigue que $\{x_1, x_2, \cdots, x_r\}$ es un conjunto LD de $\mathbb{V}$. ■

De la Proposición 3.3.1 es imediato ver que: si $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ es un $\mathbb{K}$-espacio euclídeo de dimensión finita y $\{x_1, x_2, \cdots, x_r\}$ un conjunto de vectores de $\mathbb{V}$. Entonces

$$\{x_1, x_2, \cdots, x_r\} \text{ es LI si y sólo si } \det(G_{\{x_1,x_2,\cdots,x_r\}}) \neq 0.$$

Notación para triángulos

Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{R}$-espacio euclídeo, se define el *ángulo θ entre dos vectores no nulos x e y* mediante la fórmula

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

con $\theta \in [0, \pi]$.

Sean $u, v, w \in \mathbb{V}$ (3 puntos no alineados) en el $\mathbb{R}$-espacio euclídeo $(\mathbb{V}, \langle \cdot, \cdot \rangle)$, el siguiente es el triangulo de vértices u, v, w :

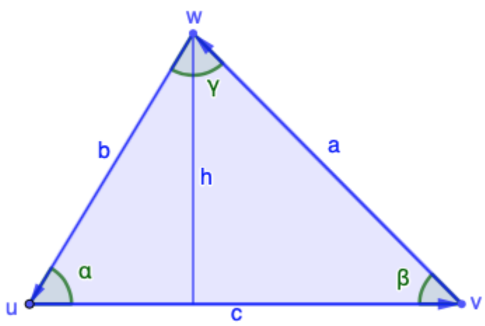

Figura 3.1: Notación triangulo.

Con la notación de la Figura 3.1, se sigue que:

$$c = v - u, \;\; a = w - v, \;\; b = u - w.$$

Como u, v, w son puntos no alineados, se sigue que $a, b, c \neq 0_{\mathbb{V}}$. Entonces, podemos considerar los vectores normalizados:

$$\tilde{a} = \frac{a}{\|a\|}, \quad \tilde{b} = \frac{b}{\|b\|}, \quad \tilde{c} = \frac{c}{\|c\|}.$$

Comprobar que $\|\tilde{a}\| = \|\tilde{b}\| = \|\tilde{c}\| = 1$. En cuánto a los ángulos interiores del triángulo, tenemos que:

$$\cos\alpha = \frac{\langle c, -b \rangle}{\|c\|\|b\|} = \left\langle \frac{c}{\|c\|}, \frac{-b}{\|b\|} \right\rangle = \langle \tilde{c}, -\tilde{b} \rangle,$$

donde usamos i.2) de la definición de producto interno. De la misma manera,

$$\cos\beta = \frac{\langle a, -c \rangle}{\|a\|\|c\|} = \langle \tilde{a}, -\tilde{c} \rangle, \quad \cos\gamma = \frac{\langle -a, b \rangle}{\|a\|\|b\|} = \langle -\tilde{a}, \tilde{b} \rangle.$$

Para obtener los ángulos interiores del triángulo, observar los signos que usamos en las fórmulas de los ángulos.

La altura del triángulo se obtiene como

$$\|h\| = \|b\| \sin(\alpha).$$

Finalmente recordemos las siguientes identidades trigonómetricas: sean $\alpha, \beta \in \mathbb{R}$ entonces

- $\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$,
- $\cos(\alpha)^2 + \sin(\alpha)^2 = 1$.

Vamos a resolver un ejercicio muy similar al **Ejercicio 3.5**.

Ejercicio 3.D. Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{R}$-espacio euclídeo de dimensión 3 y sea

$$\mathcal{B} = \{u_1, u_2, u_3\} \subset \{u \in \mathbb{V} : \|u\| = 1\},$$

una base de $\mathbb{V}$ tal que $\|u_i + u_j\|^2 = 2 + \sqrt{3}$ y $\|u_i - u_j\|^2 = 2 - \sqrt{3}$ para $i \neq j$.

a) Hallar la matriz del producto interno $\langle \cdot, \cdot \rangle$ en la base $\mathcal{B}$.

b) Hallar $\Theta := [\ \arccos(\langle u_i, u_j \rangle)]_{i \in \{1,2,3\}, j \in \{1,2,3\}}$.

c) Calcular al área del triangulo de vértices, u_1, u_2, u_3.

d) Determinar los vértices de un triángulo rectángulo T tal que $T \subset gen\{u_1, u_2\}$ y cuyos catetos midan 3 y 4.

Dem. $a)$: Vamos a calcular la matriz

$$G_{\mathcal{B}} = \begin{bmatrix} \|u_1\|^2 & \langle u_1, u_2 \rangle & \langle u_1, u_3 \rangle \\ \langle u_2, u_1 \rangle & \|u_2\|^2 & \langle u_2, u_3 \rangle \\ \langle u_3, u_1 \rangle & \langle u_3, u_2 \rangle & \|u_3\|^2 \end{bmatrix}.$$

Como $u_1, u_2, u_3 \in \{u \in \mathbb{V} : \|u\| = 1\}$, entonces $\|u_1\| = \|u_2\| = \|u_3\| = 1$.

Como estamos en un $\mathbb{R}$-espacio euclídeo podemos usar la ecuación (3.6). Entonces, si $i \neq j$,

$$\langle u_i, u_j \rangle = \frac{\|u_i + u_j\|^2 - \|u_i - u_j\|^2}{4} = \frac{2 + \sqrt{3} - (2 - \sqrt{3})}{4} = \frac{\sqrt{3}}{2}.$$

Por lo tanto

$$G_{\mathcal{B}} = \begin{bmatrix} 1 & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 1 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & 1 \end{bmatrix}.$$

$b)$: Como $\arccos(1) = 0$ y $\arccos(\frac{\sqrt{3}}{2}) = \frac{\pi}{6}$, tenemos que

$$\Theta = \begin{bmatrix} 0 & \frac{\pi}{6} & \frac{\pi}{6} \\ \frac{\pi}{6} & 0 & \frac{\pi}{6} \\ \frac{\pi}{6} & \frac{\pi}{6} & 0 \end{bmatrix}.$$

$c)$: Con la notación de la Figura 3.1, tenemos que $u = u_1$, $v = u_2$, $w = u_3$. Entonces

$$c = v - u = u_2 - u_1, \ b = u - w = u_1 - u_3.$$

Entonces $\|c\|^2 = \|u_2 - u_1\|^2 = 2 - \sqrt{3}$, $\|b\|^2 = \|u_1 - u_3\|^2 = 2 - \sqrt{3}$ y $\langle c, b \rangle = \langle u_2 - u_1, u_1 - u_3 \rangle = \langle u_2, u_1 \rangle - \langle u_2, u_3 \rangle - \|u_1\|^2 + \langle u_1, u_3 \rangle = \frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2} - 1 + \frac{\sqrt{3}}{2} = \frac{\sqrt{3}-2}{2}$. Entonces

$$\cos(\alpha)^2 = \frac{\langle c, -b \rangle^2}{\|c\|^2 \ \|b\|^2} = \frac{(\frac{2-\sqrt{3}}{2})^2}{(2 - \sqrt{3})^2} = \frac{1}{4}.$$

Entonces

$$\sin(\alpha) = \sqrt{1 - \cos(\alpha)^2} = \sqrt{1 - \frac{1}{4}} = \frac{\sqrt{3}}{2}.$$

Por lo tanto, $Area_T = \|c\| \frac{\|h\|}{2} = \|c\| \frac{\|b\| \sin(\alpha)}{2} = \sqrt{(2 - \sqrt{3})} \sqrt{(2 - \sqrt{3})} \frac{\sqrt{3}}{4} = (2 - \sqrt{3}) \frac{\sqrt{3}}{4} = \frac{\sqrt{3}}{2} - \frac{3}{4}$.

$d)$: Con la notación de la Figura 3.1, buscamos un triangulo de vértices $u, v, w \in \mathbb{V}$ tales que $\alpha = \frac{\pi}{2}$ (triángulo rectángulo) $u, v, w \in gen\{u_1, u_2\}$ y los catetos que son c y b midan 3 y 4 respectivamente, es decir $\|c\| = \|v - u\| = 3$, $\|b\| = \|u - w\| = 4$, y $0 = \langle c, -b \rangle = \langle v - u, w - u \rangle$. Como nos piden un triángulo (es decir no tenemos que hallar todos los triángulos que cumplen lo anterior), para simplificar cuentas, podemos suponer que $u = 0_{\mathbb{V}}$, $v = z_0 u_1$, para cierto $z_0 \in \mathbb{R}$ y

$w = z_1 u_1 + z_2 u_2$, para ciertos $z_1, z_2 \in \mathbb{R}$. Entonces, como $u, v, w \in gen\{u_1, u_2\}$ el triángulo T con esos vértices cumple $T \subset gen\{u_1, u_2\}$.

Vamos a obtener los escalares $z_0, z_1, z_2 \in \mathbb{R}$ para que T quede totalmente definido. Para eso vamos a usar los datos:

$$3 = \|c\| = \|v - u\| = \|z_0 u_1\| = |z_0| \|u_1\| = |z_0|.$$

Tomamos $z_0 = 3$ (podríamos haber tomado también $z_0 = -3$). Entonces

$$0 = \langle c, b \rangle = \langle v, w \rangle = \langle 3u_1, z_1 u_1 + z_2 u_2 \rangle = 3z_1 \langle u_1, u_1 \rangle + 3z_2 \langle u_1, u_2 \rangle = 3z_1 + 3z_2 \frac{\sqrt{3}}{2}.$$

Entonces $z_1 = -z_2 \frac{\sqrt{3}}{2}$. Entonces, usando que

$$\begin{aligned} 16 = \|b\|^2 = \|u - w\|^2 = \|z_1 u_1 + z_2 u_2\|^2 = z_1^2 \|u_1\|^2 + z_2^2 \|u_2\|^2 + 2 z_1 z_2 \langle u_1, u_2 \rangle = \\ = z_1^2 + z_2^2 + z_1 z_2 \sqrt{3}. \end{aligned}$$

Reemplazando $z_1 = -z_2 \frac{\sqrt{3}}{2}$, nos queda que

$$16 = z_2^2 \frac{3}{4} + z_2^2 - z_2^2 \frac{3}{2} = \frac{1}{4} z_2^2.$$

Entonces $|z_2| = \sqrt{64} = 8$. Tomamos $z_2 = 8$ y entonces $z_1 = -8\frac{\sqrt{3}}{2} = -4\sqrt{3}$.

Entonces, el triangulo T de vértices $u = 0_{\mathbb{V}}$, $v = 3u_1$ y $w = -4\sqrt{3}u_1 + 8u_2$, cumple con lo pedido. ■

Ejercicio 3.6. Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{R}$-espacio vectorial. Sean v_1, v_2 dos vectores linealmente independientes. Demostrar que la suma de los ángulos internos del triángulo de vertices $0_{\mathbb{V}}, v_1, v_2$ es π.

Dem. Con la notación de la Figura 3.1, tenemos que $u = 0_{\mathbb{V}}, v = v_1,\ w = v_2$, entonces $c = v_1,\ a = v_2 - v_1,\ b = -v_2$. Entonces,

$$a + b + c = v_2 - v_1 - v_2 + v_1 = 0_{\mathbb{V}},$$

entonces el conjunto $\{a, b, c\}$ es linealmente dependiente en $\mathbb{V}$ y por lo tanto, por la Proposición 3.3.1, su Gramiano se anula, es decir

$$\det G_{\{a,b,c\}} = \det \begin{bmatrix} \|a\|^2 & \langle a, b \rangle & \langle a, c \rangle \\ \langle b, a \rangle & \|b\|^2 & \langle b, c \rangle \\ \langle c, a \rangle & \langle c, b \rangle & \|c\|^2 \end{bmatrix} = 0.$$

Desarrollando el determinante por la primera columna, operando y usando que como estamos en un $\mathbb{R}$-espacio vectorial el producto interno es simétrico, tenemos que

$$\|a\|^2 \|b\|^2 \|c\|^2 - \|a\|^2 \langle b, c \rangle^2 - \langle a, b \rangle^2 \|c\|^2 - \langle a, c \rangle^2 \|b\|^2 + 2 \langle b, c \rangle \langle a, b \rangle \langle a, c \rangle = 0.$$

Ahora, si dividimos la expresión anterior por $\|a\|^2 \|b\|^2 \|c\|^2 \neq 0$ y normalizamos los vectores, nos queda:

$$1 - \langle \tilde{b}, \tilde{c} \rangle^2 - \langle \tilde{a}, \tilde{b} \rangle^2 - \langle \tilde{a}, \tilde{c} \rangle^2 + 2 \langle \tilde{b}, \tilde{c} \rangle \langle \tilde{a}, \tilde{b} \rangle \langle \tilde{a}, \tilde{c} \rangle = 0.$$

Entonces $(1 - \langle \tilde{b}, \tilde{c} \rangle^2)(1 - \langle \tilde{a}, \tilde{b} \rangle^2) - \langle \tilde{b}, \tilde{c} \rangle^2 \langle \tilde{a}, \tilde{b} \rangle^2 = \langle \tilde{a}, \tilde{c} \rangle^2 - 2 \langle \tilde{b}, \tilde{c} \rangle \langle \tilde{a}, \tilde{b} \rangle \langle \tilde{a}, \tilde{c} \rangle$.

Entonces,

$$\begin{aligned} (1 - \langle \tilde{b}, \tilde{c} \rangle^2)(1 - \langle \tilde{a}, \tilde{b} \rangle^2) &= \langle \tilde{b}, \tilde{c} \rangle^2 \langle \tilde{a}, \tilde{b} \rangle^2 + \langle \tilde{a}, \tilde{c} \rangle^2 - 2 \langle \tilde{b}, \tilde{c} \rangle \langle \tilde{a}, \tilde{b} \rangle \langle \tilde{a}, \tilde{c} \rangle \\ &= (- \langle \tilde{a}, \tilde{c} \rangle + \langle \tilde{b}, \tilde{c} \rangle \langle \tilde{a}, \tilde{b} \rangle)^2. \end{aligned}$$

Tomando raíz cuadrada a ambos lados, obtenemos

$$\sqrt{1-\langle\, \tilde{b}, \tilde{c}\,\rangle^2}\sqrt{1-\langle\, \tilde{a}, \tilde{b}\,\rangle^2} = |-\langle\, \tilde{a}, \tilde{c}\,\rangle + \langle\, \tilde{b}, \tilde{c}\,\rangle\langle\, \tilde{a}, \tilde{b}\,\rangle|.$$

Recordemos que: $\cos\alpha = \left\langle\, \tilde{c}, -\tilde{b}\,\right\rangle$, entonces $\sin\alpha = \sqrt{1-\cos\alpha^2} = \sqrt{1-\left\langle\, \tilde{c}, \tilde{b}\,\right\rangle^2}$. De la misma manera, tenemos que $\cos\beta = \langle\, \tilde{a}, -\tilde{c}\,\rangle$, $\sin\beta = \sqrt{1-\langle\, \tilde{a}, \tilde{c}\,\rangle^2}$, $\cos\gamma = \left\langle\, -\tilde{a}, \tilde{b}\,\right\rangle$, $\sin\gamma = \sqrt{1-\left\langle\, \tilde{a}, \tilde{b}\,\right\rangle^2}$.

Entonces, reemplazando en la ecuación anterior por las expresiones de los cosenos y senos de los ángulos correspondientes, nos queda:

$$\sin\alpha\sin\gamma = |\cos\beta + \cos\alpha\cos\gamma| = \cos\beta + \cos\alpha\cos\gamma,$$

donde usamos que como $\alpha, \beta, \gamma \in [0, \frac{\pi}{2}]$ el coseno es mayor o igual a 0 y podemos eliminar el módulo. Entonces, pasando de término, nos queda

$$-\cos\beta = \cos(\pi-\beta) = \cos\alpha\cos\gamma - \sin\alpha\sin\gamma = \cos(\alpha+\gamma).$$

Como en el intervalo $[0, \frac{\pi}{2}]$ el coseno es biyectivo, se sigue que

$$\pi-\beta = \alpha+\gamma, \text{ entonces } \alpha+\beta+\gamma = \pi$$

y probamos lo que queríamos. ■

Ejercicio 3.E (De Parcial)**.** Sea $(\mathbb{R}^2, \langle\, \cdot, \cdot\,\rangle)$ el $\mathbb{R}$-espacio euclídeo respecto del cual triángulo de vértices $u := \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $v := \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $w := \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, es un triángulo equilatero de área $\sqrt{3}$.

Hallar $\langle\, x, y\,\rangle$ para todo $x, y \in \mathbb{R}^2$.

Antes de resolver el ejercicio, notar que el producto interno $\langle\, \cdot, \cdot\,\rangle$ nunca puede ser el producto interno canónico de $\mathbb{R}^2$. Si lo fuera, el ángulo entre los vectores v y w sería $\frac{\pi}{2}$ y por lo tanto, nunca podrían ser los vértices de un triángulo equilátero.

Vamos a usar la notación de la Figura (3.1). Con esa notación $a = w - v$, $b = u - w = -w$ y $c = v - u = v$. Como su nombre lo indica, un triángulo equilátero tiene sus lados iguales. Por lo tanto

$$\|a\| = \|b\| = \|c\|.$$

Por otra parte, los ángulos de un triángulo equilátero también son iguales. De hecho, como $a + b + c = 0$, se sigue que $b = -a - c$. Entonces

$$\cos\alpha = \frac{\langle\, c, -b\,\rangle}{\|b\|\|c\|} = \frac{\langle\, c, a+c\,\rangle}{\|a\|^2} = \frac{\langle\, a, c\,\rangle}{\|a\|^2} + \frac{\|c\|^2}{\|a\|^2} = 1 + \frac{\langle\, a, c\,\rangle}{\|a\|^2},$$
$$\cos\gamma = \frac{\langle\, a, -b\,\rangle}{\|a\|\|b\|} = \frac{\langle\, a, a+c\,\rangle}{\|a\|^2} = \frac{\langle\, a, c\,\rangle}{\|a\|^2} + \frac{\|a\|^2}{\|a\|^2} = 1 + \frac{\langle\, a, c\,\rangle}{\|a\|^2}.$$

Entonces $\cos\alpha = \cos\gamma$ y como en $[0, \pi]$ el coseno es inyectivo, tenemos que $\alpha = \gamma$. De manera similar, se prueba que $\alpha = \beta$. Por lo tanto, tenemos que $\alpha = \beta = \gamma$.

Ahora sí, resolvamos el ejercicio.

Dem. Notar que el conjunto $\{v, w\}$ es una base (no ortogonal) del espacio $\mathbb{R}^2$. Entonces, dados $x, y \in \mathbb{R}^2$, vale que $x = \alpha v + \beta w$ e $y = \alpha' v + \beta' w$, con $\alpha, \beta, \alpha', \beta' \in \mathbb{R}$. Por lo tanto, usando la linealidad del producto interno, tenemos que

$$\langle\, x, y\,\rangle = \langle\, \alpha v + \beta w, \alpha' v + \beta' w'\,\rangle = \alpha\alpha'\|v\|^2 + \alpha\beta'\langle\, v, w\,\rangle + \beta\alpha'\langle\, w, v\,\rangle + \beta\beta'\|w\|^2.$$

Por lo tanto, si hallamos $\|v\|^2$, $\|w\|^2$ y $\langle\, v, w\,\rangle = \langle\, w, v\,\rangle$ tendremos $\langle\, x, y\,\rangle$ para todo $x, y \in \mathbb{R}^2$.

Como el triángulo en cuestión es equilátero, vimos que sus lados y sus ángulos son iguales. Por el **Ejercicio 3.6**, tenemos que $\alpha+\beta+\gamma=\pi$. Por lo tanto

$$\alpha=\beta=\gamma=\frac{\pi}{3}.$$

Recordemos que

$$\sqrt{3}=A_T=\frac{\|c\|\|b\|\sin\alpha}{2}=\frac{\|b\|^2\sin\frac{\pi}{3}}{2}=\frac{\|b\|^2\sqrt{3}}{4}.$$

Despejando, se sigue que $\|b\|=2$. Entonces $\|w\|=\|-b\|=2$ y $\|v\|=\|c\|=\|b\|=2$. Finalmente, tenemos que $\langle\, v,w\,\rangle=\langle\, c,-b\,\rangle=\cos(\alpha)\|b\|\|c\|=\cos(\frac{\pi}{3})4=2$.

Entonces, si $x=\alpha v+\beta w$ e $y=\alpha' v+\beta' w$, con $\alpha,\beta,\alpha',\beta'\in\mathbb{R}$, concluimos que

$$\langle\, x,y\,\rangle=\alpha\alpha'\|v\|^2+\alpha\beta'\langle\, v,w\,\rangle+\beta\alpha'\langle\, w,v\,\rangle+\beta\beta'\|w\|^2=4\alpha\alpha'+2(\alpha\beta'+\beta\alpha')+4\beta\beta'.$$

■

3.4. Subespacio ortogonal, sistemas y bases ortonormales

Sea $(\mathbb{V},\langle\,\cdot,\cdot\,\rangle)$ un $\mathbb{K}$-espacio euclídeo y sea $A\subseteq\mathbb{V}$ un conjunto no vacío de $\mathbb{V}$, el *subespacio ortogonal a A*, denotado por $A^\perp$, se define por

$$A^\perp:=\{x\in\mathbb{V}:\langle\, x,a\,\rangle=0 \text{ para todo } a\in A\}.$$

Las siguientes son algunas propiedades importantes del subespacio ortogonal.

Proposición 3.4.1. *Sea $(\mathbb{V},\langle\,\cdot,\cdot\,\rangle)$ un $\mathbb{K}$-espacio euclídeo y sea $A\subseteq\mathbb{V}$ un conjunto no vacío de $\mathbb{V}$. Entonces:*

1. *$A^\perp$ es un subespacio de $\mathbb{V}$.*
2. *$A\cap A^\perp=\{0_V\}$.*
3. *Si $B\subseteq\mathbb{V}$ es un conjunto no vacío tal que $B\subseteq A$ entonces $A^\perp\subseteq B^\perp$.*

Dem. 1. : Por un lado, es claro que $0_{\mathbb{V}}\in A^\perp$, pues para todo $a\in A$,

$$\langle\, 0_{\mathbb{V}},a\,\rangle=\langle\, 0\cdot 0_{\mathbb{V}},a\,\rangle=0\,\langle\, 0_{\mathbb{V}},a\,\rangle=0.$$

Por otra parte, sean $x,y\in A^\perp$ y $\alpha\in\mathbb{K}$, entonces vale que $\langle\, x,a\,\rangle=\langle\, y,a\,\rangle=0$, para todo $a\in A$. Por lo tanto,

$$\langle\, \alpha x+y,a\,\rangle=\alpha\,\langle\, x,a\,\rangle+\langle\, y,a\,\rangle=0,$$

para todo $a\in A$. Entonces $\alpha x+y\in A^\perp$ y con eso concluimos que $A^\perp$ es un subespacio (observar que juntamos en una misma demostración la prueba de que si $x,y\in A^\perp$ y $\alpha\in\mathbb{K}$ entonces $x+y\in A^\perp$ y $\alpha x\in A^\perp$).

2. : Sea $x\in A\cap A^\perp$ entonces $x\in A$ y $x\in A^\perp$, como $x\in A^\perp$ vale que $\langle\, x,a\,\rangle=0$ para todo $a\in A$. En particular, como $x\in A$, vale que $\langle\, x,x\,\rangle=0$ entonces, por definición de producto interno, $x=0_{\mathbb{V}}$. Por lo tanto $A\cap A^\perp=\{0_{\mathbb{V}}\}$.

3. : Sea $x\in A^\perp$, entonces $\langle\, x,a\,\rangle=0$, para todo $a\in A$. Sea $b\in B$ entonces vale que $\langle\, x,b\,\rangle=0$, pues como $B\subseteq A$, tenemos que $b\in A$. Por lo tanto $x\in B^\perp$ y concluimos que $A^\perp\subseteq B^\perp$. ■

La siguiente es una simple observación que nos resultará muy útil.

Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo y sea $\mathcal{S} \subseteq \mathbb{V}$ un subespacio de $\mathbb{V}$ tal que $\mathcal{S} = gen\{v_1, v_2, \cdots, v_r\}$. Entonces

$$w \in \mathcal{S}^\perp \text{ si y sólo si } \langle w, v_1 \rangle = \langle w, v_2 \rangle = \cdots = \langle w, v_r \rangle = 0. \tag{3.8}$$

Es decir, para ver que $w \in \mathcal{S}^\perp$ basta ver que w es ortogonal a cada generador de $\mathcal{S}$.

De hecho, si $w \in \mathcal{S}^\perp$ entonces $\langle w, s \rangle = 0$ para todo $s \in \mathcal{S}$. Entonces, como $v_1, v_2, \cdots, v_r \in \mathcal{S}$ es claro que se cumple que $\langle w, v_1 \rangle = \langle w, v_2 \rangle = \cdots = \langle w, v_r \rangle = 0$.

Recíprocamente, supongamos que $\langle w, v_1 \rangle = \langle w, v_2 \rangle = \cdots = \langle w, v_r \rangle = 0$ y sea $s \in \mathcal{S}$. Entonces $s = a_1 v_1 + a_2 v_2 + \cdots + a_r v_r$ con $a_1, a_2, \cdots, a_r \in \mathbb{K}$. Entonces

$$\langle w, s \rangle = \langle w, a_1 v_1 + a_2 v_2 + \cdots + a_r v_r \rangle = \overline{a_1} \langle w, v_1 \rangle + \overline{a_2} \langle w, v_2 \rangle + \cdots + \overline{a_r} \langle w, v_r \rangle = 0,$$

como s era cualquier vector de $\mathcal{S}$, concluimos que $w \in \mathcal{S}^\perp$ y probamos lo que queríamos.

Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo. Un conjunto de vectores no nulos $\{u_1, u_2, \cdots, u_n\} \subseteq \mathbb{V}$ se llama:

- *Sistema ortogonal* cuando

$$\langle u_i, u_j \rangle = 0 \text{ para todo } i \neq j.$$

- *Sistema ortonormal* cuando

$$\langle u_i, u_j \rangle = 0 \text{ para todo } i \neq j \text{ y } \|u_i\| = 1 \text{ para todo } i \in \{1, \cdots, n\}.$$

Finalmente, llamaremos *base ortogonal (ortonormal)* de $\mathbb{V}$ a los sistemas ortogonales (ortonormales) que generan $\mathbb{V}$.

Veamos un ejemplo:

Ejercicio 3.F. Comprobar que los siguientes sistemas de vectores son ortonormales en su correspondiente espacio euclídeo:

- El sistema $\{e^{ikt} : k \in \mathbb{Z}\}$ en el espacio $C([-\pi, \pi])$ con el producto interno definido por

$$\langle f, g \rangle := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\overline{g(t)} dt.$$

Antes de resolver este ejercicio, recordar que si $z = a + ib \in \mathbb{C}$, con $a, b \in \mathbb{R}$, entonces

$$e^z = e^{ax}(\cos(b) + i \sin(b)).$$

Es fácil ver que valen las siguientes propiedades (verificarlas):

- Si $w, z \in \mathbb{C}$ entonces $e^z e^w = e^{z+w}$.
- Si $z \in \mathbb{C}$ entonces $\overline{e^z} = e^{\overline{z}}$.
- Si $t \in \mathbb{R}$ y $z = a + ib \in \mathbb{C}$ con $a, b \in \mathbb{R}$, entonces

$$\begin{aligned} \frac{d(e^{zt})}{dt} &= (e^{(a+ib)t})' = (e^{at+ibt})' = (e^{at}(\cos(bt) + i\sin(bt)))' \\ &= ae^{at}(\cos(bt) + i\sin(bt))) + e^{at}(-b\sin(bt) + ib\cos(bt)) \\ &= (a + ib)e^{at}(\cos(bt) + i\sin(bt))) = ze^{zt}. \end{aligned}$$

- Si $t \in \mathbb{R}$ y $z = a + ib \in \mathbb{C}$ (no nulo) con $a, b \in \mathbb{R}$, usando el item anterior, se sigue que $\int e^{zt} dt = \frac{e^{zt}}{z}$.
- Si $k \in \mathbb{Z}$, entonces $e^{ik\pi} = (-1)^k$.

Ahora sí, resolvamos el ejercicio.

Dem. Observar que $\{e^{ikt} : k \in \mathbb{Z}\} = \{\cdots, e^{-3it}, e^{-2it}, e^{-it}, 1, e^{it}, e^{2it}, e^{3it}, \cdots\}$. Veamos que ese conjunto es ortonormal. Para cada $k \in \mathbb{Z}$ llamemos $f_k(t) := e^{ikt}$. Entonces, si $k \neq l$, tenemos que

$$\langle f_k, f_l \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_k(t) \overline{f_l(t)} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikt} \overline{e^{ilt}} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikt} e^{-ilt} dt =$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(k-l)t} dt = \frac{1}{2\pi} \frac{1}{i(k-l)} e^{i(k-l)t}|_{-\pi}^{\pi} = \frac{1}{2\pi} \frac{1}{i(k-l)} (e^{i(k-l)\pi} - e^{-i(k-l)\pi}) = 0.$$

Por otra parte,

$$\langle f_k, f_k \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_k(t) \overline{f_k(t)} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikt} e^{-ikt} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{0} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} dt = \frac{1}{2\pi} 2\pi = 1.$$

Por lo tanto, el sistema $\{e^{ikt} : k \in \mathbb{Z}\}$ es ortonormal. ■

Ejercicio 3.G. Sea en $\mathbb{R}_2[x]$ el producto interno $\langle p, q \rangle := \int_{-1}^{1} \frac{1}{2} p(x) q(x) dx$. Hallar una base ortogonal de $\mathcal{S} = gen\{x^2\}^{\perp}$ y descomponer cada polinomio $p \in \mathbb{R}_2[x]$ en la forma $p = p_{\mathcal{S}} + p_{\mathcal{S}^{\perp}}$, con $p_{\mathcal{S}} \in \mathcal{S}$ y $p_{\mathcal{S}^{\perp}} \in \mathcal{S}^{\perp}$.

Dem. Por definición $\mathcal{S} = gen\{x^2\}^{\perp} = \{p \in \mathbb{R}_2[x] : \langle p, x^2 \rangle = 0\}$ (observar que usamos (3.8)). Entonces, $p \in \mathcal{S}$ si $p(x) = ax^2 + bx + c$, con $a, b, c \in \mathbb{R}$ y

$$0 = \langle p, x^2 \rangle = \langle ax^2 + bx + c, x^2 \rangle = \frac{1}{2} [\int_{-1}^{1} ax^4 + bx^3 + cx^2 dx] =$$

$$= \frac{1}{2} [a\frac{x^5}{5} + b\frac{x^4}{4} + c\frac{x^3}{3}]|_{-1}^{1} = \frac{1}{2} [a\frac{2}{5} + b \cdot 0 + c\frac{2}{3}] = a\frac{1}{5} + c\frac{1}{3}.$$

Entonces, despejando nos queda $a = -\frac{5}{3}c$. Volviendo a la expresión de p tenemos que,

$$p(x) = ax^2 + bx + c = -\frac{5}{3}cx^2 + bx + c = c(-\frac{5}{3}x^2 + 1) + bx,$$

con $b, c \in \mathbb{R}$. Entonces, $\mathcal{S} = gen\{-\frac{5}{3}x^2 + 1, x\}$. Observar que afortunadamente

$$\left\langle -\frac{5}{3}x^2 + 1, x \right\rangle = 0,$$

entonces $\mathcal{B}_{\mathcal{S}} = \{-\frac{5}{3}x^2 + 1, x\}$, es una base ortogonal de $\mathcal{S}$.

Por otra parte, es claro que (sino hacer la cuenta), $\mathcal{S}^{\perp} = (gen\{x^2\}^{\perp})^{\perp} = (gen\{-\frac{5}{3}x^2 + 1, x\})^{\perp} = gen\{x^2\}$.

Por último sea $p(x) = ax^2 + bx + c$, cualquier vector de $\mathbb{R}_2[x]$ y $\alpha, \beta, \gamma \in \mathbb{R}$, tales que

$$p(x) = ax^2 + bx + c = \alpha x^2 + \beta(-\frac{5}{3}x^2 + 1) + \gamma x.$$

Es fácil ver que $\gamma = b, \beta = c$ y $\alpha = a + \frac{5}{3}c$. Entonces

$$p(x) = (a + \frac{5}{3}c)x^2 + c(-\frac{5}{3}x^2 + 1) + bx.$$

Si llamamos $p_{\mathcal{S}} := c(-\frac{5}{3}x^2 + 1) + bx \in \mathcal{S}$ y $p_{\mathcal{S}^{\perp}} := (a + \frac{5}{3}c)x^2 \in \mathcal{S}^{\perp}$, tenemos que $p = p_{\mathcal{S}} + p_{\mathcal{S}^{\perp}}$. Como vimos que $\mathcal{S} \cap \mathcal{S}^{\perp} = \{0\}$ la descomposición que acabamos de encontrar es única.

■

Ejercicio 3.H (De Parcial)**.** En el $\mathbb{R}$-espacio vectorial $\mathbb{R}_2[x]$ se define la función $\Phi : \mathbb{R}_2[x]\times\mathbb{R}_2[x] \to \mathbb{R}$ como

$$\Phi(p,q) := \int_0^2 (1-x)f(x)g(x)dx.$$

Sea $\mathcal{S} = gen\{(1-x)^2\}$. Entonces:

a) Hallar una base de $\mathcal{T} := \{p \in \mathbb{R}_2[x] :\ \Phi(p,q) = 0 \text{ para todo } q \in \mathcal{S}\}$.

b) Probar que $\mathcal{S} \subseteq \mathcal{T}$.

c) Decidir si Φ define un producto interno en $\mathbb{R}_2[x]$.

Dem. *a)* : Busquemos una base de $\mathcal{T}$. Tenemos que $p \in \mathcal{T}$ si $p \in \mathbb{R}_2[x]$, es decir $p(x) = ax^2 + bx + c$ con $a, b, c \in \mathbb{R}$ y $\Phi(p,q) = 0$ para todo $q \in \mathcal{S}$. Si $q \in \mathcal{S}$, entonces $q(x) = \alpha(1-x)^2$, con $\alpha \in \mathbb{R}$. Entonces, $\Phi(p, \alpha(1-x)^2) = 0$, para todo $\alpha \in \mathbb{R}$. En particular, lo anterior vale para $\alpha = 1$. Entonces

$$\begin{aligned}
0 = \Phi(ax^2+bx+c,(1-x)^2) &= \int_0^2 (1-x)(ax^2+bx+c)(1-x)^2 dx \\
&= \int_0^2 (1-x)^3(ax^2+bx+c)dx \\
&= a[-\frac{x}{6} + \frac{3x^5}{5} - \frac{3x^4}{4} + \frac{x^3}{3}] + b[-\frac{x}{5} + \frac{3x^4}{4} - x^3 + \frac{x^2}{2}] + c[-\frac{(1-x)^4}{4}]|_0^2 \\
&= \frac{-4a}{5} + \frac{-2b}{5} + c\cdot 0 = -\frac{4a}{5} - \frac{2b}{5}.
\end{aligned}$$

Despejando, nos queda $b = -\frac{4a}{5}\frac{5}{2} = -2a$. Por lo tanto $p(x) = ax^2 - 2ax + c = a(x^2-2x) + c\cdot 1$, con $a, c \in \mathbb{R}$. Por lo tanto,

$$\mathcal{T} = gen\{x^2 - 2x, 1\}$$

y una base de $\mathcal{T}$ puede ser $\mathcal{B}_{\mathcal{T}} = \{x^2 - 2x, 1\}$.

b) : Veamos que $\mathcal{S} \subseteq \mathcal{T}$. Si $q \in \mathcal{S}$ entonces $q(x) = \alpha(x-1)^2$, con $\alpha \in \mathbb{R}$. Observar que

$$q(x) = \alpha(x-1)^2 = \alpha(x^2-2x+1) = \alpha(x^2-2x) + \alpha 1 \in \mathcal{T}$$

y se sigue que $\mathcal{S} \subseteq \mathcal{T}$.

c) : Φ NO define un producto interno. Por ejemplo, si $p(x) = 1$, entonces

$$\Phi(p,p) = \int_0^2 (1-x)p(x)^2 dx = \int_0^2 (1-x)1dx = x - \frac{x^2}{2}|_0^2 = 0.$$

Sin embargo $p(x) = 1$ es decir $p \neq 0$ (el polinomio nulo). Por lo tanto Φ no define un producto interno en $\mathbb{R}_2[x]$. ■

3.5. Proyección ortogonal

Sean $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo, $\mathcal{S}$ un subespacio de $\mathbb{V}$ y $v \in \mathbb{V}$, decimos que $\hat{v}$ es la *proyección ortogonal* del vector v sobre $\mathcal{S}$ si verifica:

$$\hat{v} \in \mathcal{S} \text{ y } v - \hat{v} \in \mathcal{S}^{\perp}.$$

Teorema 3.5.1. *Sean $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo, $\mathcal{S}$ un subespacio de $\mathbb{V}$ (de dimensión finita) y $v \in \mathbb{V}$, entonces $\hat{v}$ (la proyección ortogonal del vector v sobre $\mathcal{S}$) existe y es única. Más aún, si $\mathcal{B}_{\mathcal{S}} = \{v_1, v_2, \cdots, v_r\}$ es una* ***base ortogonal*** *de $\mathcal{S}$. Entonces*

$$\hat{v} = \frac{\langle v, v_1 \rangle}{\|v_1\|^2}v_1 + \frac{\langle v, v_2 \rangle}{\|v_2\|^2}v_2 + \cdots + \frac{\langle v, v_r \rangle}{\|v_r\|^2}v_r.$$

Dem. Primero veamos que de existir la proyección ortogonal de v sobre $\mathcal{S}$, esta es única. De hecho, supongamos que $\hat{v}$ y $\hat{w}$ son dos proyecciones ortogonales de v sobre $\mathcal{S}$. Entonces

$$\begin{aligned}\|\hat{v}-\hat{w}\|^2 &= \langle \hat{v}-\hat{w}, \hat{v}-\hat{w} \rangle = \langle \hat{v}-\hat{w}, \hat{v}-v+v-\hat{w} \rangle \\ &= \langle \hat{v}-\hat{w}, \hat{v}-v \rangle + \langle \hat{v}-\hat{w}, v-\hat{w} \rangle = 0+0=0,\end{aligned}$$

donde usamos que como $\hat{v}, \hat{w} \in \mathcal{S}$ y $\mathcal{S}$ es un subespacio entonces $\hat{v}-\hat{w} \in \mathcal{S}$ y, como $\hat{v}$ y $\hat{w}$ son dos proyecciones ortogonales de v, entonces $v-\hat{v}$ y $v-\hat{w}$ son elementos de $\mathcal{S}^\perp$. Entonces, como $\|\hat{v}-\hat{w}\|=0$, se sigue que $\hat{v}=\hat{w}$ y la proyección ortogonal es única.

Por otra parte, notar que el vector $\frac{\langle v,v_1 \rangle}{\|v_1\|^2}v_1 + \frac{\langle v,v_2 \rangle}{\|v_2\|^2}v_2 + \cdots + \frac{\langle v,v_r \rangle}{\|v_r\|^2}v_r \in \mathcal{S}$ pues $\{v_1, v_2, \cdots, v_r\}$ es una base de $\mathcal{S}$. Además, para $1 \le i \le r$, tenemos que

$$\begin{aligned}\left\langle v - \frac{\langle v,v_1 \rangle}{\|v_1\|^2}v_1 + \frac{\langle v,v_2 \rangle}{\|v_2\|^2}v_2 + \cdots + \frac{\langle v,v_r \rangle}{\|v_r\|^2}v_r, v_i \right\rangle &= \langle v, v_i \rangle - \frac{\langle v,v_i \rangle}{\|v_i\|^2}\langle v_i, v_i \rangle \\ &= \langle v, v_i \rangle - \frac{\langle v,v_i \rangle}{\|v_i\|^2}\|v_i\|^2 = 0,\end{aligned}$$

donde usamos que, como $\{v_1, v_2, \cdots, v_r\}$ es una **base ortogonal** de $\mathcal{S}$, entonces $\langle v_i, v_j \rangle = 0$ para $i \neq j$. Entonces, por (3.8), el vector $v - \frac{\langle v,v_1 \rangle}{\|v_1\|^2}v_1 + \frac{\langle v,v_2 \rangle}{\|v_2\|^2}v_2 + \cdots + \frac{\langle v,v_r \rangle}{\|v_r\|^2}v_r \in \mathcal{S}^\perp$. Por lo tanto, $\hat{v} = \frac{\langle v,v_1 \rangle}{\|v_1\|^2}v_1 + \frac{\langle v,v_2 \rangle}{\|v_2\|^2}v_2 + \cdots + \frac{\langle v,v_r \rangle}{\|v_r\|^2}v_r$ es la (única) proyección ortogonal de v sobre $\mathcal{S}$. ∎

Sean $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo, $\mathcal{S} \subseteq \mathbb{V}$ un subespacio de dimensión finita y $\mathcal{B}_\mathcal{S} = \{v_1, v_2, \cdots, v_r\}$ una **base ortogonal** de $\mathcal{S}$. Definimos la función $P_\mathcal{S} : \mathbb{V} \to \mathbb{V}$ como la función que a cada elemento $v \in \mathbb{V}$ le asigna su proyección ortogonal sobre $\mathcal{S}$, es decir

$$P_\mathcal{S}(v) := \hat{v} = \frac{\langle v,v_1 \rangle}{\|v_1\|^2}v_1 + \frac{\langle v,v_2 \rangle}{\|v_2\|^2}v_2 + \cdots + \frac{\langle v,v_r \rangle}{\|v_r\|^2}v_r.$$

La función $P_\mathcal{S}$ está bien definida porque a cada elemento $v \in \mathbb{V}$ le corresponde un único elemento de $\mathbb{V}$ que es $\hat{v}$. Por otra parte, $P_\mathcal{S}$ es una transformación lineal, pues si $v, w \in \mathbb{V}$ y $\alpha, \beta \in \mathbb{K}$ entonces

$$\begin{aligned}P_\mathcal{S}(\alpha v + \beta w) &= \sum_{i=1}^r \frac{\langle \alpha v + \beta w, v_i \rangle}{\|v_i\|^2}v_i = \alpha \sum_{i=1}^r \frac{\langle v, v_i \rangle}{\|v_i\|^2}v_i + \beta \sum_{i=1}^r \frac{\langle w, v_i \rangle}{\|v_i\|^2}v_i \\ &= \alpha P_\mathcal{S}(v) + \beta P_\mathcal{S}(w).\end{aligned}$$

Notar que como $P_\mathcal{S}(v) = \sum_{i=1}^r \frac{\langle v,v_i \rangle}{\|v_i\|^2}v_i \in \mathcal{S}$, entonces

$$\begin{aligned}P_\mathcal{S}(P_\mathcal{S}(v)) &= P_\mathcal{S}(\sum_{i=1}^r \frac{\langle v,v_i \rangle}{\|v_i\|^2}v_i) = \sum_{j=1}^r\sum_{i=1}^r \left\langle \frac{\langle v,v_i \rangle v_i}{\|v_i\|^2}, v_j \right\rangle \frac{v_j}{\|v_j\|^2} \\ &= \sum_{j=1}^r\sum_{i=1}^r \frac{\langle v,v_i \rangle}{\|v_i\|^2}\langle v_i, v_j \rangle \frac{v_j}{\|v_j\|^2} = \sum_{i=1}^r \frac{\langle v,v_i \rangle}{\|v_i\|^2}v_i = P_\mathcal{S}(v),\end{aligned}$$

donde usamos que $\langle v_i, v_j \rangle = 0$ si $i \neq j$. Por lo tanto $P_\mathcal{S}^2 = P_\mathcal{S} \circ P_\mathcal{S} = P_\mathcal{S}$, entonces $P_\mathcal{S}$ es un proyector. Finalmente, observar que

$$\mathrm{Im}(P_\mathcal{S}) = \mathcal{S} \text{ y } \mathrm{Nu}(P_\mathcal{S}) = \mathcal{S}^\perp.$$

De hecho, como $P_\mathcal{S}$ es un proyector, vimos que $v \in \mathrm{Im}(P_\mathcal{S})$ si y sólo si $v = P_\mathcal{S}(v) = \hat{v} \in \mathcal{S}$, donde para la última igualdad usamos la definición de proyección ortogonal del vector v sobre el subespacio $\mathcal{S}$. Meditar por qué esto demuestra que $\mathrm{Im}(P_\mathcal{S}) = \mathcal{S}$.

Por último, es claro que si $v \in \mathcal{S}^\perp$, entonces $P_\mathcal{S}(v) = 0_\mathbb{V}$, porque $\langle v, v_i \rangle = 0$ para $i = 1, 2, \cdots, r$. Por otra parte, si $v \in \mathrm{Nu}(P_\mathcal{S})$ entonces $P_\mathcal{S}(v) = \hat{v} = 0_\mathbb{V}$, entonces, por definición de proyección ortogonal del vector v sobre el subespacio $\mathcal{S}$, tenemos que $v = v - 0_\mathbb{V} = v - \hat{v} \in \mathcal{S}^\perp$. Concluimos entonces que $\mathrm{Nu}(P_\mathcal{S}) = \mathcal{S}^\perp$.

Vamos a remarcar todas estas propiedades que probamos para tenerlas presentes:

Sean $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo, $\mathcal{S} \subseteq \mathbb{V}$ un subespacio de dimensión finita y $\mathcal{B}_{\mathcal{S}} = \{v_1, v_2, \cdots, v_r\}$ una **base ortogonal** de $\mathcal{S}$. La función $P_{\mathcal{S}} : \mathbb{V} \to \mathbb{V}$ definida como

$$P_{\mathcal{S}}(v) := \hat{v} = \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle v, v_2 \rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle v, v_r \rangle}{\|v_r\|^2} v_r, \tag{3.9}$$

es una transformación lineal que cumple:

- $P_{\mathcal{S}}^2 = P_{\mathcal{S}}$, es decir $P_{\mathcal{S}}$ es un proyector.
- $\mathrm{Im}(P_{\mathcal{S}}) = \mathcal{S}$ y $\mathrm{Nu}(P_{\mathcal{S}}) = \mathcal{S}^{\perp}$. Entonces, claramente $\mathbb{V} = \mathrm{Im}(T) \oplus \mathrm{Nu}(T)$ y además $\mathrm{Im}(P_{\mathcal{S}}) \perp \mathrm{Nu}(P_{\mathcal{S}})$.
- Como $P_{\mathcal{S}}$ es un proyector y tenemos que $\mathrm{Im}(P_{\mathcal{S}}) = \mathcal{S}$ y $\mathrm{Nu}(P_{\mathcal{S}}) = \mathcal{S}^{\perp}$, entonces

$$P_{\mathcal{S}^{\perp}} = I_{\mathbb{V}} - P_{\mathcal{S}}. \tag{3.10}$$

Observar que si T es un proyector, tenemos que $\mathrm{Im}(T)$ y $\mathrm{Nu}(T)$ están en suma directa (no necesariamente son ortogonales). Por otra parte, si T es un proyector ortogonal entonces T es un proyector y además $\mathrm{Im}(T) = \mathrm{Nu}(T)^{\perp}$, es decir $\mathrm{Im}(T)$ y $\mathrm{Nu}(T)$ son ortogonales entre sí (en particular están en suma directa). Todo proyector ortogonal es un proyector, pero obviamente no todo proyector es un proyector ortogonal. Por ejemplo, en $\mathbb{R}^2$ con el producto interno canónico, definimos $T : \mathbb{R}^2 \to \mathbb{R}^2$, como

$$T(\begin{bmatrix} 1 \\ 1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad T(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Entonces T es un proyector (verificarlo). Pero como $\mathrm{Nu}(T) = gen\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$ NO es ortogonal a $\mathrm{Im}(T) = gen\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$, concluimos que T NO es un proyector ortogonal.

A partir de la existencia de la proyección ortogonal podemos enunciar las siguientes propiedades:

Proposición 3.5.2. *Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo de dimensión finita y $\mathcal{S} \subseteq \mathbb{V}$ un subespacio. Entonces:*

1. $\mathbb{V} = \mathcal{S} \oplus \mathcal{S}^{\perp}$.
2. $\dim(\mathcal{S}^{\perp}) + \dim(\mathcal{S}) = \dim(\mathbb{V})$.
3. $(\mathcal{S}^{\perp})^{\perp} = \mathcal{S}$.

Notar que la Proposición 3.5.2 sólo vale cuando $\mathcal{S}$ es un subespacio y $\mathbb{V}$ es de dimensión finita.

Dem. 1. : Es claro que $\mathcal{S} \cap \mathcal{S}^{\perp} = \{0_{\mathbb{V}}\}$. Por otra parte, si $v \in \mathbb{V}$ tenemos que

$$v = P_{\mathcal{S}}(v) + (v - P_{\mathcal{S}}(v)).$$

Por definición de proyección ortogonal, $P_{\mathcal{S}}(v) \in \mathcal{S}$ y $v - P_{\mathcal{S}}(v) \in \mathcal{S}^{\perp}$. Por lo tanto $v \in \mathcal{S} \oplus \mathcal{S}^{\perp}$ y se sigue que $\mathbb{V} \subseteq \mathcal{S} \oplus \mathcal{S}^{\perp}$. Como la otra inclusión siempre vale, tenemos que $\mathbb{V} = \mathcal{S} \oplus \mathcal{S}^{\perp}$.

2. : Usando el item 1. y el Teorema de la dimensión para suma de subespacios, tenemos que

$$\dim(\mathbb{V}) = \dim(\mathcal{S} \oplus \mathcal{S}^{\perp}) = \dim(\mathcal{S}) + \dim(\mathcal{S}^{\perp}).$$

3. : Es claro que $\mathcal{S} \subseteq (\mathcal{S}^\perp)^\perp$, pues si $s \in \mathcal{S}$ entonces $\langle s, t \rangle = 0$ para todo $t \in \mathcal{S}^\perp$ entonces $s \in (\mathcal{S}^\perp)^\perp$. Por otra parte, usando el item 2. dos veces, tenemos que

$$\dim((\mathcal{S}^\perp)^\perp) = \dim(\mathbb{V}) - \dim(\mathcal{S}^\perp) = \dim(\mathcal{S}).$$

Entonces, como $\mathcal{S} \subseteq (\mathcal{S}^\perp)^\perp$ y $\dim(\mathcal{S}) = \dim((\mathcal{S}^\perp)^\perp)$ se sigue que $(\mathcal{S}^\perp)^\perp = \mathcal{S}$. ■

Recordemos la definición de distancia de un punto a un subespacio: sean $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo, $\mathcal{S}$ un subespacio de $\mathbb{V}$ y $v \in \mathbb{V}$. Se define la *distancia del punto v al subespacio $\mathcal{S}$* como

$$d(v, \mathcal{S}) := \text{ínf}\{\|v - s\| : s \in \mathcal{S}\}.$$

Donde ínf denota el ínfimo del conjunto en cuestión. Notar que como $\mathcal{S}$ es un subespacio, entonces el conjunto $\{\|v - s\| : s \in \mathcal{S}\} \neq \emptyset$ y además, $\|v - s\| \geq 0$ para todo $s \in \mathcal{S}$. Por lo tanto, el conjunto en cuestión es no vacío y acotado inferiormente, entonces existe el ínfimo y $d(v, \mathcal{S})$ queda bien definida.

En el siguiente teorema probaremos que si $\mathcal{S}$ es de dimensión finita, ese ínfimo en realidad es un mínimo y el vector que realiza el mínimo es justamente $\hat{v} = P_{\mathcal{S}}(v)$.

Teorema 3.5.3. *Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo de dimensión finita y $\mathcal{S} \subseteq \mathbb{V}$ un subespacio. Entonces*

$$d(v, \mathcal{S}) = \|v - P_{\mathcal{S}}(v)\| = \|P_{\mathcal{S}^\perp}(v)\|.$$

Dem. Sea $s \in \mathcal{S}$, entonces como $v - P_{\mathcal{S}}(v) \in \mathcal{S}^\perp$ y $P_{\mathcal{S}}(v) - s \in \mathcal{S}$ podemos aplicar el Teorema de Pitágoras:

$$\|v - s\|^2 = \|v - P_{\mathcal{S}}(v) + P_{\mathcal{S}}(v) - s\|^2 = \|v - P_{\mathcal{S}}(v)\|^2 + \|P_{\mathcal{S}}(v) - s\|^2 \geq \|v - P_{\mathcal{S}}(v)\|^2.$$

Tomando raíz en la ecuación anterior, se sigue que $\|v - s\| \geq \|v - P_{\mathcal{S}}(v)\|$ para todo $s \in \mathcal{S}$. Además como $P_{\mathcal{S}}(v) \in \mathcal{S}$, concluimos que $P_{\mathcal{S}}(v)$ realiza el mínimo del conjunto $\{\|v - s\| : s \in \mathcal{S}\}$.

Por lo tanto

$$d(v, \mathcal{S}) = \|v - P_{\mathcal{S}}(v)\| = \|P_{\mathcal{S}^\perp}(v)\|,$$

donde para la última igualdad usamos que por (3.10), tenemos que $P_{\mathcal{S}^\perp} = I_{\mathbb{V}} - P_{\mathcal{S}}$. Entonces

$$P_{\mathcal{S}^\perp}(v) = I_{\mathbb{V}}(v) - P_{\mathcal{S}}(v) = v - P_{\mathcal{S}}(v).$$

■

El siguiente es un ejemplo de cálculo de la proyección ortogonal $P_{\mathcal{S}}$.

Ejercicio 3.I. En $\mathbb{R}^{2\times 2}$ con el producto interno $\langle A, B \rangle = \text{tr}(B^T A)$, consideramos el subespacio $\mathcal{S}$ de todas las matrices simétricas.

a) Dada $X \in \mathbb{R}^{2\times 2}$, hallar la expresión de $P_{\mathcal{S}}(X)$.

b) Hallar la proyección ortogonal de $B = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ sobre $\mathcal{S}$. Calcular $d(B, \mathcal{S})$.

Dem. *a)* : El subespacio $\mathcal{S}$ en cuestión es $\mathcal{S} = \{A \in \mathbb{R}^{2\times 2} : A^T = A\}$. Si $A \in \mathcal{S}$, entonces $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ con $a, b, c, d \in \mathbb{R}$ y $A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} = A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Entonces, $b = c$. Por lo tanto $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ y tenemos que

$$\mathcal{S} = gen\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \}.$$

Observar que

$$\left\langle \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\rangle = \operatorname{tr}(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}) = \operatorname{tr}(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}) = 0,$$

$$\left\langle \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle = \operatorname{tr}(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}) = \operatorname{tr}(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}) = 0,$$

$$\left\langle \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle = \operatorname{tr}(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}) = \operatorname{tr}(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}) = 0.$$

Entonces, $B_{\mathcal{S}} = \{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \}$ es una base ortogonal de $\mathcal{S}$. Y podemos usar la ecuación (3.9). Entonces

$$P_{\mathcal{S}}(X) = \frac{\left\langle X, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\rangle}{\| \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \|^2} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{\left\langle X, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\rangle}{\| \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \|^2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$+ \frac{\left\langle X, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle}{\| \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \|^2} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Haciendo cuentas, tenemos que $\| \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \|^2 = 1$, $\| \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \|^2 = 1$ y $\| \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \|^2 = 2$. Por otra parte, si $X = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} \in \mathbb{R}^{2\times 2}$, se sigue que

$$\left\langle \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\rangle = \operatorname{tr}(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}) = \operatorname{tr}(\begin{bmatrix} x_1 & x_2 \\ 0 & 0 \end{bmatrix}) = x_1.$$

Operando de manera similar, tenemos que $\left\langle \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\rangle = x_2 + x_3$ y $\left\langle \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle = x_4$. Entonces, reemplazando, nos queda:

$$P_{\mathcal{S}}(X) = x_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + (\frac{x_2 + x_3}{2}) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} x_1 & \frac{x_2+x_3}{2} \\ \frac{x_2+x_3}{2} & x_4 \end{bmatrix}.$$

Observar que (obviamente) $P_{\mathcal{S}}(X)$ siempre es simétrica pues por definición $P_{\mathcal{S}}(X) \in \mathcal{S}$.

Otra manera de resolver este ejercicio es observar que por un lado $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathcal{S}^\perp$ (hacer la cuenta) y como, por la Proposición 3.5.2, $\dim(\mathcal{S}^\perp) = \dim(\mathbb{R}^{2\times 2}) - \dim(\mathcal{S}) = 4 - 3 = 1$, tenemos que $\mathcal{S}^\perp = gen\{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \}$. Una base de $\mathcal{S}^\perp$ (ortogonal) es $B_{\mathcal{S}^\perp} = \{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \}$. Entonces (verificarlo haciendo las cuentas),

$$P_{\mathcal{S}^\perp}(X) = \frac{\left\langle X, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\rangle}{\| \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \|^2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{x_2-x_3}{2} \\ \frac{x_3-x_2}{2} & 0 \end{bmatrix}.$$

Entonces, por la ecuación (3.10),

$$P_{\mathcal{S}}(X) = X - P_{\mathcal{S}^\perp}(X) = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} - \begin{bmatrix} 0 & \frac{x_2-x_3}{2} \\ \frac{x_3-x_2}{2} & 0 \end{bmatrix} = \begin{bmatrix} x_1 & \frac{x_2+x_3}{2} \\ \frac{x_2+x_3}{2} & x_4 \end{bmatrix},$$

y (por suerte) obtuvimos el mismo resultado.

$b)$: Usando la fórmula que acabamos de probar, tenemos que

$$P_{\mathcal{S}}(B) = P_{\mathcal{S}}(\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}) = \begin{bmatrix} 1 & \frac{-1+1}{2} \\ \frac{-1+1}{2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Finalmente,

$$d(B, \mathcal{S}) = \|B - P_{\mathcal{S}}(B)\| = \| \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \| = \| \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \| =$$

$$[\mathrm{tr}(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix})]^{1/2} = [\mathrm{tr}(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix})]^{1/2} = \sqrt{2}.$$

■

3.6. Matriz de la proyección ortogonal y distancia mínima

Si $\mathbb{V}$ es dimensión finita y B es cualquier base de $\mathbb{V}$, podemos estar interesados en calcular la matriz de la transformación lineal $P_{\mathcal{S}}$ en dicha base B. Recuerden que si $B = \{w_1, w_2, \cdots, w_n\}$ es una base del espacio vectorial $\mathbb{V}$ entonces, las columnas de $[P_{\mathcal{S}}]_B^B$ nos quedan

$$[P_{\mathcal{S}}]_B^B = [\ [P_{\mathcal{S}}(w_1)]^B\ [P_{\mathcal{S}}(w_2)]^B\ \cdots\ [P_{\mathcal{S}}(w_n)]^B] \in \mathbb{K}^{n\times n},$$

donde para $i = 1, 2, \cdots, n$ podemos calcular $P_{\mathcal{S}}(w_i)$ usando la fórmula (3.9).

Veamos dos ejemplos.

Ejercicio 3.J.

a) En $\mathbb{R}^{2\times 2}$ con el producto interno canónico de $\mathbb{R}^{2\times 2}$, hallar la matriz con respecto a la base canónica de $\mathbb{R}^{2\times 2}$ de la proyección ortogonal de $\mathbb{R}^{2\times 2}$ sobre $\mathcal{S}$ el subespacio de todas las matrices simétricas.

b) En $\mathbb{R}^4$ con el producto interno canónico de $\mathbb{R}^4$, hallar la matriz con respecto a la base canónica de $\mathbb{R}^4$ de la proyección ortogonal de $\mathbb{R}^4$ sobre $\mathcal{S} = gen\{[1\ -1\ 0\ 0]^T, [1\ 1\ -2\ 0]^T\}$.

Dem. $a)$: En el **Ejercicio 3.I**, hallamos la expresión de $P_{\mathcal{S}} : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ la proyección ortogonal de $\mathbb{R}^{2\times 2}$ sobre $\mathcal{S}$ (el subespacio de todas las matrices simétricas) y nos quedó

$$P_{\mathcal{S}}(X) = \begin{bmatrix} x_1 & \frac{x_2+x_3}{2} \\ \frac{x_2+x_3}{2} & x_4 \end{bmatrix}.$$

Recordar que $E = \{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \}$ es la base canónica de $\mathbb{R}^{2\times 2}$. Entonces, haciendo cuentas tenemos que:

$$P_{\mathcal{S}}(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad [P_{\mathcal{S}}(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix})]^E = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

$$P_{\mathcal{S}}(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}) = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} \quad \text{y} \quad [P_{\mathcal{S}}(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix})]^E = \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}.$$

$$P_{\mathcal{S}}(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}) = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} \quad \text{y} \quad [P_{\mathcal{S}}(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix})]^E = \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}.$$

$$P_{\mathcal{S}}(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad [P_{\mathcal{S}}(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix})]^E = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Por lo tanto

$$[P_{\mathcal{S}}]_E^E = [\ [P_{\mathcal{S}}(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix})]^E\ [P_{\mathcal{S}}(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix})]^E\ [P_{\mathcal{S}}(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix})]^E\ [P_{\mathcal{S}}(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix})]^E]$$
$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

b) : Para calcular $P_{\mathcal{S}}(x)$ en cualquier vector $x \in \mathbb{R}^4$ con la fórmula (3.9), necesitamos una base ortogonal de $\mathcal{S}$. Afortunadamente, $B := \{[1\ \ -1\ \ 0\ \ 0]^T, [1\ \ 1\ \ -2\ \ 0]^T\}$ es una base de $\mathcal{S}$ que además es ortogonal (en el producto interno canónico de $\mathbb{R}^4$). Recordar que $E = \{[1\ 0\ 0\ 0]^T, [0\ 1\ 0\ 0]^T, [0\ 0\ 1\ 0]^T, [0\ 0\ 0\ 1]^T\}$ es la base canónica de $\mathbb{R}^4$. Entonces, haciendo cuentas tenemos que:

$$\begin{aligned} P_{\mathcal{S}}([1\ 0\ 0\ 0]^T) &= \frac{\left\langle [1\ 0\ 0\ 0]^T, [1\ \ -1\ 0\ 0]^T \right\rangle}{\|[1\ \ -1\ 0\ 0]^T\|^2}[1\ \ -1\ 0\ 0]^T \\ &+ \frac{\left\langle [1\ 0\ 0\ 0]^T, [1\ 1\ \ -2\ 0]^T \right\rangle}{\|[1\ 1\ \ -2\ 0]^T\|^2}[1\ 1\ \ -2\ 0]^T \\ &= \frac{1}{2}[1\ \ -1\ 0\ 0]^T + \frac{1}{6}[1\ 1\ \ -2\ 0]^T = [\frac{2}{3}\ \ -\frac{1}{3}\ \ -\frac{1}{3}\ \ 0]^T, \end{aligned}$$

y $[P_{\mathcal{S}}([1\ 0\ 0\ 0]^T)]^E = \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \\ 0 \end{bmatrix}.$

Operando de manera similar, nos queda que:

$$P_{\mathcal{S}}([0\ 1\ 0\ 0]^T) = [-\tfrac{1}{3}\ \tfrac{2}{3}\ \ -\tfrac{1}{3}\ 0]^T \quad \text{y} \quad [P_{\mathcal{S}}(0\ 1\ 0\ 0]^T)]^E = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ -\frac{1}{3} \\ 0 \end{bmatrix}.$$

$$P_{\mathcal{S}}([0\ 0\ 1\ 0]^T) = [-\tfrac{1}{3}\ \ -\tfrac{1}{3}\ \tfrac{2}{3}\ 0]^T \quad \text{y} \quad [P_{\mathcal{S}}(0\ 0\ 1\ 0]^T)]^E = \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \\ 0 \end{bmatrix}.$$

$$P_{\mathcal{S}}([0\ 0\ 0\ 1]^T) = [0\ 0\ 0\ 0]^T \quad \text{y} \quad [P_{\mathcal{S}}(0\ 0\ 0\ 1]^T)]^E = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Por lo tanto

$$[P_{\mathcal{S}}]_E^E = [\ [P_{\mathcal{S}}([1\ 0\ 0\ 0]^T)]^E\ [P_{\mathcal{S}}([0\ 1\ 0\ 0]^T)]^E\ [P_{\mathcal{S}}([0\ 0\ 1\ 0]^T)]^E\ [P_{\mathcal{S}}([0\ 0\ 0\ 1]^T)]^E]$$
$$= \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

■

Veamos a continuación un ejemplo de cálculo de matriz de la proyección ortogonal en la base canónica de $\mathbb{R}^3$ donde NO usamos el producto interno canónico.

Ejercicio 3.8. En $\mathbb{R}^3$ con el producto interno $\langle \cdot, \cdot \rangle$ definido por

$$\langle x, y \rangle = y^T \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} x$$

se consideran los subespacios $\mathcal{S}_1 = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$ y $\mathcal{S}_2\{x \in \mathbb{R}^3 : x_1 - x_3 = 0\}$.

a) Hallar las matrices con respecto a la base canónica de las proyecciones ortogonales de $\mathbb{R}^3$ sobre $\mathcal{S}_1^\perp$ y sobre $\mathcal{S}_2^\perp$.

b) Sea $b = [1\ 1\ 2]^T$. Hallar la distancia de b al subespacio $\mathcal{S}_1^\perp$ y la distancia de b al subespacio $\mathcal{S}_2$.

c) Hallar el conjunto de todos los $x \in \mathbb{R}^3$ cuya distancia a $\mathcal{S}_1$ coincide con su distancia a $\mathcal{S}_2$.

Dem. $a)$: Notar que $\mathcal{S}_1 = gen\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \}$. Por otra parte, por (3.8),

$$\mathcal{S}_1^\perp = \{x \in \mathbb{R}^3 : \left\langle x, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\rangle = \left\langle x, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\rangle = 0\}.$$

Por lo tanto, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathcal{S}_1^\perp$ si

$$\begin{aligned} 0 = \left\langle x, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\rangle &= [1\ \ -1\ 0] \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [1\ \ -1\ 0] \begin{bmatrix} 2x_1 - 2x_2 \\ -2x_1 + 5x_2 + 4x_3 \\ 4x_2 + 6x_3 \end{bmatrix} \\ &= 2x_1 - 2x_2 - (-2x_1 + 5x_2 + 4x_3) = 4x_1 - 7x_2 - 4x_3. \end{aligned}$$

y

$$\begin{aligned} 0 = \left\langle x, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\rangle &= [0\ 1\ \ -1] \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [0\ 1\ \ -1] \begin{bmatrix} 2x_1 - 2x_2 \\ -2x_1 + 5x_2 + 4x_3 \\ 4x_2 + 6x_3 \end{bmatrix} \\ &= -2x_1 + 5x_2 + 4x_3 - (4x_2 + 6x_3) = -2x_1 + x_2 - 2x_3. \end{aligned}$$

Resolviendo el sistema $4x_1 - 7x_2 - 4x_3 = -2x_1 + x_2 - 2x_3 = 0$. Nos queda que

$$\mathcal{S}_1^\perp = gen\{ \begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix} \}.$$

Consideremos el conjunto $B := \{ \begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \}$. Entonces, como $\mathcal{S}_1 \oplus \mathcal{S}_1^\perp = \mathbb{R}^3$, vale que B es una base (no ortogonal) de $\mathbb{R}^3$. Sea P la proyección ortogonal sobre $\mathcal{S}_1^\perp$, es decir el proyector que actúa sobre $\mathcal{S}_1^\perp$ en la dirección $\mathcal{S}_1$. Entonces, como $P(\begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix}) = \begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix}$, $P(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, $P(\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, es claro que

$$[P]_B^B = [[P(\begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix})]^B\ [P(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix})]^B\ [P(\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix})]^B] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Haciendo el cambio de base correspondiente, vale que

$$[P]_E^E = [M]_B^E\ [P]_B^B\ [P]_E^B.$$

Es fácil ver que la matriz de cambio de base de B a E es $[M]_B^E = \begin{bmatrix} -9 & 1 & 0 \\ -8 & -1 & 1 \\ 5 & 0 & -1 \end{bmatrix}$. Además, $[M]_E^B = ([M]_B^E)^{-1} = \frac{1}{12}\begin{bmatrix} -1 & -1 & -1 \\ 3 & -9 & -9 \\ -5 & -5 & -17 \end{bmatrix}$. Por lo tanto

$$[P]_E^E = \begin{bmatrix} -9 & 1 & 0 \\ -8 & -1 & 1 \\ 5 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{1}{12}\begin{bmatrix} -1 & -1 & -1 \\ 3 & -9 & -9 \\ -5 & -5 & -17 \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 9 & 9 & 9 \\ 8 & 8 & 8 \\ -5 & -5 & -5 \end{bmatrix}.$$

Por otra parte, notar que $\mathcal{S}_2 = gen\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\}$. Además, por (3.8),

$$\mathcal{S}_2^{\perp} = \{x \in \mathbb{R}^3 : \left\langle x, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\rangle = \left\langle x, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\rangle = 0\}.$$

Por lo tanto, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathcal{S}_2^{\perp}$ si

$$\begin{aligned} 0 &= \left\langle x, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\rangle = [1\ 0\ 1] \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [1\ 0\ 1] \begin{bmatrix} 2x_1 - 2x_2 \\ -2x_1 + 5x_2 + 4x_3 \\ 4x_2 + 6x_3 \end{bmatrix} \\ &= 2x_1 - 2x_2 + (4x_2 + 6x_3) = 2x_1 + 2x_2 + 6x_3. \end{aligned}$$

y

$$\begin{aligned} 0 &= \left\langle x, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\rangle = [0\ 1\ 0] \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [0\ 1\ 0] \begin{bmatrix} 2x_1 - 2x_2 \\ -2x_1 + 5x_2 + 4x_3 \\ 4x_2 + 6x_3 \end{bmatrix} \\ &= -2x_1 + 5x_2 + 4x_3. \end{aligned}$$

Resolviendo el sistema $2x_1 + 2x_2 + 6x_3 = -2x_1 + 5x_2 + 4x_3 = 0$. Nos queda que

$$\mathcal{S}_2^{\perp} = gen\{\begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix}\}.$$

Consideremos el conjunto $C := \{\begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\}$. Entonces, como $\mathcal{S}_2 \oplus \mathcal{S}_2^{\perp} = \mathbb{R}^3$, vale que C es una base (no ortogonal) de $\mathbb{R}^3$. Sea Q la proyección ortogonal sobre $\mathcal{S}_2^{\perp}$, es decir el proyector que actúa sobre $\mathcal{S}_2^{\perp}$ en la dirección $\mathcal{S}_2$. Entonces, como $Q(\begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix}) = \begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix}$, $Q(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, $Q(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, es claro que

$$[Q]_C^C = [[Q(\begin{bmatrix}-11\\-10\\7\end{bmatrix})]^C \ [Q(\begin{bmatrix}1\\0\\1\end{bmatrix})]^C \ [Q(\begin{bmatrix}0\\1\\0\end{bmatrix})]^C] = \begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}.$$

Haciendo el cambio de base correspondiente, vale que

$$[Q]_E^E = [M]_C^E \ [Q]_C^C \ [P]_E^C.$$

Es fácil ver que la matriz de cambio de base de C a E es $[M]_C^E = \begin{bmatrix}-11&1&0\\-10&0&1\\7&1&0\end{bmatrix}$. Además, $[M]_E^C = ([M]_C^E)^{-1} = \frac{1}{18}\begin{bmatrix}-1&0&1\\7&0&11\\-10&18&10\end{bmatrix}$. Por lo tanto

$$[Q]_E^E = \begin{bmatrix}-11&1&0\\-10&0&1\\7&1&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}\frac{1}{18}\begin{bmatrix}-1&0&1\\7&0&11\\-10&18&10\end{bmatrix} = \frac{1}{18}\begin{bmatrix}11&0&-11\\10&0&-10\\-7&0&7\end{bmatrix}.$$

b): Notar que el conjunto $\{\begin{bmatrix}-9\\-8\\5\end{bmatrix}\}$ es una base (ortogonal) de $\mathcal{S}_1^\perp$. Entonces, aplicando (3.9),

$$P_{\mathcal{S}_1^\perp}(b) = P_{\mathcal{S}_1^\perp}(\begin{bmatrix}1\\1\\2\end{bmatrix}) = \frac{\left\langle \begin{bmatrix}1\\1\\2\end{bmatrix}, \begin{bmatrix}-9\\-8\\5\end{bmatrix}\right\rangle}{\| \begin{bmatrix}-9\\-8\\5\end{bmatrix} \|^2}\begin{bmatrix}-9\\-8\\5\end{bmatrix}$$

$$= \left(\frac{[-9 \ -8 \ 5]\begin{bmatrix}2&-2&0\\-2&5&4\\0&4&6\end{bmatrix}\begin{bmatrix}1\\1\\2\end{bmatrix}}{[-9 \ -8 \ 5]\begin{bmatrix}2&-2&0\\-2&5&4\\0&4&6\end{bmatrix}\begin{bmatrix}-9\\-8\\5\end{bmatrix}}\right)\begin{bmatrix}-9\\-8\\5\end{bmatrix} = \frac{-8}{24}\begin{bmatrix}-9\\-8\\5\end{bmatrix} = -\frac{1}{3}\begin{bmatrix}-9\\-8\\5\end{bmatrix}.$$

Entonces

$$d(b,\mathcal{S}_1^\perp)^2 = \|b - P_{\mathcal{S}_1^\perp}b\|^2 = \| \begin{bmatrix}1\\1\\2\end{bmatrix} + \frac{1}{3}\begin{bmatrix}-9\\-8\\5\end{bmatrix} \|^2 = \|\frac{1}{3}\begin{bmatrix}-6\\-5\\11\end{bmatrix}\|^2$$

$$= \frac{1}{9}[-6 \ -5 \ 11]\begin{bmatrix}2&-2&0\\-2&5&4\\0&4&6\end{bmatrix}\begin{bmatrix}-6\\-5\\11\end{bmatrix} = \frac{363}{9} = \frac{121}{3}.$$

Por lo tanto, $d(b,\mathcal{S}_1^\perp) = \sqrt{\frac{121}{3}} = \frac{11}{\sqrt{3}}$.

De manera similar, notar que $\{\begin{bmatrix}-11\\-10\\7\end{bmatrix}\}$ es una base (ortogonal) de $\mathcal{S}_2^\perp$. Entonces, por (3.9),

$$P_{\mathcal{S}_2^\perp}(b) = P_{\mathcal{S}_2^\perp}(\begin{bmatrix}1\\1\\2\end{bmatrix}) = \frac{\left\langle \begin{bmatrix}1\\1\\2\end{bmatrix}, \begin{bmatrix}-11\\-10\\7\end{bmatrix}\right\rangle}{\| \begin{bmatrix}-11\\-10\\7\end{bmatrix} \|^2}\begin{bmatrix}-11\\-10\\7\end{bmatrix}$$

$$= \left(\frac{[-11\ \ -10\ 7] \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}}{[-11\ \ -10\ 7] \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix}} \right) \begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix} = \frac{2}{36} \begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix} = \frac{1}{18} \begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix}.$$

Entonces

$$d(b,\mathcal{S}_2)^2 = \|b - P_{\mathcal{S}_2} b\|^2 = \|P_{\mathcal{S}_2^\perp} b\|^2 = \| \frac{1}{18} \begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix} \|^2$$

$$= \frac{1}{324} [-11\ \ -10\ 7] \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix} = \frac{1}{9}.$$

Por lo tanto, $d(b,\mathcal{S}_2) = \frac{1}{3}$.

$c)$: Sea $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$ tal que

$$d(x,\mathcal{S}_1) = \|P_{\mathcal{S}_1^\perp} x\| = d(x, S_2) = \|P_{\mathcal{S}_2^\perp} x\|.$$

Entonces, aplicando la fórmula (3.9), tenemos

$$P_{\mathcal{S}_1^\perp}(x) = \left(\frac{[-9\ \ -8\ 5] \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}{[-9\ \ -8\ 5] \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix}} \right) \begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix} =$$

$$= \frac{-2(x_1 + x_2 + x_3)}{24} \begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix} = -\frac{x_1 + x_2 + x_3}{12} \begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix}.$$

$$P_{\mathcal{S}_2^\perp}(x) = \left(\frac{[-11\ \ -10\ 7] \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}{[-11\ \ -10\ 7] \begin{bmatrix} 2 & -2 & 0 \\ -2 & 5 & 4 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix}} \right) \begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix} =$$

$$= \frac{2(-x_1 + x_3)}{36} \begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix} = \frac{-x_1 + x_3}{18} \begin{bmatrix} -11 \\ -10 \\ 7 \end{bmatrix}.$$

Entonces

$$\|P_{\mathcal{S}_1^\perp}(x)\| = \| - \frac{x_1 + x_2 + x_3}{12} \begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix} \| = \frac{|x_1 + x_2 + x_3|}{12} \| \begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix} \|$$

$$= \frac{|x_1 + x_2 + x_3|}{12} \sqrt{24} = \frac{\sqrt{6}}{6} |x_1 + x_2 + x_3|.$$

$$\|P_{\mathcal{S}_2^\perp}(x)\| = \|\frac{-x_1+x_3}{18}\begin{bmatrix}-11\\-10\\7\end{bmatrix}\| = \frac{|-x_1+x_3|}{18}\|\begin{bmatrix}-11\\-10\\7\end{bmatrix}\|$$
$$= \frac{|-x_1+x_3|}{18}6 = \frac{1}{3}|-x_1+x_3|.$$

Igualando, nos queda resolver el sistema

$$|-x_1+x_3| = \frac{\sqrt{6}}{2}|x_1+x_2+x_3|.$$

Supongamos que $x_1+x_2+x_3>0$ y $x_3-x_1>0$ ó $x_1+x_2+x_3<0$ y $x_3-x_1<0$. Entonces, tenemos que $-x_1+x_3=\frac{\sqrt{6}}{2}(x_1+x_2+x_3)$. Resolviendo, nos queda que

$$x_2 = x_1(-\frac{\sqrt{6}}{3}-1)+x_3(\frac{\sqrt{6}}{3}-1).$$

Notar que, con lo que acabamos de resolver tenemos que

$$x_1+x_2+x_3 = x_1(1-\frac{\sqrt{6}}{3}-1)+x_3(1+\frac{\sqrt{6}}{3}-1) = \frac{\sqrt{6}}{3}(x_3-x_1).$$

Entonces, en este caso, $x_1+x_2+x_3>0$ si y sólo si $x_3-x_1>0$ y $x_1+x_2+x_3<0$ si y sólo si $x_3-x_1<0$.

Ahora, supongamos que $x_1+x_2+x_3>0$ y $x_3-x_1<0$ ó $x_1+x_2+x_3<0$ y $x_3-x_1>0$. Entonces, tenemos que $-x_1+x_3=-\frac{\sqrt{6}}{2}(x_1+x_2+x_3)$. Resolviendo, nos queda que

$$x_2 = x_1(\frac{\sqrt{6}}{3}-1)+x_3(-\frac{\sqrt{6}}{3}-1).$$

Notar que, con lo que acabamos de resolver tenemos que

$$x_1+x_2+x_3 = x_1(1+\frac{\sqrt{6}}{3}-1)+x_3(1-\frac{\sqrt{6}}{3}-1) = \frac{\sqrt{6}}{3}(x_1-x_3).$$

Observar que si $x_1+x_2+x_3>0$ no puede ocurrir que $x_3-x_1<0$. De la misma manera, si $x_1+x_2+x_3<0$ no puede ocurrir que $x_3-x_1>0$. Por lo tanto, este caso nunca ocurre y lo descartamos.

Conclusión:

$$x = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}x_1\\x_1(-\frac{\sqrt{6}}{3}-1)+x_3(\frac{\sqrt{6}}{3}-1)\\x_3\end{bmatrix} = x_1\begin{bmatrix}1\\-\frac{\sqrt{6}}{3}-1\\0\end{bmatrix}+x_3\begin{bmatrix}0\\\frac{\sqrt{6}}{3}-1\\1\end{bmatrix},$$

con $x_1,x_3\in\mathbb{R}$. Por lo tanto $d(x,\mathcal{S}_1)=d(x,S_2)$ si y sólo si

$$x\in gen\{\begin{bmatrix}1\\-\frac{\sqrt{6}}{3}-1\\0\end{bmatrix},\begin{bmatrix}0\\\frac{\sqrt{6}}{3}-1\\1\end{bmatrix}\}.$$

■

En el siguiente ejercicio vamos a aplicar el concepto de distancia mínima de un punto a un subespacio.

Ejercicio 3.10. En $\mathbb{R}_2[x]$ se considera el producto interno definido por

$$\langle\, p, q\,\rangle := \int_0^1 p(x)q(x)dx.$$

Calcular

$$\min_{a,b\in\mathbb{R}} \int_0^1 (x^2 - ax - b)^2 dx.$$

Dem. Primero, observemos que

$$\int_0^1 (x^2 - ax - b)^2 dx = \|x^2 - ax - b\|^2.$$

Donde (abusando de la notación) con $\|x^2 - ax - b\|^2$ queremos decir $\|p\|^2$ con $p(x) = x^2 - ax - b$.
Entonces

$$\begin{aligned}
\min_{a,b\in\mathbb{R}} \int_0^1 (x^2 - ax - b)^2 dx &= \min\{\|x^2 - ax - b\|^2 : a, b \in \mathbb{R}\} \\
&= \min\{\|x^2 - p\|^2 : p \in gen\{1, x\}\} \\
&= d(x^2, gen\{1, x\})^2 \\
&= \|x^2 - P_{gen\{1,x\}}(x^2)\|^2 = \|P_{gen\{1,x\}^\perp}(x^2)\|^2,
\end{aligned}$$

donde usamos el Teorema 3.5.3 .

Entonces, busquemos una base del subespacio $gen\{1, x\}^\perp$. De hecho, $q \in gen\{1, x\}^\perp$ si y sólo si $q \in \mathbb{R}_2[x]$, es decir $q(x) = \alpha x^2 + \beta x + \gamma$ para ciertos $\alpha, \beta, \gamma \in \mathbb{R}$ y

$$\langle\, q, 1\,\rangle = \langle\, q, x\,\rangle = 0.$$

Entonces

$$0 = \langle\, q, 1\,\rangle = \int_0^1 [1 \cdot (\alpha x^2 + \beta x + \gamma)]dx = \alpha\frac{x^3}{3} + \beta\frac{x^2}{2} + \gamma x|_0^1 = \frac{\alpha}{3} + \frac{\beta}{2} + \gamma.$$

$$0 = \langle\, q, x\,\rangle = \int_0^1 [x \cdot (\alpha x^2 + \beta x + \gamma)]dx = \alpha\frac{x^4}{4} + \beta\frac{x^3}{3} + \gamma\frac{x^2}{2}|_0^1 = \frac{\alpha}{4} + \frac{\beta}{3} + \frac{\gamma}{2}.$$

Resolviendo el sistema $0 = \frac{\alpha}{3} + \frac{\beta}{2} + \gamma = \frac{\alpha}{4} + \frac{\beta}{3} + \frac{\gamma}{2}$. Nos queda que $\alpha = -\beta$ y $\gamma = -\frac{\beta}{6}$. Entonces, volviendo a la expresión de q, tenemos que $q(x) = -\beta x^2 + \beta x - \frac{\beta}{6}$ con $\beta \in \mathbb{R}$. Por lo tanto

$$gen\{1, x\}^\perp = gen\{-6x^2 + 6x - 1\}.$$

Una base (ortogonal) del subspacio $gen\{1, x\}^\perp$ puede ser $\{-6x^2 + 6x - 1\}$. Entonces, por (3.9),

$$\begin{aligned}
P_{gen\{1,x\}^\perp}(x^2) &= \left(\frac{\langle\, x^2, -6x^2 + 6x - 1\,\rangle}{\| - 6x^2 + 6x - 1\|^2}\right)(-6x^2 + 6x - 1) \\
&= \left(\frac{\int_0^1 x^2(-6x^2 + 6x - 1)dx}{\int_0^1 (-6x^2 + 6x - 1)^2 dx}\right)(-6x^2 + 6x - 1) = \frac{5}{30}(-6x^2 + 6x - 1) \\
&= \frac{1}{6}(-6x^2 + 6x - 1).
\end{aligned}$$

Por lo tanto,

$$\|P_{gen\{1,x\}^\perp}(x^2)\|^2 = \|\frac{1}{6}(-6x^2 + 6x - 1)\|^2 = \frac{1}{36}\int_0^1 (-6x^2 + 6x - 1)^2 dx = \frac{1}{36 \cdot 5} = \frac{1}{180}.$$

Conclusión

$$\min_{a,b\in\mathbb{R}} \int_0^1 (x^2-ax-b)^2dx = \frac{1}{180}.$$

■

En el próximo ejercicio veremos la diferencia entre un proyector ortogonal y un proyector que no es ortogonal.

Ejercicio 3.12. Sean $\Pi_1, \Pi_2 \in \mathcal{L}(\mathbb{R}^2)$ las transformaciones lineales definidas por

$$\Pi_1(x) = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x, \quad \Pi_2(x) = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} x.$$

a) Comprobar que Π_1 y Π_2 son dos proyecciones, identificando en cada caso el subespacio sobre el que proyectan y la dirección en la qué lo hacen.

b) ¿Es Π_1 una proyección ortogonal?

b) ¿Es Π_2 una proyección ortogonal?

d) ¿Existe $x \in \mathbb{R}^2$ tal que $\|\Pi_1(x)\| > \|x\|$? Si la respuesta es afirmativa, exhibir al menos uno.

e) ¿Existe $x \in \mathbb{R}^2$ tal que $\|\Pi_2(x) > \|x\|$? Si la respuesta es afirmativa, exhibir al menos uno.

Dem. En este ejercicio consideramos $\mathbb{R}^2$ con el producto interno canónico.

a) : Sea $x \in \mathbb{R}^2$, entonces

$$\begin{aligned}\Pi_1^2(x) = \Pi_1(\Pi_1(x)) &= \Pi_1\left(\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x\right) \\ &= \left(\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right) x = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x = \Pi_1(x).\end{aligned}$$

Por lo tanto $\Pi_1^2 = \Pi_1$ y por **Ejercicio 2.23**, Π_1 es un proyector sobre $\mathrm{Im}(\Pi_1) = gen\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$ en la dirección $\mathrm{Nu}(\Pi_1) = gen\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\}$.

Por otra parte, sea $x \in \mathbb{R}^2$, entonces

$$\begin{aligned}\Pi_2^2(x) = \Pi_2(\Pi_2(x)) &= \Pi_2\left(\frac{1}{3}\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} x\right) \\ &= \left(\frac{1}{3}\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}\frac{1}{3}\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}\right) x = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} x = \Pi_2(x).\end{aligned}$$

Por lo tanto $\Pi_2^2 = \Pi_2$ y por **Ejercicio 2.23**, Π_2 es un proyector sobre $\mathrm{Im}(\Pi_2) = gen\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$ en la dirección $\mathrm{Nu}(\Pi_1) = gen\{\begin{bmatrix} 1 \\ -2 \end{bmatrix}\}$.

b) : Como $\Pi_1^2 = \Pi_1$ y además $\mathrm{Im}(\Pi_1) \perp \mathrm{Nu}(\Pi_1)$ entonces Π_1 es una proyección ortogonal.

c) : Como $\mathrm{Im}(\Pi_2) \not\perp \mathrm{Nu}(\Pi_2)$ entonces Π_2 NO es una proyección ortogonal.

d) : No existe $x \in \mathbb{R}^2$ tal que $\|\Pi_1(x)\| > \|x\|$. De hecho, como Π_1 es una proyección ortogonal, para cada $x \in \mathbb{R}^2$, vale que $\Pi_1(x) \perp (x - \Pi_1(x))$. Entonces, aplicando el Teorema de Pitágoras, se sigue que

$$\|x\|^2 = \|\Pi_1(x) + (x - \Pi_1(x))\|^2 = \|\Pi_1(x)\|^2 + \|x - \Pi_1(x)\|^2 \geq \|\Pi_1(x)\|^2.$$

Entonces, tomando raíz en la ecuación anterior, tenemos que para todo $x \in \mathbb{R}^2$,

$$\|x\| \geq \|\Pi_1(x)\|.$$

e) : Notar que $\|\begin{bmatrix} 2 \\ 1 \end{bmatrix}\| = \sqrt{5}$. Además, $\Pi_2(\begin{bmatrix} 2 \\ 1 \end{bmatrix}) = \frac{1}{3}\begin{bmatrix} 5 \\ 5 \end{bmatrix}$. Entonces $\|\frac{1}{3}\begin{bmatrix} 5 \\ 5 \end{bmatrix}\| = \frac{5\sqrt{2}}{3}$. Entonces

$$\|\begin{bmatrix} 1 \\ 2 \end{bmatrix}\| = \sqrt{5} < \frac{5\sqrt{2}}{3} = \|\Pi_2(\begin{bmatrix} 1 \\ 2 \end{bmatrix})\|.$$

■

El siguiente resultado permite caracterizar a los proyectores ortogonales.

Teorema 3.6.1. *Sea* $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ *un* $\mathbb{K}$*-espacio euclídeo y sea* $T \in \mathcal{L}(\mathbb{V})$ *un proyector, es decir* $T^2 = T$. *Verificar que las siguientes afirmaciones son equivalentes:*

a) T *es una proyección ortogonal,*

b) *para todo* $x, y \in \mathbb{V}$ *vale* $\langle T(x), y \rangle = \langle x, T(y) \rangle$ *,*

c) *para todo* $x \in \mathbb{V}$ *vale* $\|T(x)\| \leq \|x\|$.

Dem. *a*) $\Rightarrow$ *b*) : Supongamos que T es una proyección ortogonal y veamos que para todo $x, y \in \mathbb{V}$ vale $\langle T(x), y \rangle = \langle x, T(y) \rangle$.

Sean $x, y \in \mathbb{V}$, entonces como T es un proyector, vimos que vale que $\mathbb{V} = \mathrm{Im}(T) \oplus \mathrm{Nu}(T)$. Entonces $x = x_1 + x_2$ e $y = y_1 + y_2$, con (únicos) $x_1, y_1 \in \mathrm{Im}(T)$ y $x_2, y_2 \in \mathrm{Nu}(T)$. Entonces

$$T(x) = T(x_1 + x_2) = T(x_1) + T(x_2) = T(x_1) + 0_{\mathbb{V}} = T(x_1) = x_1 \text{ y}$$

$$T(y) = T(y_1 + y_2) = T(y_1) + T(y_2) = T(y_1) + 0_{\mathbb{V}} = T(y_1) = y_1,$$

donde usamos que como T es un proyector $T(z) = z$, para todo $z \in \mathrm{Im}(T)$ y como $x_1, y_1 \in \mathrm{Im}(T)$, tenemos que $T(x_1) = x_1$ y $T(y_1) = y_1$. Como T es una proyección ortogonal tenemos también que $\mathrm{Im}(T) \perp \mathrm{Nu}(T)$, entonces $\langle x_1, y_2 \rangle = \langle x_2, y_1 \rangle = 0$. Por lo tanto

$$\begin{aligned} \langle T(x), y \rangle &= \langle x_1, y_1 + y_2 \rangle = \langle x_1, y_1 \rangle + \langle x_1, y_2 \rangle \\ &= \langle x_1, y_1 \rangle + 0 = \langle x_1, y_1 \rangle + \langle x_2, y_1 \rangle = \langle x_1 + x_2, y_1 \rangle = \langle x, T(y) \rangle . \end{aligned}$$

b) $\Rightarrow$ *c*) : Supongamos que para todo $x, y \in \mathbb{V}$ vale $\langle T(x), y \rangle = \langle x, T(y) \rangle$ y veamos que para todo $x \in \mathbb{V}$ vale $\|T(x)\| \leq \|x\|$. Sea $x \in \mathbb{V}$, usando la hipótesis (es decir que vale a)) y que T es proyector, tenemos que

$$\langle T(x), x - T(x) \rangle = \langle x, T(x - T(x)) \rangle = \left\langle x, T(x) - T^2(x) \right\rangle = \langle x, T(x) - T(x) \rangle = \langle x, 0_{\mathbb{V}} \rangle = 0.$$

Por lo tanto, por el Teorema de Pitágoras (o haciendo la cuenta),

$$\|x\|^2 = \|T(x) + [x - T(x)]\|^2 = \|T(x)\|^2 + \|x - T(x)\|^2 \geq \|T(x)\|^2,$$

donde usamos que $\|x - T(x)\|^2 \geq 0$. Finalmente, tomando raíz (que es una función creciente), tenemos que para todo $x \in \mathbb{V}$ vale $\|T(x)\| \leq \|x\|$.

c) $\Rightarrow$ *d*) : Supongamos que para todo $x \in \mathbb{V}$, $\|T(x)\| \leq \|x\|$ y veamos que T es una proyección ortogonal. Como T es un proyector (por hipótesis), para probar que T es un proyector ortogonal sólo nos resta probar que $\mathrm{Nu}(T)^{\perp} = \mathrm{Im}(T)$.

Veamos primero que $\mathrm{Nu}(T)^{\perp} \subseteq \mathrm{Im}(T)$. Sea $x \in \mathrm{Nu}(T)^{\perp}$ queremos ver que $x \in \mathrm{Im}(T)$ es decir $T(x) = x$ (recordemos que como T es proyector, $x \in \mathrm{Im}(T)$ si y sólo si $T(x) = x$). Como $x - T(x) \in \mathrm{Nu}(T)$ pues $T(x - T(x)) = T(x) - T^2(x) = T(x) - T(x) = 0_{\mathbb{V}}$ y $x \in \mathrm{Nu}(T)^{\perp}$ se sigue que

$$0 = \langle x, x - T(x) \rangle = \langle x, x \rangle - \langle x, T(x) \rangle = \|x\|^2 - \langle x, T(x) \rangle .$$

Entonces, nos queda que

$$\|x\|^2 = \langle x, T(x) \rangle .$$

Entonces

$$\|x - T(x)\|^2 = \|x\|^2 - 2\,\mathrm{Re}\,\langle\, x, T(x)\,\rangle + \|T(x)\|^2 = \|x\|^2 - 2\|x\|^2 + \|T(x)\|^2 = \|T(x)\|^2 - \|x\|^2.$$

Usando la hipótesis, tenemos que

$$\|x - T(x)\|^2 = \|T(x)\|^2 - \|x\|^2 \leq 0.$$

Entonces, como siempre tenemos que $\|x - T(x)\|^2 \geq 0$, se sigue que $\|x - T(x)\|^2 = 0$, entonces $x - T(x) = 0_{\mathbb{V}}$, por lo tanto $T(x) = x$ y $x \in \mathrm{Im}(T)$.

Recíprocamente, supongamos que $x \in \mathrm{Im}(T)$. Por la Proposición 3.5.2, $\mathrm{Nu}(T) \oplus \mathrm{Nu}(T)^{\perp} = \mathbb{V}$ y tenemos que $x = x_1 + x_2$, con (únicos) $x_1 \in \mathrm{Nu}(T)$ y $x_2 \in \mathrm{Nu}(T)^{\perp}$. Entonces, por un lado $x = T(x)$ pues $x \in \mathrm{Im}(T)$. Por otro lado, acabamos de probar que $\mathrm{Nu}(T)^{\perp} \subseteq \mathrm{Im}(T)$, entonces, como $x_2 \in \mathrm{Nu}(T)^{\perp}$ tenemos que $x_2 \in \mathrm{Im}(T)$ y (nuevamente) como T es proyector, se sigue que $T(x_2) = x_2$. Por lo tanto

$$x = T(x) = T(x_1 + x_2) = T(x_1) + T(x_2) = 0_{\mathbb{V}} + T(x_2) = T(x_2) = x_2.$$

En conclusión,

$$x = x_2 \in \mathrm{Nu}(T)^{\perp},$$

entonces $x \in \mathrm{Nu}(T)^{\perp}$ y tenemos que $\mathrm{Im}(T) \subseteq \mathrm{Nu}(T)^{\perp}$. Por lo tanto, como probamos la doble inclusión, se sigue que $\mathrm{Nu}(T)^{\perp} = \mathrm{Im}(T)$ y T es un proyector ortogonal.

Como probamos *a*) $\Rightarrow$ *b*) $\Rightarrow$ *c*) $\Rightarrow$ *a*), se sigue que *a*), *b*) y *c*) son equivalentes. ■

3.7. Mínimos cuadrados

Sean $\mathbb{V}, \mathbb{W}$ dos $\mathbb{K}$-espacios vectoriales y sea $T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$. Dado $w_0 \in \mathbb{W}$, recordemos que en la guía de transformaciones lineales vimos que:

$$\text{existe } v_0 \in \mathbb{V} \text{ tal que } T(v_0) = w_0 \text{ si y sólo si } w_0 \in \mathrm{Im}(T).$$

Cuando $w_0 \notin \mathrm{Im}(T)$, sabemos que la ecuación $T(v) = w_0$ NO tiene solución. En ese caso, nos interesa obtener una solución aproximada (en algún sentido) de dicha ecuación.

Si $\mathbb{W}, \langle \cdot, \cdot \rangle$ es un $\mathbb{K}$-espacio euclídeo, el problema de *mínimos cuadrados* consiste en encontrar (si existen) todos los vectores $\hat{v} \in \mathbb{V}$ tales que la distancia de $T(\hat{v})$ a w_0 sea mínima, es decir, buscamos $\hat{v} \in \mathbb{V}$ tal que

$$\|w_0 - T(\hat{v})\| = \text{mín}\{\|w_0 - T(v)\| : v \in \mathbb{V}\}.$$

Notar que

$$\{\|w_0 - T(v)\| : v \in \mathbb{V}\} = \{\|w_0 - w\| : w \in \mathrm{Im}(T)\}$$

(verificarlo probando la doble inclusión).

Entonces,

$$\|w_0 - T(\hat{v})\| = \text{mín}\{\|w_0 - w\| : w \in \mathrm{Im}(T)\}.$$

Ya vimos en el Teorema 3.5.3, que si $\mathrm{Im}(T)$ es de dimensión finita, entonces el vector $P_{\mathrm{Im}(T)}(w_0)$ es el (único) vector de $\mathrm{Im}(T)$ que minimiza la distancia a w_0. Por lo tanto los vectores $\hat{v}$ que buscamos son las soluciones del sistema

$$T(v) = P_{\mathrm{Im}(T)}(w_0),$$

como $P_{\mathrm{Im}(T)}(w_0) \in \mathrm{Im}(T)$, la ecuación $T(v) = P_{\mathrm{Im}(T)}(w_0)$ siempre tiene solución y será única si y sólo si T es un monomorfismo.

El siguiente ejercicio y las siguientes propiedades las usaremos para resolver el **Ejercicio 3.13**.

Ejercicio 3.K (De Parcial)**.** En $\mathbb{R}^n$ con el producto interno canónico consideramos $\mathrm{nul}(A)$ donde A es una matriz de $\mathbb{R}^{m\times n}$. Probar que

$$\mathrm{nul}(A)^\perp = \mathrm{fil}(A)$$

y expresar las dimensiones de $\mathrm{nul}(A)$ y $\mathrm{nul}(A)^\perp$ en función del rango de A. ¿Qué forma tomaría el problema si en lugar de $\mathbb{R}$ apareciese $\mathbb{C}$?

Dem. Veamos que $\mathrm{nul}(A)^\perp = \mathrm{fil}(A) = \mathrm{col}(A^T)$.

Supongamos que $y \in \mathrm{fil}(A) = \mathrm{col}(A^T)$ entonces existe $x \in \mathbb{R}^m$ tal que $y = A^T x$ (estamos usando que como $A \in \mathbb{R}^{m\times n}$, $A^T \in \mathbb{R}^{n\times m}$). Recordemos que el producto interno canónico se define como $\langle u, v \rangle = v^T u = u^T v$, para $u, v \in \mathbb{R}^n$. Sea $z \in \mathrm{nul}(A)$, entonces $Az = 0$ y tenemos que

$$\langle y, z \rangle = \langle A^T x, z \rangle = (A^T x)^T z = x^T (A^T)^T z = x^T (Az) = 0.$$

Es decir, $\langle y, z \rangle = 0$ para todo $z \in \mathrm{nul}(A)$. Por lo tanto $y \in \mathrm{nul}(A)^\perp$ y se sigue que

$$\mathrm{fil}(A) = \mathrm{col}(A^T) \subseteq \mathrm{nul}(A)^\perp.$$

En la Proposición 3.5.2, probamos que

$$\dim(\mathrm{nul}(A)) + \dim(\mathrm{nul}(A)^\perp) = \dim(\mathbb{R}^n) = n.$$

Entonces, aplicando el Teorema de la Dimensión y recordando que $\mathrm{rg}(A) = \mathrm{rg}(A^T)$, se sigue que

$$\dim(\mathrm{nul}(A)^\perp) = n - \dim(\mathrm{nul}(A)) = \mathrm{rg}(A) = \mathrm{rg}(A^T) = \dim(\mathrm{col}(A^T)).$$

Como ya probamos que $\mathrm{col}(A^T) \subseteq \mathrm{nul}(A)^\perp$ y los subespacios tienen la misma dimensión, concluimos que $\mathrm{fil}(A) = \mathrm{col}(A^T) = \mathrm{nul}(A)^\perp$.

Si ahora consideramos $\mathbb{C}^n$ como $\mathbb{C}$-espacio euclídeo, en este caso el producto interno canónico se define como $\langle u, v \rangle = v^* u$, para $u, v \in \mathbb{C}^n$. Entonces, operando de la misma manera que en $\mathbb{R}^n$, nos quedaría que

$$\mathrm{nul}(A)^\perp = \mathrm{col}(A^*).$$

∎

Proposición 3.7.1. *Sea $A \in \mathbb{R}^{m\times n}$ y consideremos el espacio euclídeo canónico correspondiente. Entonces:*

1. $\mathrm{nul}(A^T A) = \mathrm{nul}(A)$.
2. $\mathrm{col}(A^T A) = \mathrm{col}(A^T)$.
3. $\mathrm{nul}(A) = \mathrm{col}(A^T)^\perp$.
4. $\mathrm{nul}(A^T) = \mathrm{col}(A)^\perp$.

Dem. 1. : Vamos a probar $\mathrm{nul}(A^T A) = \mathrm{nul}(A)$ viendo la doble inclusión. Ya sabemos que siempre vale $\mathrm{nul}(A) \subseteq \mathrm{nul}(A^T A)$ (verificarlo). Por otra parte, si $x \in \mathrm{nul}(A^T A)$ entonces $A^T A x = 0$. Entonces,

$$\|Ax\|^2 = (Ax)^T (Ax) = x^T A^T (Ax) = x^T (A^T A x) = x^T 0 = 0.$$

Entonces $Ax = 0$, $x \in \mathrm{nul}(A)$ y tenemos que $\mathrm{nul}(A^T A) \subseteq \mathrm{nul}(A)$. Como probamos la doble inclusión, se sigue que $\mathrm{nul}(A^T A) = \mathrm{nul}(A)$.

2. : Recordemos que en el **Ejercicio 3.K**, probamos que si $A \in \mathbb{R}^{m\times n}$ y $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ es el espacio euclídeo canónico entonces

$$\mathrm{nul}(A)^\perp = \mathrm{col}(A^T).$$

Por lo tanto, usando esa propiedad y que (acabamos de ver que) $\text{nul}(A^T A) = \text{nul}(A)$, tenemos que

$$\text{col}(A^T A) = \text{col}((A^T A)^T) = \text{nul}(A^T A)^\perp = \text{nul}(A)^\perp = \text{col}(A^T).$$

3. : Como $\text{nul}(A)$ es un subespacio, por la Proposición 3.5.2, tenemos que

$$\text{nul}(A) = (\text{nul}(A)^\perp)^\perp$$

y usando que $\text{nul}(A)^\perp = \text{col}(A^T)$ se sigue que

$$\text{nul}(A) = (\text{nul}(A)^\perp)^\perp = \text{col}(A^T)^\perp.$$

4. : Finalmente, por 3.,

$$\text{nul}(A^T) = \text{col}((A^T)^T)^\perp = \text{col}(A)^\perp.$$

■

En el siguiente ejercicio vamos a estudiar un caso particular del problema de mínimos cuadrados cuando consideramos matrices $A \in \mathbb{R}^{m\times n}$ y el espacio $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ euclídeo canónico.

Ejercicio 3.13. Sean $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ y $(\mathbb{R}^m, \langle\cdot,\cdot\rangle)$ los espacios euclídeos canónicos de dimensión n y m, respectivamente. Sean $A \in \mathbb{R}^{m\times n}$ y $b \in \mathbb{R}^m$.

Por definición, resolver la ecuación $Ax = b$ por mínimos cuadrados significa determinar el conjunto

$$\arg \min_{x\in\mathbb{R}^n} \|Ax - b\| := \{x \in \mathbb{R}^n : \|Ax - b\| = \min_{x\in\mathbb{R}^n} \|Ax - b\|\}.$$

En palabras, $\arg \min_{x\in\mathbb{R}^n} \|Ax - b\|$ es el conjunto de todos los $x \in \mathbb{R}^n$ cuyas imágenes por A minimizan la distancia al vector b.

a) Explicar por qué
$$\arg \min_{x\in\mathbb{R}^n} \|Ax - b\| = \{x \in \mathbb{R}^n : Ax = P_{\text{col}(A)}b\}.$$

b) Observar que $\{x \in \mathbb{R}^n : Ax = P_{\text{col}(A)}b\} \neq \emptyset$. Motivo por el cual para cualquier $b \in \mathbb{R}^m$ existe al menos una solución por mínimos cuadrados de la ecuación $Ax = b$.

c) Verificar que $\text{nul}(A^T) = \text{col}(A)^\perp$.

d) Utilizando que $b - P_{\text{col}(A)}(b) \perp \text{col}(A)$, concluir que
$$\arg \min_{x\in\mathbb{R}^n} \|Ax - b\| = \{x \in \mathbb{R}^n : A^T Ax = A^T b\}.$$

e) Verificar que $\text{nul}(A)^\perp = \text{fil}(A) = \text{col}(A^T)$.

f) Utilizando que todo $x \in \mathbb{R}^n$ se descompone de manera única en la forma $x = x_f + x_h$, con $x_f \in \text{fil}(A)$ y $x_h \in \text{nul}(A)$, deducir que la aplicación $x_f \to Ax_f$ es una biyección de $\text{fil}(A)$ en $\text{col}(A)$.

g) Concluir que para cada $b \in \mathbb{R}^m$ existe un único $x_f(b) \in \text{fil}(A)$ tal que
$$\arg \min_{x\in\mathbb{R}^n} \|Ax - b\| = \{x_f(b) + x_h : x_h \in \text{nul}(A)\}.$$

Utilizando el Teorema de Pitágoras observar que de todas las soluciones por mínimos cuadrados de la ecuación $Ax = b$, la solución $x_f(b)$ es la que tiene norma mínima.

h) Observar que cuando $\dim(\text{col}(A)) = n$ la matriz $A^T A$ es inversible y que por lo tanto, $\hat{x} = (A^T A)^{-1} A^T b$ es la única solución por mínimos cuadrados de la ecuación $Ax = b$. Como $A\hat{x} = P_{\text{col}(A)}b$ se deduce que $A(A^T A)^{-1} A$ es la matriz en base canónica de la proyección ortogonal sobre $\text{col}(A)$.

Dem. $a)$: Observar que

$$\min_{x\in\mathbb{R}^n}\|Ax-b\| = \min_{y\in\mathrm{col}(A)}\|y-b\| = d(b,\mathrm{col}(A)).$$

Por el Teorema 3.5.3, como $\mathrm{col}(A)$ es un subespacio de dimensión finita, entonces el vector $P_{\mathrm{col}(A)}b$ es el (único) vector de $\mathrm{col}(A)$ que minimiza la distancia a b. Por lo tanto, los vectores $x\in\mathbb{R}^n$ que buscamos son las soluciones de la ecuación

$$Ax = P_{\mathrm{col}(A)}b.$$

Conclusión, $\arg\min_{x\in\mathbb{R}^n}\|Ax-b\| = \{x\in\mathbb{R}^n : Ax = P_{\mathrm{col}(A)}b\}$.

$b)$: Como $P_{\mathrm{col}(A)}b \in \mathrm{col}(A)$, el sistema $Ax = P_{\mathrm{col}(A)}b$ siempre tiene solución. Por lo tanto, $\{x\in\mathbb{R}^n : Ax = P_{\mathrm{col}(A)}b\} \neq \emptyset$.

$c)$: Lo probamos en la Proposición 3.7.1.

$d)$: Por definición de proyección ortogonal, tenemos que

$$b - P_{\mathrm{col}(A)}(b) \in \mathrm{col}(A)^{\perp} = \mathrm{nul}(A^T),$$

donde usamos $c)$. Entonces

$$A^T(b - P_{\mathrm{col}(A)}b) = 0. \tag{3.11}$$

Recordar que los vectores $x\in\mathbb{R}^n$ que buscamos son las soluciones de la ecuación $Ax = P_{\mathrm{col}(A)}b$. Entonces, por (3.11), los vectores $x\in\mathbb{R}^n$ que buscamos son las soluciones de la ecuación $A^TAx = A^T(P_{\mathrm{col}(A)}b) = A^Tb$.

$e)$: Demostrado en el **Ejercicio 3.K**.

$f)$: Llamemos $\Phi : \mathrm{fil}(A) \to \mathrm{col}(A)$ a la aplicación $\Phi(x_f) = Ax_f$. Veamos que Φ es una biyección de $\mathrm{fil}(A)$ en $\mathrm{col}(A)$. Supongamos que $\Phi(x_f) = Ax_f = \Phi(x'_f) = Ax'_f$, para $x_f, x'_f \in \mathrm{fil}(A)$. Entonces

$$A(x_f - x'_f) = 0.$$

Por lo tanto $x_f - x'_f \in \mathrm{nul}(A)\cap\mathrm{fil}(A) = \mathrm{nul}(A)\cap\mathrm{nul}(A)^{\perp} = \{0\}$, donde usamos el ítem $e)$. Por lo tanto $x_f = x'_f$ y Φ es inyectiva. Por otra parte, sea $y\in\mathrm{col}(A)$, entonces existe $x\in\mathbb{R}^n$ tal que $Ax = y$. Descompongamos x de la forma $x = x_f + x_h$, con $x_f\in\mathrm{fil}(A) = \mathrm{nul}(A)^{\perp}$ y $x_h\in\mathrm{nul}(A)$. Entonces $Ax_h = 0$ y tenemos que

$$\Phi(x_f) = Ax_f = A(x_f + x_h) = Ax = y.$$

Por lo tanto Φ es sobreyectiva y entonces Φ es biyectiva.

$g)$: Los vectores $x\in\mathbb{R}^n$ que buscamos son las soluciones de la ecuación $Ax = P_{\mathrm{col}(A)}b$. Sea $x\in\mathbb{R}^n$ una solución de $Ax = P_{\mathrm{col}(A)}b$. Como $\mathbb{R}^n = \mathrm{nul}(A)^{\perp}\oplus\mathrm{nul}(A) = \mathrm{fil}(A)\oplus\mathrm{nul}(A)$, podemos escribir de manera única al vector x como

$$x = x_f + x_h,$$

con $x_f\in\mathrm{fil}(A)$ y $x_h\in\mathrm{nul}(A)$. Entonces como $Ax_h = 0$, tenemos que

$$Ax_f = Ax = P_{\mathrm{col}(A)}b.$$

Por lo tanto $x_f\in\mathrm{fil}(A)$ es una solución de cuadrados mínimos de $Ax = b$. Si existe otra solución de cuadrados mínimos $x'_f\in\mathrm{fil}(A)$ entonces $Ax'_f = P_{\mathrm{col}(A)}b = Ax_f$. Entonces $A(x_f - x'_f) = 0$ y tal como hicimos en $f)$, se sigue que $x_f - x'_f \in \mathrm{nul}(A)\cap\mathrm{fil}(A) = \mathrm{nul}(A)\cap\mathrm{nul}(A)^{\perp} = \{0\}$. Entonces $x_f = x'_f$ y x_f es la única solución de cuadrados mínimos de $Ax = b$ que pertenece a $\mathrm{fil}(A)$. Por lo tanto, concluimos que para cada $b\in\mathbb{R}^m$ existe un único $x_f(b)\in\mathrm{fil}(A)$ (el vector x_f depende de b) tal que

$$\arg\min_{x\in\mathbb{R}^n}\|Ax-b\| = \{x_f(b) + x_h : x_h\in\mathrm{nul}(A)\}.$$

Por último, sea $\hat{x}$ una solución de cuadrados mínimos de $Ax = b$. Entonces $A\hat{x} = P_{\text{col}(A)}b$. Entonces $A(\hat{x} - x_f) = P_{\text{col}(A)}b - P_{\text{col}(A)}b = 0$ y, tenemos que $\hat{x} - x_f \in \text{nul}(A)$. Además, como $x_f \in \text{nul}(A)^\perp$, podemos aplicar el Teorema de Pitátogoras y concluir que

$$\|\hat{x}\|^2 = \|\hat{x} - x_f + x_f\|^2 = \|\hat{x} - x_f\|^2 + \|x_f\|^2 \geq \|x_f\|^2.$$

Por lo tanto $\|\hat{x}\| \geq \|x_f\|$ y x_f es la solución de cuadrados mínimos de $Ax = b$ que tiene norma mínima.

h) : Por la Proposición 3.7.1, $\text{col}(A^TA) = \text{col}(A^T)$. Como $A^TA \in \mathbb{R}^{n\times n}$ y

$$\dim(\text{col}(A^TA)) = \dim(\text{col}(A^T)) = \text{rg}(A^T) = \text{rg}(A) = n,$$

tenemos que A^TA es inversible (es de rango completo) y entonces existe la matriz $(A^TA)^{-1}$. Por lo tanto, la única solución del sistema $A^TAx = A^Tb$ es

$$\hat{x} := (A^TA)^{-1}A^Tb.$$

y $\hat{x}$ es la única solución por cuadrados mínimos de $Ax = b$. Entonces, por *a*) vale que, para cada $b \in \mathbb{R}^m$,

$$A(A^TA)^{-1}A^Tb = A\hat{x} = P_{\text{col}(A)}b = [P_{\text{col}(A)}b]^E = [P_{\text{col}(A)}]_E^E\,[b]^E = [P_{\text{col}(A)}]_E^E\,b,$$

donde E es la base canónica de $\mathbb{R}^m$. Como las matrices $A(A^TA)^{-1}A^T$ y $[P_{\text{col}(A)}]_E^E$ coinciden en todo $b \in \mathbb{R}^m$, se sigue que $A(A^TA)^{-1}A^T = [P_{\text{col}(A)}]_E^E$. ■

Vamos a remarcar las conclusiones del ejercicio anterior para tenerlas presentes:

Sean $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ y $(\mathbb{R}^m, \langle\cdot,\cdot\rangle)$ los espacios euclídeos canónicos de dimensión n y m respectivamente. Sean $A \in \mathbb{R}^{m\times n}$ y $b \in \mathbb{R}^m$. Entonces, son equivalentes:

a) $\hat{x} \in \mathbb{R}^n$ es una solución de mínimos cuadrados de la ecuación $Ax = b$.

b) $\|A\hat{x} - b\| = \text{mín}\{\|Ax - b\| : x \in \mathbb{R}^n\}$ ó, equivalentemente, $\|A\hat{x} - b\| \leq \|Ax - b\|$ para todo $x \in \mathbb{R}^n$.

c) $\hat{x} \in \mathbb{R}^n$ es una solución de la ecuación $Ax = P_{\text{col}(A)}(b)$, es decir $A\hat{x} = P_{\text{col}(A)}(b)$.

d) $\hat{x} \in \mathbb{R}^n$ es una solución de la *ecuación normal* $A^TAx = A^Tb$, es decir $A^TA\hat{x} = A^Tb$.

En este caso, si $\hat{x}_p$ es una solución de mínimos cuadrados de $Ax = b$ entonces todas las soluciones de mínimos cuadrados de $Ax = b$ son

$$\hat{x} = \hat{x}_p + z,$$

con $z \in \text{nul}(A)$. En particular, existe una única solución de mínimos cuadrados de $Ax = b$ si y sólo si $\text{nul}(A) = \{0\}$.

Para hacer alguna cuenta, consideremos el siguiente ejercicio muy similar al **Ejercicio 3.15**.

Ejercicio 3.L. La siguiente tabla muestra, para algunos instantes t (en segundos), la posición de un móvil $p(t)$ (en metros).

t	10	20	30	40
p(t)	2,5	5,2	6,6	9,4

a) Suponiendo velocidad constante. Hallar la recta $p(t)$ que mejor ajusta los datos (en el sentido de mínimos cuadrados).

b) Determinar en qué posición se estima que estará el móvil en $t = 50$ seg.

c) Calcular el error de estimación.

Dem. $a)$: Buscamos una recta que mejor ajuste los datos. La recta será de la forma

$$p(t) = p_0 + vt,$$

con p_0 y v a determinar. Además, la recta que buscamos, queremos que cumpla:

$$2,5 = p(10) = p_0 + v \cdot 10.$$
$$5,2 = p(20) = p_0 + v \cdot 20.$$
$$6,6 = p(30) = p_0 + v \cdot 30.$$
$$9,4 = p(40) = p_0 + v \cdot 40.$$

Si llamamos $b := \begin{bmatrix} 2,5 \\ 5,2 \\ 6,6 \\ 9,4 \end{bmatrix}$ y $A := \begin{bmatrix} 1 & 10 \\ 1 & 20 \\ 1 & 30 \\ 1 & 40 \end{bmatrix}$. Buscamos la solución del sistema $Ax = b$, con $x := \begin{bmatrix} p_0 \\ v \end{bmatrix}$. Como claramente ese sistema no tiene solución, vamos a buscar una solución aproximada aplicando mínimos cuadrados. Por lo que vimos arriba, $\hat{x}$ es una solución de mínimos cuadrados de $Ax = b$ si y sólo si $A^T A\hat{x} = A^T b$. Entonces, como $A^T A = \begin{bmatrix} 3000 & 100 \\ 100 & 4 \end{bmatrix}$ y $A^T b = \begin{bmatrix} 703 \\ 23,7 \end{bmatrix}$, buscamos las soluciones del sistema

$$\begin{bmatrix} 3000 & 100 \\ 100 & 4 \end{bmatrix} x = \begin{bmatrix} 703 \\ 23,7 \end{bmatrix}.$$

Como $\operatorname{rg}(A) = \operatorname{rg}(A^T A) = 2$, el sistema tiene única solución y es $\hat{x} = \begin{bmatrix} 0,4 \\ 0,221 \end{bmatrix}$. Entonces, la recta que mejor ajusta los datos (en el sentido de mínimos cuadrados) es $p(t) = 0,4 + 0,221 \cdot t$.

$b)$: Usando la recta que mejor ajusta los datos (en el sentido de mínimos cuadrados) tenemos que $p(50) = 11,45$ m.

$c)$: El error cometido se obtiene calculando

$$e = \|A\hat{x} - b\| = \| \begin{bmatrix} 2,61 \\ 4,82 \\ 7,03 \\ 9,24 \end{bmatrix} - \begin{bmatrix} 2,5 \\ 5,2 \\ 6,6 \\ 9,4 \end{bmatrix} \| = [(0,11)^2 + (-0,38)^2 + (0,43)^2 + (0,16)^2]^{1/2}$$
$$= 0,613.$$

■

En el siguiente ejemplo vamos a hallar soluciones de cuadrados mínimos de la ecuación $T(x) = b$, con $T \in \mathcal{L}(\mathbb{R}_3[x], \mathbb{R}^4)$.

Ejercicio 3.M (De Parcial)**.** Sea la transformación lineal $T : \mathbb{R}_3[x] \to \mathbb{R}^4$ definida por

$$T(p) = \begin{bmatrix} p(-1) \\ p(0) \\ p(1) \\ \int_{-1}^{1} p(x)dx \end{bmatrix}.$$

Consideremos en $\mathbb{R}^3$ el producto interno canónico y la ecuación $T(p) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}$.

a) Decidir si existe $p \in \mathbb{R}_3[x]$ tal que $T(p) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}$.

b) Si la respuesta de a) es negativa, hallar todas las soluciones aproximadas (en el sentido de mínimos cuadrados) de dicha ecuación y estimar el error cometido.

Dem. $a)$: Primero vamos a obtener una base de $\mathrm{Im}(T)$. Recordemos que como $\{1, x, x^2, x^3\}$ es un sistema de generadores de $\mathbb{R}_3[x]$, tenemos que $\mathrm{Im}(T) = gen\{T(1), T(x), T(x^2), T(x^3)\}$.

Observar que $\int_{-1}^{1} 1dx = 2$, $\int_{-1}^{1} xdx = 0$, $\int_{-1}^{1} x^2dx = \frac{2}{3}$ $\int_{-1}^{1} x^3dx = 0$. Entonces

$$T(1) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \ T(x) = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \ T(x^2) = \begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{2}{3} \end{bmatrix} \text{ y } T(x^3) = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Entonces, $\mathrm{Im}(T) = gen\{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{2}{3} \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}\}$. Una base de $\mathrm{Im}(T)$ puede ser (verificarlo)

$$\mathcal{B}_{\mathrm{Im}(T)} = \{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{2}{3} \end{bmatrix}\}.$$

Como $\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix} \notin \mathrm{Im}(T)$ (verificarlo). No existe p tal que $T(p) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}$.

$b)$: Como vimos arriba, todas las soluciones aproximadas (en el sentido de mínimos cuadrados) de dicha ecuación son las soluciones del sistema (ahora compatible) $T(p) = P_{\mathrm{Im}(T)}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix})$.

Para ahorrar cuentas, en vez de calcular $P_{\mathrm{Im}(T)}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix})$, vamos a calcular $P_{\mathrm{Im}(T)^\perp}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix})$ ya que sabemos que vale que

$$P_{\mathrm{Im}(T)}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix} - P_{\mathrm{Im}(T)^\perp}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}).$$

Observar que $\mathrm{Im}(T)^\perp = gen\{\begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix}\}$. De hecho, $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \in \mathrm{Im}(T)^\perp$ si y sólo si

$$0 = \left\langle y, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix} \right\rangle = \left\langle y, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\rangle = \left\langle y, \begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{2}{3} \end{bmatrix} \right\rangle.$$

Entonces, operando, nos queda que

$$0 = y_1 + y_2 + y_3 + 2y_4 = -y_1 + y_2 = y_1 + y_3 + \frac{2}{3}y_4.$$

Resolviendo el sistema, se sigue que $y_3 = -\frac{1}{3}y_4,\ y_2 = -\frac{4}{3}y_4,\ y_1 = -\frac{1}{3}y_4$. Entonces,

$$y = \begin{bmatrix} -\frac{1}{3}y_4 \\ -\frac{4}{3}y_4 \\ -\frac{1}{3}y_4 \\ y_4 \end{bmatrix} = y_4 \begin{bmatrix} -\frac{1}{3} \\ -\frac{4}{3} \\ -\frac{1}{3} \\ 1 \end{bmatrix} \text{ con } y_4 \in \mathbb{R}.$$

Por lo tanto, $\operatorname{Im}^{\perp} = gen\{ \begin{bmatrix} -\frac{1}{3} \\ -\frac{4}{3} \\ -\frac{1}{3} \\ 1 \end{bmatrix} \} = gen\{ \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} \}$. Como $\{ \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} \}$ es una base (ortogonal) de $\operatorname{Im}(T)^{\perp}$ (tiene un sólo elemento), por lo que vimos, tenemos que

$$P_{\operatorname{Im}(T)^{\perp}}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}) = \frac{\left\langle \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} \right\rangle}{\| \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} \|^2} \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} = \frac{-27}{27} \begin{bmatrix} -1 \\ -4 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 1 \\ -3 \end{bmatrix}.$$

Entonces,

$$P_{\operatorname{Im}(T)}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix} - P_{\operatorname{Im}(T)^{\perp}}(\begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 \\ 4 \\ 1 \\ -3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}.$$

Por lo tanto, buscamos $p \in \mathbb{R}_3[x]$ tales que $T(p) = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}$.

Si $p \in \mathbb{R}_3[x]$ es tal que $T(p) = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}$ entonces, $p(x) = a + bx + cx^2 + dx^3$ con $a, b, c, d \in \mathbb{R}$, tal que $T(p) = aT(1) + bT(x) + cT(x^2) + dT(x^3) = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}$. Entonces,

$$a \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix} + b \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{2}{3} \end{bmatrix} + d \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}.$$

Resolviendo el sistema, nos queda que $a = 1, b = 1 - d, c = 0$. Volviendo a la expresión de p nos queda que $p(x) = 1 + x + d(x^3 - x)$, con $d \in \mathbb{R}$. Por lo tanto, todas las soluciones de mínimos cuadrados de la ecuación $T(p) = \begin{bmatrix} 1 \\ 5 \\ 3 \\ 1 \end{bmatrix}$ son

$$\hat{p}(x) = 1 + x + d(x^3 - x), \text{ con } d \in \mathbb{R}.$$

Finalmente, el error cometido, se obtiene como

$$e = \|T(\hat{p}) - \begin{bmatrix} 1 \\ 5 \\ 3 \\ 1 \end{bmatrix}\| = \| \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 5 \\ 3 \\ 1 \end{bmatrix}\| = \| \begin{bmatrix} -1 \\ -4 \\ -1 \\ 1 \end{bmatrix}\| = \sqrt{19}.$$

■

Proposición 3.7.2. *Sean $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo de dimensión n y $B = \{v_1, v_2, \cdots, v_n\}$ una **base ortogonal** de $\mathbb{V}$. Para cada $v \in \mathbb{V}$, vale que $[v]^B = \begin{bmatrix} \frac{\langle v, v_1 \rangle}{\|v_1\|^2} \\ \frac{\langle v, v_2 \rangle}{\|v_2\|^2} \\ \vdots \\ \frac{\langle v, v_n \rangle}{\|v_n\|^2} \end{bmatrix}$.*

En particular, si B es una base ortonormal de $\mathbb{V}$, $[v]^B = \begin{bmatrix} \langle v, v_1 \rangle \\ \langle v, v_2 \rangle \\ \vdots \\ \langle v, v_n \rangle \end{bmatrix}$.

Dem. Supongamos que $[v]^B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{K}^n$. Entonces $v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$. Por lo tanto como $\langle v_i, v_j \rangle = 0$ si $i \neq j$, tenemos que para cada $1 \leq i \leq n$,

$$\langle v, v_i \rangle = \langle a_1 v_1 + a_2 v_2 + \cdots + a_n v_n, v_i \rangle = a_i \langle v_i, v_i \rangle = a_i \|v_i\|^2.$$

Finalmente, como para cada $1 \leq i \leq n$, $\|v_i\|^2 \neq 0$, se sigue que

$$a_i = \frac{\langle v, v_i \rangle}{\|v_i\|^2}.$$

Si B es una base ortonormal de $\mathbb{V}$, tenemos que $\|v_i\|^2 = 1$ para todo $1 \leq i \leq n$. Entonces, para cada $1 \leq i \leq n$,

$$a_i = \langle v, v_i \rangle.$$

■

Ejercicio 3.19 (ver **Ejercicio 1.26**)**.** Se considera el $\mathbb{R}$-espacio euclídeo $(\mathbb{R}_2[x], \langle \cdot, \cdot \rangle)$ con el producto interno definido por

$$\langle p, q \rangle := p(0)q(0) + p(1)q(1) + p(2)q(2).$$

Comprobar que

$$B := \{\frac{1}{2}(x-1)(x-2), -x(x-2), \frac{1}{2}x(x-1)\}$$

es una base ortonormal de $(\mathbb{R}_2[x], \langle \cdot, \cdot \rangle)$ y utilizar ese resultado para hallar el vector de coordenadas del polinomio $p(x) = 1 + x + x^2$ en base B.

Dem. Llamemos $p_1(x) = \frac{1}{2}(x-1)(x-2)$, $p_2(x) = -x(x-2)$ y $p_3(x) = \frac{1}{2}x(x-1)$. Entonces, $p_1(0) = 1, p_1(1) = 0, p_1(2) = 0$, $p_2(0) = 0, p_2(1) = 1, p_2(2) = 0$ y $p_3(0) = 0, p_3(1) = 0, p_3(2) = 1$. Entonces

$$\langle p_1, p_1 \rangle = p_1(0)p_1(0) + p_1(1)p_1(1) + p_1(2)p_1(2) = 1 + 0 + 0 = 1.$$

$$\langle p_2, p_2 \rangle = p_2(0)p_2(0) + p_2(1)p_2(1) + p_2(2)p_2(2) = 0 + 1 + 0 = 1.$$

$$\langle p_3, p_3 \rangle = p_3(0)p_3(0) + p_3(1)p_3(1) + p_3(2)p_3(2) = 0 + 0 + 1 = 1.$$

$$\langle p_1, p_2 \rangle = p_1(0)p_2(0) + p_1(1)p_2(1) + p_1(2)p_2(2) = 1 \cdot 0 + 0 \cdot 1 + 0 \cdot 0 = 0.$$

$$\langle p_1, p_3 \rangle = p_1(0)p_3(0) + p_1(1)p_3(1) + p_1(2)p_3(2) = 1 \cdot 0 + 0 \cdot 0 + 0 \cdot 1 = 0.$$

$$\langle p_2, p_3 \rangle = p_2(0)p_3(0) + p_2(1)p_3(1) + p_2(2)p_3(2) = 0 \cdot 0 + 1 \cdot 0 + 0 \cdot 1 = 0.$$

Por lo tanto, como $\langle p_i, p_j \rangle = \delta_{ij}$, se sigue que B es una base ortonormal de $(\mathbb{R}_2[x], \langle \cdot, \cdot \rangle)$. Finalmente, por la Proposición 3.7.2, tenemos que si $p(x) = 1 + x + x^2$ entonces

$$[p]^B = \begin{bmatrix} \langle p, p_1 \rangle \\ \langle p, p_2 \rangle \\ \langle p, p_3 \rangle \end{bmatrix}.$$

Como

$$\langle p, p_1 \rangle = p(0)p_1(0) + p(1)p_1(1) + p(2)p_1(2) = 1 \cdot 1 + 3 \cdot 0 + 7 \cdot 0 = 1.$$

$$\langle p, p_2 \rangle = p(0)p_2(0) + p(1)p_2(1) + p(2)p_2(2) = 1 \cdot 0 + 3 \cdot 1 + 7 \cdot 0 = 3.$$

$$\langle p, p_3 \rangle = p(0)p_3(0) + p(1)p_3(1) + p(2)p_3(2) = 1 \cdot 0 + 3 \cdot 0 + 7 \cdot 1 = 7.$$

Entonces,

$$[p]^B = \begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix}.$$

■

Ejercicio 3.20. Sean $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{R}$-espacio euclídeo de dimensión n y sea $B = \{u_1, u_2, \cdots, u_n\}$ una **base ortonormal** de $\mathbb{V}$.

a) Comprobar que para toda pareja de vectores $v_1, v_2 \in \mathbb{V}$ vale que

$$d(v_1, v_2) = \sqrt{\sum_{j=1}^{n} (\langle v_2, u_j \rangle - \langle v_1, u_j \rangle)^2}.$$

b) Deducir que para cada $v_0 \in \mathbb{V}$ y cada $r > 0$, el conjunto

$$\{v \in \mathbb{V} : d(v, v_0) = r\} = \{\sum_{j=1}^{n} x_j u_j : \; x_1, \cdots, x_n \in \mathbb{R}, \; \sum_{j=1}^{n} (x_j - a_j)^2 = r^2\},$$

donde $[a_1 \; \cdots \; a_n]^T$ es el vector de coordenadas de v_0 respecto de la base B.

Dem. Para resolver este ejercicio vamos a usar la Proposición 3.7.2.

a) : Recordar que $d(v_1, v_2) = \|v_1 - v_2\|^2$. Como B es una base ortonormal, por la Proposición 3.7.2, tenemos que

$$[v_1]^B = \begin{bmatrix} \langle v_1, u_1 \rangle \\ \langle v_1, u_2 \rangle \\ \vdots \\ \langle v_1, u_n \rangle \end{bmatrix} \text{ y } [v_2]^B = \begin{bmatrix} \langle v_2, u_1 \rangle \\ \langle v_2, u_2 \rangle \\ \vdots \\ \langle v_2, u_n \rangle \end{bmatrix}.$$

Por lo tanto,

$$v_1 = \langle\, v_1, u_1 \,\rangle\, u_1 + \langle\, v_1, u_2 \,\rangle\, u_2 + \cdots + \langle\, v_1, u_n \,\rangle\, u_n \text{ y } v_2 = \langle\, v_2, u_1 \,\rangle\, u_1 + \langle\, v_2, u_2 \,\rangle\, u_2 + \cdots + \langle\, v_2, u_n \,\rangle\, u_n.$$

Entonces, usando propiedades de producto interno y aplicando el Teorema de Pitágoras (recordar que el conjunto $\{u_1, u_2, \cdots, u_n\}$es ortonormal), se sigue que

$$\begin{aligned}
\|v_1 - v_2\|^2 &= \|(\langle\, v_1, u_1 \,\rangle - \langle\, v_2, u_1 \,\rangle)u_1 + (\langle\, v_1, u_2 \,\rangle - \langle\, v_2, u_2 \,\rangle)u_2 + \cdots + (\langle\, v_1, u_n \,\rangle - \langle\, v_2, u_n \,\rangle)u_n\|^2 \\
&= \|\, \langle\, v_1 - v_2, u_1 \,\rangle\, u_1 + \langle\, v_1 - v_2, u_2 \,\rangle\, u_2 + \cdots + \langle\, v_1 - v_2, u_n \,\rangle\, u_n\|^2 \\
&= \langle\, v_1 - v_2, u_1 \,\rangle^2 \|u_1\|^2 + \langle\, v_1 - v_2, u_2 \,\rangle^2 \|u_2\|^2 + \cdots + \langle\, v_1 - v_2, u_n \,\rangle^2 \|u_n\|^2 \\
&= \langle\, v_1 - v_2, u_1 \,\rangle^2 + \langle\, v_1 - v_2, u_2 \,\rangle^2 + \cdots + \langle\, v_1 - v_2, u_n \,\rangle^2.
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
d(v_1, v_2) &= \sqrt{\langle\, v_1 - v_2, u_1 \,\rangle^2 + \langle\, v_1 - v_2, u_2 \,\rangle^2 + \cdots + \langle\, v_1 - v_2, u_n \,\rangle^2} \\
&= \sqrt{\sum_{j=1}^{n} (\langle\, v_2, u_j \,\rangle - \langle\, v_1, u_j \,\rangle)^2}
\end{aligned}$$

y probamos lo que queríamos.

b) : Vamos a probar la doble inclusión. Supongamos que $v \in \mathbb{V}$ es tal que $d(v, v_0) = r$. Entonces, por un lado como $v \in \mathbb{V}$ y B es una base ortonormal, por la Proposición 3.7.2, se sigue que $[v]^B = [\langle\, v, u_1 \,\rangle \;\; \langle\, v, u_2 \,\rangle \;\; \cdots \;\; \langle\, v, u_n \,\rangle]^T := [x_1 \;\; \cdots \;\; x_n]^T$. Entonces,

$$v = \sum_{j=1}^{n} x_j u_j.$$

Por la Proposición 3.7.2, $[v_0]^B = [\langle\, v_0, u_1 \,\rangle \;\; \langle\, v_0, u_2 \,\rangle \;\; \cdots \;\; \langle\, v_0, u_n \,\rangle]^T := [a_1 \;\; \cdots \;\; a_n]^T$.

Además, por *a*),

$$\begin{aligned}
r^2 = d(v, v_0)^2 &= \langle\, v - v_0, u_1 \,\rangle^2 + \langle\, v - v_0, u_2 \,\rangle^2 + \cdots + \langle\, v - v_0, u_n \,\rangle^2 \\
&= (\langle\, v, u_1 \,\rangle - \langle\, v_0, u_1 \,\rangle)^2 + (\langle\, v, u_2 \,\rangle - \langle\, v_0, u_2 \,\rangle)^2 + \cdots + (\langle\, v, u_n \,\rangle - \langle\, v_0, u_n \,\rangle)^2 \\
&= (x_1 - a_1)^2 + (x_2 - a_2)^2 + \cdots + (x_n - a_n)^2
\end{aligned}$$

y entonces $v \in \{\sum_{j=1}^{n} x_j u_j : \; x_1, \cdots, x_n \in \mathbb{R}, \; \sum_{j=1}^{n} (x_j - a_j)^2 = r^2\}$.

Recíprocamente, supongamos que $v = \sum_{j=1}^{n} x_j u_j$ con $x_1, \cdots, x_n \in \mathbb{R}$, y $\sum_{j=1}^{n} (x_j - a_j)^2 = r^2$. Donde $[a_1 \;\; \cdots \;\; a_n]^T = [v_0]^B = [\langle\, v_0, u_1 \,\rangle \;\; \langle\, v_0, u_2 \,\rangle \;\; \cdots \;\; \langle\, v_0, u_n \,\rangle]^T$, es el vector de coordenadas de v_0 respecto de la base B. Entonces, $[v]^B = [x_1 \;\; \cdots \;\; x_n]^T = [\langle\, v, u_1 \,\rangle \;\; \langle\, v, u_2 \,\rangle \;\; \cdots \;\; \langle\, v, u_n \,\rangle]^T$ y, por *a*), tenemos que

$$\begin{aligned}
d(v, v_0)^2 &= \langle\, v - v_0, u_1 \,\rangle^2 + \langle\, v - v_0, u_2 \,\rangle^2 + \cdots + \langle\, v - v_0, u_n \,\rangle^2 \\
&= (\langle\, v, u_1 \,\rangle - \langle\, v_0, u_1 \,\rangle)^2 + (\langle\, v, u_2 \,\rangle - \langle\, v_0, u_2 \,\rangle)^2 + \cdots + (\langle\, v, u_n \,\rangle - \langle\, v_0, u_n \,\rangle)^2 \\
&= (x_1 - a_1)^2 + (x_2 - a_2)^2 + \cdots + (x_n - a_n)^2 = r^2.
\end{aligned}$$

Entonces $v \in \{v \in \mathbb{V} : d(v, v_0) = r\}$. Como probamos la doble inclusión, se sigue que $\{v \in \mathbb{V} : d(v, v_0) = r\} = \{\sum_{j=1}^{n} x_j u_j : \; x_1, \cdots, x_n \in \mathbb{R}, \; \sum_{j=1}^{n} (x_j - a_j)^2 = r^2\}$. ■

Ejercicio 3.21. Se considera $\mathbb{R}^2$ con el producto interno $\langle\, \cdot, \cdot \,\rangle$ definido por

$$\langle\, x, y \,\rangle = y^T \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} x.$$

a) Describir el significado geométrico del conjunto

$$\Sigma = \{x \in \mathbb{R}^2 : 2x_1^2 + 2x_1x_2 + 2x_2^2 = 4\}.$$

b) Hallar dos vectores $v_1, v_2 \in \mathbb{R}^2$ tales que

$$\Sigma = \{y_1v_1 + y_2v_2 : y_1, y_2 \in \mathbb{R}, y_1^2 + y_2^2 = 4\} = \{2\cos(\theta)v_1 + 2\sin(\theta)v_2 : \theta \in [0, 2\pi)\}.$$

Dem. *a*) : Notar que, para cada $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$, vale que

$$\|x\|^2 = \langle x, x \rangle = x^T \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} x = [x_1 \ x_2]^T \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 2x_1x_2 + 2x_2^2.$$

Por lo tanto, $\Sigma = \{x \in \mathbb{R}^2 : \|x\|^2 = 4\} = \{x \in \mathbb{R}^2 : \|x\| = 2\}$. Entonces Σ es el conjunto de todos los vectores de $\mathbb{R}^2$ con norma 2. Observar que la norma considerada NO es la inducidad por el producto interno canónico de $\mathbb{R}^2$ sino la inducida por el producto interno dado al principio del ejercicio.

b) : Vamos a completar cuadrados en la expresión $2x_1^2 + 2x_1x_2 + 2x_2^2 = 4$. Notar que

$$(\sqrt{2}x_1 + \frac{\sqrt{2}}{2}x_2)^2 + (\frac{\sqrt{3}}{\sqrt{2}}x_2)^2 = 2x_1^2 + 2x_1x_2 + 2x_2^2 = 4.$$

A continuación, proponemos el siguiente cambio de variables

$$y_1 := \sqrt{2}x_1 + \frac{\sqrt{2}}{2}x_2, \ y_2 := \frac{\sqrt{3}}{\sqrt{2}}x_2 = \frac{\sqrt{6}}{2}x_2.$$

Entonces $y_1^2 + y_2^2 = 4$ y además,

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2\sqrt{2} & \sqrt{2} \\ 0 & \sqrt{6} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Notar que la matriz $\frac{1}{2} \begin{bmatrix} 2\sqrt{2} & \sqrt{2} \\ 0 & \sqrt{6} \end{bmatrix}$ es inversible (su determinante no es nulo). Por lo tanto,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (\frac{1}{2} \begin{bmatrix} 2\sqrt{2} & \sqrt{2} \\ 0 & \sqrt{6} \end{bmatrix})^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 3\sqrt{2} & -\sqrt{6} \\ 0 & 2\sqrt{6} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Sean $v_1 := \frac{1}{6} \begin{bmatrix} 3\sqrt{2} \\ 0 \end{bmatrix}$ y $v_2 := \frac{1}{6} \begin{bmatrix} -\sqrt{6} \\ 2\sqrt{6} \end{bmatrix}$. Entonces, si $x \in \Sigma$, tenemos que $x = y_1v_1 + y_2v_2$ y además, $y_1^2 + y_2^2 = 4$. Por lo tanto $\Sigma \subseteq \{y_1v_1 + y_2v_2 : y_1, y_2 \in \mathbb{R}, y_1^2 + y_2^2 = 4\}$.

Recíprocamente, supongamos que $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = y_1v_1 + y_2v_2 = y_1\frac{1}{6} \begin{bmatrix} 3\sqrt{2} \\ 0 \end{bmatrix} + y_2\frac{1}{6} \begin{bmatrix} -\sqrt{6} \\ 2\sqrt{6} \end{bmatrix}$ con $y_1^2 + y_2^2 = 4$. Entonces $x_1 = \frac{1}{6}(3\sqrt{2}y_1 - \sqrt{6}y_2)$ y $x_2 = \frac{2}{6}\sqrt{6}y_2$. Entonces

$$\begin{aligned} 2x_1^2 + 2x_1x_2 + 2x_2^2 &= \frac{2}{36}(18y_1^2 - 6\sqrt{12}y_1y_2 + 6y_2^2) + \frac{4}{36}(y_1y_2 3\sqrt{12} - 6y_2^2) + \frac{48}{36}y_2^2 \\ &= y_1^2 + y_2^2 = 4. \end{aligned}$$

Por lo tanto $x \in \Sigma$ y tenemos que $\{y_1v_1 + y_2v_2 : y_1, y_2 \in \mathbb{R}, y_1^2 + y_2^2 = 4\} \subseteq \Sigma$. Como probamos la doble inclusión, se sigue que $\Sigma = \{y_1v_1 + y_2v_2 : y_1, y_2 \in \mathbb{R}, y_1^2 + y_2^2 = 4\}$.

Por último, es claro que si $y_1 := 2\cos(\theta), \ y_2 := 2\sin(\theta)$ con $\theta \in [0, 2\pi)$, se sigue que $y_1^2 + y_2^2 = 4$. Recíprocamente, si $y_1^2 + y_2^2 = 4$, entonces el conjunto $\{y \in \mathbb{R}^2 : y_1^2 + y_2^2 = 4\}$ es la circunferencia de radio 2 cuya parametrización es justamente, $y_1 = 2\cos(\theta), y_2 = 2\sin(\theta)$ con $\theta \in [0, 2\pi)$ y por lo tanto tenemos la otra igualdad de conjuntos, $\{y_1v_1 + y_2v_2 : y_1, y_2 \in \mathbb{R}, y_1^2 + y_2^2 = 4\} = \{2\cos(\theta)v_1 + 2\sin(\theta)v_2 : \theta \in [0, 2\pi)\}$. ∎

3.8. Distancia mínima

El siguiente ejercicio nos permite demostrar que si $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ es un espacio euclídeo (no necesariamente de dimensión finita), para cada punto $v \in \mathbb{V}$, siempre existe un único vector de un subespacio (de dimensión finita) que es "más cercano" (en el sentido que minimiza la distancia) a v.

Lo que probemos en el siguiente ejercicio también se puede usar para definir la proyección ortogonal de un punto a un subespacio (de dimensión finita). Antes de resolver el ejercicio, repasemos el concepto de sucesiones en espacios normados.

Sucesiones en espacios normados

Sea $(\mathbb{V}, \|\cdot\|)$ un espacio normado y sea $(v_n)_{n\in\mathbb{N}}$ una sucesión de elementos de $\mathbb{V}$. Diremos que la sucesión $(v_n)_{n\in\mathbb{N}}$ es *convergente* si existe $v_0 \in \mathbb{V}$ tal que

$$\lim_{n\to\infty} \|v_n - v_0\| = 0,$$

ó, equivalentemente, si para cada $\varepsilon > 0$ existe $N_0 \in \mathbb{N}$ tal que

$$\|v_n - v_0\| < \varepsilon,$$

para todo $n \geq N_0$. En este caso, escribimos

$$\lim_{n\to\infty} v_n = v_0.$$

Ejercicio 3.23. Sea $\{u_i : i \in \mathbb{N}\}$ un sistema ortonormal de vectores en un $\mathbb{R}$-espacio euclídeo $(\mathbb{V}, \langle \cdot, \cdot \rangle)$. Dados $v \in \mathbb{V}$ y $n \in \mathbb{N}$, se considera el problema de hallar el vector $\hat{v}_n \in gen\{u_i : i \in \{1, 2, \cdots, n\}\} =: \mathcal{U}_n$ más cercano a v.

a) Mostrar que para todo $[a_1\ a_2\ \cdots\ a_n]^T \in \mathbb{R}^n$ vale que

$$\|v - \sum_{i=1}^{n} a_i v_i\|^2 = \|v\|^2 - \sum_{i=1}^{n} \langle v, u_i \rangle^2 + \sum_{i=1}^{n} (a_i - \langle v, v_i \rangle)^2,$$

y deducir de allí que $\min_{w\in\mathcal{U}_n} \|v - w\|$ se realiza en

$$\hat{v}_n = \sum_{i=1}^{n} \langle v, u_i \rangle u_i,$$

y que su valor es

$$\min_{w\in\mathcal{U}_n} \|v - w\| = \|v - \hat{v}_n\| = \sqrt{\|v\|^2 - \sum_{i=1}^{n} \langle v, u_i \rangle^2}.$$

b) Notar que $\lim_{n\to\infty} \hat{v_n} = v$, si y sólo si, $\sum_{i=1}^{\infty} \langle v, u_i \rangle^2 = \|v\|^2$.

c) Observar que para todo $v \in \mathbb{V}$ y todo $n \in \mathbb{N}$, el vector $v - \hat{v}_n \in \mathcal{U}_n^\perp$ y deducir de allí que $\mathbb{V} = \mathcal{U}_n \oplus \mathcal{U}_n^\perp$ para todo $n \in \mathbb{N}$.

Antes de resolver el ejercicio, entendamos qué es lo queremos probar. En primer lugar, con elemento de $\mathcal{U}_n$ "más cercano" a v, nos referimos a aquel elemento de $\mathcal{U}_n$ (si existe) que minimiza la distancia a v. En este caso, como estamos en un $\mathbb{K}$-espacio euclídeo vamos a tomar como distancia la inducida por el producto interno. Recordar que si $x, y \in \mathbb{V}$, la distancia de x a y está definida por

$$d(x, y) := \langle x - y, x - y \rangle^{1/2} = \|x - y\|.$$

Por otra parte, sea $v \in \mathbb{V}$ y fijemos un valor de $n \in \mathbb{N}$. Entonces $\mathcal{U}_n = gen\{u_1, \cdots, u_n\}$.

Consideremos el siguiente conjunto de números reales

$$A := \{\|v - w\| : w \in \mathcal{U}_n\} \subseteq \mathbb{R}.$$

Claramente el conjunto A es no vacío y además como $\|v - w\| \geq 0$ para todo $w \in \mathcal{U}_n$, el conjunto A está acotado inferiormente por 0. Entonces, existe el ínfimo de dicho conjunto, y además vale que ínf $\{\|v - w\| : w \in \mathcal{U}_n\} \geq 0$. Lo que nos está pidiendo el ejercicio es probar que ese ínfimo (que siempre existe) se realiza y entonces tenemos un mínimo.

Recordemos que $m \in \mathbb{R}$ es el mínimo del conjunto de números reales de A si pasan dos cosas:

- $\|v - w\| \geq m$, para todo $w \in \mathcal{U}_n$. Es decir m es una *cota inferior* de A.
- Existe $\hat{w} \in \mathcal{U}_n$ tal que $\|v - \hat{w}\| = m$, es decir hay un elemento del conjunto A que *realiza* el mínimo.

En conclusión, el ejercicio nos pide probar que si $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ es un $\mathbb{K}$-espacio euclídeo, existe un (único) vector de $\mathcal{U}_n$ que minimiza A. De hecho, el ejercicio no sólo nos pide probar que tal mínimo existe y que es único sino que ese mínimo es $\hat{w} := \hat{v}_n = \sum_{i=1}^n \langle v, u_i \rangle u_i \in \mathcal{U}_n$. Es decir, vamos a probar que

$$\|v - w\| \geq \|v - \hat{v}_n\|$$

para todo $w \in \mathcal{U}_n$ y que el vector que realiza el mínimo es único. Entendido esto, resolvamos el ejercicio.

Dem. $a)$: Vimos que si $v, w \in \mathbb{V}$ con $\mathbb{V}$ un $\mathbb{R}$-espacio euclídeo, entonces

$$\|v - w\|^2 = \|v\|^2 - 2 \langle v, w \rangle + \|w\|^2.$$

Entonces, para el vector $w = \sum_{i=1}^n a_i u_i \in \mathcal{U}_n$ tenemos que

$$\|v - \sum_{i=1}^n a_i u_i\|^2 = \|v\|^2 - 2 \sum_{i=1}^n a_i \langle v, u_i \rangle + \| \sum_{i=1}^n a_i u_i\|^2$$

Por otra parte, como $\{u_1, u_2, \cdots, u_n\}$ es un conjunto ortonormal, tenemos que

$$\| \sum_{i=1}^n a_i u_i\|^2 = \left\langle \sum_{i=1}^n a_i u_i, \sum_{i=j}^n a_j u_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \langle u_i, u_j \rangle = \sum_{i=1}^n a_i^2,$$

donde usamos que $\langle u_i, u_j \rangle = 0$ si $i \neq j$ y $\langle u_i, u_i \rangle = 1$.

Entonces, asociando, sumando y restando el término $\sum_{i=1}^n \langle v, u_i \rangle^2$ se sigue que

$$\begin{aligned}
\|v - \sum_{i=1}^n a_i u_i\|^2 &= \|v\|^2 - \sum_{i=1}^n 2a_i \langle v, u_i \rangle + \sum_{i=1}^n a_i^2 = \|v\|^2 + \sum_{i=1}^n (a_i^2 - 2a_i \langle v, u_i \rangle) \\
&= \|v\|^2 + \sum_{i=1}^n (a_i^2 - 2a_i \langle v, u_i \rangle) + \sum_{i=1}^n \langle v, u_i \rangle^2 - \sum_{i=1}^n \langle v, u_i \rangle^2 \\
&= \|v\|^2 + \sum_{i=1}^n (a_i^2 - 2a_i \langle v, u_i \rangle + \langle v, u_i \rangle^2) - \sum_{i=1}^n \langle v, u_i \rangle^2 \\
&= \|v\|^2 + \sum_{i=1}^n (a_i - \langle v, u_i \rangle)^2 - \sum_{i=1}^n \langle v, u_i \rangle^2 .
\end{aligned}$$

y probamos lo que queríamos.

Por otra parte, observar que $\hat{v}_n := \sum_{i=1}^n \langle v, u_i \rangle u_i \in \mathcal{U}_n$ (porque es una CL de elementos de $\mathcal{U}_n$), además, operando de manera similar que arriba y usando que $\{u_1, \cdots, u_n\}$ es un conjunto ortonormal, tenemos que

$$\begin{aligned}
\|v - \hat{v}_n\|^2 &= \|v\|^2 - 2\langle v, \hat{v}_n \rangle + \|\hat{v}_n\|^2 \\
&= \|v\|^2 - 2\sum_{i=1}^n \langle v, \langle v, u_i \rangle u_i \rangle + \|\sum_{i=1}^n \langle v, u_i \rangle u_i\|^2 \\
&= \|v\|^2 - 2\sum_{i=1}^n \langle v, u_i \rangle^2 + \sum_{i=1}^n \langle v, u_i \rangle^2 \\
&= \|v\|^2 - \sum_{i=1}^n \langle v, u_i \rangle^2 .
\end{aligned}$$

Entonces, usando que $\sum_{i=1}^n (a_i - \langle v, u_i \rangle)^2 > 0$, tenemos que para todo $[a_1\ a_2\ \cdots\ a_n]^T \in \mathbb{R}^n$,

$$\|v - \sum_{i=1}^n a_i u_i\|^2 = \|v\|^2 + \sum_{i=1}^n (a_i - \langle v, u_i \rangle)^2 - \sum_{i=1}^n \langle v, u_i \rangle^2 \geq \|v\|^2 - \sum_{i=1}^n \langle v, u_i \rangle^2 = \|v - \hat{v}_n\|^2.$$

Tomando raíz cuadrada (que es una función creciente) se sigue que, para todo $[a_1\ a_2\ \cdots\ a_n]^T \in \mathbb{R}^n$,

$$\|v - \sum_{i=1}^n a_i u_i\| \geq \|v - \hat{v}_n\|.$$

Como cualquier elemento $w \in \mathcal{U}_n$ se escribe como $w = \sum_{i=1}^n a_i u_i$ para ciertos $[a_1\ a_2\ \cdots\ a_n]^T \in \mathbb{R}^n$, la ecuación anterior nos asegura que $\hat{v}_n$ realiza el mínimo que buscamos, es decir

$$\text{mín}\{\|v - w\|^2 : w \in \mathcal{U}_n\} = \text{mín}\{\|v - \sum_{i=1}^n a_i u_i\| : [a_1\ a_2\ \cdots\ a_n]^T \in \mathbb{R}^n\} = \|v - \hat{v}_n\|.$$

Finalmente, como $\|v - \hat{v}_n\| = \sqrt{\|v\|^2 - \sum_{i=1}^n \langle v, u_i \rangle^2}$, concluimos que

$$\text{mín}\{\|v - w\| : w \in \mathcal{U}_n\} = \|v - \hat{v}_n\| = \sqrt{\|v\|^2 - \sum_{i=1}^n \langle v, u_i \rangle^2}.$$

Por último, veamos que el vector de $\mathcal{U}_n$ que es "más cercano" a v (en el sentido que minimiza la distancia a v) es único. De hecho, supongamos que existe otro elemento $\hat{w} \in \mathcal{U}_n$ que es "más cercano" a v. Entonces, claramente

$$\|v - \hat{w}\| = \|v - \hat{v}_n\|,$$

esto es así porque como $\hat{v}_n$ realiza el mínimo y $\hat{w} \in \mathcal{U}_n$ entonces $\|v - \hat{v}_n\| \leq \|v - \hat{w}\|$. De la misma manera, como estamos suponiendo que $\hat{w}$ realiza el mínimo y $\hat{v}_n \in \mathcal{U}_n$ entonces $\|v - \hat{w}\| \leq \|v - \hat{v}_n\|$ y tenemos la igualdad.

Por otra parte, recordar que en (3.7), vimos que para todo $x, y \in \mathbb{V}$,

$$\|x - y\|^2 + \|x + y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Observar que (obviamente) $\hat{w} - \hat{v}_n = \hat{w} - v - (\hat{v}_n - v)$. Entonces, aplicando la igualdad anterior a $x = (\hat{w} - v)$ e $y = (\hat{v}_n - v)$, nos queda

$$\|\hat{w} - \hat{v}_n\|^2 = \|(\hat{w} - v) - (\hat{v}_n - v)\|^2 = 2\|\hat{w} - v\|^2 + 2\|\hat{v}_n - v\|^2 - \|(\hat{w} - v) + (\hat{v}_n - v)\|^2$$

$$= 2\|\hat{w} - v\|^2 + 2\|\hat{v}_n - v\|^2 - \|\hat{w} + \hat{v}_n - 2v\|^2 = 4\|\hat{w} - v\|^2 - 4\|\frac{\hat{w} + \hat{v}_n}{2} - v\|^2$$
$$= 4(\|\hat{w} - v\|^2 - \|\frac{\hat{w} + \hat{v}_n}{2} - v\|^2) \leq 0,$$

donde usamos que como $\hat{w}$ realiza el mínimo, vale que $\|\hat{w} - v\|^2 \leq \|\frac{\hat{w}+\hat{v}_n}{2} - v\|^2$, pues como $\hat{v}_n, \hat{w} \in \mathcal{U}_n$ y $\mathcal{U}_n$ es un subespacio entonces $\frac{\hat{w}+\hat{v}_n}{2} \in \mathcal{U}_n$.

Entonces, nos quedó que $0 \leq \|\hat{w} - \hat{v}_n\|^2 \leq 0$, por lo tanto $\|\hat{w} - \hat{v}_n\| = 0$, entonces $\hat{w} - \hat{v}_n = 0_{\mathbb{V}}$ y se sigue que $\hat{w} = \hat{v}_n$. Conclusión : el vector que minimiza la distancia es único y es $\hat{v}_n$.

b) : Por definición de límite de sucesiones en espacios normados sabemos que $\lim_{n\to\infty} \hat{v}_n = v$ si y sólo si $\lim_{n\to\infty} \|v - \hat{v}_n\| = 0$. Por el ítem *a*), sabemos que

$$\|v - \hat{v}_n\|^2 = \|v\|^2 - \sum_{i=1}^{n} \langle v, u_i \rangle^2 .$$

Entonces, existe $\lim_{n\to\infty} \|v - \hat{v}_n\|$ si y sólo si existe $\lim_{n\to\infty} (\|v\|^2 - \sum_{i=1}^{n} \langle v, u_i \rangle^2)$. Más aún, en este caso,

$$0 = \lim_{n\to\infty} \|v - \hat{v}_n\|^2 = \lim_{n\to\infty} (\|v\|^2 - \sum_{i=1}^{n} \langle v, u_i \rangle^2) = \|v\|^2 - \lim_{n\to\infty} \sum_{i=1}^{n} \langle v, u_i \rangle^2 ,$$

donde usamos que $\|v\|^2$ es una constante que no depende de n. Finalmente, recordemos que $\sum_{i=1}^{\infty} \langle v, u_i \rangle^2 := \lim_{n\to\infty} \sum_{i=1}^{n} \langle v, u_i \rangle^2$. Por lo tanto, $\lim_{n\to\infty} \|v - \hat{v}_n\| = 0$ si y sólo si

$$\sum_{i=1}^{\infty} \langle v, u_i \rangle^2 = \lim_{n\to\infty} \sum_{i=1}^{n} \langle v, u_i \rangle^2 = \|v\|^2.$$

c) : Ya vimos que para todo $v \in \mathbb{V}$ y todo $n \in \mathbb{N}$ el vector $\hat{v}_n \in \mathcal{U}_n$ (es el único que) realiza el mínimo del conjunto $\{\|v - w\|^2 : w \in \mathcal{U}_n\}$. Por otra parte, para cada $j \in \{1, 2, \cdots, n\}$, tenemos que

$$\langle v - \hat{v}_n, u_j \rangle = \left\langle v - \sum_{i=1}^{n} \langle v, u_i \rangle u_i, u_j \right\rangle =$$

$$= \langle v, u_j \rangle - \sum_{i=1}^{n} \langle v, u_i \rangle \langle u_i, u_j \rangle = \langle v, u_j \rangle - \langle v, u_j \rangle = 0,$$

donde usamos que $\langle u_i, u_j \rangle = 0$ si $i \neq j$ y $\langle u_i, u_i \rangle = 1$. Por lo tanto $v - \hat{v}_n \perp u_j$ para cada $j = \{1, 2, \cdots, n\}$. Por lo tanto, si $u \in \mathcal{U}_n$, tenemos que $u = a_1 u_1 + \cdots + a_n u_n$, para ciertos $a_1, \cdots, a_n \in \mathbb{R}$. Entonces, usando que $v - \hat{v}_n \perp u_j$ para cada $j = \{1, 2, \cdots, n\}$, tenemos que

$$\langle v - \hat{v}_n, u \rangle = \langle v - \hat{v}_n, a_1 u_1 + \cdots + a_n u_n \rangle = a_1 \langle v - \hat{v}_n, u_1 \rangle + \cdots + a_n \langle v - \hat{v}_n, u_n \rangle = 0,$$

entonces $v - \hat{v}_n \in \mathcal{U}_n^{\perp}$.

Finalmente, es claro que $\mathcal{U}_n \cap \mathcal{U}_n^{\perp} = \{0\}$ y que $\mathcal{U}_n \oplus \mathcal{U}_n^{\perp} \subseteq \mathbb{V}$. Entonces, sea $v \in \mathbb{V}$, acabamos de probar que existe $\hat{v}_n \in \mathcal{U}_n$ tal que $v - \hat{v}_n \in \mathcal{U}_n^{\perp}$ y como

$$v = \hat{v}_n + (v - \hat{v}_n)$$

se sigue que $v \in \mathcal{U}_n \oplus \mathcal{U}_n^{\perp}$ con lo que concluimos que $\mathcal{U}_n \oplus \mathcal{U}_n^{\perp} = \mathbb{V}$. ∎

Ejercicio 3.N (De Parcial)**.** En $\mathbb{R}_2[x]$ se considera el producto interno definido por

$$\langle\, p, q \,\rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1).$$

Sea $\mathcal{U} := \{p \in \mathbb{R}_2[x] : p'(0) = 0\}$ y sea $q(x) := -x^2 + x$. Hallar todos los $p \in \mathcal{U}$ tales que $p \perp q$ y $\|p\| = 20$.

Dem. Si $p \in \mathcal{U}$, entonces $p \in \mathbb{R}_2[x]$, es decir $p(x) = ax^2 + bx + c$ con $a, b, c \in \mathbb{R}$ y $p'(0) = 0$. Entoces, como $p'(x) = 2ax + b$, tenemos que $p'(0) = b = 0$ y, volviendo a la expresión de p, nos queda que $p(x) = ax^2 + c$, con $a, c \in \mathbb{R}$.

Por otra parte, $p \perp q$, entonces $0 = \langle\, p, q \,\rangle = \left\langle\, ax^2 + c, -x^2 + x \,\right\rangle$. Entonces, como $p(-1) = a + c,\ p(0) = c,\ p(1) = a + c,\ q(-1) = -2,\ q(0) = 0,\ q(1) = 0$. Tenemos que

$$0 = \left\langle\, ax^2 + b, -x^2 + x \,\right\rangle = (a + c)(-2) + c \cdot 0 + (a + c) \cdot 0 = -2(a + c).$$

Entonces $a = -c$ y, volviendo a la expresión de p nos queda que $p(x) = ax^2 - a = a(x^2 - 1)$, con $a \in \mathbb{R}$. Finalmente

$$20 = \|p\| = \|a(x^2 - 1)\| = |a|\|x^2 - 1\|.$$

Como $\|x^2 - 1\|^2 = 0 \cdot 0 + -1 \cdot -1 + 0 \cdot 0 = 1$. Entonces $\|x^2 - 1\| = 1$ y

$$20 = |a|\|x^2 - 1\| = |a|,$$

por lo tanto $a = 20$ ó $a = -20$.

Conclusión, los polinomios $p \in \mathcal{U}$ tales que $p \perp q$ y $\|p\| = 20$, son $p(x) = 20(x^2 - 1)$ y $p(x) = -20(x^2 - 1)$.

■

El siguiente ejercicio es un ejercicio similar al **Ejercicio 3.24**, ambos ejercicios se resuelven aplicando las conclusiones obtenidas del **Ejercicio 3.23**.

Ejercicio 3.Ñ (De Parcial)**.** Sea $C([0, 1], \mathbb{R}) = \{f : [0, 1] \to \mathbb{R} : f \text{ es continua } \}$ con el producto interno

$$\langle\, f, g \,\rangle = \int_0^1 f(t)g(t)dt.$$

Dada $h(t) = t^4 + 1 \in C([0, 1], \mathbb{R})$, hallar la función cuadrática de expresión $g(t) = a + bt^2$, que minimiza la distancia a h.

Dem. Si llamamos $\mathcal{U}_2 = gen\{1, t^2\} \subseteq C([0, 1], \mathbb{R})$, el ejercicio nos pide el elemento de $\mathcal{U}_2$ más cercano a h. Es decir, si $g \in \mathcal{U}_2$, entonces $g(t) = a + bt^2$ con $a, b \in \mathbb{R}$ y buscamos los valores de $a, b \in \mathbb{R}$ tales que la g que resulte minimiza la distancia a h. Pero eso es justamente lo que hicimos en el **Ejercicio 3.23**. Si $\mathcal{B}_{\mathcal{U}_2} = \{u_1, u_2\}$ es una base ortonormal de $\mathcal{U}_2$, entonces (como vimos en el ejercicio anterior) la función g que buscamos es

$$g := \hat{h}_2 = \langle\, h, u_1 \,\rangle\, u_1 + \langle\, h, u_2 \,\rangle\, u_2.$$

Entonces, primero obtengamos una base ortonormal de $\mathcal{U}_2$. Buscamos $u_1, u_2 \in \mathcal{U}_2$ ortonormales tales que $gen\{u_1, u_2\} = \mathcal{U}_2$. Podemos tomar $u_1(t) := 1$, y busquemos $\alpha \in \mathbb{R}$ tal que $u_2(t) := t^2 + \alpha 1$ sea ortogonal a u_1. Entonces

$$0 = \langle\, u_2, u_1 \,\rangle = \left\langle\, t^2 + \alpha 1, 1 \,\right\rangle = \left\langle\, t^2, 1 \,\right\rangle + \alpha \,\langle\, 1, 1 \,\rangle\,.$$

Entonces

$$\alpha = \frac{-\left\langle\, t^2, 1 \,\right\rangle}{\langle\, 1, 1 \,\rangle}.$$

Haciendo, cuentas tenemos que

$$\langle 1, 1 \rangle = \int_0^1 1dt = t|_0^1 = 1,$$

$$\langle t^2, 1 \rangle = \int_0^1 t^2 dt = \frac{t^3}{3}|_0^1 = \frac{1}{3}.$$

Entonces $\alpha = -\frac{1}{3}$ y $u_2(t) = t^2 - \frac{1}{3}$. Sólo nos falta normalizar los vectores, para eso calculemos las normas de u_1 y u_2 entonces

$$\|u_1\|^2 = \langle 1, 1 \rangle = 1,$$

$$\|u_2\|^2 = \left\langle t^2 - \frac{1}{3}, t^2 - \frac{1}{3} \right\rangle = \int_0^1 (t^2 - \frac{1}{3})^2 dt = \frac{t^5}{5} - \frac{2}{9}t^3 + \frac{1}{9}t|_0^1 = \frac{4}{45}.$$

Entonces, $u_1(t) = 1$ (ya normalizado) y $u_2(t) = \frac{t^2 - \frac{1}{3}}{\sqrt{\frac{4}{45}}} = \frac{3\sqrt{5}}{2}(t^2 - \frac{1}{3})$. Por lo tanto

$$\begin{aligned} g(t) &= \langle h, u_1 \rangle u_1 + \langle h, u_2 \rangle u_2 = \langle t^4 + 1, 1 \rangle 1 + \left\langle t^4 + 1, \frac{3\sqrt{5}}{2}(t^2 - \frac{1}{3}) \right\rangle \frac{3\sqrt{5}}{2}(t^2 - \frac{1}{3}) \\ &= \langle t^4 + 1, 1 \rangle 1 + \left\langle t^4 + 1, t^2 - \frac{1}{3} \right\rangle \frac{45}{4}(t^2 - \frac{1}{3}). \end{aligned}$$

Haciendo, cuentas tenemos que

$$\langle t^4 + 1, 1 \rangle = \int_0^1 (t^4 + 1)dt = \frac{t^5}{5} + t|_0^1 = \frac{6}{5},$$

$$\left\langle t^4 + 1, t^2 - \frac{1}{3} \right\rangle = \int_{-1}^1 (t^4 + 1)(t^2 - \frac{1}{3})dt = \frac{t^7}{7} - \frac{t^5}{15} + \frac{t^3}{3} - \frac{t}{3}|_0^1 = \frac{8}{105}.$$

Por lo tanto,

$$g(t) = \frac{6}{5} \cdot 1 + \frac{8}{105}\frac{45}{4}(t^2 - \frac{1}{3}) = \frac{6}{5} \cdot 1 + \frac{6}{7}(t^2 - \frac{1}{3}) = \frac{32}{35} \cdot 1 + \frac{6}{7}t^2.$$

■

3.9. Algoritmo de Gram-Schmidt

Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo y sea $B = \{v_1, v_2, \cdots, v_n\}$ una base de $\mathbb{V}$, el *algoritmo de Gram-Schmidt* nos permite obtener una base ortonormal $B' = \{w_1, w_2, \cdots, w_n\}$ de $\mathbb{V}$ a partir de la base B con la siguiente propiedad:

$$\begin{aligned} gen\{v_1\} &= gen\{w_1\}, \\ gen\{v_1, v_2\} &= gen\{w_1, w_2\}, \\ gen\{v_1, v_2, v_3\} &= gen\{w_1, w_2, w_3\}, \\ &\vdots \\ gen\{v_1, v_2, v_3, \cdots, v_n\} &= gen\{w_1, w_2, w_3, \cdots, w_n\}. \end{aligned}$$

Veamos un ejemplo de aplicación del algoritmo de Gram-Schmidt.

Ejercicio 3.27. Se considera el $\mathbb{R}$-espacio euclídeo $(C([-1, 1], \mathbb{R}), \langle \cdot, \cdot \rangle)$ con el producto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x)dx$$

y el subespacio $\mathbb{R}_2[x]$.

a) Utilizar Gram-Schmidt para producir una base ortonormal $\{p_0, p_1, p_2\}$ de $\mathbb{R}_2[x]$ a partir de la base $\{q_0, q_1, q_2\}$ donde $q_0(x) = 1,\ q_1(x) = x,\ q_2(x) = x^2$.

b) Hallar las siguientes proyecciones ortogonales $P_{\mathbb{R}_2[x]}(\sin(x)),\ P_{\mathbb{R}_2[x]}(\cos(x))$.

c) Calcular las distancias $d(\sin(x), \mathbb{R}_2[x])$ y $d(\cos(x), \mathbb{R}_2[x])$

Dem. Vamos a aplicar el algoritmo de Gram-Schmidt. Primero vamos a obtener una base ortogonal y al finalizar normalizaremos todos los vectores obtenidos para producir una base ortonormal.

$a)$: Paso 1:

$$r_0 = q_0.$$

Paso 2:

$$r_1 = q_1 - \frac{\langle\, q_1, r_0 \,\rangle}{\|r_0\|^2} r_0.$$

Haciendo cuentas

$$\langle\, q_1, r_0 \,\rangle = \int_{-1}^{1} q_1(x) r_0(x) dx = \int_{-1}^{1} x \cdot 1 dx = \frac{x^2}{2}\Big|_{-1}^{1} = 0.$$

Entonces $r_1(x) = q_1(x) = x$.

Paso 3:

$$r_2 = q_2 - \frac{\langle\, q_2, r_1 \,\rangle}{\|r_1\|^2} r_1 - \frac{\langle\, q_2, r_0 \,\rangle}{\|r_0\|^2} r_0.$$

Haciendo cuentas

$$\langle\, q_2, r_0 \,\rangle = \int_{-1}^{1} q_2(x) r_0(x) dx = \int_{-1}^{1} x^2 \cdot 1 dx = \frac{x^3}{3}\Big|_{-1}^{1} = \frac{2}{3},$$

$$\langle\, q_2, r_1 \,\rangle = \int_{-1}^{1} q_2(x) r_1(x) dx = \int_{-1}^{1} x^2\ x\ dx = \frac{x^4}{4}\Big|_{-1}^{1} = 0.$$

Por otra parte, $\|r_0\|^2 = \int_{-1}^{1} r_0(x)^2 dx = \int_{-1}^{1} 1 dx = 2$.

Por lo tanto,

$$r_2(x) = x^2 - \frac{2}{3} \cdot \frac{1}{2} = x^2 - \frac{1}{3}.$$

La base ortogonal del subespacio $\mathbb{R}_2[x]$ nos quedó

$$\{r_0, r_1, r_2\} = \{1, x, x^2 - \frac{1}{3}\}.$$

Ahora calculemos la norma de cada vector de la base ortogonal que obtuvimos.

$$\|r_0\|^2 = \int_{-1}^{1} r_0(x)^2 dx = \int_{-1}^{1} 1 dx = x|_{-1}^{1} = 2,$$

$$\|r_1\|^2 = \int_{-1}^{1} r_1(x)^2 dx = \int_{-1}^{1} x^2 dx = \frac{x^3}{3}\Big|_{-1}^{1} = \frac{2}{3},$$

$$\|r_2\|^2 = \int_{-1}^{1} r_2(x)^2 dx = \int_{-1}^{1} (x^2 - \frac{1}{3})^2 dx = \frac{8}{45}.$$

Normalizando, nos que

$$p_0(x) = \frac{r_0}{\|r_0\|} = \frac{1}{\sqrt{2}},$$

$$p_1(x) = \frac{r_1(x)}{\|r_1\|} = \sqrt{\frac{3}{2}} x,$$

$$p_2(x) = \frac{r_2(x)}{\|r_2\|} = \sqrt{\frac{45}{8}}(x^2 - \frac{1}{3}).$$

Entonces, la base ortonormal nos queda

$$\{p_0, p_1, p_2\} = \{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{45}{8}}(x^2 - \frac{1}{3})\}.$$

Recuerden que siempre podemos verificar si hicimos bien las cuentas viendo que efectivamente $\langle\, p_0, p_1 \,\rangle = \langle\, p_0, p_2 \,\rangle = \langle\, p_1, p_2 \,\rangle = 0$ y $\|p_0\| = \|p_1\| = \|p_2\| = 1$.

b) : Como en el paso a) calculamos una base ortonormal del subespacio $\mathbb{R}_2[x]$ podemos aplicar la fórmula de la proyección ortogonal. Recuerden que esa fórmula sólo se puede usar cuando tenemos una **base ortogonal** del subespacio en cuestión, en este caso tenemos una base ortonormal (en particual es ortogonal). Entonces, usando la base ortonormal $\{p_0, p_1, p_2\}$ del item a), nos queda

$$P_{\mathbb{R}_2[x]}(\sin(x)) = \langle\, \sin(x), p_0 \,\rangle\, p_0 + \langle\, \sin(x), p_1 \,\rangle\, p_1 + \langle\, \sin(x), p_2 \,\rangle\, p_2.$$

Haciendo cuentas, tenemos que

$$\langle\, \sin(x), p_0 \,\rangle = \int_{-1}^{1} \sin(x) p_0(x)\ dx = \int_{-1}^{1} \sin(x) \frac{1}{\sqrt{2}}\ dx = 0.$$

$$\langle\, \sin(x), p_1 \,\rangle = \int_{-1}^{1} \sin(x) p_1(x)\ dx = \int_{-1}^{1} \sin(x) \sqrt{\frac{3}{2}}\ x\ dx = \sqrt{\frac{3}{2}}(2\sin(1) - 2\cos(1)).$$

$$\langle\, \sin(x), p_2 \,\rangle = \int_{-1}^{1} \sin(x) p_2(x)\ dx = \int_{-1}^{1} \sin(x) \sqrt{\frac{45}{8}}(x^2 - \frac{1}{3})\ dx = 0.$$

Entonces,

$$\begin{aligned} P_{\mathbb{R}_2[x]}(\sin(x)) &= \sqrt{\frac{3}{2}}(2\sin(1) - 2\cos(1))(\sqrt{\frac{3}{2}}\ x) \\ &= 3(\sin(1) - \cos(1))\ x. \end{aligned}$$

De manera similar,

$$P_{\mathbb{R}_2[x]}(\cos(x)) = \langle\, \cos(x), p_0 \,\rangle\, p_0 + \langle\, \cos(x), p_1 \,\rangle\, p_1 + \langle\, \cos(x), p_2 \,\rangle\, p_2.$$

Haciendo cuentas, tenemos que

$$\langle\, \cos(x), p_0 \,\rangle = \int_{-1}^{1} \cos(x) p_0(x)\ dx = \int_{-1}^{1} \cos(x) \frac{1}{\sqrt{2}}\ dx = \frac{1}{\sqrt{2}}\ 2\ \sin(1).$$

$$\langle\, \cos(x), p_1 \,\rangle = \int_{-1}^{1} \cos(x) p_1(x)\ dx = \int_{-1}^{1} \cos(x) \sqrt{\frac{3}{2}}\ x\ dx = 0.$$

$$\langle\, \cos(x), p_2 \,\rangle = \int_{-1}^{1} \cos(x) p_2(x)\ dx = \int_{-1}^{1} \cos(x) \sqrt{\frac{45}{8}}(x^2 - \frac{1}{3})\ dx = \sqrt{\frac{45}{8}}(4\cos(1) - \frac{8\sin(1)}{3}).$$

Entonces

$$\begin{aligned} P_{\mathbb{R}_2[x]}(\cos(x)) &= \frac{1}{\sqrt{2}}\ 2\ \sin(1)(\frac{1}{\sqrt{2}}\ 1) + \sqrt{\frac{45}{8}}(4\cos(1) - \frac{8\sin(1)}{3})(\sqrt{\frac{45}{8}}(x^2 - \frac{1}{3})) \\ &= \sin(1)\ 1 + \frac{15}{2}\ (3\cos(1) - 2\sin(1))\ (x^2 - \frac{1}{3}). \end{aligned}$$

c) : Recordando la fórmula de la distancia de un punto a un subespacio, tenemos que

$$d(\sin(x), \mathbb{R}_2[x]) = \| \sin(x) - P_{\mathbb{R}_2[x]}(\sin(x)\| = \| \sin(x) - 3(\sin(1) - \cos(1))\ x\|.$$

Haciendo la cuenta, tenemos que

$$\| \sin(x) - 3(\sin(1) - \cos(1))\ x\|^2 = \int_{-1}^{1} (\sin(x) - 3(\sin(1) - \cos(1))\ x)^2\ dx = \frac{11\sin(2)}{2} - 5.$$

Entonces

$$d(\sin(x), \mathbb{R}_2[x]) = \sqrt{\frac{11\sin(2)}{2} - 5} \approx 0,0337.$$

$$d(\cos(x), \mathbb{R}_2[x]) = \| \cos(x) - P_{\mathbb{R}_2[x]}(\cos(x)\| = \| \cos(x) - \sin(1)\ 1 - \frac{15}{2}\ (3\cos(1) - 2\sin(1))\ (x^2 - \frac{1}{3})\|.$$

Haciendo la cuenta, tenemos que

$$\begin{aligned} &\| \cos(x) - \sin(1)\ 1 - \frac{15}{2}\ (3\cos(1) - 2\sin(1))\ (x^2 - \frac{1}{3})\|^2 \\ &= \int_{-1}^{1} (\cos(x) - \sin(1)\ 1 - \frac{15}{2}\ (3\cos(1) - 2\sin(1))\ (x^2 - \frac{1}{3}))^2\ dx \\ &= \frac{121\sin(2)}{2} - 65 - 24\cos(2). \end{aligned}$$

Entonces

$$d(\cos(x), \mathbb{R}_2[x]) = \sqrt{\frac{121\sin(2)}{2} - 65 - 24\cos(2)} \approx 0,0043.$$

■

Como se ve la aproximación de las funciones $\sin(x)$ y $\cos(x)$ por los 3 primeros polinomios de Legendre normalizados es bastante buena.

Ejercicio 3.O (De Parcial)**.** Se considera $\mathbb{R}_2[x]$ con el producto interno definido por

$$\langle\, p, q\,\rangle = \int_0^1 p(x)q(x)dx.$$

Sea $C = \{1; x - \frac{1}{2}; x^2 - x + \frac{1}{6}\}$ una base (ordenada) ortogonal de $\mathbb{R}_2[x]$ que se obtuvo aplicando el algoritmo de Gram-Schmidt a la base ordenada $B = \{p_1; p_2; p_3\}$ de $\mathbb{R}_2[x]$. Calcular

$$d(q, gen\{p_1, p_2\}),$$

donde $q(x) = 1 + x^2$.

Dem. Recordar que como $gen\{p_1, p_2\}$ es un subespacio de dimensión finita, tenemos que

$$d(q, gen\{p_1, p_2\}) = \|q - P_{gen\{p_1,p_2\}}(q)\|.$$

No sabemos quiénes son los vectores p_1 y p_2. Sin embargo, sabemos que la base (ordenada) ortogonal C se obtuvo aplicando el algoritmo de Gram-Schmidt a la base ordenada B. Vimos al principio de esta Sección que el algoritmo de Gram-Schmidt, asegura que:

$$\begin{aligned} gen\{p_1\} &= gen\{1\}, \\ gen\{p_1, p_2\} &= gen\{1, x - \frac{1}{2}\}, \\ gen\{p_1, p_2, p_3\} &= gen\{1, x - \frac{1}{2}, x^2 - x + \frac{1}{6}\}. \end{aligned}$$

Por lo tanto, $gen\{p_1, p_2\} = gen\{1, x-\frac{1}{2}\}$ y además $\{1, x-\frac{1}{2}\}$ es una **base ortogonal** de $gen\{p_1, p_2\}$. Entonces, podemos aplicar la fórmula de la proyección ortogonal y tenemos que

$$\begin{aligned}
P_{gen\{p_1,p_2\}}(q) = P_{gen\{p_1,p_2\}}(1+x^2) &= \frac{\langle\, 1+x^2, 1\,\rangle}{\|1\|^2} 1 + \frac{\langle\, 1+x^2, x-\frac{1}{2}\,\rangle}{\|x-\frac{1}{2}\|^2}(x-\frac{1}{2}) \\
&= \frac{\frac{4}{3}}{1} 1 + \frac{\frac{1}{12}}{\frac{1}{12}}(x-\frac{1}{2}) \\
&= \frac{4}{3} + (x-\frac{1}{2}) = x + \frac{5}{6}.
\end{aligned}$$

Entonces

$$\begin{aligned}
d(1+x^2, gen\{p_1,p_2\}) &= \|1+x^2 - P_{gen\{p_1,p_2\}}(1+x^2)\| = \|1+x^2-(x+\frac{5}{6})\| \\
&= \|x^2 - x + \frac{1}{6}\| = \sqrt{\frac{1}{180}}.
\end{aligned}$$

■

3.10. Descomposición QR

Sea $A \in \mathbb{K}^{m\times n}$ de rango n. Una *descomposición QR* de A es una factorización

$$A = QR \text{ con } Q \in \mathbb{K}^{m\times n} \text{ y } R \in \mathbb{K}^{n\times n},$$

tales que las columnas de Q son una base ortonormal de $\operatorname{col}(A)$ considerando el producto interno canónico de $\mathbb{R}^m$ y R es triangular superior con números positivos en la diagonal principal.

Ejercicio 3.28. Sea $A \in \mathbb{R}^{m\times n}$ de rago n.

a) Comprobar que existe una descomposición QR de A.

b) Demostrar que la descomposición QR de A es única.

Dem. $a)$: Lo vamos a demostrar usando el algoritmo de Gram-Schmidt.

Supongamos que $a_1, a_2, \cdots a_n \in \mathbb{R}^m$ son las columnas de A, es decir $A = [a_1\ a_2\ a_3\ \cdots\ a_n]$. Como $\dim(\operatorname{col}(A)) = \operatorname{rg}(A) = n$, tenemos que el conjunto $\{a_1, a_2, \cdots, a_n\}$ es LI, entonces podemos aplicar el algoritmo de Gram-Schimdt a dicho conjunto usando el producto interno canónico de $\mathbb{R}^m$. Si hacemos eso, vamos a obtener una base ortogonal $\{w_1, w_2, \cdots, w_n\}$ de $\operatorname{col}(A)$, donde:

$$\begin{aligned}
w_1 &= a_1 \\
w_2 &= a_2 - \alpha_{12} w_1, \\
w_3 &= a_3 - \alpha_{13} w_1 - \alpha_{23} w_2, \\
&\vdots \\
w_n &= a_n - \alpha_{1n} w_1 - \alpha_{2n} w_2 - \cdots - \alpha_{(n-1)n} w_{n-1}.
\end{aligned}$$

Con

$$\alpha_{ij} = \frac{\langle\, a_j, w_i\,\rangle}{\|w_i\|^2} = \frac{w_i^T a_j}{\|w_i\|^2} \text{ para } 1 \le i < j.$$

Despejando cada a_i de la ecuación anterior, obtenemos:

$$\begin{aligned}
a_1 &= w_1, \\
a_2 &= \alpha_{12}w_1 + w_2, \\
a_3 &= \alpha_{13}w_1 + \alpha_{23}w_2 + w_3, \\
&\vdots \\
a_n &= \alpha_{1n}w_1 + \alpha_{2n}w_2 + \cdots + \alpha_{(n-1)n}w_{n-1} + w_n.
\end{aligned}$$

Entonces, si escribimos lo anterior de manera matricial, nos queda

$$A = [a_1\ a_2\ a_3\ \cdots\ a_n] = [w_1\ w_2\ w_3\ \cdots\ w_n] \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ 0 & 1 & \alpha_{23} & \cdots & \alpha_{2n} \\ 0 & 0 & 1 & \cdots & \alpha_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Llamemos $Q_0 = [w_1\ w_2\ \cdots\ w_n] \in \mathbb{R}^{m\times n}$ y $R_0 = \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ 0 & 1 & \alpha_{23} & \cdots & \alpha_{2n} \\ 0 & 0 & 1 & \cdots & \alpha_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}$.

Entonces $A = Q_0R_0$ y ya casi estamos. Notar que las columnas de Q_0 forman una base ortogonal de $\mathbb{R}^m$ (con el producto interno canónico de $\mathbb{R}^m$). Queremos que las columnas formen una base ortonormal, eso lo arreglamos definiendo

$$Q := [q_1\ q_2\ q_3\ \cdots\ q_n] \text{ con } q_i = \frac{w_i}{\|w_i\|}.$$

Para que el producto nos siga dando A necesitamos modificar R_0 de la siguiente manera:

$$R := \begin{bmatrix} \|w_1\| & \alpha_{12}\|w_1\| & \alpha_{13}\|w_1\| & \cdots & \alpha_{1n}\|w_1\| \\ 0 & \|w_2\| & \alpha_{23}\|w_2\| & \cdots & \alpha_{2n}\|w_2\| \\ 0 & 0 & \|w_3\| & \cdots & \alpha_{3n}\|w_3\| \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \|w_n\| \end{bmatrix}.$$

Entonces

$$A = QR,$$

donde R es triangular superior con elementos positivos (no nulos) en su diagonal principal y Q es una matriz cuyas columnas forman una base ortonormal de $\mathbb{R}^m$. Sólo nos resta ver que $\operatorname{col}(Q) = \operatorname{col}(A)$. De hecho, como $A = QR$, entonces es claro que $\operatorname{col}(A) = \operatorname{col}(QR) \subseteq \operatorname{col}(Q)$. Por otra parte, como $R \in \mathbb{R}^{n\times n}$ es una matriz triangular superior con elementos positivos (no nulos) en su diagonal principal, R es una matriz inversible. Por lo tanto, $\det(R) > 0$, entonces

$$AR^{-1} = (QR)R^{-1} = QI = Q.$$

Entonces, $\operatorname{col}(Q) = \operatorname{col}(AR^{-1}) \subseteq \operatorname{col}(A)$. Como probamos la doble inclusión, concluimos que $\operatorname{col}(Q) = \operatorname{col}(A)$.

Entonces $A = QR$ es una descomposición QR de A.

b) : Veamos que la factorización $A = QR$ del item a) es única. Pero antes de hacer la prueba, vamos a remarcar los siguientes hechos que se pueden verificar fácilmente:

- Si $R \in \mathbb{R}^{n\times n}$ es una matriz **triangular superior** entonces R^T es una matriz **triangular inferior**.
- Si $R \in \mathbb{R}^{n\times n}$ es una matriz **triangular superior** inversible entonces R^{-1} también es una matriz **triangular superior** inversible.
- Si $R, R' \in \mathbb{R}^{n\times n}$ son dos matrices **triangulares superiores** inversibles entonces RR'^{-1} y $R'R^{-1}$ también son matrices **triangulares superiores** inversibles.
- Si $Q \in \mathbb{R}^{m\times n}$ es una matriz cuyas columnas forman una base ortonormal (tomando el producto interno canónico de $\mathbb{R}^m$). Entonces

$$\boxed{Q^TQ = I_{\mathbb{R}^n}.}$$

De hecho, si $Q = [q_1\ q_2\ \cdots\ q_n]$, donde $q_1, q_2, \cdots, q_n \in \mathbb{R}^m$ son las columnas de Q. Entonces

$$Q^TQ = \begin{bmatrix} q_1^T \\ q_2^T \\ \cdots \\ q_n^T \end{bmatrix} [q_1\ q_2\ \cdots\ q_n] = \begin{bmatrix} q_1^Tq_1 & q_1^Tq_2 & \cdots & q_1^Tq_n \\ q_2^Tq_1 & q_2^Tq_2 & \cdots & q_2^Tq_n \\ \vdots & \vdots & & \vdots \\ q_n^Tq_1 & q_n^Tq_2 & \cdots & q_n^Tq_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_{\mathbb{R}^n}.$$

Donde usamos que las columnas de Q forman una base ortonormal (con el producto interno canónico de $\mathbb{R}^m$). Entonces $0 = \langle\, q_i, q_j \,\rangle = q_j^T q_i$ para todo $i \neq j$ y $1 = \|q_i\|^2 = q_i^T q_i$, para todo $i = 1, 2, \cdots, n$.

Ahora sí, probemos el item b). Supongamos que $A = Q'R'$ es otra descomposición QR de A. Entonces, $R' \in \mathbb{R}^{n\times n}$ es triangular superior con elementos positivos (no nulos) en su diagonal principal y $Q' \in \mathbb{R}^{m\times n}$ es una matriz cuyas columnas forman una base ortonormal de $\mathrm{col}(A)$. Como $A = QR$ tenemos que

$$QR = Q'R'. \tag{3.12}$$

Multipliquemos la ecuación (3.12) a izquierda por Q'^T y a derecha por R^{-1}, entonces usando las propiedades que acabamos de ver, tenemos por un lado

$$Q'^T(QR)R^{-1} = Q'^T(Q'R')R^{-1} = (Q'^TQ')R'R^{-1} = R'R^{-1}$$

y por el otro

$$Q'^T(QR)R^{-1} = Q'^TQ(RR^{-1}) = Q'^TQ.$$

Entonces, probamos que

$$Q'^TQ = Q'^T(QR)R^{-1} = R'R^{-1}. \tag{3.13}$$

De (3.13) tenemos que como $R'R^{-1}$ es una matriz triangular superior, entonces Q'^TQ es una matriz triangular superior.

Ahora, multipliquemos la ecuación (3.12) a izquierda por Q^T y a derecha por R'^{-1}, entonces usando las propiedades que acabamos de ver, tenemos por un lado

$$Q^T(QR)R'^{-1} = Q^T(Q'R')R'^{-1} = Q^TQ'(R'R'^{-1}) = Q^TQ'$$

y por el otro lado

$$Q^T(QR)R'^{-1} = (Q^TQ)RR'^{-1} = RR'^{-1}.$$

Entonces, probamos que

$$RR'^{-1} = Q^T(QR)R'^{-1} = Q^TQ'. \tag{3.14}$$

De (3.14) tenemos que como RR'^{-1} es una matriz triangular superior, entonces Q^TQ' es una matriz triangular superior. Entonces, como Q^TQ' es una matriz triangular superior, al transponer, se sigue que $(Q^TQ')^T$ es triangular inferior. Por el otro lado, como $(Q^TQ')^T = Q'^TQ$, por (3.13), tenemos que $(Q^TQ')^T$ es triangular superior. Entonces $(Q^TQ')^T = Q'^TQ$, es a la vez triangular superior y triangular inferior. Entonces Q'^TQ es una matriz diagonal. A esa matriz diagonal, la llamamos D y entonces tenemos que

$$Q'^TQ = R'R^{-1} = D := \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix},$$

con $d_1, d_2, \cdots, d_n \in \mathbb{R}$. Multiplicando (3.12) a izquierda por Q'^T, tenemos por un lado que

$$Q'^T(QR) = (Q'^TQ)R = DR$$

y por el otro lado

$$Q'^T(Q'R') = (Q'^TQ')R' = R'.$$

Entonces, nos queda que

$$R' = DR.$$

Observar que como R y R' son dos matrices **triangulares superiores** con elementos positivos (no nulos) en su diagonal principal, no queda otra (hacer la cuenta) que $d_1, d_2, \cdots, d_n$ sean todos números positivos (no nulos).

Ahora, multiplicando (3.12) a derecha por R'^{-1}, tenemos por un lado que

$$(QR)R'^{-1} = Q(RR'^{-1}) = QD$$

y por el otro lado

$$(Q'R')R'^{-1} = Q'(R'R'^{-1}) = Q'.$$

Entonces, tenemos que $Q' = QD$ y

$$I = Q'^TQ' = (QD)^T(QD) = D^T(Q^TQ)D = D^TID = D^TD = D^2,$$

recordar que como D es diagonal $D^T = D$. Entonces

$$\begin{bmatrix} d_1^2 & 0 & \cdots & 0 \\ 0 & d_2^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Por lo tanto $|d_1| = |d_2| = \cdots = |d_n| = 1$. Como probamos que $d_1, d_2, \cdots, d_n$ son todos positivos (no nulos), se sigue que $d_1 = d_2 = \cdots = d_n = 1$, entonces $D = I$. Por lo tanto $R'R^{-1} = D = I$ y multiplicando esa ecuación a derecha por R, nos queda que

$$R' = R.$$

De la misma manera, teníamos que $Q' = QD = QI = Q$, entonces

$$Q' = Q.$$

Con esto probamos que la descomposición QR de A es única. ∎

Antes de hacer ejemplos de cálculo de descomposición QR, vamos a notar dos propiedades importantes.

Primero observar que si $A = QR$ es la descomposición QR de A. Entonces, usando que $Q^T Q = I$, se sigue que

$$Q^T A = Q^T QR = R, \tag{3.15}$$

y tenemos una forma fácil de hallar R una vez encontrada Q.

Por otra parte,

considerando el producto interno canónico, notar que

$$QQ^T = [P_{\mathrm{col}(A)}]_E^E, \tag{3.16}$$

donde E es la base canónica de $\mathbb{R}^m$.

De hecho, observar que $(QQ^T)(QQ^T) = Q(Q^T Q)Q^T = QI_{\mathbb{R}^n}Q^T = QQ^T$. Entonces

$$(QQ^T)^2 = QQ^T. \tag{3.17}$$

Además,

$$(QQ^T)^T = (Q^T)^T Q^T = QQ^T. \tag{3.18}$$

Sea $T : \mathbb{C}^m \to \mathbb{C}^m$, la transformación lineal definida por

$$T(x) := QQ^T(x).$$

Entonces, claramente T es un proyector de $\mathbb{C}^m$. De hecho, por (3.17), se sigue que $T^2(x) = (QQ^T)^2(x) = QQ^T(x) = T(x)$. Entonces $T^2 = T$.

Por otra parte, considerando el **producto interno canónico** de $\mathbb{C}^m$, por (3.18), para cada $x, y \in \mathbb{C}^m$, tenemos que

$$\langle T(x), y \rangle = (T(x))^T y = (QQ^T(x))^T y = x^T (QQ^T)^T y = x^T QQ^T(y) = x^T T(y) = \langle x, T(y) \rangle .$$

Entonces, tenemos que T es un proyector y

$$\langle T(x), y \rangle = \langle x, T(y) \rangle ,$$

para todo $x, y \in \mathbb{C}^m$. Por lo tanto, por el Teorema 3.6.1, T es un proyector ortogonal.

Además, por un lado es claro que $\mathrm{Im}(T) = \mathrm{col}(QQ^T)$ y por el otro lado, recordemos que probamos que $\mathrm{col}(QQ^T) = \mathrm{nul}((QQ^T)^T)^\perp = \mathrm{nul}(QQ^T)^\perp = \mathrm{nul}(Q^T)^\perp = \mathrm{col}(Q)$. Entonces

$$\mathrm{Im}(T) = \mathrm{col}(QQ^T) = \mathrm{col}(Q) = \mathrm{col}(A).$$

Por lo tanto, probamos que

$$T = P_{\mathrm{col}(A)}.$$

Finalmente, si E es la base canónica de $\mathbb{R}^m$, es inmediato ver que $[T]_E^E = QQ^T$. Por lo tanto

$$[P_{\mathrm{col}(A)}]_E^E = [T]_E^E = QQ^T$$

y probamos lo que queríamos.

Ejercicio 3.29. Hallar la descomposición QR de

$$A_1 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{bmatrix}.$$

Sea $\mathcal{S} = gen\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}\}$. Considerando el producto interno canónico de $\mathbb{R}^3$ calcular $[P_{\mathcal{S}}]_E^E$, donde E es la base canónica de $\mathbb{R}^3$.

Dem. Como $A_1 \in \mathbb{R}^{3\times 2}$ y $\operatorname{rg}(A_1) = 2$ podemos usar lo que probamos en el **Ejercicio 3.28**. Sólo basta encontrar una base ortonormal (con el producto interno canónico de $\mathbb{R}^3$) del subespacio

$$\operatorname{col}(A_1) = gen\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}\}.$$

Usando Gram-Schmidt (hacer la cuenta), se sigue que

$$\{\frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{210}}\begin{bmatrix} 5 \\ -8 \\ 11 \end{bmatrix}\}$$

es una base ortonormal de $\operatorname{col}(A_1)$. Entonces, tomamos

$$Q := \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{5}{\sqrt{210}} \\ \frac{2}{\sqrt{6}} & \frac{-8}{\sqrt{210}} \\ \frac{1}{\sqrt{6}} & \frac{11}{\sqrt{210}} \end{bmatrix}.$$

Entonces, por (3.15), tenemos que

$$R = Q^T A_1 = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{5}{\sqrt{210}} & \frac{-8}{\sqrt{210}} & \frac{11}{\sqrt{210}} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} \sqrt{6} & \frac{7}{\sqrt{6}} \\ 0 & \frac{35}{\sqrt{210}} \end{bmatrix}.$$

Las matrices Q y R que definimos cumplen que $A_1 = QR$, donde $Q \in \mathbb{R}^{3\times 2}$ es una matriz cuyas columnas forman una base ortonormal de $\operatorname{col}(A_1)$ y $R \in \mathbb{R}^{2\times 2}$ es una matriz triangular superior con elementos positivos (no nulos) en la diagonal. Esa es la descomposición QR de A_1.

Finalmente, observar que $\mathcal{S} = gen\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}\} = \operatorname{col}(A_1)$. Entonces, usando la ecuación (3.16), que vale porque estamos considerando **el producto interno canónico de** $\mathbb{R}^3$ **y la base canónica de** $\mathbb{R}^3$ tenemos

$$[P_{\mathcal{S}}]_E^E = QQ^T = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{5}{\sqrt{210}} \\ \frac{2}{\sqrt{6}} & \frac{-8}{\sqrt{210}} \\ \frac{1}{\sqrt{6}} & \frac{11}{\sqrt{210}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{5}{\sqrt{210}} & -\frac{8}{\sqrt{210}} & \frac{11}{\sqrt{210}} \end{bmatrix} = \begin{bmatrix} \frac{2}{7} & \frac{1}{7} & \frac{3}{7} \\ \frac{1}{7} & \frac{34}{35} & -\frac{3}{35} \\ \frac{3}{7} & -\frac{3}{35} & \frac{26}{35} \end{bmatrix}.$$

■

3.11. Teorema de representación de Riesz

Finalizamos la Unidad con una teorema muy importante y varias aplicaciones del mismo.

Teorema de representación de Riesz: sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo de dimensión finita y $\phi : \mathbb{V} \to \mathbb{K}$ una funcional lineal, entonces existe un único vector $u \in \mathbb{V}$ tal que

$$\phi(v) = \langle v, u \rangle,$$

para todo $v \in \mathbb{V}$.

El siguiente ejercicio es una demostración del Teorema de representación de Riesz.

Ejercicio 3.30. Sea $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ un $\mathbb{K}$-espacio euclídeo finito dimensional.

a) Sea $\phi \in \mathcal{L}(\mathbb{V}, \mathbb{K})$ tal que $\phi \neq 0$. Demostrar que

$$\phi(v) = \left\langle v, \overline{\phi(w)}w \right\rangle,$$

para cualquier $w \in \mathrm{Nu}(\phi)^{\perp}$ tal que $\|w\| = 1$.

b) Notar que $\overline{\phi(w)}w$ no depende de la elección de w.

Observar que si lo que afirma el ejercicio es correcto, podemos tomar $u := \overline{\phi(w)}w$, entonces u es el único vector de $\mathbb{V}$ tal que

$$\phi(v) = \langle v, u \rangle,$$

para todo $v \in \mathbb{V}$. Por lo tanto, si resolvemos el **Ejercicio 3.30**, estamos demostrando el Teorema de representación de Riesz.

Dem. $a)$: Como $\phi \neq 0$, existe $v \neq 0_{\mathbb{V}}$ tal que $\phi(v) \neq 0$ (¿por qué?) y entonces $\mathrm{Im}(\phi) = \mathbb{K}$ y, como $\mathbb{V}$ es de dimensión finita, por el teorema de la dimensión, se sigue que

$$\dim(\mathrm{Nu}(\phi)) = \dim(\mathbb{V}) - \dim(\mathrm{Im}(\phi)) = \dim(\mathbb{V}) - \dim(\mathbb{K}) = \dim(\mathbb{V}) - 1.$$

De la misma manera, como $\mathrm{Nu}(\phi)$ es un subespacio y $\mathbb{V}$ es de dimensión finita, tenemos que

$$\dim(\mathrm{Nu}(\phi)^{\perp}) = \dim(\mathbb{V}) - \dim(\mathrm{Nu}(\phi)) = \dim(\mathbb{V}) - (\dim(\mathbb{V}) - 1) = 1.$$

Entonces, tomemos cualquier $w \in \mathrm{Nu}(\phi)^{\perp}$ tal que $\|w\| = 1$. Como $\dim(\mathrm{Nu}(\phi)^{\perp}) = 1$, tenemos que

$$\mathrm{Nu}(\phi)^{\perp} = gen\{w\}.$$

Sea $v \in \mathbb{V}$, como $\mathrm{Nu}(\phi)$ es un subespacio y $\mathbb{V}$ es de dimensión finita, vale que $\mathbb{V} = \mathrm{Nu}(\phi) \oplus \mathrm{Nu}(\phi)^{\perp}$, entonces existen (únicos) $v_1 \in \mathrm{Nu}(\phi)$ y $v_2 \in \mathrm{Nu}(\phi)^{\perp}$ tales que $v = v_1 + v_2$. Como $v_2 \in \mathrm{Nu}(\phi)^{\perp} = gen\{w\}$, tenemos que existe $\alpha \in \mathbb{K}$ tal que $v_2 = \alpha w$. Es más, podemos ver quién es α. De hecho

$$\begin{aligned}\langle v, w \rangle &= \langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle = \langle v_1, w \rangle + \langle \alpha w, w \rangle = 0 + \alpha \langle w, w \rangle \\ &= \alpha \|w\|^2 = \alpha \cdot 1 = \alpha.\end{aligned}$$

Entonces $\alpha = \langle v, w \rangle$. Por lo tanto,

$$\begin{aligned}\phi(v) &= \phi(v_1 + v_2) = \phi(v_1) + \phi(v_2) = 0 + \phi(v_2) = \phi(v_2) = \phi(\alpha w) = \alpha\phi(w) \\ &= \langle v, w \rangle \phi(w) = \phi(w) \langle v, w \rangle = \left\langle v, \overline{\phi(w)}w \right\rangle,\end{aligned}$$

donde para la última igualdad usamos que $\phi(w) \in \mathbb{K}$ (es decir es un escalar) y que el producto interno cumple que $c \langle u, v \rangle = \langle u, \overline{c}v \rangle$.

Entonces, probamos que para todo $v \in \mathbb{V}$ vale que

$$\phi(v) = \left\langle v, \overline{\phi(w)}w \right\rangle,$$

para cualquier $w \in \mathrm{Nu}(\phi)^{\perp}$ tal que $\|w\| = 1$.

$b)$: Tomemos otro vector $w' \in \mathrm{Nu}(\phi)^{\perp}$ tal que $\|w'\| = 1$. Veamos que $\overline{\phi(w')}w' = \overline{\phi(w)}w$. Observar que como $w, w' \in \mathrm{Nu}(\phi)^{\perp}$ (y no son nulos) entonces $\mathrm{Nu}(\phi)^{\perp} = gen\{w\} = gen\{w'\}$. Entonces, como $\overline{\phi(w')}w' \in \mathrm{Nu}(\phi)^{\perp} = gen\{w\}$, existe $\alpha \in \mathbb{K}$ tal que

$$\overline{\phi(w')}w' = \alpha w. \tag{3.19}$$

Veamos a qué es igual α. Por un lado

$$\langle \alpha w, w \rangle = \alpha \langle w, w \rangle = \alpha \|w\|^2 = \alpha.$$

Por el otro lado, por lo que acabamos de probar en el item a), como w' es cualquier vector en $\mathrm{Nu}(\phi)^\perp$ tal que $\|w'\| = 1$, vale que

$$\phi(w) = \left\langle w, \overline{\phi(w')}w' \right\rangle.$$

Entonces, usando (3.19) y la observación que acabamos de hacer se sigue que

$$\alpha = \langle \alpha w, w \rangle = \left\langle \overline{\phi(w')}w', w \right\rangle = \overline{\left\langle w, \overline{\phi(w')}w' \right\rangle} = \overline{\phi(w)}.$$

Entonces $\alpha = \overline{\phi(w)}$ y volviendo a (3.19), se sigue que $\overline{\phi(w')}w' = \overline{\phi(w)}w$ y probamos que $\overline{\phi(w)}w$ no depende de la elección de w. ■

Con la resolución del **Ejercicio 3.30** probamos el Teorema de representación de Riesz. Veamos ahora un ejemplo de aplicación de lo que vimos en dicho ejercicio.

Ejercicio 3.33. En $\mathbb{R}_n[x]$ con el producto interno

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)dx.$$

Se considera $\delta : \mathbb{R}^n[x] \to \mathbb{R}$ la funcional lineal definida por

$$\delta(p) = p(0).$$

Para cada $n \in \{1, 2, 3, 4, 5, 6\}$ hallar el polinomio $p_n \in \mathbb{R}_n[x]$ tal que

$$\delta(\cdot) = \langle \cdot, p_n \rangle.$$

Por si no se entiende la notación, lo que nos pide el ejercicio es hallar hallar el polinomio $p_n \in \mathbb{R}_n[x]$ tal que

$$\delta(p) = \langle p, p_n \rangle,$$

para todo $p \in \mathbb{R}_n[x]$. Vamos a demostrarlo para el caso $n = 3$, los demás casos son similares.

Dem. Consideremos $n = 3$. Por el **Ejercicio 3.30**, para todo $p \in \mathbb{R}_n[x]$, se sigue que para cualquier $q \in \mathrm{Nu}(\delta)^\perp$ tal que $\|q\| = 1$, vale que

$$\delta(p) = \left\langle p, \overline{\delta(q)}q \right\rangle.$$

Si tomamos $p_3(x) := \overline{\delta(q)}q(x)$ entonces encontramos el polinomio p_3 que queríamos. Busquemos entonces algún polinomio q que cumpla lo que buscamos.

Recordar que $\mathrm{Nu}(\delta) = \{p \in \mathbb{R}_3[x] : \delta(p) = 0\}$, entonces $p \in \mathrm{Nu}(\delta)$ si $p(x) = ax^3 + bx^2 + cx + d$ con $a, b, c, d \in \mathbb{R}$ y

$$0 = \delta(p) = p(0) = a(0)^3 + b(0)^2 + c(0) + d = d.$$

Entonces $d = 0$ y volviendo a la expresión de p tenemos que $p(x) = ax^3 + bx^2 + cx$, con $a, b, c \in \mathbb{R}$. Entonces

$$Nu(\delta) = gen\{x, x^2, x^3\}.$$

A continuación, vamos a hallar

$$\mathrm{Nu}(\delta)^\perp = \{r \in \mathbb{R}_3[x] : \langle r, p \rangle = 0 \text{ para todo } p \in \mathrm{Nu}(\delta)\}.$$

Recordemos que $r \in \mathrm{Nu}(\delta)^\perp$ si r es ortogonal a cada generador de $\mathrm{Nu}(\delta)$. Entonces, $r \in \mathrm{Nu}(\delta)^\perp$ si $r(x) = ex^3 + fx^2 + gx + h$ con $e, f, g, h \in \mathbb{R}$ y

$$0 = \langle r, x \rangle = \left\langle r, x^2 \right\rangle = \left\langle r, x^3 \right\rangle.$$

Haciendo cuentas, tenemos que

$$0 = \langle r, x \rangle = \int_{-1}^{1} (ex^3 + fx^2 + gx + h)xdx = e\frac{x^5}{5} + f\frac{x^4}{4} + g\frac{x^3}{3} + h\frac{x^2}{2}|_{-1}^{1} = e\frac{2}{5} + g\frac{2}{3}.$$

$$0 = \left\langle r, x^2 \right\rangle = \int_{-1}^{1} (ex^3 + fx^2 + gx + h)x^2dx = e\frac{x^6}{6} + f\frac{x^5}{5} + g\frac{x^4}{4} + h\frac{x^3}{3}|_{-1}^{1} = f\frac{2}{5} + h\frac{2}{3}.$$

$$0 = \left\langle r, x^3 \right\rangle = \int_{-1}^{1} (ex^3 + fx^2 + gx + h)x^3dx = e\frac{x^7}{7} + f\frac{x^6}{6} + g\frac{x^5}{5} + h\frac{x^4}{4}|_{-1}^{1} = e\frac{2}{7} + g\frac{2}{5}.$$

Nos quedaron las ecuaciones

$$e\frac{2}{5} + g\frac{2}{3} = f\frac{2}{5} + h\frac{2}{3} = e\frac{2}{7} + g\frac{2}{5} = 0.$$

Resolviendo el sistema, obtenemos que

$$e = g = 0 \text{ y } f = -\frac{5}{3}h.$$

Entonces, volviendo a la expresión de r, tenemos que $r(x) = ex^3 + fx^2 + gx + h = h(1 - \frac{5}{3}x^2)$ con $h \in \mathbb{R}$. Entonces

$$\text{Nu}(\delta)^{\perp} = gen\{1 - \frac{5}{3}x^2\}.$$

Notar que

$$\|1 - \frac{5}{3}x^2\|^2 = \int_{-1}^{1} (1 - \frac{5}{3}x^2)^2 dx = \frac{5x^5}{9} - \frac{10x^3}{9} + x|_{-1}^{1} = \frac{8}{9}.$$

Tomando $q(x) := \frac{(1-\frac{5}{3}x^2)}{\sqrt{\frac{8}{9}}} = \sqrt{\frac{9}{8}}(1 - \frac{5}{3}x^2) = \frac{3\sqrt{2}}{4}(1 - \frac{5}{3}x^2)$, tenemos que $q \in \text{Nu}(\delta)^{\perp}$ y $\|q\| = 1$. Observar que $\delta(q) = q(0) = \frac{3\sqrt{2}}{4}$.
Entonces,

$$p_3(x) := \overline{\delta(q)}q(x) = \overline{q(0)}q(x) = \frac{3\sqrt{2}}{4}\,\frac{3\sqrt{2}}{4}(1 - \frac{5}{3}x^2) = \frac{9}{8}(1 - \frac{5}{3}x^2) = \frac{9}{8} - \frac{15}{8}x^2.$$

El polinomio $p_3(x) = \frac{9}{8} - \frac{15}{8}x^2$ que encontramos cumple que, para todo $p \in \mathbb{R}_3[x]$,

$$\delta(p) = \left\langle p, \frac{9}{8} - \frac{15}{8}x^2 \right\rangle.$$

Además p_3 es el único polinomio que cumple la igualdad anterior por el item b) del **Ejercicio 3.30**. ■

3.12. Aplicaciones del Teorema de representación de Riesz

En este sección, estudiaremos varias aplicaciones del Teorema de representación de Riesz. Antes de resolver el **Ejercicio 3.34** repasemos algunas propiedades del determinante.

Sea $A := [A_1 \ A_2 \ \cdots \ A_n] \in \mathbb{K}^{n\times n}$ donde A_i son las n columnas de A. Entonces el determinante es una función multilineal alternada por columnas, es decir:

- $\det([A_1 \ \cdots \ \lambda A_i + A_i' \ \cdots \ A_n]) = \lambda \det([A_1 \ \cdots \ A_i \ \cdots \ A_n]) + \det([A_1 \ \cdots \ A_i' \ \cdots \ A_n])$ para todo $i = 1, \cdots, n$ y $\lambda \in \mathbb{K}$. Es decir, para cada columna i, una vez fijados los valores de las columnas restantes, det es una función lineal en la columna i.
- $\det([A_1 \ \cdots \ A_i \ \cdots \ A_i \ \cdots \ A_n]) = 0$ para cada $i \neq j$, $j = 1, \cdots, n$. El determinante se anula si la matriz tiene dos columnas iguales.

En la siguiente proposición damos algunas de las propiedades básicas que verifica el determinante y que se deducen fácilmente del hecho que el determinante es una función multilineal alternada por columnas.

Proposición 3.12.1. *Sea* $A := [A_1\ A_2\ \cdots\ A_n] \in \mathbb{K}^{n\times n}$ *donde* A_i *son las* n *columnas de* A. *Entonces*

- $\det(A) = \det(A^T)$.
- $\det([A_1\ \cdots\ 0\ \cdots\ A_n]) = 0$. *Es decir, si alguna columna (o fila) es nula el determinante da* 0.
- $\det([A_1\ \cdots\ A_i\ \cdots\ A_j\ \cdots\ A_n]) = -\det([A_1 \cdots\ A_j\ \cdots\ A_i\ \cdots\ A_n])$. *Es decir, si intercambiamos columnas (o filas), el determinante cambia de signo.*
- $\det([A_1\ \cdots\ A_i + \lambda A_j\ \cdots\ A_j\ \cdots\ A_n]) = \det([A_1 \cdots\ A_i\ \cdots\ A_j\ \cdots\ A_n])$. *Es decir, si a una columna (o fila) le sumamos un múltiplo de otra, el determinante no cambia.*
- *Si* $A_i = \sum_{j=1,\ j\neq i}^{n} \alpha_j A_j$ *para ciertos* $a_j \in \mathbb{K}$ *entonces* $\det([A_1\ \cdots\ A_i\ \cdots\ A_n]) = 0$. *Es decir, si una columna (o fila) es combinación lineal de las demás columas (o filas) el determinante da* 0.

A continuación veamos una aplicación interesante del Teorema de representación de Riesz.

Ejercicio 3.34. Sea $(\mathbb{V}, \langle\,\cdot,\cdot\,\rangle)$ un $\mathbb{R}$-espacio euclídeo de dimensión 3. Sea $E = \{e_1, e_2, e_3\}$ una base ortonormal fija. Para $v_1, v_2, v_3 \in \mathbb{V}$ se define

$$\det(v_1, v_2, v_3) := \det([\ [v_1]^E\ [v_2]^E\ [v_3]^E]).$$

a) Sean $v_1, v_2 \in \mathbb{V}$. Comprobar que existe un único vector $v_1 \wedge v_2 \in \mathbb{V}$ tal que

$$\det(v_1, v_2, v) = \langle\, v, v_1 \wedge v_2\,\rangle \text{ para cada } v \in \mathbb{V}.$$

b) Probar que $v_1 \wedge v_2$ es ortogonal a v_1 y v_2.

c) Comprobar que v_1, v_2 son linealmente dependientes si y sólo si $v_1 \wedge v_2 = 0$.

d) Sea $A \in \mathbb{R}^{3\times 2}$ la matriz definida por $A := [\ [v_1]^E\ [v_2]^E\ ]$. Utilizando la descomposición

$$v_1 \wedge v_2 = \langle\, e_1, v_1 \wedge v_2\,\rangle\, e_1 + \langle\, e_2, v_1 \wedge v_2\,\rangle\, e_2 + \langle\, e_3, v_1 \wedge v_2\,\rangle\, e_3$$

probar que

$$v_1 \wedge v_2 = \sum_{i=1}^{3} (-1)^{3-i} \Delta_i e_i$$

donde Δ_i es el determinante de la matriz de 2×2 que se obtiene de A eliminando su i-ésima fila.

e) Utilizando el resultado anterior, observar que

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \wedge \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_2 \end{bmatrix}.$$

f) Probar que si $\{u_1, u_2, u_3\}$ es una base ortonormal de $\mathbb{V}$, entonces vale que $\det(u_1, u_2, u_3) \in \{-1, 1\}$. Si $\det(u_1, u_2, u_3) = 1$, se dice que $\{u_1, u_2, u_3\}$ es orientada positivamente con respecto a E.

g) Probar que si v_1, v_2 es un sistema ortonormal, entonces $v_1, v_2, v_1 \wedge v_2$ es una base ortonormal orientada positivamente con respecto a E.

Dem. $a)$: Llamemos $\phi(v) := \det(v_1, v_2, v) = \det([\ [v_1]^E\ [v_2]^E\ [v]^E])$. Veamos que ϕ es una funcional lineal. De hecho, sean $v, w \in \mathbb{V}$ y $\lambda \in \mathbb{K}$, como tomar coordenadas es lineal, entonces $[\lambda v + w]^E = \lambda[v]^E + [w]^E$. Entonces, $\det([\ [v_1]^E\ [v_2]^E\ [\lambda v+w]^E]) = \det([\ [v_1]^E\ [v_2]^E\ \lambda[v]^E + [w]^E])$. Además, vimos arriba que det es lineal para cada columna fijando el resto de las columnas (como en este caso), entonces se sigue que

$$\begin{aligned}\phi(\lambda v + w) &= \det([\ [v_1]^E\ [v_2]^E\ \lambda[v]^E + [w]^E]) \\ &= \lambda \det([\ [v_1]^E\ [v_2]^E\ [v]^E]) + \det([\ [v_1]^E\ [v_2]^E\ [w]^E]) = \lambda\phi(v) + \phi(w).\end{aligned}$$

Por lo tanto, como ϕ es una funcional lineal de $\mathbb{V}$, por el **Ejercicio 3.30**, existe un único vector $u \in \mathbb{V}$ tal que

$$\phi(v) = \langle\, v, u \,\rangle\,.$$

Sea $v_1 \wedge v_2 := u$, entonces,

$$\det(v_1, v_2, v) = \phi(v) = \langle\, v, v_1 \wedge v_2 \,\rangle$$

y probamos lo que queríamos.

Notar que esto además prueba que la definición de $\det(v_1, v_2, v)$ **NO depende de la base ortonormal** E **elegida, ¿por qué?**.

$b)$: Veamos que $v_1 \wedge v_2$ es ortogonal a v_1 y v_2. De hecho, por el item $a)$ tenemos que,

$$\langle\, v_1, v_1 \wedge v_2 \,\rangle = \phi(v_1) = \det(v_1, v_2, v_1) = \det([\ [v_1]^E\ [v_2]^E\ [v_1]^E]) = 0,$$

pues estamos calculando el determinante de una matriz cuya primera y tercera columna es la misma. De la misma manera,

$$\langle\, v_2, v_1 \wedge v_2 \,\rangle = \phi(v_2) = \det(v_1, v_2, v_2) = \det([\ [v_1]^E\ [v_2]^E\ [v_2]^E]) = 0.$$

$c)$ $\Rightarrow$: Supongamos que v_1 y v_2 son linealmente dependientes. Entonces existe $\lambda \in \mathbb{K}$ tal que $v_1 = \lambda v_2$ (un vector es múltiplo del otro). Entonces, por el item $a)$, tenemos que para cada $v \in \mathbb{V}$,

$$\langle\, v, v_1 \wedge v_2 \,\rangle = \det([\ [v_1]^E\ [v_2]^E\ [v]^E]) = \det([\ [v_1]^E\ [\lambda v_1]^E\ [v]^E]) = \lambda \det([\ [v_1]^E\ [v_1]^E\ [v]^E]) = 0,$$

donde usamos que $[v_2]^E = [\lambda v_1]^E = \lambda[v_1]^E$, que el determinante es lineal para cada columna fijando el resto de las columnas y que el determinante de una matriz cuya primera y segunda columna es la misma da 0. Conclusión, tenemos que para todo $v \in \mathbb{V}$, vale que

$$\langle\, v, v_1 \wedge v_2 \,\rangle = 0.$$

En particular, si tomamos $v = v_1 \wedge v_2$, tenemos que

$$\|v_1 \wedge v_2\| = \langle\, v_1 \wedge v_2, v_1 \wedge v_2 \,\rangle^{1/2} = 0.$$

Por lo tanto, $v_1 \wedge v_2 = 0$.

La recíproca del item $c)$ **la probamos al final**.

$d)$: Como $E = \{e_1, e_2, e_3\}$ es una bon de $\mathbb{V}$ ya vimos en la Proposición 3.7.2, que $[v_1 \wedge v_2]^E = \begin{bmatrix} \langle\, e_1, v_1 \wedge v_2 \,\rangle \\ \langle\, e_2, v_1 \wedge v_2 \,\rangle \\ \langle\, e_3, v_1 \wedge v_2 \,\rangle \end{bmatrix}$. Entonces,

$$v_1 \wedge v_2 = \langle\, e_1, v_1 \wedge v_2 \,\rangle\, e_1 + \langle\, e_2, v_1 \wedge v_2 \,\rangle\, e_2 + \langle\, e_3, v_1 \wedge v_2 \,\rangle\, e_3.$$

Por el item $a)$,

$$\langle\, e_1, v_1 \wedge v_2 \,\rangle = \det([\ [v_1]^E\ [v_2]^E\ [e_1]^E]) = \det([\ [v_1]^E\ [v_2]^E\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}]) = \Delta_1,$$

donde Δ_1 es el determinante de la matriz de 2×2 que se obtiene de A eliminando su primera fila.

De manera similar, $\langle\, e_2, v_1\wedge v_2\,\rangle = \det([\,[v_1]^E\ [v_2]^E\ [e_2]^E]) = \det([\,[v_1]^E\ [v_2]^E\ \begin{bmatrix}0\\1\\0\end{bmatrix}]) = -\Delta_2.$

$\langle\, e_3, v_1\wedge v_2\,\rangle = \det([\,[v_1]^E\ [v_2]^E\ [e_3]^E]) = \det([\,[v_1]^E\ [v_2]^E\ \begin{bmatrix}0\\0\\1\end{bmatrix}]) = \Delta_3.$

Entonces,

$$\begin{aligned}v_1\wedge v_2 &= \langle\, e_1, v_1\wedge v_2\,\rangle e_1 + \langle\, e_2, v_1\wedge v_2\,\rangle e_2 + \langle\, e_3, v_1\wedge v_2\,\rangle e_3\\ &= \Delta_1 e_1 - \Delta_2 e_2 + \Delta_3 e_3\end{aligned}$$

y probamos lo que queríamos.

e) : **Notar que hay un error en la guía.** Apliquemos lo que probamos en el item *f*). Supongamos que $[v_1]^E = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ y $[v_2]^E = \begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$, entonces por el item *f*).

$$v_1\wedge v_2 = \Delta_1 e_1 - \Delta_2 e_2 + \Delta_3 e_3.$$

Donde

$$\begin{aligned}\Delta_1 &= \det([\,[v_1]^E\ [v_2]^E\ \begin{bmatrix}1\\0\\0\end{bmatrix}]) = \det[\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\ \begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}\ \begin{bmatrix}1\\0\\0\end{bmatrix}]) = x_2y_3 - x_3y_2.\\ \Delta_2 &= \det([\,[v_1]^E\ [v_2]^E\ \begin{bmatrix}0\\1\\0\end{bmatrix}]) = \det[\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\ \begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}\ \begin{bmatrix}0\\1\\0\end{bmatrix}]) = x_3y_1 - x_1y_3.\\ \Delta_3 &= \det([\,[v_1]^E\ [v_2]^E\ \begin{bmatrix}0\\0\\1\end{bmatrix}]) = \det[\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\ \begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}\ \begin{bmatrix}0\\0\\1\end{bmatrix}]) = x_1y_2 - x_2y_1.\end{aligned}$$

Por lo tanto,

$$\begin{aligned}v_1\wedge v_2 &= \Delta_1 e_1 - \Delta_2 e_2 + \Delta_3 e_3 = (x_2y_3 - x_3y_2)e_1 + (x_3y_1 - x_1y_3)e_2 + (x_1y_2 - x_2y_1)e_3\\ &= \begin{bmatrix}x_2y_3 - x_3y_2\\ x_3y_1 - x_1y_3\\ x_1y_2 - x_2y_2\end{bmatrix}.\end{aligned}$$

g) : En el item *a*) vimos que el cálculo de $\det(u_1, u_2, u_3)$ no depende de la base ortonormal elegida. Entonces, como por hipótesis $\{u_1, u_2, u_3\}$ es una base ortonormal de $\mathbb{V}$, vamos a tomar $E = \{u_1, u_2, u_3\}$. Por lo tanto, $[u_1]^E = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $[u_2]^E = \begin{bmatrix}0\\1\\0\end{bmatrix}$, $[u_3]^E = \begin{bmatrix}0\\0\\1\end{bmatrix}$.

En el item *f*), probamos que

$$u_1\wedge u_2 = \Delta_1 u_1 - \Delta_2 u_2 + \Delta_3 u_3 = (x_2y_3 - x_3y_2)u_1 + (x_3y_1 - x_1y_3)u_2 + (x_1y_2 - x_2y_1)u_3,$$

donde $[u_1]^E = \begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ y $[u_2]^E = \begin{bmatrix}0\\1\\0\end{bmatrix} = \begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$.

Por lo tanto,

$$u_1\wedge u_2 = (x_2y_3 - x_3y_2)u_1 + (x_3y_1 - x_1y_3)u_2 + (x_1y_2 - x_2y_1)u_3 = 0u_1 + 0u_2 + 1u_3 = u_3.$$

Entonces, por el item a), se sigue que

$$\det(u_1, u_2, u_3) = \langle u_3, u_1 \wedge v_2 \rangle = \langle u_3, u_3 \rangle = 1,$$

pues u_3 es un vector de norma 1.

Notar que si la base la ordenamos de otra manera, por ejemplo $\{u_2, u_1, u_3\}$, tendríamos que

$$\det(u_2, u_1, u_3) = \langle u_3, u_2 \wedge v_1 \rangle = -\langle u_3, u_3 \rangle = -1,$$

en ese caso la base NO estaría orientada positivamente.

g) : Supongamos que $\{v_1, v_2\}$ es un sistema ortonormal. Veamos que $\{v_1, v_2, v_1 \wedge v_2\}$ es una bon de $\mathbb{V}$. En el item b), ya vimos que $v_1 \wedge v_2$ es ortogonal a v_1 y v_2. Veamos que $v_1 \wedge v_2$ tiene norma 1.

Supongamos que $[v_1]^E = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ y $[v_2]^E = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$. Como E es una bon de $\mathbb{V}$, se sigue que

$$\langle v_1, v_2 \rangle = \langle x_1e_1 + x_2e_2 + x_3e_3, y_1e_1 + y_2e_2 + y_3e_3 \rangle = x_1y_1 + x_2y_2 + x_3y_3.$$

De la misma manera, $\|v_1\|^2 = x_1^2 + x_2^2 + x_3^2$ y $\|v_2\|^2 = y_1^2 + y_2^2 + y_3^2$. Por otra parte, por el item d), tenemos que

$$v_1 \wedge v_2 = (x_2y_3 - x_3y_2)e_1 + (x_3y_1 - x_1y_3)e_2 + (x_1y_2 - x_2y_1)e_3.$$

Entonces, usando el item a),

$$\begin{aligned}
\|v_1 \wedge v_2\|^2 &= \langle v_1 \wedge v_2, v_1 \wedge v_2 \rangle = \det([\ [v_1]^E\ [v_2]^E\ [v_1 \wedge v_2]^E]) \\
&= \det(\begin{bmatrix} x_1 & y_1 & x_2y_3 - x_3y_2 \\ x_2 & y_2 & x_3y_1 - x_1y_3 \\ x_3 & y_3 & x_1y_2 - x_2y_1 \end{bmatrix}) \\
&= x_1[y_2(x_1y_2 - x_2y_1) - y_3(x_3y_1 - x_1y_3)] - x_2[y_1(x_1y_2 - x_2y_1) - y_3(x_2y_3 - x_3y_2)] \\
&+ x_3[y_1(x_3y_1 - x_1y_3) - y_2(x_2y_3 - x_3y_2)] = x_1^2y_2^2 + x_1^2y_3^2 - x_1y_1x_2y_2 - x_1y_1x_3y_3 \\
&+ x_2^2y_1^2 + x_2^2y_3^2 - x_1y_1x_2y_2 - x_2y_2x_3y_3 + x_3^2y_1^2 + x_3^2y_2^2 - x_1y_1x_3y_3 - x_2y_2x_3y_3 \\
&= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - x_1^2y_1^2 - x_2^2y_2^2 - x_3^2y_3^2 \\
&- x_1y_1x_2y_2 - x_1y_1x_3y_3 - x_2y_2x_1y_1 - x_2y_2x_3y_3 - x_3y_3x_1y_1 - x_3y_3x_2y_2 \\
&= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2 \\
&= \|v_1\|^2\|v_2\|^2 - \langle v_1, v_2 \rangle^2
\end{aligned}$$

Acabamos de probar que:

si $v_1, v_2 \in \mathbb{V}$ (no es necesario que formen un sistema ortonormal) entonces

$$\|v_1 \wedge v_2\|^2 = \|v_1\|^2\|v_2\|^2 - \langle v_1, v_2 \rangle^2. \tag{3.20}$$

Ahora sí, usando que $\{v_1, v_2\}$ es un sistema ortonormal (por hipótesis), se sigue que

$$\|v_1 \wedge v_2\|^2 = \|v_1\|^2\|v_2\|^2 - \langle v_1, v_2 \rangle^2 = 1 - 0 = 1.$$

Entonces $\{v_1, v_2, v_1 \wedge v_2\}$ es una sistema ortonormal. Finalmente, como $\dim(\mathbb{V}) = 3$ concluimos que $\{v_1, v_2, v_1 \wedge v_2\}$ es una bon de $\mathbb{V}$.

c) $\Leftarrow$: Ahora sí, veamos la vuelta del item c). Supongamos que $v_1 \wedge v_2 = 0$. Entonces, por (3.20), se sigue que $\|v_1\|^2\|v_2\|^2 = \langle v_1, v_2 \rangle^2$. Entonces el Gramiano, nos queda

$$\det(G_{\{v_1,v_2\}}) = \det \begin{bmatrix} \|v_1\|^2 & \langle v_1, v_2 \rangle \\ \langle v_2, v_1 \rangle & \|v_2\|^2 \end{bmatrix} = \|v_1\|^2\|v_2\|^2 - \langle v_1, v_2 \rangle^2 = 0.$$

Por lo tanto, por la Proposición 3.3.1, el conjunto $\{v_1, v_2\}$ es LD. ■

Ejercicio 3.35. Sea $\mathbb{V}$ un $\mathbb{R}$-espacio euclídeo de dimensión 3. Sean $v_1, v_2 \in \mathbb{V}$ dos vectores linealmente independientes y sea $\mathcal{S} = gen\{v_1, v_2\}$.

a) Probar que para todo $v \in \mathbb{V}$ vale que $P_{\mathcal{S}}(v) = v - \frac{\langle v, v_1 \wedge v_2 \rangle}{\|v_1 \wedge v_2\|^2}(v_1 \wedge v_2)$ y $d(v, \mathcal{S}) = \frac{|\langle v, v_1 \wedge v_2 \rangle|}{\|v_1 \wedge v_2\|}$.

b) Comprobar que $\mathcal{S}^{\perp} = gen\{v_1 \wedge v_2\}$.

Dem. En el **Ejercicio 3.34** vimos que dados dos vectores $v_1, v_2 \in \mathbb{V}$ tenemos que $v_1 \wedge v_2$ es ortogonal a v_1 y v_2. Entonces, $v_1 \wedge v_2 \in \mathcal{S}^{\perp}$. Por otra parte, como v_1 y v_2 son linealmente independientes y $\mathcal{S} = gen\{v_1, v_2\}$, se sigue que $\dim(\mathcal{S}) = 2$. Por lo tanto, $\dim(\mathcal{S}^{\perp}) = \dim(\mathbb{V}) - \dim(\mathcal{S}) = 3 - 2 = 1$. Entonces, $\mathcal{S}^{\perp} = gen\{v_1 \wedge v_2\}$.

Claramente $\{v_1 \wedge v_2\}$ es una base ortogonal de $\mathcal{S}^{\perp}$ (tiene sólo 1 elemento no nulo). Usando la fórmula de la proyección ortogonal, vale que para cada $v \in \mathbb{V}$,

$$P_{\mathcal{S}^{\perp}}(v) = \frac{\langle v, v_1 \wedge v_2 \rangle}{\|v_1 \wedge v_2\|^2}(v_1 \wedge v_2).$$

Como $v = P_{\mathcal{S}}(v) + P_{\mathcal{S}^{\perp}}(v)$, tenemos que $P_{\mathcal{S}}(v) = v - \frac{\langle v, v_1 \wedge v_2 \rangle}{\|v_1 \wedge v_2\|^2}(v_1 \wedge v_2)$.

Finalmente,

$$d(v, \mathcal{S}) = \|P_{\mathcal{S}^{\perp}}(v)\| = \|\frac{\langle v, v_1 \wedge v_2 \rangle}{\|v_1 \wedge v_2\|^2}(v_1 \wedge v_2)\| = \frac{|\langle v, v_1 \wedge v_2 \rangle|}{\|v_1 \wedge v_2\|^2}\|v_1 \wedge v_2\| = \frac{|\langle v, v_1 \wedge v_2 \rangle|}{\|v_1 \wedge v_2\|}.$$

■

Ejercicio 3.36. Sea $\mathbb{V}$ un $\mathbb{R}$-espacio euclídeo de dimensión 3.

a) Probar que para todo $v_1, v_2 \in \mathbb{V}$ vale que

$$\langle v_1, v_2 \rangle^2 + \|v_1 \wedge v_2\|^2 = \|v_1\|^2 \|v_2\|^2.$$

b) Probar que si θ es el ángulo entre v_1 y v_2, entonces

$$\|v_1 \wedge v_2\| = \|v_1\| \|v_2\| \sin(\theta).$$

Dem. *a)* : En el item *f)* del **Ejercicio 3.34**, probamos que $\|v_1 \wedge v_2\|^2 = \|v_1\|^2 \|v_2\|^2 - \langle v_1, v_2 \rangle^2$. Por lo tanto, la igualdad que queremos probar se sigue pasando de término.

b) : Si θ es el ángulo entre v_1 y v_2, recordemos que

$$\langle v_1, v_2 \rangle = \|v_1\| \|v_2\| \cos(\theta).$$

Entonces, por el item *a)*, se sigue que

$$\begin{aligned} \|v_1 \wedge v_2\|^2 &= \langle v_1, v_2 \rangle^2 - \|v_1\|^2 \|v_2\|^2 = \|v_1\|^2 \|v_2\|^2 \cos(\theta)^2 - \|v_1\|^2 \|v_2\|^2 \\ &= \|v_1\|^2 \|v_2\|^2 (1 - \cos(\theta)^2) = \|v_1\|^2 \|v_2\|^2 \sin(\theta)^2. \end{aligned}$$

Tomando raíz, $\|v_1 \wedge v_2\| = \|v_1\| \|v_2\| \sin(\theta)$ y probamos lo que queríamos. ■

Otras aplicaciones del Teorema de Riesz

Veamos otra aplicación muy importante del Teorema de representación de Riesz.

Ejercicio 3.P. Sean $(\mathbb{V}, \langle \cdot, \cdot \rangle_{\mathbb{V}})$ y $(\mathbb{W}, \langle \cdot, \cdot \rangle_{\mathbb{W}})$ dos $\mathbb{K}$-espacios euclídeos finito dimensionales y $T : \mathbb{V} \to \mathbb{W}$ una transformación lineal. Demostrar que existe una única transformación lineal $S : \mathbb{W} \to \mathbb{V}$ tal que

$$\langle T(v), w \rangle_{\mathbb{W}} = \langle v, S(w) \rangle_{\mathbb{V}},$$

para todo $v \in \mathbb{V}$ y $w \in \mathbb{W}$.

Ayuda: usar el Teorema de representación de Riesz para cada $w \in \mathbb{W}$ fijo con la funcional lineal $\phi_w(v) := \langle T(v), w \rangle_{\mathbb{W}}$.

Dem. Tomemos $w \in \mathbb{W}$ arbitrario pero fijo y consideremos la funcional lineal de la ayuda, $\phi_w : \mathbb{V} \to \mathbb{K}$ definida por

$$\phi_w(v) := \langle T(v), w \rangle_{\mathbb{W}} .$$

Ver **Ejercicio 3.1**.

Efectivamente ϕ_w es una funcional lineal. De hecho, si $v, v' \in \mathbb{V}$ y $\alpha, \beta \in \mathbb{K}$ entonces, usando que T es una transformación lineal y los axiomas de producto interno tenemos que

$$\begin{aligned}\phi_w(\alpha v + \beta v') &= \langle T(\alpha v + \beta v'), w \rangle_{\mathbb{W}} = \langle \alpha T(v) + \beta T(v'), w \rangle_{\mathbb{W}} = \alpha \langle T(v), w \rangle_{\mathbb{W}} + \beta \langle T(v'), w \rangle_{\mathbb{W}} \\ &= \alpha \phi_w(v) + \beta \phi_w(v').\end{aligned}$$

Como $\phi_w \in \mathcal{L}(\mathbb{V}, \mathbb{K})$, por el Teorema de representación de Riesz que probamos en el **Ejercicio 3.30**, existe un único vector $u \in \mathbb{V}$ tal que

$$\phi_w(v) = \langle v, u \rangle_{\mathbb{V}} .$$

Tenemos todos los ingredientes para definir la transformación lineal $S : \mathbb{W} \to \mathbb{V}$ que pide el ejercicio. De hecho, vamos a definir S como

$$S(w) := u.$$

Notar que S está bien definida, porque para cada $w \in \mathbb{W}$ que fijamos, por el Teorema de representación de Riesz obtuvimos un único $u \in \mathbb{V}$ tal que $\phi_w(v) = \langle v, u \rangle_{\mathbb{V}}$. Además,

$$\phi_w(v) = \langle T(v), w \rangle_{\mathbb{W}} = \langle v, u \rangle_{\mathbb{V}} = \langle v, S(w) \rangle_{\mathbb{V}} . \tag{3.21}$$

Por otra parte, S es una transformación lineal, pues, si $w, w' \in \mathbb{W}$, por el Teorema de representación de Riesz, existen únicos $u, u' \in \mathbb{V}$ tales que

$$\phi_w(v) = \langle T(v), w \rangle_{\mathbb{W}} = \langle v, u \rangle_{\mathbb{V}} \quad \text{y} \quad \phi_{w'}(v) = \langle T(v), w' \rangle_{\mathbb{W}} = \langle v, u' \rangle_{\mathbb{V}} .$$

Sean $\alpha, \beta \in \mathbb{K}$ entonces, por los axiomas del producto interno tenemos que

$$\begin{aligned}\phi_{\alpha w + \beta w'}(v) &= \langle T(v), \alpha w + \beta w' \rangle_{\mathbb{W}} = \overline{\alpha} \langle T(v), w \rangle_{\mathbb{W}} + \overline{\beta} \langle T(v), w' \rangle_{\mathbb{W}} \\ &= \overline{\alpha} \phi_w(v) + \overline{\beta} \phi_{w'}(v) = \overline{\alpha} \langle v, u \rangle_{\mathbb{V}} + \overline{\beta} \langle v, u' \rangle_{\mathbb{V}} = \langle v, \alpha u + \beta u' \rangle_{\mathbb{V}} .\end{aligned}$$

Entonces, por la unicidad del Teorema de representación de Riesz, se sigue que $\alpha u + \beta u'$ es el único vector de $\mathbb{V}$ tal que

$$\phi_{\alpha w + \beta w'}(v) = \langle v, \alpha u + \beta u' \rangle_{\mathbb{V}} .$$

Entonces, por definición de S, tenemos que

$$S(\alpha w + \beta w') = \alpha u + \beta u' = \alpha S(w) + \beta S(w')$$

y concluimos que S es una transformación lineal.

Finalmente, volviendo a (3.21) tenemos que, para cada $v \in \mathbb{V}$ y $w \in \mathbb{W}$

$$\langle T(v), w \rangle_{\mathbb{W}} = \phi_w(v) = \langle v, u \rangle_{\mathbb{V}} = \langle v, S(w) \rangle_{\mathbb{V}} .$$

Por lo tanto la transformación lineal S cumple lo que queríamos.

Veamos que S es la única transformación que lo cumple. Supongamos que existe $S' \in \mathcal{L}(\mathbb{W}, \mathbb{V})$ tal que para todo $v \in \mathbb{V}$ y $w \in \mathbb{W}$

$$\langle T(v), w \rangle_{\mathbb{W}} = \langle v, S(w) \rangle_{\mathbb{V}} = \langle v, S'(w) \rangle_{\mathbb{V}} .$$

Entonces, $\langle v, S(w) \rangle_{\mathbb{V}} = \langle v, S'(w) \rangle_{\mathbb{V}}$, para todo $v \in \mathbb{V}$ y $w \in \mathbb{W}$. Entonces, usando la linealidad del producto interno, se sigue que para todo $v \in \mathbb{V}$ y $w \in \mathbb{W}$

$$\langle v, S(w) - S'(w) \rangle_{\mathbb{V}} = \langle v, S(w) \rangle_{\mathbb{V}} - \langle v, S'(w) \rangle_{\mathbb{V}} = 0.$$

Es decir $\langle v, S(w) - S'(w) \rangle_{\mathbb{V}} = 0$, para todo $v \in \mathbb{V}$ y $w \in \mathbb{W}$. En particular podemos tomar $v := S(w) - S'(w)$, entonces,

$$0 = \langle v, S(w) - S'(w) \rangle_{\mathbb{V}} = \langle S(w) - S'(w), S(w) - S'(w) \rangle_{\mathbb{V}} = \|S(w) - S'(w)\|_{\mathbb{V}}^2,$$

para todo $w \in \mathbb{W}$. Entonces,

$$S(w) - S(w') = 0_{\mathbb{V}},$$

para todo $w \in \mathbb{W}$. Por lo tanto

$$S(w) = S(w'),$$

para todo $w \in \mathbb{W}$. Entonces $S = S'$ y probamos lo que queríamos. ∎

A la única transformación lineal $S : \mathbb{W} \to \mathbb{V}$ que cumple

$$\langle T(v), w \rangle_{\mathbb{W}} = \langle v, S(w) \rangle_{\mathbb{V}},$$

para todo $v \in \mathbb{V}$ y $w \in \mathbb{W}$, se la llama *la transformacioń lineal adjunta de* T y se la suele notar $T^* := S$. No confundir con la notación que usamos para matrices, recordar que si $A \in \mathbb{C}^{m\times n}$ es una matriz, definimos $A^* = \overline{A^T}$. De todas maneras, el próximo ejemplo muestra que la notación para la transformación lineal adjunta, no es casual.

Ejemplo 3.a. Tomemos $\mathbb{V} = \mathbb{C}^n$ y $\mathbb{W} = \mathbb{C}^m$ con el producto interno canónico y definamos $T : \mathbb{C}^n \to \mathbb{C}^m$ como

$$T(x) = Ax,$$

donde $A \in \mathbb{C}^{m\times n}$ es una matriz. Recordemos que $A^* = \overline{A^T}$ y el producto interno canónico se define como $\langle u, v \rangle = v^* u$. Entonces, observar que para todo $x \in \mathbb{C}^n$ e $y \in \mathbb{C}^m$, tenemos que

$$\langle T(x), y \rangle_{\mathbb{C}^m} = y^*(T(x)) = y^*(A(x)) = (y^*A)x = (A^*y)^*x = \langle x, A^*y \rangle_{\mathbb{C}^n}.$$

Por lo tanto, por definición de transformación lineal adjunta, tenemos que $T^* : \mathbb{C}^m \to \mathbb{C}^n$ es la transformación lineal definida como

$$T^*(y) = A^*y.$$

Ejemplo 3.b. Supongamos que $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ es un $\mathbb{K}$-espacio euclídeo y $\mathcal{S} \subseteq \mathbb{V}$ es un subespacio de dimensión finita. Sea

$$T := P_{\mathcal{S}}$$

(el proyector ortogonal de $\mathbb{V}$ sobre $\mathcal{S}$). Por un lado, T es un proyector, por lo tanto

$$T^2 = T.$$

Además, en el Teorema 3.6.1, probamos que como T es un proyector ortogonal, para todo $x, y \in \mathbb{V}$ se cumple que

$$\langle T(x), y \rangle = \langle x, T(y) \rangle.$$

Entonces, por definición de transformación lineal adjunta, se sigue que en este caso $T^* = T$. Es más, a partir del Teorema 3.6.1, podemos decir que T es un proyector ortogonal si y sólo si $T^2 = T$ y $T^* = T$.

Juntando el **Ejemplo 3.a** y el **Ejemplo 3.b**, podemos ver que: si consideramos $\mathbb{C}^n$, con **producto interno canónico** y E es la **base canónica** de $\mathbb{C}^n$, entonces, para cualquier subespacio $\mathcal{S} \subseteq \mathbb{V}$ se sigue que

$$[P_{\mathcal{S}}]_E^E = ([P_{\mathcal{S}}]_E^E)^2 = ([P_{\mathcal{S}}]_E^E)^*.$$

De hecho, como $P_{\mathcal{S}}$ es una proyector, ya habíamos visto que $[P_{\mathcal{S}}]_E^E = ([P_{\mathcal{S}}]_E^E)^2$. Por otra parte, notar que para cada $x \in \mathbb{C}^n$, tenemos que $P_{\mathcal{S}}(x) = [P_{\mathcal{S}}]_E^E(x)$. Entonces, como $P_{\mathcal{S}}$ es un proyector ortogonal, por el **Ejemplo 3.b**, $P_{\mathcal{S}}^* = P_{\mathcal{S}}$, y por el **Ejemplo 3.a**, $P_{\mathcal{S}}^*(x) = ([P_{\mathcal{S}}]_E^E)^*(x)$. Entonces, para todo $x \in \mathbb{C}^m$, nos queda que

$$([P_{\mathcal{S}}]_E^E)^*(x) = ([P_{\mathcal{S}}]_E^E)(x).$$

Eso implica que

$$[P_{\mathcal{S}}]_E^E = ([P_{\mathcal{S}}]_E^E)^*.$$

De manera similar, si consideramos $\mathbb{R}^n$, con **producto interno canónico** y E es la **base canónica** de $\mathbb{R}^n$, entonces:

$$[P_{\mathcal{S}}]_E^E = ([P_{\mathcal{S}}]_E^E)^2 = ([P_{\mathcal{S}}]_E^E)^T.$$

CAPÍTULO 4

Diagonalización

En este capítulo, comenzaremos a estudiar la estructura de los endomorfismos de un espacio vectorial de dimensión finita.

4.1. Autovalores y autovectores

Definición. Sea $A \in \mathbb{K}^{n \times n}$, el escalar $\lambda \in \mathbb{K}$ (notar que λ es un elemento del cuerpo considerado) es un *autovalor* de A si existe un vector no nulo $v \in \mathbb{K}^n \setminus \{0\}$ tal que

$$Av = \lambda v.$$

Al vector v lo llamamos *autovector asociado a* λ.

El conjunto de todos los autovalores de A se denomina *espectro de* A y se denota por $\sigma(A)$.

Definimos el *autoespacio asociado a* λ como

$$\mathcal{S}_\lambda := \{v \in \mathbb{K}^n : Av = \lambda v\} = \mathrm{nul}(A - \lambda I).$$

La siguiente propiedad la vamos a usar para obtener los autovalores de una matriz $A \in \mathbb{K}^{n \times n}$:

Proposición 4.1.1. *Sea $A \in \mathbb{K}^{n \times n}$. Entonces, son equivalentes:*

i) $\lambda \in \mathbb{K}$ *es autovalor de* A,

ii) $\mathrm{nul}(A - \lambda I) \neq \{0\}$,

iii) $\mathrm{rg}(A - \lambda I) < n$,

iv) $\det(A - \lambda I) = 0$.

Dem. *i)* $\Rightarrow$ *ii)* : Si $\lambda \in \mathbb{K}$ es un autovalor de A entonces existe un vector no nulo $v \in \mathbb{K}^n \setminus \{0\}$ tal que $Av = \lambda v$. Entonces,

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

Por lo tanto $v \in \mathrm{nul}(A - \lambda I)$ y como $v \neq 0$, tenemos que $\mathrm{nul}(A - \lambda I) \neq \{0\}$.

ii) $\Rightarrow$ *iii)* : Si $\mathrm{nul}(A - \lambda I) \neq \{0\}$, entonces $\dim(\mathrm{nul}(A - \lambda I)) > 0$. Entonces, por el Teorema de la Dimensión, tenemos que

$$\mathrm{rg}(A - \lambda I) = \dim(\mathrm{col}(A - \lambda I)) = n - \dim(\mathrm{nul}(A - \lambda I)) < n.$$

iii) $\Rightarrow$ *iv)* : Como $A - \lambda I \in \mathbb{K}^{n \times n}$, si $\mathrm{rg}(A - \lambda I) < n$, entonces $A - \lambda I$ no es inversible y por lo tanto $\det(A - \lambda I) = 0$.

iv) $\Rightarrow$ *i)* : Si $\det(A - \lambda I) = 0$, entonces la matriz $A - \lambda I \in \mathbb{K}^{n \times n}$ no es inversible. Por lo tanto $\mathrm{nul}(A - \lambda I) \neq \{0\}$ y existe $v \neq 0$ tal que $(A - \lambda I)v = 0$. Entonces $Av = \lambda v$ y $\lambda \in \mathbb{K}$ es autovalor de A. ∎

Sea $A \in \mathbb{K}^{n\times n}$. Llamamos *polinomio característico* de A a

$$\chi_A(x) := \det(A - xI) \in \mathbb{K}_n[x].$$

A partir de la Proposición 4.1.1 podemos ver que $\lambda \in \mathbb{K}$ es un autovalor de A si y sólo si $\lambda \in \mathbb{K}$ es una raíz de χ_A, es decir

$$\sigma(A) = \{\lambda \in \mathbb{K} : \chi_A(\lambda) = 0\}.$$

Sea $\lambda \in \mathbb{K}$ un autovalor de A entonces:

- La *multiplicidad algebraica* de λ es la multiplicidad de λ como raíz del polinomio característico y se denota $m(\lambda)$.
- La *multiplicidad geométrica* de λ es la dimensión del autoespacio asociado de λ y se denota $\mu(\lambda)$. Es decir, $\mu(\lambda) := \dim(\mathrm{nul}(A - \lambda I))$.

Recordar que siempre tenemos que

$$1 \leq m(\lambda) \leq \mu(\lambda) \leq n, \text{ para todo } \lambda \in \mathbb{K} \text{ autovalor de } A. \tag{4.1}$$

De hecho, sea λ un autovalor de A tal que $r \in \mathbb{N}$ es la multiplicidad algebraica de λ (multiplicidad de λ como raíz de χ_A). Supongamos que $\dim(S_\lambda) = \dim(\mathrm{nul}(A - \lambda I)) = s \geq 1$. Consideremos $\{v_1, v_2, \cdots, v_s\}$ una base de S_λ y sea $B = \{v_1, v_2, \cdots, v_s, \cdots, v_n\}$ una base de $\mathbb{K}^n$, es decir completamos la base de S_λ a una base de $\mathbb{K}^n$.

Consideremos la transformación lineal $T : \mathbb{K}^n \to \mathbb{K}^n$ definida por $T(v) = Av$. Entonces, si E es la base canónica de $\mathbb{K}^n$, tenemos que $[T]_E^E = A$ y, por otra parte, como $T(v_i) = \lambda v_i$ para $i = 1, \cdots, s$, se sigue que

$$[T]_B^B = [[T(v_1)]^B \ [T(v_2)]^B \ \cdots \ [T(v_s)]^B \ \cdots \ [T(v_n)]^B] = \begin{bmatrix} \lambda I_{s\times s} & N \\ 0_{n-s\times s} & M \end{bmatrix},$$

donde M y N son ciertas matrices. Entonces, como $A = [T]_E^E = P[T]_B^B P^{-1}$, donde $P = [M]_B^E$ es la matriz de cambio de base de B a E, tenemos que

$$\chi_A(x) = \det(A - xI) = \det([T]_B^B - xI) = \det(\begin{bmatrix} \lambda I_{s\times s} - xI_{s\times s} & N \\ 0 & M - xI \end{bmatrix}) = (\lambda - x)^s \det(M - xI).$$

Por otra parte, como λ es un autovalor de A con multiplicidad algebraica $r \in \mathbb{N}$, tenemos que $\chi_A(x) = (x - \lambda)^r R(x)$ con R un polinomio tal que $R(\lambda) \neq 0$. Entonces

$$(\lambda - x)^s \det(M - xI) = \chi_A(x) = (x - \lambda)^r R(x),$$

con $R(\lambda) \neq 0$. Entonces $s \leq r$, porque si $s > r$ entonces $R(x) = (\lambda - x)^{s-r} \det(M - xI)$ y $R(\lambda) = 0$, lo cual es absurdo. Por lo tanto

$$\mu(\lambda) = \dim(\mathrm{nul}(A - \lambda I)) = s \leq r = m(\lambda).$$

Propiedades de autovalores y autovectores

Las siguientes son algunas propiedades de autovalores y autovectores a tener en cuenta. Sea $A \in \mathbb{K}^{n\times n}$, entonces:

- $\lambda = 0$ es autovalor de A si y sólo si A es no inversible (o singular).

  De hecho, $\lambda = 0$ es autovalor de A si y sólo si $\lambda = 0$ es una raíz de $\chi_A(x) = \det(A - xI)$. Es decir,
$$\det(A) = \det(A - 0I) = \chi_A(0) = 0$$
  si y sólo si A es no inversible (o singular).

- Si $\operatorname{rg}(A) = k < n$ entonces $\lambda = 0$ es autovalor de A con multiplicidad geométrica $n - k$.

  De hecho, por la Proposición 4.1.1, como $\operatorname{rg}(A) = \operatorname{rg}(A - 0I) = k < n$ entonces $\lambda = 0$ es autovalor de A. Entonces, por el Teorema de la Dimensión,
$$\mu(\lambda = 0) = \dim(\operatorname{nul}(A - 0I)) = \dim(\operatorname{nul}(A)) = n - \operatorname{rg}(A) = n - k.$$

- Si $\lambda \in \mathbb{K}$ es autovalor de A entonces:

  - $r\lambda$ es autovalor de rA para todo $r \in \mathbb{K}$.
  - λ^k es autovalor de A^k para todo $k \in \mathbb{N}$.
  - si existe A^{-1} entonces λ^{-1} es autovalor de A^{-1}.

  De hecho, si $\lambda \in \mathbb{K}$ es autovalor de A entonces existe $v \neq 0$ tal que
$$Av = \lambda v. \tag{4.2}$$
  Entonces, multiplicando la ecuación (4.2) a ambos lados por $r \in \mathbb{K}$. Tenemos que
$$(rA)v = (r\lambda)v.$$
  Entonces $r\lambda \in \mathbb{K}$ es autovalor de rA con el mismo autovector asociado.

  Por otra parte, multiplicando la ecuación (4.2) a ambos lados por A. Tenemos que
$$A^2v = AAv = A\lambda v = \lambda Av = \lambda\lambda v = \lambda^2 v.$$
  Entonces λ^2 es autovalor de A^2 con el mismo autovector asociado.

  Vamos a probar la propiedad para todo $k \in \mathbb{N}$ por inducción: supongamos que para $k = n \in \mathbb{N}$ vale que $A^n v = \lambda^n v$ (HI). Entonces, multiplicamos esa ecuación a ambos lados por A y tenemos que
$$A^{n+1}v = AA^n v = A\lambda^n v = \lambda^n Av = \lambda^n \lambda v = \lambda^{n+1}v.$$
  Entonces, a partir de la hipótesis inductiva deducimos la validez de la propiedad para $k = n+1$. Por lo tanto, por inducción, λ^k es autovalor de A^k para todo $k \in \mathbb{N}$ con el mismo autovector asociado.

  Finalmente, si existe A^{-1} entonces A es inversible y por lo que vimos arriba $\lambda \neq 0$. Entonces, multiplicando la ecuación (4.2) a ambos lados por A^{-1},
$$v = A^{-1}Av = A^{-1}\lambda v = \lambda A^{-1}v.$$
  Multiplicando la ecuación anterior por λ^{-1} tenemos que
$$\lambda^{-1}v = \lambda^{-1}\lambda A^{-1}v = A^{-1}v.$$
  Por lo tanto λ^{-1} es autovalor de A^{-1} con el mismo autovector asociado.

La siguiente es una propiedad útil para matrices $A \in \mathbb{K}^{2\times 2}$. Sea $A \in \mathbb{K}^{2\times 2}$ entonces

$$\chi_A(x) = \det(A - xI) = x^2 - \operatorname{tr}(A)x + \det(A). \tag{4.3}$$

En consecuencia las raíces de χ_A son

$$\lambda_\pm = \frac{1}{2}\big(\operatorname{tr}(A) \pm \sqrt{\operatorname{tr}(A)^2 - 4\det(A)}\big). \tag{4.4}$$

De hecho, supongamos que $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, con $a, b, c, d \in \mathbb{K}$. Entonces

$$\begin{aligned} \chi_A(x) = \det(A - xI) = \det \begin{bmatrix} a-x & b \\ c & d-x \end{bmatrix} = (a-x)(d-x) - bc \\ = x^2 - (a+d)x + (ad - bc) = x^2 - \operatorname{tr}(A)x + \det(A). \end{aligned}$$

Por lo tanto, las raíces de χ_A son las que aparecen en (4.4).

Sea $A \in \mathbb{K}^{n\times n}$ y sean $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$ los autovalores de A (no necesariamente distintos). Entonces

$$\det(A) = \lambda_1\ \lambda_2\ \cdots\ \lambda_n, \tag{4.5}$$

es decir el producto de los autovalores de A es igual al determinante de A.

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n, \tag{4.6}$$

es decir la suma de los autovalores de A es igual a la traza A.

De hecho, supongamos que $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$ son los autovalores de A (no necesariamente distintos). Entonces $\lambda_1, \cdots, \lambda_n$ son las n raíces de χ_A y podemos factorizar χ_A de la siguiente manera:

$$\det(A - xI) = \chi_A(x) = (-1)^n \det(xI - A) = (-1)^n (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n).$$

Evaluando el polinomio anterior en $x = 0$ nos queda que

$$\begin{aligned} \det(A) = \det(A - 0I) = \chi_A(0) = (-1)^n(0 - \lambda_1)(0 - \lambda_2)\cdots(0 - \lambda_n) = (-1)^n(-1)^n\lambda_1\ \lambda_2\ \cdots\ \lambda_n \\ = \lambda_1\ \lambda_2\ \cdots\ \lambda_n. \end{aligned}$$

La demostración de que la suma de los autovalores de A es igual a $\operatorname{tr}(A)$ la veremos al final de la Sección de Formas de Jordan 4.5.

A continuación veremos un ejemplo de cálculo de autovectores y autovalores. Antes recordemos el siguiente resultado que puede resultar útil para encontrar raíces racionales de polinomios con coeficientes enteros:

Teorema de Gauss: Sea

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}_n[x]$$

es decir un polinomio con coeficientes enteros con a_0 y a_n no nulos. Entonces, si $\lambda := \frac{p}{q}$ es una raíz racional de dicho polinomio (donde $\frac{p}{q}$ es una fracción irreducible o que no se puede simplificar) es porque p es un divisor de a_0 y q es un divisor de a_n.

Por ejemplo, consideremos el polinomio

$$r(x) = 4x^4 - x^2 - 6x + 3 \in \mathbb{Z}_4[x].$$

Por el Teorema de Gauss, si r tiene alguna raíz racional de la forma $\lambda = \frac{p}{q}$, entonces los posibles p serían: $1, -1, 3, -3$ (es decir todos los divisores de $a_0 = 3$) y los posibles q serían, $1, -1, 2, -2, 4, -4$ (es decir todos los divisores de $a_4 = 4$). Por lo tanto, si r tiene alguna raíz racional λ, no queda otra que:

$$\lambda = \frac{p}{q} \in \{1, -1, \frac{1}{2}, -\frac{1}{2}, \frac{1}{4}, -\frac{1}{4}, 3, -3, \frac{3}{2}, -\frac{3}{2}, \frac{3}{4}, -\frac{3}{4}\}.$$

Haciendo la cuenta (es decir evaluando r en los valores de arriba y viendo si se anula), tenemos que 1 y $\frac{1}{2}$ son raíces de r (no son las únicas, pero sí son todas las raíces racionales de r).

Por ejemplo, si ahora consideramos el polinomio

$$r(x) = x^3 + 2 \in \mathbb{Z}_3[x].$$

Por el Teorema de Gauss, si r tiene alguna raíz racional de la forma $\lambda = \frac{p}{q}$, entonces los posibles p serían: $1, -1, 2, -2$ (es decir todos los divisores de $a_0 = 2$) y los posibles q serían, $1, -1$ (es decir todos los divisores de $a_3 = 1$). Es decir, si r tiene alguna raíz racional λ, no queda otra que:

$$\lambda = \frac{p}{q} \in \{1, -1, 2, -2\}.$$

Si evaluamos el polinomio r en los valores de arriba vemos que nunca se anula, por lo tanto, podemos concluir que r no tiene ninguna raíz racional.

Supongamos ahora que tenemos una matriz $A \in \mathbb{Z}^{n \times n}$ (con coeficientes enteros), entonces el polinomio característico de A es mónico (es decir $a_n = 1$, meditar por qué) y tiene la forma

$$\chi_A(x) = \det(A - xI) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}_n[x].$$

Por el Teorema de Gauss, si $\lambda = \frac{p}{q}$ es una raíz racional de χ_A entonces los valores posibles de q son 1 y -1 (pues $a_n = 1$) y los posibles valores de p son los divisores de a_0. Observar además que

$$a_0 = \chi_A(0) = \det(A - 0I) = \det(A).$$

Por lo tanto, si $\chi_A \in \mathbb{Z}_n[x]$ tiene alguna raíz racional, entonces sólo puede ser algún divisor entero del $\det(A)$.

Ejercicio 4.A. Encontrar los autovalores de la siguiente matriz, sus multiplicidades algebraicas y geométricas y tantos autovectores linealmente independientes como sea posible.

$$A = \begin{bmatrix} -3 & 1 & -3 \\ 20 & 3 & 10 \\ 2 & -2 & 4 \end{bmatrix}.$$

Dem. Primero, notar que (hacer la cuenta) $\det(A) = -18$.

Ahora, calculemos el polinomio característico de A:

$$\chi_A(x) = \det(A - xI) = \det \begin{bmatrix} -3-x & 1 & -3 \\ 20 & 3-x & 10 \\ 2 & -2 & 4-x \end{bmatrix} = -x^3 + 4x^2 + 3x - 18.$$

Por el Teorema de Gauss, si χ_A tiena alguna raíz racional λ, entonces sólo puede ser algún divisor entero del $\det(A) = -18$. En ese caso, las posibles raíces racionales podrían ser

$$\{1, -1, 2, -2, 3, -3, 6, -6, 9, -9, 18, -18\}.$$

Haciendo la cuenta (es decir evaluando χ_A en los valores de arriba) se puede ver que tuvimos suerte y que $\lambda_1 = -2$ y $\lambda_2 = 3$ son raíces de χ_A. Además, como $-18 = \det(A) = \lambda_1 \cdot \lambda_2 \cdot \lambda_3 = -6 \cdot \lambda_3$, tenemos que $\lambda_3 = 3$, donde usamos la propiedad (4.1).

Por lo tanto, los autovalores de A son $\lambda_1 = -2$ con multiplicidad algebraica $m(\lambda = -2) = 1$ y $\lambda_{2,3} = 3$ con multiplicidad algebraica $m(\lambda = 3) = 2$.

Para obtener los autoespacios asociados, calculamos:

$$\mathcal{S}_{\lambda=2} = \text{nul}(A - (-2)I) = \text{nul}(\begin{bmatrix} -1 & 1 & -3 \\ 20 & 5 & 10 \\ 2 & -2 & 6 \end{bmatrix}) = \text{nul}(\begin{bmatrix} -1 & 1 & -3 \\ 0 & 25 & -50 \\ 0 & 0 & 0 \end{bmatrix}) = gen\{\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}\}.$$

$$\mathcal{S}_{\lambda=3} = \text{nul}(A - 3I) = \text{nul}(\begin{bmatrix} -6 & 1 & -3 \\ 20 & 0 & 10 \\ 2 & -2 & 1 \end{bmatrix}) = \text{nul}(\begin{bmatrix} -6 & 1 & -3 \\ 0 & 10 & 0 \\ 0 & -5 & 0 \end{bmatrix}) = gen\{\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}\}.$$

Por lo tanto, la multiplicidad geométrica de $\lambda_1 = -2$ es $\mu(\lambda = -2) = 1$ (lo cual es obvio porque $\lambda = -2$ es un autovalor simple) y la multiplicidad geométrica de $\lambda_{2,3} = 3$ es $\mu(\lambda = 3) = 1$.

Finalmente, si $v_1 \in gen\{\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}\} \setminus \{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\}$ entonces v_1 es un autovector asociado a $\lambda_1 = -2$ y si $v_2 \in gen\{\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}\} \setminus \{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\}$ entonces v_2 es un autovector asociado a $\lambda_{2,3} = 3$. Por lo tanto, siempre podemos encontrar a lo sumo 2 autovectores de A que sean linealmente independientes. ■

4.2. Diagonalización

Sea $A \in \mathbb{K}^{n\times n}$ diremos que A es *diagonalizable* si existen $P \in \mathbb{K}^{n\times n}$ inversible y $\Lambda \in \mathbb{K}^{n\times n}$ diagonal tales que

$$A = P\Lambda P^{-1}.$$

Antes de enunciar el teorema central de la sección, veamos una propiedad que será muy importante:

Proposición 4.2.1. *Sea $A \in \mathbb{K}^{n\times n}$ y sean $\lambda_1, \cdots, \lambda_r \in \mathbb{K}$ autovalores distintos de A. Entonces los autoespacios asociados $\mathcal{S}_{\lambda_1}, \cdots, \mathcal{S}_{\lambda_r}$ están en suma directa.*

Dem. Lo vamos a probar por inducción en r.

Para $r = 1$, la propiedad es obvia.

Para $r = 2$, sean $\lambda_1 \neq \lambda_2$ dos autovalores de A. Si $v \in \mathcal{S}_{\lambda_1} \cap \mathcal{S}_{\lambda_2}$ entonces, $Av = \lambda_1 v$ y $Av = \lambda_2 v$. Por lo tanto $(\lambda_1 - \lambda_2)v = Av - Av = 0$ y como $\lambda_1 - \lambda_2 \neq 0$, se sigue que $v = 0$ y $\mathcal{S}_{\lambda_1}, \mathcal{S}_{\lambda_2}$ están en suma directa.

Ahora, supongamos que el resultado vale para r autovalores distintos (HI) y vamos a probarlo para $r + 1$ autovalores distintos: sean $\lambda_1, \cdots, \lambda_r, \lambda_{r+1}$ autovalores distintos de A. Veamos que los autoespacios $\mathcal{S}_{\lambda_1}, \cdots, \mathcal{S}_{\lambda_r}, \mathcal{S}_{r+1}$ están en suma directa. Primero, veamos que $\mathcal{S}_{\lambda_{r+1}} \cap (\mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_r) = \{0\}$. De hecho, si $v \in \mathcal{S}_{\lambda_{r+1}} \cap (\mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_r)$, entonces $v \in \mathcal{S}_{\lambda_{r+1}}$ y, como $v \in (\mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_r)$, existen únicos $v_i \in \mathcal{S}_{\lambda_i}$ con $i = 1, 2, \cdots, r$, tales que

$$v = v_1 + v_2 + \cdots + v_r \in \mathcal{S}_{\lambda_{r+1}}.$$

Multiplicando la igualdad anterior por A, como $v \in \mathcal{S}_{\lambda_{r+1}}$, tenemos que

$$\lambda_{r+1} v = Av = Av_1 + \cdots + Av_r = \lambda_1 v_1 + \cdots + \lambda_r v_r. \tag{4.7}$$

Por otra parte,

$$\lambda_{r+1} v = \lambda_{r+1} v_1 + \cdots \lambda_{r+1} v_r. \tag{4.8}$$

Restando las ecuaciones (4.7) y (4.8), tenemos que

$$0 = (\lambda_{r+1} - \lambda_1)v_1 + \cdots + (\lambda_{r+1} - \lambda_r)v_r.$$

Por (HI), los autoespacios $\mathcal{S}_1, \cdots, \mathcal{S}_r$ están en suma directa, entonces el vector 0 se escribe únicamente como suma de ceros. Entonces, $(\lambda_{r+1} - \lambda_1)v_1 = \cdots = (\lambda_{r+1} - \lambda_r)v_r = 0$. Como $\lambda_{r+1} - \lambda_i \neq 0$, para cada $i = 1, 2, \cdots, r$, se sigue que $v_1 = v_2 = \cdots = v_r = 0$. Entonces $v = v_1 + v_2 + \cdots + v_r = 0$ y $\mathcal{S}_{\lambda_{r+1}} \cap (\mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_r) = \{0\}$. Veamos que esto implica que $\mathcal{S}_{\lambda_1}, \cdots, \mathcal{S}_{\lambda_r}, \mathcal{S}_{r+1}$ están en suma directa. Sea $w \in \mathcal{S}_{\lambda_1} + \cdots + \mathcal{S}_{\lambda_r} + \mathcal{S}_{r+1}$ y supongamos que

$$w = w_1 + w_2 + \cdots + w_{r+1} = w_1' + w_2' + \cdots + w_{r+1}',$$

para ciertos $w_i, w_i' \in \mathcal{S}_{\lambda_i}$ con $i = 1, 2, \cdots, r+1$. Entonces

$$w_{r+1} - w_{r+1}' = (w_1' - w_1) + \cdots + (w_r' - w_r) \in \mathcal{S}_{\lambda_{r+1}} \cap (\mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_r) = \{0\},$$

donde usamos lo que acabamos de ver. Por lo tanto, $w_{r+1} - w_{r+1}' = 0$. Por otra parte, $0 = (w_1' - w_1) + \cdots + (w_r' - w_r)$ con $w_i, w_i' \in \mathcal{S}_{\lambda_i}$ con $i = 1, 2, \cdots, r$. Nuevamente, por (HI), los autoespacios $\mathcal{S}_1, \cdots, \mathcal{S}_r$ están en suma directa, entonces el vector 0 se escribe únicamente como suma de ceros. Entonces

$$0 = w_1' - w_1 = \cdots = w_r' - w_r = w_{r+1} - w_{r+1}'$$

y el vector w se escribe de manera única. Por lo tanto los autoespacios $\mathcal{S}_{\lambda_1}, \cdots, \mathcal{S}_{\lambda_r}, \mathcal{S}_{r+1}$ están en suma directa. ■

El siguiente teorema nos brinda herramientas que nos permiten determinar si una matriz es diagonalizable:

Teorema 4.2.2. *Sea $A \in \mathbb{K}^{n\times n}$. Entonces son equivalentes:*

i) *A es diagonalizable,*

ii) *existe una base de $\mathbb{K}^n$ compuesta por autovectores de A,*

iii) *$\mathbb{K}^n = \mathcal{S}_{\lambda_1} \oplus \cdots \oplus \mathcal{S}_{\lambda_r}$, donde $\sigma(A) = \{\lambda_1, \cdots, \lambda_r\}$ son los r autovalores distintos de A,*

iv) *las multiplicidades algebraicas y geométricas de cada autovalor distinto de A coinciden.*

En este caso,

$$\bigoplus_{\lambda\in\sigma(A),\ \lambda\neq\lambda_i} \mathrm{nul}(A - \lambda I) = \mathrm{col}(A - \lambda_i I).$$

Dem. *i)* $\Leftrightarrow$ *ii)* : Supongamos que A es diagonalizable, entonces existen $P \in \mathbb{K}^{n\times n}$ inversible y $\Lambda \in \mathbb{K}^{n\times n}$ diagonal tales que $A = P\Lambda P^{-1}$. Sea $P = [v_1\ v_2\ \cdots\ v_n]$ donde $\{v_1, v_2, \cdots, v_n\}$ (el conjunto de las columnas de P) es una base de $\mathbb{K}^n$, pues P es inversible. Supongamos que $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ donde $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{K}$ (no necesariamente todos distintos). Entonces, como $A = P\Lambda P^{-1}$, multiplicando a izquierda por P, se sigue que

$$AP = P\Lambda.$$

Por lo tanto, como

$$AP = A[v_1\ v_2\ \cdots v_n] = [Av_1\ Av_2\ \cdots\ Av_n] \text{ y}$$

$$P\Lambda = [v_1\ v_2\ \cdots\ v_n]\ \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) = [\lambda_1 v_1\ \lambda_2 v_2\ \cdots\ \lambda_n v_n].$$

Tenemos que

$$Av_i = \lambda_i v_i,$$

para $i = 1, 2, \cdots, n$. Como $v_i \neq 0$, para todo $i = 1, 2, \cdots, n$, v_i es un autovector de A para todo $i = 1, 2, \cdots, n$. Por lo tanto, $\{v_1, v_2, \cdots, v_n\}$ es una base de $\mathbb{K}^n$ formada por autovectores de A.

Recíprocamente, si $\{v_1, v_2, \cdots, v_n\}$ es una base de $\mathbb{K}^n$ formada por autovectores de A. Entonces, existen $\lambda_i \in \mathbb{K}$ (no necesariamente distintos entre sí) tales que

$$Av_i = \lambda_i v_i$$

para todo $i = 1, 2, \cdots, n$. Sea $P := [v_1\ v_2\ \cdots\ v_n]$ y $\Lambda := \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ entonces P es inversible pues sus columnas forman un base de $\mathbb{K}^n$ y además

$$AP = [Av_1\ Av_2\ \cdots\ Av_n] = [\lambda_1 v_1\ \lambda_2 v_2\ \cdots\ \lambda_n v_n] = P\Lambda.$$

Entonces

$$A = P\Lambda P^{-1}$$

y por lo tanto A resulta diagonalizable.

$ii) \Rightarrow iii)$: Supongamos que $\lambda_1, \cdots, \lambda_r \in \mathbb{K}$ son los r autovalores distintos de A y que existe una base $B = \{v_1, \cdots, v_n\}$ de $\mathbb{K}^n$ formada por autovectores de A. Entonces, para cada $v_j \in B$ existe $i \in \{1, \cdots, r\}$ tal que v_j es un autovector de A asociado al autovalor λ_i. Entonces $v_j \in \mathcal{S}_{\lambda_1} \oplus \cdots \oplus \mathcal{S}_{\lambda_r}$. En consecuencia, $\mathbb{K}^n = \mathcal{S}_{\lambda_1} \oplus \cdots \oplus \mathcal{S}_{\lambda_r}$, donde usamos la Proposición 4.2.1, para afirmar que los autoespacios $\mathcal{S}_{\lambda_1}, \cdots, \mathcal{S}_{\lambda_r}$, están en suma directa.

$iii) \Rightarrow iv)$: Supongamos que $\mathbb{K}^n = \mathcal{S}_{\lambda_1} \oplus \cdots \oplus \mathcal{S}_{\lambda_r}$, donde $\lambda_1, \cdots, \lambda_r \in \mathbb{K}$ son los r autovalores distintos de A. Entonces

$$\begin{aligned} n = \dim(\mathbb{K}^n) &= \dim(\mathcal{S}_{\lambda_1} \oplus \cdots \oplus \mathcal{S}_{\lambda_r}) \\ &= \dim(\mathcal{S}_{\lambda_1}) + \cdots + \dim(\mathcal{S}_{\lambda_r}) = \mu(\lambda_1) + \cdots \mu(\lambda_r) \\ &\leq m(\lambda_1) + \cdots + m(\lambda_r) = n, \end{aligned}$$

donde usamos (4.1). Entonces, en la cadena anterior, son todas igualdades. Es decir,

$$\mu(\lambda_1) + \cdots + \mu(\lambda_r) = m(\lambda_1) + \cdots + m(\lambda_r) = n.$$

Entonces, $(m(\lambda_1) - \mu(\lambda_1)) + \cdots + (m(\lambda_r) - \mu(\lambda_r)) = 0$. Pero, por (4.1), tenemos que $\mu(\lambda_i) \leq m(\lambda_i)$ para cada $i = 1, \cdots, r$. Por lo tanto,

$$m(\lambda_i) = \mu(\lambda_i),$$

para cada $i = 1, \cdots, r$ y probamos que las multiplicidades algebraicas y geométricas de cada autovalor distinto de A coinciden.

$iv) \Rightarrow ii)$: Supongamos que las multiplicidades algebraicas y geométricas de cada autovalor distinto de A coinciden. Sea B_i una base de $\mathcal{S}_{\lambda_i}$ para cada $i = 1, \cdots, r$. Por la Proposición 4.2.1, los autoespacios $\mathcal{S}_{\lambda_1}, \cdots, \mathcal{S}_{\lambda_r}$ están en suma directa. Entonces $B := B_1 \cup \cdots \cup B_r$ es una base de $\mathcal{S}_{\lambda_1} \oplus \cdots \oplus \mathcal{S}_{\lambda_r} \subseteq \mathbb{K}^n$. Dado un conjunto X, $\#X$ denota el cardinal (número de elementos) de X. Entonces

$$\begin{aligned} \#B = \#B_1 + \cdots + \#B_1 &= \dim(\mathcal{S}_{\lambda_1}) + \cdots + \dim(\mathcal{S}_{\lambda_r}) \\ &= \mu(\lambda_1) + \cdots + \mu(\lambda_r) = m(\lambda_1) + \cdots + m(\lambda_r) = n. \end{aligned}$$

Por lo tanto, $\dim(\mathcal{S}_{\lambda_1} \oplus \cdots \oplus \mathcal{S}_{\lambda_r}) = \#B = n$ y entonces $\mathcal{S}_{\lambda_1} \oplus \cdots \oplus \mathcal{S}_{\lambda_r} = \mathbb{K}^n$. Entonces B es una base de $\mathbb{K}^n$ formada por autovectores de A.

Como probamos $i) \Leftrightarrow ii)$, $ii) \Rightarrow iii)$, $iii) \Rightarrow iv)$ y $iv) \Rightarrow ii)$, se sigue que $i), ii), iii)$ y $iv)$ son equivalentes.

En este caso, veamos que $\mathcal{S}_{\lambda_2} \oplus \cdots \oplus \mathcal{S}_{\lambda_r} = \operatorname{col}(A - \lambda_1 I)$. De hecho, si $y \in \operatorname{col}(A - \lambda_1 I)$ existe $x \in \mathbb{K}^n$ tal que $y = Ax - \lambda_1 x$. Como $x \in \mathbb{K}^n$ existen únicos $v_i \in \mathcal{S}_{\lambda_i}$, $i = 1, \cdots, r$ tales que $x = v_1 + v_2 + \cdots + v_r$. Entonces

$$y = Ax - \lambda_1 x = Av_1 + Av_2 + \cdots + Av_r - \lambda_1 v_1 - \lambda_1 v_2 - \cdots - \lambda_1 v_r$$

$$= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r - \lambda_1 v_1 - \lambda_1 v_2 - \cdots - \lambda_1 v_r$$
$$= (\lambda_2 - \lambda_1) v_2 + \cdots + (\lambda_r - \lambda_1) v_r \in \mathcal{S}_{\lambda_2} \oplus \cdots \oplus \mathcal{S}_{\lambda_r}$$

donde usamos que $(\lambda_2 - \lambda_1) \neq 0, \cdots, (\lambda_r - \lambda_1) \neq 0$. Por lo tanto, $\operatorname{col}(A - \lambda_1 I) \subseteq \mathcal{S}_{\lambda_2} \oplus \cdots \oplus \mathcal{S}_{\lambda_r}$. Pero, como

$$\dim(\operatorname{col}(A - \lambda_1 I)) = n - \dim(\operatorname{nul}(A - \lambda_1 I)) = n - \dim(\mathcal{S}_{\lambda_1}) = \dim(\mathcal{S}_{\lambda_2} \oplus \cdots \oplus \mathcal{S}_{\lambda_r}),$$

se sigue que $\mathcal{S}_{\lambda_2} \oplus \cdots \oplus \mathcal{S}_{\lambda_r} = \bigoplus_{\lambda \in \sigma(A),\ \lambda \neq \lambda_1} \operatorname{nul}(A - \lambda I) = \operatorname{col}(A - \lambda_1 I)$.

De manera similar, se prueba que $\bigoplus_{\lambda \in \sigma(A),\ \lambda \neq \lambda_i} \operatorname{nul}(A - \lambda I) = \operatorname{col}(A - \lambda_i I)$. ■

Notar que, por el Teorema 4.2.2, tenemos que

> Si $A \in \mathbb{K}^{n \times n}$ tiene n autovalores distintos, entonces A es diagonalizable.

El siguiente ejercicio (muy similar al **Ejercicio 4.1**) es un ejemplo de cálculo de autovectores y autovalores y de diagonalización de una matriz de 3×3.

Ejercicio 4.B. Sea $A \in \mathbb{R}^{3 \times 3}$ la matriz

$$A = \begin{bmatrix} 1 & 2 & -2 \\ 0 & 1 & 0 \\ 1 & 3 & 4 \end{bmatrix}.$$

a) Hallar los autovalores y autoespacios de A.

b) Verificar que los autovectores de A forman una base $\mathcal{B}$ de $\mathbb{R}^3$.

c) Hallar la matriz de cambio de coordenadas de la base $\mathcal{B}$ en la base canónica $\mathcal{E}$ de $\mathbb{R}^3$, $P = M_{\mathcal{B}}^{\mathcal{E}}$, y comprobar, que
$$A = P \Lambda P^{-1},$$
donde Λ es una matriz diagonal.

Dem. $a)$: Calculamos el polinomio característico de A :

$$\chi_A(\lambda) = \det(A - xI) = \det \begin{bmatrix} 1-x & 2 & -2 \\ 0 & 1-x & 0 \\ 1 & 3 & 4-x \end{bmatrix} = (1-x)[(1-x)(4-x) - (-2)]$$
$$= (1-x)[(1-x)(4-x) - (-2)] = (1-x)(x^2 - 5x + 6).$$

Los autovalores de A son las raíces de χ_A y son $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$.
Para obtener los autoespacios asociados, calculamos:

$$\mathcal{S}_{\lambda=1} = \operatorname{nul}(A - I) = \operatorname{nul}(\begin{bmatrix} 0 & 2 & -2 \\ 0 & 0 & 0 \\ 1 & 3 & 3 \end{bmatrix}) = gen\{\begin{bmatrix} -6 \\ 1 \\ 1 \end{bmatrix}\}.$$

$$\mathcal{S}_{\lambda=2} = \operatorname{nul}(A - 2I) = \operatorname{nul}(\begin{bmatrix} -1 & 2 & -2 \\ 0 & -1 & 0 \\ 1 & 3 & 2 \end{bmatrix}) = \operatorname{nul}(\begin{bmatrix} -1 & 2 & -2 \\ 0 & -1 & 0 \\ 0 & 5 & 0 \end{bmatrix}) = gen\{\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}\}.$$

$$\mathcal{S}_{\lambda=3} = \operatorname{nul}(A - 3I) = \operatorname{nul}(\begin{bmatrix} -2 & 2 & -2 \\ 0 & -2 & 0 \\ 1 & 3 & 1 \end{bmatrix}) = \operatorname{nul}(\begin{bmatrix} -2 & 2 & -2 \\ 0 & -2 & 0 \\ 0 & 8 & 0 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\}.$$

$b)$: Si tomamos $\mathcal{B} = \{ \begin{bmatrix} -6 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \}$. Entonces, $\mathcal{B}$ es una base de $\mathbb{R}^3$ (ver Teorema 4.2.2) compuesta por autovectores de A. Como obtuvimos una base de autovectores de A, tenemos que A es diagonalizable.

$c)$: Si $\mathcal{E}$ es la base canónica de $\mathbb{R}^3$ entonces

$$M_{\mathcal{B}}^{\mathcal{E}} = \begin{bmatrix} -6 & 2 & 1 \\ 1 & 0 & 0 \\ 1 & -1 & -1 \end{bmatrix} := P.$$

Por lo tanto $P^{-1} = (M_{\mathcal{B}}^{\mathcal{E}})^{-1} = M_{\mathcal{E}}^{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 5 & 1 \\ -1 & -4 & -2 \end{bmatrix}$.

Si $\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, se puede comprobar haciendo la cuenta que $A = P\Lambda P^{-1}$.

Si no queremos hacer la cuenta, podemos pensar lo siguiente: llamemos $T_A : \mathbb{R}^3 \to \mathbb{R}^3$ a la transformación lineal $T_A(x) = Ax$. Entonces, claramente $[T_A]_{\mathcal{E}}^{\mathcal{E}} = A$. Por otra parte, como

$$A(\begin{bmatrix} -6 \\ 1 \\ 1 \end{bmatrix}) = \begin{bmatrix} -6 \\ 1 \\ 1 \end{bmatrix}, \; A(\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}) = 2(\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}) \text{ y } A(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}) = 3(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}).$$

Tenemos que

$$\begin{aligned}
[T_A]_{\mathcal{B}}^{\mathcal{B}} &= [\; [T_A(\begin{bmatrix} -6 \\ 1 \\ 1 \end{bmatrix})]^{\mathcal{B}} \; [T_A(\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix})]^{\mathcal{B}} \; [T_A(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix})]^{\mathcal{B}} \;] \\
&= [\; [A(\begin{bmatrix} -6 \\ 1 \\ 1 \end{bmatrix})]^{\mathcal{B}} \; [A(\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix})]^{\mathcal{B}} \; [A(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix})]^{\mathcal{B}} \;] \\
&= [\; [\begin{bmatrix} -6 \\ 1 \\ 1 \end{bmatrix}]^{\mathcal{B}} \; [2(\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix})]^{\mathcal{B}} \; [3(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix})]^{\mathcal{B}} \;] \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \Lambda.
\end{aligned}$$

Entonces,

$$A = [T_A]_{\mathcal{E}}^{\mathcal{E}} = M_{\mathcal{B}}^{\mathcal{E}} \; [T_A]_{\mathcal{B}}^{\mathcal{B}} \; M_{\mathcal{E}}^{\mathcal{B}} = P\Lambda P^{-1}.$$

∎

Ejercicio 4.2. Sea $A \in \mathbb{R}^{3\times 3}$ que tiene autovalores $\lambda_1 = 0$, $\lambda_2 = 2$, $\lambda_3 = 5$ con autovectores

$$v_1 = [1\ 1\ -1]^T,\; v_2 = [2\ 2\ -1]^T,\; v_3 = [1\ 2\ -1]^T,$$

respectivamente.

a) Hallar una base de de $\mathrm{nul}(A)$ y una base de $\mathrm{col}(A)$.

b) Hallar una solución particular de $Ax = v_2 + v_3$.

c) Hallar todas las soluciones de $Ax = v_2 + v_3$.

d) Explicar por qué la ecuación $Ax = v_1$ no tiene solución.

Dem. $a)$: Recordemos que $v \in \mathrm{nul}(A)$ si y sólo si $Av = 0$. Por lo tanto, $\mathrm{nul}(A)$ es el autoespacio asociado a $\lambda_1 = 0$. Entonces

$$\mathrm{nul}(A) = \mathcal{S}_{\lambda_1=0} = gen\{v_1\} = gen\{[1\ 1\ \ -1]^T\}$$

y una base de $\mathrm{nul}(A)$ puede ser $B_{\mathrm{nul}(A)} = \{[1\ 1\ \ -1]^T\}$.

Notar que $\{v_1, v_2, v_3\}$ es una base de $\mathbb{R}^3$ (esto se puede probar por definición o viendo que v_1, v_2 y v_3 son autovectores asociados a autovalores distintos y entonces, por la Proposición 4.2.1, son linealmente independientes) y recordemos que $y \in \mathrm{col}(A)$ si y sólo si existe $x \in \mathbb{R}^3$ tal que $y = Ax$. Entonces, veamos que que

$$\mathrm{col}(A) = gen\{v_2, v_3\}.$$

De hecho, si $y \in \mathrm{col}(A)$, entonces $y = Ax$ para cierto $x \in \mathbb{R}^3$. Como $x \in \mathbb{R}^3$ y $\{v_1, v_2, v_3\}$ es una base de $\mathbb{R}^3$, existen $a, b, c \in \mathbb{R}$ tales que $x = av_1 + bv_2 + cv_3$. Por lo tanto

$$y = Ax = A(av_1 + bv_2 + cv_3) = aAv_1 + bAv_2 + cAv_3 = 0 + 2bv_2 + 5cv_3 = 2bv_2 + 5cv_3,$$

con $b, c \in \mathbb{R}$. Por lo tanto $y \in gen\{v_2, v_3\}$ y tenemos que $\mathrm{col}(A) \subseteq gen\{v_2, v_3\}$. Finalmente, por el Teorema de la Dimensión, vale que

$$\dim(\mathrm{col}(A)) = 3 - \dim(\mathrm{nul}(A)) = 3 - 1 = 2 = \dim(gen\{v_2, v_3\}).$$

Entonces, $\mathrm{col}(A) = gen\{v_2, v_3\}$ y una base de $\mathrm{col}(A)$ puede ser $B_{\mathrm{col}(A)} = \{[2\ 2\ \ -1]^T, [1\ 2\ \ -1]^T\}$.

$b)$ Sabemos que $Av_2 = 2v_2$ y $Av_3 = 5v_3$. Entonces, si tomamos $x_p := \frac{1}{2}v_2 + \frac{1}{5}v_3$, vale que

$$Ax_p = A(\frac{1}{2}v_2 + \frac{1}{5}v_3) = \frac{1}{2}Av_2 + \frac{1}{5}Av_3 = \frac{1}{2}2v_2 + \frac{1}{5}5v_3 = v_2 + v_3.$$

$c)$: Todas las soluciones x_s de la ecuación $Ax = v_2 + v_3$ se construyen de la forma $x_s = x_p + x_h$, donde x_p es una solución particular y $x_h \in \mathrm{nul}(A)$. Entonces, por los ítems $a)$ y $b)$, tenemos que todas las soluciones de la ecuación $Ax = v_2 + v_3$ son

$$x_s = \frac{1}{2}v_2 + \frac{1}{5}v_3 + \alpha v_1,$$

con $\alpha \in \mathbb{R}$.

$d)$ En el item $a)$ vimos que $gen\{v_2, v_3\} = \mathrm{col}(A)$. Entonces, como el conjunto $\{v_1, v_2, v_3\}$ es linealmente independiente, es claro que $v_1 \notin gen\{v_2, v_3\} = \mathrm{col}(A)$. Por lo tanto, no existe solución al sistema $Ax = v_1$. ∎

Ejercicio 4.C. Explicar por qué las siguientes matrices no son diagonalizables en $\mathbb{R}^{2\times 2}$:

$$\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \lambda \in \mathbb{R}; \quad \rho \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \rho > 0,\ \theta \in \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}.$$

Dem. Calculamos el polinomio característico de la matriz $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ con $\lambda \in \mathbb{R}$:

$$\chi(x) = \det(\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} - xI) = \det \begin{bmatrix} \lambda - x & 1 \\ 0 & \lambda - \mu \end{bmatrix} = (\lambda - x)^2.$$

Los autovalores de la matriz $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ con $\lambda \in \mathbb{R}$, son las raíces del polinomio característico. En este caso, las raíces son $\lambda_1 = \lambda_2 = \lambda$. Veamos la dimensión del autoespacio asociado al autovalor $\lambda_{1,2} = \lambda$. Para eso, calculamos

$$\mathrm{nul}(\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} - \lambda I) = nul(\begin{bmatrix} \lambda - \lambda & 1 \\ 0 & \lambda - \lambda \end{bmatrix}) = \mathrm{nul}(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\}.$$

Entonces, la multiplicidad geométrica asociada al autovalor λ es 1, mientras que la multiplicidad algebraica asociada al autovalor λ era 2, es decir $\mu(\lambda) < m(\lambda)$. Por lo tanto, por el Teorema 4.2.2, la matriz $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ con $\lambda \in \mathbb{R}$ nunca es diagonalizable.

Ahora, calculamos el polinomio característico de la matriz $\rho \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ con $\rho > 0$ y $\theta \in \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}$:

$$\begin{aligned}\chi(x) &= \det(\rho \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} - xI) = \det \begin{bmatrix} \rho\cos\theta - x & -\rho\sin\theta \\ \rho\sin\theta & \rho\cos\theta - x \end{bmatrix} \\ &= (\rho\cos\theta - x)(\rho\cos\theta - x) + \rho^2 \sin\theta^2 = \rho^2\cos\theta^2 - 2\rho x\cos\theta + x^2 + \rho^2\sin\theta^2 \\ &= x^2 - 2x\rho\cos\theta + \rho^2(\cos\theta^2 + \sin\theta^2) = x^2 - 2x\rho\cos\theta + \rho^2.\end{aligned}$$

Los autovalores de la matriz $\rho \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ con $\rho > 0$ y $\theta \in \mathbb{R}\setminus\{k\pi : k \in \mathbb{Z}\}$ son las raíces en $\mathbb{R}$ (estamos considerando cuerpo real) del polinomio característico. Es decir, buscamos $\lambda \in \mathbb{R}$ tal que $\chi(\lambda) = \lambda^2 - \lambda 2\rho\cos\theta + \rho^2 = 0$. Entonces, despejando y recordando que $\rho > 0$, nos queda que

$$\lambda_{1,2} = \frac{2\rho\cos\theta \pm \sqrt{4\rho^2\cos\theta^2 - 4\rho^2}}{2} = \frac{2\rho\cos\theta \pm 2\rho\sqrt{\cos\theta^2 - 1}}{2}.$$

Observar que como $\theta \in \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}$ entonces $\cos\theta^2 < 1$. Por lo tanto, $\cos\theta^2 - 1 < 0$ y entonces $\sqrt{\cos\theta^2 - 1} \in \mathbb{C} \setminus \mathbb{R}$. Por lo tanto, como $\rho > 0$, tenemos que

$$\lambda_{1,2} = \rho(\cos\theta \pm i\sqrt{|\cos\theta^2 - 1|}) \in \mathbb{C} \setminus \mathbb{R},$$

entonces la matriz $\rho \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ con $\rho > 0$ y $\theta \in \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}$ no tiene autovalores (en $\mathbb{R}$) y concluimos que dicha matriz no es diagonalizable. ■

Ejercicio 4.4. Sea $A \in \mathbb{R}^{3\times3}$ la matriz dependiente de los parámetros reales a, b, c definida por:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ c & b & a \end{bmatrix}.$$

a) Hallar los valores $a, b, c \in \mathbb{R}$ tales que $\chi_A(x) = \det(A - xI) = 9x - x^3$.¿$A$ es diagonalizable? Si la respuesta es afirmativa hallar $P \in \mathbb{R}^{3\times3}$ inversible y $\Lambda \in \mathbb{R}^{3\times3}$ diagonal tales que $A = P\Lambda P^{-1}$.

b) Para $c = 0$, hallar el conjunto de las parejas $a, b \in \mathbb{R}$ tales que A es diagonalizable.

Dem. El polinomio característico de A es

$$\begin{aligned}\chi_A(x) &= \det(A - xI) = \det \begin{bmatrix} -x & 1 & 0 \\ 0 & -x & 1 \\ c & b & a - x \end{bmatrix} = -x[-x(a - x) - b] + c[1 - 0] \\ &= -x^3 + ax^2 + bx - c = -x^3 + 9x.\end{aligned}$$

Por lo tanto $a = 0, b = 9$ y $c = 0$. En ese caso, como $\chi_A(x) = -x^3 + 9x = x(9 - x^2)$. Los autovalores de A (que son las raíces de χ_A) son $\lambda_1 = 0$, $\lambda_2 = 3$ $\lambda_3 = -3$. Como A tiene 3 autovalores distintos, por el Teorema 4.2.2, A es diagonalizable. Calculemos los autoespacios asociados:

$$\mathcal{S}_{\lambda=0} = \text{nul}(A) = \text{nul}(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 9 & 0 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\},$$

$$\mathcal{S}_{\lambda=3} = \text{nul}(A - 3I) = \text{nul}(\begin{bmatrix} -3 & 1 & 0 \\ 0 & -3 & 1 \\ 0 & 9 & -3 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix}\},$$

$$\mathcal{S}_{\lambda=-3} = \text{nul}(A + 3I) = \text{nul}(\begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 9 & 3 \end{bmatrix}) = gen\{\begin{bmatrix} -1 \\ 3 \\ -9 \end{bmatrix}\}.$$

Si tomamos $P := \begin{bmatrix} 1 & 1 & -1 \\ 0 & 3 & 3 \\ 0 & 9 & -9 \end{bmatrix}$ y $\Lambda := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$, entonces P es inversible (sus columnas son una base de $\mathbb{R}^3$), Λ es diagonal y $A = P\Lambda P^{-1}$.

b) Sea $c = 0$, entonces $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & b & a \end{bmatrix}$ y el polinomio característico de A es

$$\begin{aligned} \chi_A(x) = \det(A - xI) = \det \begin{bmatrix} -x & 1 & 0 \\ 0 & -x & 1 \\ 0 & b & a - x \end{bmatrix} &= -x[-x(a - x) - b] \\ &= -x[x^2 - ax - b]. \end{aligned}$$

Entonces, los autovalores de A (que son las raíces de χ_A) son $\lambda_1 = 0$ y

$$\lambda_{2,3} = \frac{a \pm \sqrt{a^2 + 4b}}{2}.$$

Veamos 4 casos posibles:

- **Caso 1:** $a^2 + 4b = 0$ y $a = 0$. Entonces $b = -\frac{a^2}{4} = 0$ y $\lambda_{2,3} = \frac{a}{2} = 0$. Por lo tanto $\lambda = 0$ es un autovalor triple de A, es decir $m(\lambda = 0) = 3$. Por otra parte,

$$\text{nul}(A) = \text{nul}(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\}.$$

  Entonces $1 = \mu(\lambda = 0) < 3 = m(\lambda = 0)$ y entonces A no es diagonalizable.

- **Caso 2:** $a^2 + 4b = 0$ y $a \neq 0$. Entonces $b = -\frac{a^2}{4} \neq 0$. Por lo tanto $\lambda = \frac{a}{2} \neq 0$ es un autovalor doble de A, es decir $m(\lambda = \frac{a}{2}) = 2$. Por otra parte,

$$\text{nul}(A - \frac{a}{2}I) = \text{nul}(\begin{bmatrix} -\frac{a}{2} & 1 & 0 \\ 0 & -\frac{a}{2} & 1 \\ 0 & -\frac{a^2}{4} & \frac{a}{2} \end{bmatrix}) = gen\{\begin{bmatrix} \frac{2}{a} \\ 1 \\ \frac{a}{2} \end{bmatrix}\}.$$

  Entonces $1 = \mu(\lambda = \frac{a}{2}) < 2 = m(\lambda = \frac{a}{2})$ y entonces A no es diagonalizable.

- **Caso 3:** $a^2 + 4b > 0$. Entonces $\lambda_1 = 0$, $\lambda_2 = \frac{a+\sqrt{a^2+4b}}{2}$ y $\lambda_3 = \frac{a-\sqrt{a^2+4b}}{2}$ son 3 autovalores distintos de A y por lo tanto, por el Teorema 4.2.2, A es diagonalizable.

- **Caso 4:** $a^2 + 4b < 0$. Entonces $\lambda_{2,3} = \frac{a \pm i\sqrt{|a^2+4b|}}{2} \in \mathbb{C}$. Como estamos considerando como cuerpo $\mathbb{R}$, concluimos que A no es diagonalizable.

Conclusión, A es diagonalizable si y sólo si $a^2 + 4b > 0$. ■

Las mismas nociones de autovalores, autovectores y diagonalización que vimos para matrices se pueden definir para transformaciones lineales.

Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial y sea $T : \mathbb{V} \to \mathbb{V}$ una transformación lineal. El escalar $\lambda \in \mathbb{K}$ es un *autovalor* de T si existe un vector no nulo $v \in \mathbb{V} \setminus \{0_{\mathbb{V}}\}$ tal que

$$T(v) = \lambda v.$$

Al vector v lo llamamos *autovector asociado a* λ y definimos el *autoespacio asociado a* λ como

$$\mathcal{S}_\lambda := \{v \in \mathbb{V} : T(v) = \lambda v\} = \mathrm{Nu}(T - \lambda I).$$

Finalmente, diremos que $T \in \mathcal{L}(\mathbb{V})$ es diagonalizable si existe una base de $\mathbb{V}$ formada por autovectores de T.

Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial de dimensión finita y $T \in \mathcal{L}(\mathbb{V})$, se define el *polinomio característico de* T como

$$\chi_T(x) := \det([T]_{\mathcal{B}}^{\mathcal{B}} - xI),$$

donde $\mathcal{B}$ es cualquier base de $\mathbb{V}$.

Observar que si $\mathcal{B}'$ es otra base de $\mathbb{V}$, entonces

$$[T]_{\mathcal{B}'}^{\mathcal{B}'} = (M_{\mathcal{B}'}^{\mathcal{B}})^{-1}\ [T]_{\mathcal{B}}^{\mathcal{B}}\ M_{\mathcal{B}'}^{\mathcal{B}}.$$

Entonces

$$\begin{aligned}
\det([T]_{\mathcal{B}'}^{\mathcal{B}'} - xI) &= \det((M_{\mathcal{B}'}^{\mathcal{B}})^{-1}\ [T]_{\mathcal{B}}^{\mathcal{B}}\ M_{\mathcal{B}'}^{\mathcal{B}} - x(M_{\mathcal{B}'}^{\mathcal{B}})^{-1}M_{\mathcal{B}'}^{\mathcal{B}}) = \det((M_{\mathcal{B}'}^{\mathcal{B}})^{-1}([T]_{\mathcal{B}}^{\mathcal{B}} - xI)M_{\mathcal{B}'}^{\mathcal{B}}) \\
&= \det((M_{\mathcal{B}'}^{\mathcal{B}})^{-1})\ \det([T]_{\mathcal{B}}^{\mathcal{B}} - xI)\ \det(M_{\mathcal{B}'}^{\mathcal{B}}) \\
&= \det(M_{\mathcal{B}'}^{\mathcal{B}})^{-1}\ \det(M_{\mathcal{B}'}^{\mathcal{B}})\ \det([T]_{\mathcal{B}}^{\mathcal{B}} - xI) \\
&= \det([T]_{\mathcal{B}}^{\mathcal{B}} - xI).
\end{aligned}$$

Por lo tanto el polinomio característico de T está bien definido (no depende de la base $\mathcal{B}$ elegida).

Entonces, de la misma manera que vimos para matrices tenemos que:

Sea $\mathbb{V}$ un $\mathbb{K}$-espacio vectorial de dimensión finita y $T \in \mathcal{L}(\mathbb{V})$, entonces $\lambda \in \mathbb{K}$ es un autovalor de T si y sólo si $\lambda \in \mathbb{K}$ es una raíz de χ_T.

Veamos un ejemplo para aplicar todos estos conceptos a transformaciones lineales:

Ejercicio 4.D (De Parcial)**.** Dada $T \in \mathcal{L}(\mathbb{V})$ definida por $T(v_1) = 7v_1 + 2v_2$ y $T(v_2) = -4v_1 + v_2$, con $\mathcal{B} = \{v_1; v_2\}$ una base de $\mathbb{V}$. Entonces

a) Hallar, si existe, una base $\mathcal{C}$ de $\mathbb{V}$ tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ sea diagonal.

b) Calcular $[T^k]_{\mathcal{B}}^{\mathcal{B}}$ para todo $k \in \mathbb{N}$.

Dem. $a)$: Veamos si T es diagonalizable. Para eso elijamos una base conveniente y calculemos el polinomio característico de T. Si tomamos la base $\mathcal{B}$, entonces $[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 7 & -4 \\ 2 & 1 \end{bmatrix}$. Por lo tanto

$$\chi_T(x) = \det([T]_{\mathcal{B}}^{\mathcal{B}} - xI) = \det(\begin{bmatrix} 7-x & -4 \\ 2 & 1-x \end{bmatrix}) = (7-x)(1-x) + 8 = x^2 - 8x + 15.$$

Cuyas raíces son $\lambda_1 = 3$ y $\lambda_2 = 5$.

Calculemos los autoespacios asociados:

$$\mathrm{nul}([T]_{\mathcal{B}}^{\mathcal{B}} - 3I) = \mathrm{nul}(\begin{bmatrix} 4 & -4 \\ 2 & -2 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}.$$

Por lo tanto, $S_{\lambda=3} = \mathrm{Nu}(T - 3I_{\mathbb{V}}) = gen\{v_1 + v_2\}$.

$$\mathrm{nul}([T]_{\mathcal{B}}^{\mathcal{B}} - 5I) = \mathrm{nul}(\begin{bmatrix} 2 & -4 \\ 2 & -4 \end{bmatrix}) = gen\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}\}.$$

Por lo tanto, $S_{\lambda=5} = \mathrm{Nu}(T - 5I_{\mathbb{V}}) = gen\{2v_1 + v_2\}$.

Sea $\mathcal{C} := \{v_1 + v_2; 2v_1 + v_2\}$ entonces $\mathcal{C}$ es una base de $\mathbb{V}$ formada por autovectores de T. Entonces, como $T(v_1 + v_2) = 3(v_1 + v_2)$ y $T(2v_1 + v_2) = 5(2v_1 + v_2)$, tenemos que

$$[T]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$$

y T resulta diagonalizable.

b): En la Proposición 2.4.3, vimos que para todo $k \in \mathbb{N}$, vale

$$[T^k]_{\mathcal{B}}^{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}})^k.$$

Por otra parte, observar que si $\mathcal{B}$ y $\mathcal{C}$ son bases de $\mathbb{V}$ entonces

$$[T]_{\mathcal{B}}^{\mathcal{B}} = M_{\mathcal{C}}^{\mathcal{B}}\ [T]_{\mathcal{C}}^{\mathcal{C}}\ M_{\mathcal{B}}^{\mathcal{C}}.$$

Llamemos $P := M_{\mathcal{C}}^{\mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$, entonces $P^{-1} := M_{\mathcal{B}}^{\mathcal{C}} = (M_{\mathcal{C}}^{\mathcal{B}})^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$. Por lo tanto, $[T]_{\mathcal{B}}^{\mathcal{B}} = P \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} P^{-1}$. Observar que

$$\begin{aligned} ([T]_{\mathcal{B}}^{\mathcal{B}})^2 &= P \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} P^{-1}P \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} P^{-1} = P\ (\begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix})^2\ P^{-1} \\ &= P \begin{bmatrix} 3^2 & 0 \\ 0 & 5^2 \end{bmatrix} P^{-1}. \end{aligned}$$

A partir de la observación anterior, vamos a probar por inducción que

$$([T]_{\mathcal{B}}^{\mathcal{B}})^k = P \begin{bmatrix} 3^k & 0 \\ 0 & 5^k \end{bmatrix} P^{-1}, \tag{4.9}$$

para todo $k \in \mathbb{N}$.

Si $k = 1$, claramente (4.9) vale.

Supongamos que para $k = n \in \mathbb{N}$ vale que $([T]_{\mathcal{B}}^{\mathcal{B}})^n = P \begin{bmatrix} 3^n & 0 \\ 0 & 5^n \end{bmatrix} P^{-1}$ (HI).

Entonces, usando la HI,

$$\begin{aligned} ([T]_{\mathcal{B}}^{\mathcal{B}})^{n+1} &= [T]_{\mathcal{B}}^{\mathcal{B}}([T]_{\mathcal{B}}^{\mathcal{B}})^n = [T]_{\mathcal{B}}^{\mathcal{B}}\ (P \begin{bmatrix} 3^n & 0 \\ 0 & 5^n \end{bmatrix} P^{-1}) \\ &= P \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} P^{-1}P \begin{bmatrix} 3^n & 0 \\ 0 & 5^n \end{bmatrix} P^{-1} = P \begin{bmatrix} 3^{n+1} & 0 \\ 0 & 5^{n+1} \end{bmatrix} P^{-1}. \end{aligned}$$

Como a partir de la validez de la ecuación (4.9) para $k = n$ deducimos la validez de dicha ecuación para $k = n + 1$, por inducción, se sigue que,

$$([T]_{\mathcal{B}}^{\mathcal{B}})^k = P \begin{bmatrix} 3^k & 0 \\ 0 & 5^k \end{bmatrix} P^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3^k & 0 \\ 0 & 5^k \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix},$$

para todo $k \in \mathbb{N}$. ■

Entonces, para todo $k \in \mathbb{N}$,

$$([T]_{\mathcal{B}}^{\mathcal{B}})^k = \begin{bmatrix} -3^k + 2 \cdot 5^k & 2 \cdot (3^k - 5^k) \\ -3^k + 5^k & 2 \cdot 3^k - 5^k \end{bmatrix}.$$

Ejercicio 4.E (De Parcial)**.** Sea

$$A = \begin{bmatrix} 2 & 2a-4 & 2a^2-8 \\ 0 & 4 & 0 \\ 0 & 2a-4 & 2a-2 \end{bmatrix}.$$

a) Obtener todos los valores de $a \in \mathbb{R}$ para los cuales A es diagonalizable.

b) Diagonalizar A para $a = 1$.

Dem. $a)$: Calculamos el polinomio característico de A :

$$\begin{aligned}
\chi_A(x) = \det(A - xI) &= \det(\begin{bmatrix} 2-x & 2a-4 & 2a^2-8 \\ 0 & 4-x & 0 \\ 0 & 2a-4 & 2a-2-x \end{bmatrix}) \\
&= (4-\lambda)[(2-x)(2a-2-x) - (2a^2-8)\cdot 0] \\
&= (4-x)(x^2 - 2ax + 4(a-1)).
\end{aligned}$$

Los autovalores de A son las raíces de su polinomio característico. Es decir, buscamos $\lambda \in \mathbb{R}$ tales que $\chi_A(x) = (4-x)(x^2-2ax+4(a-1)) = 0$. En este caso, las raíces son $\lambda_1 = 4$ y

$$\lambda_{2,3} = \frac{2a \pm \sqrt{4a^2 - 16(a-1)}}{2} = \frac{2a \pm \sqrt{(2a-4)^2}}{2} = \frac{2a \pm (2a-4)}{2} = a \pm (a-2).$$

Entonces, $\lambda_2 = 2a-2$ y $\lambda_3 = 2$.

Observar que si $2a-2 \neq 2$ y $2a-2 \neq 4$ entonces, A tiene 3 autovalores distintos y por ende es diagonalizable. Es decir si $a \neq 2$ y $a \neq 3$ entonces A es diagonalizable. Veamos qué pasa si $a = 2$ ó $a = 3$.

Si $a = 2$, entonces los autovalores de A son $\lambda_1 = 4$ y $\lambda_{2,3} = 2$. En este caso, $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.
Entonces A es una matriz diagonal y en ese caso (trivialmente) A es diagonalizable.

Si $a = 3$, entonces los autovalores de A son $\lambda_{1,2} = 4$ y $\lambda_3 = 2$. En este caso, $A = \begin{bmatrix} 2 & 2 & 10 \\ 0 & 4 & 0 \\ 0 & 2 & 4 \end{bmatrix}$.
Veamos la multiplicidad geométrica del autovalor $\lambda_{2,3} = 4$:

$$\text{nul}(A - 4I) = \text{nul}(\begin{bmatrix} -2 & 2 & 10 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}) = gen\{ \begin{bmatrix} 5 \\ 0 \\ 1 \end{bmatrix} \}.$$

Como la multiplicidad geométrica de $\lambda = 4$ es 1, pero la multiplicidad algebraica de dicho autovalor es 2, entonces $\mu(\lambda = 4) < m(\lambda = 4)$. Por lo tanto, concluimos que en este caso A NO es diagonalizable.

En conclusión, A es diagonalizable para todo $a \in \mathbb{R} \setminus \{3\}$.

$b)$: Si $a = 1$, entonces $A = \begin{bmatrix} 2 & -2 & -6 \\ 0 & 4 & 0 \\ 0 & -2 & 0 \end{bmatrix}$. En este caso, los autovalores de A son $\lambda_1 = 4$, $\lambda_2 = 0$ y $\lambda_3 = 2$. Calculemos los autoespacios asociados:

$$\mathcal{S}_{\lambda=4} = \text{nul}(A - 4I) = \text{nul}(\begin{bmatrix} -2 & -2 & -6 \\ 0 & 0 & 0 \\ 0 & -2 & -4 \end{bmatrix}) = gen\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \}.$$

$$\mathcal{S}_{\lambda=0} = \text{nul}(A) = \text{nul}(\begin{bmatrix} 2 & -2 & -6 \\ 0 & 4 & 0 \\ 0 & -2 & 0 \end{bmatrix}) = gen\{ \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \}.$$

$$\mathcal{S}_{\lambda=2} = \mathrm{nul}(A-2I) = \mathrm{nul}(\begin{bmatrix} 0 & -2 & -6 \\ 0 & 2 & 0 \\ 0 & -2 & -2 \end{bmatrix}) = \mathrm{nul}(\begin{bmatrix} 0 & -2 & -6 \\ 0 & 0 & -6 \\ 0 & 0 & 4 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\}.$$

Si tomamos $P := \begin{bmatrix} 1 & 3 & 1 \\ 2 & 0 & 0 \\ -1 & 1 & 0 \end{bmatrix}$ y $\Lambda := \begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, entonces P es inversible, Λ es diagonal y

$$A = P\Lambda P^{-1}.$$

■

4.3. Polinomios matriciales

Sean $q(x) = a_k x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0 \in \mathbb{K}_k[x]$ y $A \in \mathbb{K}^{n\times n}$. Definimos

$$q(A) := a_k A^k + a_{k-1}A^{k-1} + \cdots + a_1 A + a_0 I.$$

Por ejemplo, sean $q(x) = 2x^2 + 1$ y $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Entonces,

$$q(A) = 2A^2 + I = 2\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^2 + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 0 & 3 \end{bmatrix}.$$

A continuación veremos algunas propiedades muy importantes de los polinomios matriciales.

Proposición 4.3.1. *Sean $q(x) = a_k x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0 \in \mathbb{K}_k[x]$ y $A \in \mathbb{K}^{n\times n}$. Si $\lambda \in \mathbb{K}$ es un autovalor de A con autovector asociado v entonces $q(\lambda)$ es un autovalor de $q(A)$ con mismo autovector asociado.*

Dem. Sea $\lambda \in \mathbb{K}$ un autovalor de A, entonces existe $v \neq 0$ tal que $Av = \lambda v$. Entonces (lo vimos en la Sección 4.1): λ^k es autovalor de A^k para todo $k \in \mathbb{N}$ con el mismo autovector asociado v. Por lo tanto,

$$\begin{aligned} q(A)v &= (a_k A^k + a_{k-1}A^{k-1} + \cdots + a_1 A + a_0 I)v = a_k A^k v + a_{k-1}A^{k-1}v + \cdots + a_1 Av + a_0 v \\ &= a_k \lambda^k v + a_{k-1}\lambda^{k-1} v + \cdots + a_1 \lambda v + a_0 v = (a_k\lambda^k + a_{k-1}\lambda^{k-1} + \cdots + a_1\lambda + a_0)v \\ &= q(\lambda)v. \end{aligned}$$

Entonces, $q(\lambda)$ es un autovalor de $q(A)$ con mismo autovector asociado. ■

Proposición 4.3.2. *Sean $q(x) = a_k x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0 \in \mathbb{K}_k[x]$ con $a_k \neq 0$ y $A \in \mathbb{K}^{n\times n}$. Si $\lambda \in \mathbb{K}$ es un autovalor de $q(A)$ entonces existe $\nu \in \mathbb{K}$ autovalor de A tal que $\lambda = q(\nu)$.*

Dem. Sea $\lambda \in \mathbb{K}$ un autovalor de $q(A)$, entonces existe $v \neq 0$ tal que $q(A)v = \lambda v$. Sea $p(x) := q(x) - \lambda$, entonces como $a_k \neq 0$ tenemos que p es de grado k. Entonces, por el Teorema fundamental del álgebra, existen $\nu_1, \nu_2, \cdots, \nu_k \in \mathbb{K}$ tales que

$$p(x) = q(x) - \lambda = a_k(x-\nu_1)(x-\nu_2)\cdots(x-\nu_k).$$

Entonces

$$p(A) = q(A) - \lambda I = a_k(A-\nu_1 I)(A-\nu_2 I)\cdots(A-\nu_k I).$$

Como $p(A)v = q(A)v - \lambda v = 0$ y $v \neq 0$, se sigue que

$$\begin{aligned} 0 = \det(p(A)) &= a_k^k \det[(A-\nu_1 I)(A-\nu_2 I)\cdots(A-\nu_k I)] \\ &= a_k^k \det(A-\nu_1 I)\det(A-\nu_2 I)\cdots\det(A-\nu_k I). \end{aligned}$$

Entonces, como $a_k \neq 0$, existe $\nu \in \{\nu_1, \nu_2, \cdots, \nu_k\}$ tal que $\det(A-\nu I) = 0$. Entonces ν es autovalor de A y además, $0 = p(\nu) = q(\nu) - \lambda$. Por lo tanto $q(\nu) = \lambda$. ■

Proposición 4.3.3. *Sea $A \in \mathbb{K}^{n\times n}$ diagonalizable y sea $A = P\Lambda P^{-1}$ con $P \in \mathbb{K}^{n\times n}$ inversible y $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) \in \mathbb{K}^{n\times n}$ donde $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{K}$ son los autovalores de A. Entonces, para cada $n \in \mathbb{N}$,*

$$A^n = P\Lambda^n P^{-1} = P \operatorname{diag}(\lambda_1^n, \lambda_2^n, \cdots, \lambda_n^n)P^{-1}.$$

Más aún, si $q(x) = a_k x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0 \in \mathbb{K}_k[x]$. Entonces

$$q(A) = Pq(\Lambda)P^{-1} = P \operatorname{diag}(q(\lambda_1), q(\lambda_2), \cdots, q(\lambda_n))P^{-1}.$$

Dem. Vamos a probar la propiedad para todo $n \in \mathbb{N}$ por inducción: Para $n = 1$ es claro que la propiedad vale.

Supongamos que para $n = k \in \mathbb{N}$ vale que $A^k = P\Lambda^k P^{-1}$ (HI). Entonces, multiplicamos esa ecuación a ambos lados por A y tenemos que

$$A^{k+1} = AA^k = P\Lambda P^{-1}P\Lambda^k P^{-1} = P\Lambda\Lambda^k P^{-1} = P\Lambda^{k+1}P^{-1}.$$

Entonces, a partir de la hipótesis inductiva deducimos la validez de la propiedad para $n = k + 1$. Por lo tanto, por inducción, concluimos que para cada $n \in \mathbb{N}$,

$$A^n = P\Lambda^n P^{-1} = P \operatorname{diag}(\lambda_1^n, \lambda_2^n, \cdots, \lambda_n^n)P^{-1}.$$

Más aún,

$$\begin{aligned}
q(A) &= a_k A^k + a_{k-1}A^{k-1} + \cdots + a_1 A + a_0 I \\
&= a_k P\Lambda^k P^{-1} + a_{k-1}P\Lambda^{k-1}P^{-1} + \cdots + a_1 P\Lambda P^{-1} + a_0 PP^{-1} \\
&= P(a_k\Lambda^k + a_{k-1}\Lambda^{k-1} + \cdots + a_1\Lambda + a_0 I)P^{-1} \\
&= Pq(\Lambda)P^{-1} \\
&= P \operatorname{diag}(q(\lambda_1), q(\lambda_2), \cdots, q(\lambda_n))P^{-1}.
\end{aligned}$$

■

Corolario 4.3.4. *Supongamos que $A \in \mathbb{K}^{n\times n}$ es diagonalizable y sea χ_A es el polinomio característico de A, entonces*

$$\chi_A(A) = 0_{n\times n},$$

donde $0_{n\times n}$ es la matriz nula de $n \times n$.

La propiedad anterior es el conocido *Teorema de Cayley-Hamilton* y dicho teorema sigue valiendo aún cuando A no es diagonalizable.

Dem. Recordemos que si $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{K}$ son los autovalores de A entonces

$$\chi_A(\lambda_1) = \chi_A(\lambda_2) = \cdots = \chi_A(\lambda_n) = 0$$

porque los autovalores de A son las raíces del polinomio característico de A. Sea $A = P\Lambda P^{-1}$ con $P \in \mathbb{K}^{n\times n}$ inversible y $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) \in \mathbb{K}^{n\times n}$. Entonces, por la Proposición 4.3.3, tenemos que

$$\chi_A(A) = P\chi_A(\Lambda)P^{-1} = P \operatorname{diag}(\chi_A(\lambda_1), \chi_A(\lambda_2), \cdots, \chi_A(\lambda_n))P^{-1} = P0_{n\times n}P^{-1} = 0_{n\times n}.$$

■

El siguiente ejercicio es una aplicación de todos los conceptos que vimos arriba:

Ejercicio 4.6. Hallar una matriz $A \in \mathbb{R}^{2\times 2}$ que posea las siguientes propiedades:

a) $A^2 - 3A + 2I = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}$.

a) $A^2-3A+2I=\begin{bmatrix}3&3\\3&3\end{bmatrix}$ y $\det(A)=-1$.

a) $A^2-3A+2I=\begin{bmatrix}3&3\\3&3\end{bmatrix}$ y $\operatorname{tr}(A)=6$.

Sea $B:=\begin{bmatrix}3&3\\3&3\end{bmatrix}$. Entonces $B=q(A)$ con $q(x)=x^2-3x+2$.

El polinomio característico de B es $\chi_B(x)=\det(B-xI)=(3-x)^2-9$. Entonces, $\lambda_1=6$ y $\lambda_2=0$ son los autovalores de B con los respectivos autoespacios asociados: $\mathcal{S}_{\lambda=6}=\operatorname{nul}(B-6I)=gen\{\begin{bmatrix}1\\1\end{bmatrix}\}$ y $\mathcal{S}_{\lambda=0}=\operatorname{nul}(B)=gen\{\begin{bmatrix}1\\-1\end{bmatrix}\}$.

Entonces B es diagonalizable y, si $P:=\begin{bmatrix}1&1\\1&-1\end{bmatrix}$, entonces P es inversible y

$$B=P\begin{bmatrix}6&0\\0&0\end{bmatrix}P^{-1}.$$

Por la Proposición 4.3.2, existe $\nu_1\in\mathbb{R}$ autovalor de A tal que $q(\nu_1)=\nu_1^2-3\nu_1+2=6$ y existe $\nu_2\in\mathbb{R}$ autovalor de A tal que $q(\nu_2)=\nu_1^2-3\nu_1+2=0$. Resolviendo las ecuaciones cuadráticas nos queda que $\nu_1=4$ ó $\nu_1=-1$ y $\nu_2=2$ ó $\nu_2=1$.

a) Si definimos $A:=P\begin{bmatrix}4&0\\0&1\end{bmatrix}P^{-1}$. Entonces, por la Proposición 4.3.3,

$$A^2-3A+2I=q(A)=P\begin{bmatrix}q(4)&0\\0&q(1)\end{bmatrix}P^{-1}=P\begin{bmatrix}6&0\\0&0\end{bmatrix}P^{-1}=B.$$

b) Si definimos $A:=P\begin{bmatrix}-1&0\\0&1\end{bmatrix}P^{-1}$. Entonces, $\det(A)=-1$ y, por la Proposición 4.3.3,

$$A^2-3A+2I=q(A)=P\begin{bmatrix}q(-1)&0\\0&q(1)\end{bmatrix}P^{-1}=P\begin{bmatrix}6&0\\0&0\end{bmatrix}P^{-1}=B.$$

c) Si definimos $A:=P\begin{bmatrix}4&0\\0&2\end{bmatrix}P^{-1}$. Entonces, $\operatorname{tr}(A)=6$ y, por la Proposición 4.3.3,

$$A^2-3A+2I=q(A)=P\begin{bmatrix}q(4)&0\\0&q(2)\end{bmatrix}P^{-1}=P\begin{bmatrix}6&0\\0&0\end{bmatrix}P^{-1}=B.$$

Veamos otra aplicación de los resultados que probamos:

Ejercicio 4.F. Sea $A=\begin{bmatrix}0&\frac12&\frac12\\\frac12&0&\frac12\\\frac12&\frac12&0\end{bmatrix}$. Calcular A^k para $k\in\mathbb{N}$. Sea $q(x)=x^5+3x^2+1$, calcular $q(A)$.

Dem. Calculamos el polinomio característico de A :

$$\chi_A(x)=\det(A-xI)=\det(\begin{bmatrix}-x&\frac12&\frac12\\\frac12&-x&\frac12\\\frac12&\frac12&-x\end{bmatrix})=-x^3+\frac34x+\frac14.$$

Por lo tanto, los autovalores de A son $\lambda_1=1,\lambda_{2,3}=-\frac12$. Calculemos los autoespacios asociados:

$$\mathcal{S}_{\lambda_1}=\operatorname{nul}(A-I)=\operatorname{nul}(\begin{bmatrix}-1&\frac12&\frac12\\\frac12&-1&\frac12\\\frac12&\frac12&-1\end{bmatrix})=gen\{\begin{bmatrix}1\\1\\1\end{bmatrix}\}.$$

$$\mathcal{S}_{\lambda_{2,3}} = \text{nul}(A + \frac{1}{2}I) = \text{nul}(\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}) = gen\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\}.$$

Por lo tanto, como las multiplicidades geométricas y algebraicas de los autovalores coinciden A es diagonalizable. Sea $P := \begin{bmatrix} 1 & -1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ y $\Lambda := \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$. Entonces $P^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 2 \\ -1 & 2 & -1 \end{bmatrix}$ y

$$A = P\Lambda P^{-1}.$$

Entonces, por la Proposición 4.3.3,

$$\begin{aligned} A^k = P\Lambda^k P^{-1} &= P(\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix})^k P^{-1} = P\begin{bmatrix} 1^k & 0 & 0 \\ 0 & (-\frac{1}{2})^k & 0 \\ 0 & 0 & (-\frac{1}{2})^k \end{bmatrix} P^{-1} \\ &= \frac{1}{3}\begin{bmatrix} 1 & -1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & (-\frac{1}{2})^k & 0 \\ 0 & 0 & (-\frac{1}{2})^k \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 2 \\ -1 & 2 & -1 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} 1+2(-\frac{1}{2})^k & 1-(-\frac{1}{2})^k & 1-(-\frac{1}{2})^k \\ 1-(-\frac{1}{2})^k & 1+2(-\frac{1}{2})^k & 1-(-\frac{1}{2})^k \\ 1-(-\frac{1}{2})^k & 1-(-\frac{1}{2})^k & 1+2-\frac{1}{2})^k \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \frac{1}{3}(-\frac{1}{2})^k\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}. \end{aligned}$$

Finalmente, por la Proposición 4.3.3, como $q(1) = 5$ y $q(-\frac{1}{2}) = \frac{55}{32}$, tenemos que

$$\begin{aligned} q(A) = A^5 + 3A^2 + I = Pq(\Lambda)P^{-1} &= P\begin{bmatrix} q(1) & 0 & 0 \\ 0 & q(-\frac{1}{2}) & 0 \\ 0 & 0 & q(-\frac{1}{2}) \end{bmatrix} P^{-1} \\ &= P\begin{bmatrix} 5 & 0 & 0 \\ 0 & \frac{55}{32} & 0 \\ 0 & 0 & \frac{55}{32} \end{bmatrix} P^{-1}. \end{aligned}$$

■

4.4. Teorema espectral

Sea $A \in \mathbb{K}^{n\times n}$ entonces A es diagonalizable si y sólo si A admite descomposición espectral:

Teorema 4.4.1. *Sea $A \in \mathbb{K}^{n\times n}$ una matriz con espectro $\sigma(A) = \{\lambda_1, \lambda_2, \cdots, \lambda_r\}$ (los r autovalores distintos de A). Entonces A es diagonalizable si y sólo si existen matrices no nulas $G_1, G_2, \cdots, G_r \in \mathbb{K}^{n\times n}$ tales que*

$$A = \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_r G_r,$$

donde

1. *$G_i = [\Pi_i]_E^E$, es decir G_i es la matriz en base canónica (denotada E) de Π_i que es el proyector de $\mathbb{K}^n$ sobre $\mathcal{S}_{\lambda_i} = \text{nul}(A - \lambda_i I)$ en la dirección $\text{col}(A - \lambda_i I)$.*

2. *$G_i G_j = 0$ para $i \neq j$.*

3. *$G_1 + G_2 + \cdots + G_r = I_{n\times n}$.*

La descomposición $A = \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_r G_r$ se llama la *descomposición espectral* de A.

Dem. Como A es diagonalizable, entonces, por el Teorema 4.2.2,

$$\mathbb{K}^n = \mathcal{S}_{\lambda_1} \oplus \mathcal{S}_{\lambda_2} \oplus \cdots \oplus \mathcal{S}_{\lambda_r} = \mathrm{nul}(A - \lambda_1 I) \oplus \mathrm{nul}(A - \lambda_2 I) \oplus \cdots \oplus \mathrm{nul}(A - \lambda_r I)$$

con $\sigma(A) = \{\lambda_1, \lambda_2, \cdots, \lambda_r\}$ los r distintos autovalores de A. Para cada $i \in \{1, \cdots, r\}$, definimos la transformación lineal $\Pi_i : \mathbb{K}^n \to \mathbb{K}^n$ por

$$\Pi_i(v) = v_i$$

si $v = v_1 + v_2 + \cdots + v_r$, con $v_i \in \mathcal{S}_{\lambda_i}$, $i = 1, \cdots, r$. Entonces, para cada $v \in \mathbb{K}^n$, tenemos que $\Pi_i^2(v) = \Pi_i(\Pi_i(v)) = \Pi_i(v_i) = v_i = \Pi_i(v)$. Por lo tanto $\Pi_i^2 = \Pi_i$ y Π_i es un proyector de $\mathbb{K}^n$ sobre $\mathrm{Im}(\Pi_i) = \mathcal{S}_{\lambda_i} = \mathrm{nul}(A - \lambda_i I)$ en la dirección $\mathrm{Nu}(\Pi_i) = \bigoplus_{\lambda \in \sigma(A),\ \lambda \neq \lambda_i} \mathrm{nul}(A - \lambda I) = \mathrm{col}(A - \lambda_i I)$, donde usamos el Teorema 4.2.1. Si $i \neq j$, entonces

$$\Pi_i \circ \Pi_j(v) = \Pi_i(\Pi_j(v)) = \Pi_i(v_j) = 0,$$

entonces $\Pi_i \circ \Pi_j = 0$ si $i \neq j$. Además, si $v \in \mathbb{K}^n$, existen únicos $v_i \in \mathcal{S}_{\lambda_i}$, $i = 1, \cdots, r$ tales que $v = v_1 + v_2 + \cdots + v_r$, entonces

$$v = id_{\mathbb{K}^n}(v) = \Pi_1(v) + \Pi_2(v) + \cdots \Pi_r(v) = (\Pi_1 + \Pi_2 + \cdots + \Pi_r)(v).$$

Por lo tanto, $id_{\mathbb{K}^n} = \Pi_1 + \Pi_2 + \cdots + \Pi_r$. Sea $G_i := [\Pi_i]_E^E$, es decir G_i es la matriz (no nula) en la base canónica de $\mathbb{K}^n$ (denotada E) de Π_i. Entonces, es claro que

$$\Pi_i(v) = G_i v = v_i, \text{ para todo } v \in \mathbb{K}^n$$

y además, se cumplen 1., 2. y 3.

Finalmente, si $v \in \mathbb{K}^n$ entonces $v = v_1 + v_2 + \cdots + v_r$, con $v_i \in \mathcal{S}_{\lambda_i}$, $i = 1, \cdots, r$. Por lo tanto

$$v = \Pi_1(v) + \Pi_2(v) + \cdots + \Pi_r(v) = G_1 v + G_2 v + \cdots + G_r v.$$

Entonces, usando que $Av_i = \lambda_i v_i = \lambda_i G_i v$ para $i = 1, \cdots, r$, se sigue que

$$Av = A(G_1 v + G_2 v + \cdots + G_r v) = \lambda_1 G_1 v + \lambda_2 G_2 v + \cdots + \lambda_r G_r v = (\lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_r G_r)v.$$

Por lo tanto,

$$A = \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_r G_r.$$

Recíprocamente, supongamos que existen $\lambda_1, \lambda_2, \cdots, \lambda_r \in \mathbb{K}$ distintos y matrices no nulas $G_1, \cdots, G_r$, que verifican 1., 2. y 3 tales que $A = \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_r G_r$. Veamos que A es diagonalizable. Sea $V_i := \mathrm{col}(G_i)$, por 1., 2. y 3., tenemos que

$$\mathbb{K}^n = V_1 \oplus V_2 \oplus \cdots \oplus V_r.$$

Además, si $v = v_1 + v_2 + \cdots + v_r$, con $v_i \in V_i$ entonces, como G_i es la matriz en base canónica de Π_i que es el proyector de $\mathbb{K}^n$ sobre $\mathcal{S}_{\lambda_i} = \mathrm{nul}(A - \lambda_i I)$ en la dirección $\mathrm{col}(A - \lambda_i I)$, se sigue que

$$G_i v = v_i \text{ para todo } i = 1, \cdots, r.$$

Como $G_i \neq 0$, tenemos que $V_i \neq \{0\}$ para todo $i = 1, \cdots, r$. Entonces, sea $v \in V_i$, tenemos que $G_i v = v$ y entonces

$$\begin{aligned} Av &= (\lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_r G_r)v = (\lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_r G_r)(G_i v) \\ &= \lambda_1 G_1 G_i v + \lambda_2 G_2 G_i v + \cdots + \lambda_r G_r G_i v = \lambda_i G_i^2 v = \lambda_i G_i v = \lambda_i v, \end{aligned}$$

donde usamos que $G_i G_j = 0$ para $i \neq j$ y que $G_i^2 = G_i$, para todo $i \in \{1, 2, \cdots, r\}$. Como $V_i \neq \{0\}$, lo anterior implica que λ_i es un autovalor de A y además, vale que $V_i \subseteq \mathcal{S}_{\lambda_i}$. Además, tenemos que

$$\mathbb{K}^n = V_1 \oplus V_2 \oplus \cdots \oplus V_r \subseteq \mathcal{S}_{\lambda_1} \oplus \mathcal{S}_{\lambda_2} \oplus \cdots \oplus \mathcal{S}_{\lambda_r} \subseteq \mathbb{K}^n.$$

Por lo tanto, $\mathcal{S}_{\lambda_1} \oplus \mathcal{S}_{\lambda_2} \oplus \cdots \oplus \mathcal{S}_{\lambda_r} = \mathbb{K}^n$ y como $\lambda_1, \lambda_2, \cdots, \lambda_r \in \mathbb{K}$ son los distintos autovalores de A por el Teorema 4.2.2, A es diagonalizable. ■

Corolario 4.4.2. *Sea $A \in \mathbb{K}^{n\times n}$ una matriz diagonalizable con descomposición espectral $A = \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_r G_r$. Entonces, para cada $n \in \mathbb{N}$,*

$$A^n = \lambda_1^n G_1 + \lambda_2^n G_2 + \cdots + \lambda_r^n G_r.$$

Más aún, si $q(x) = a_k x^k + \cdots + a_1 x + a_0 \in \mathbb{K}_k[x]$ entonces

$$q(A) = q(\lambda_1)G_1 + q(\lambda_2)G_2 + \cdots + q(\lambda_r)G_r.$$

Dem. Vamos a probar la propiedad para todo $n \in \mathbb{N}$ por inducción: Para $n = 1$ es claro que la propiedad vale.

Supongamos que para $n = k \in \mathbb{N}$ vale que $A^k = \lambda_1^k G_1 + \lambda_2^k G_2 + \cdots + \lambda_r^k G_r$ (HI). Entonces, multiplicamos esa ecuación a ambos lados por A y tenemos que

$$\begin{aligned} A^{k+1} &= A^k A = [\lambda_1^k G_1 + \lambda_2^k G_2 + \cdots + \lambda_r^k G_r](\lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_r G_r) = \\ &= \lambda_1^{k+1} G_1^2 + \lambda_1^k \lambda_2 G_1 G_2 + \cdots + \lambda_1^k \lambda_r G_1 G_r + \cdots + \lambda_r^k \lambda_1 G_r G_1 + \lambda_r^k \lambda_2 G_r G_2 + \cdots + \lambda_r^{k+1} G_r^2 \\ &= \lambda_1^{k+1} G_1 + \lambda_2^{k+1} G_2 + \cdots + \lambda_r^{k+1} G_r, \end{aligned}$$

donde usamos que las matrices $G_1, G_2, \cdots, G_r$ son idempotentes y que cumplen 1., 2. y 3. del Teorema 4.4.1.

Entonces, a partir de la hipótesis inductiva deducimos la validez de la propiedad para $n = k+1$. Por lo tanto, por inducción, concluimos que para cada $n \in \mathbb{N}$,

$$A^n = \lambda_1^n G_1 + \lambda_2^n G_2 + \cdots + \lambda_r^n G_r.$$

Más aún, si $q(x) = a_k x^k + \cdots + a_1 x + a_0$, entonces como para $k \in \mathbb{N}$ vale que $A^k = \lambda_1^k G_1 + \lambda_2^k G_2 + \cdots + \lambda_r^k G_r$ se sigue que

$$\begin{aligned} q(A) &= a_k A^k + a_{k-1} A^{k-1} + \cdots + a_1 A + a_0 I = a_k(\lambda_1^k G_1 + \lambda_2^k G_2 + \cdots + \lambda_r^k G_r) + \\ &+ \cdots + a_1(\lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_r G_r) + a_0(G_1 + G_2 + \cdots + G_r) \\ &= q(\lambda_1)G_1 + q(\lambda_2)G_2 + \cdots + q(\lambda_r)G_r. \end{aligned}$$

■

Veamos un ejemplo de aplicación del Teorema espectral:

Ejercicio 4.G. Sea $A = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}$, hallar (si existe) la descomposición espectral de A.

Dem. Calculamos el polinomio característico de A :

$$\chi_A(x) = \det(A - xI) = \det \begin{bmatrix} 4-x & 0 & 1 \\ 2 & 3-x & 2 \\ 1 & 0 & 4-x \end{bmatrix} = (3-x)[(4-x)^2 - 1].$$

Por lo tanto, los autovalores de A son $\lambda_{1,2} = 3, \lambda_3 = 5$. Calculemos los autoespacios asociados:

$$\mathcal{S}_{\lambda_{1,2}} = \text{nul}(A - 3I) = \text{nul}(\begin{bmatrix} 3 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix}) = gen\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\}.$$

$$\mathcal{S}_{\lambda_3} = \text{nul}(A - 5I) = \text{nul}(\begin{bmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\}.$$

Por lo tanto, como las multiplicidades geométricas y algebraicas de los autovalores coinciden, A es diagonalizable. Además, $B := \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$ es una base $\mathbb{R}^3$ compuesta por autovectores de A.

Para construir la descomposición espectral de A recordemos que G_1 es la matriz en base canónica del proyector Π_1 de $\mathbb{R}^3$ sobre $\mathcal{S}_{\lambda_{1,2}}$ en la dirección $\mathcal{S}_3$. Asimísmo, G_2 es la matriz en base canónica del proyector Π_2 de $\mathbb{R}^3$ sobre $\mathcal{S}_{\lambda_3}$ en la dirección $\mathcal{S}_{1,2}$. Es claro que $[\Pi_1]_B^B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ y $[\Pi_2]_B^B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Por lo tanto, usando que la matriz de cambio de base de B a E es $[M]_B^E = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix}$, tenemos que

$$
\begin{aligned}
G_1 &= [\Pi_1]_E^E = [M]_B^E[\Pi_1]_B^B[M]_E^B = [M]_B^E[\Pi_1]_B^B([M]_B^E)^{-1} \\
&= \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{1}{2} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 2 & -2 \\ 1 & 0 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 & -1 \\ -2 & 2 & -2 \\ -1 & 0 & 1 \end{bmatrix}.
\end{aligned}
$$

De manera similar, tenemos que

$$
\begin{aligned}
G_2 &= [\Pi_2]_E^E = [M]_B^E[\Pi_2]_B^B[M]_E^B = [M]_B^E[\Pi_1]_B^B([M]_B^E)^{-1} \\
&= \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \frac{1}{2} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 2 & -2 \\ 1 & 0 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix}.
\end{aligned}
$$

Notar que $G_1G_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $G_1^2 = G_1$, $G_2^2 = G_2$, $G_1 + G_2 = I_{3\times 3}$ y

$$A = 3G_1 + 5G_2.$$

■

Recordemos que si $(\mathbb{V}, \|\cdot\|)$ un $\mathbb{K}$-espacio normado y $(v_n)_{n\in\mathbb{N}}$ es una sucesión de elementos de $\mathbb{V}$. La sucesión $(v_n)_{n\in\mathbb{N}}$ es *convergente* si existe $v_0 \in \mathbb{V}$ tal que

$$\lim_{n\to\infty} \|v_n - v_0\| = 0.$$

En este caso, escribimos

$$\lim_{n\to\infty} v_n = v_0.$$

Recordemos que si en $\mathbb{K}^{n\times n}$ definimos la función $\langle\, \cdot, \cdot \,\rangle : \mathbb{K}^{n\times n} \times \mathbb{K}^{n\times n} \to \mathbb{K}$, por

$$\langle\, A, B \,\rangle := \operatorname{tr}(B^*A)$$

entonces $(\mathbb{K}^{n\times n}, \langle\, \cdot, \cdot \,\rangle)$ resulta un $\mathbb{K}$-espacio euclídeo. La norma inducida por dicho producto interno

$$\|A\|_F := \langle\, A, A \,\rangle^{1/2} = \operatorname{tr}(A^*A)^{1/2},$$

se denomina *norma de Frobenius*. Recordemos que como $\|A\|_F$ es una norma, entonces vale:

- $\|A\|_F \geq 0$ para toda $A \in \mathbb{K}^{n\times n}$ y $\|A\|_F = 0$ si y sólo si $A = 0_{n\times n}$ (la matriz nula de $n \times n$).
- $\|\alpha\ A\|_F = |\alpha|\|A\|_F$, para toda $A \in \mathbb{K}^{n\times n}$ y $\alpha \in \mathbb{K}$.
- $\|A + B\|_F \leq \|A\|_F + \|B\|_F$, para toda $A, B \in \mathbb{K}^{n\times n}$. De esta propiedad se deduce que $|\|A\|_F - \|B\|_F| \leq \|A - B\|_F$, para toda $A, B \in \mathbb{K}^{n\times n}$.

Sea $(A_n)_{n\in\mathbb{N}}$ una sucesión de matrices, diremos que la sucesión $(A_n)_{n\in\mathbb{N}}$ es *convergente* si existe $A_0 \in \mathbb{K}^{n\times n}$ tal que

$$\lim_{n\to\infty} \|A_n - A_0\|_F = 0. \tag{4.10}$$

En ese caso, escribimos $\lim_{n\to\infty} A_n = A_0$.

$$\text{Si } \lim_{n\to\infty} A_n = A_0 \text{ entonces } \lim_{n\to\infty} \|A_n\|_F = \|A_0\|_F. \tag{4.11}$$

De hecho, notar que para todo $n \in \mathbb{N}$, vale que

$$0 \leq |\|A_n\|_F - \|A_0\|_F| \leq \|A_n - A_0\|_F$$

(propiedad de la norma). Entonces, si $\lim_{n\to\infty} A_n = A_0$ vale que $\lim_{n\to\infty} \|A_n - A_0\|_F = 0$. Por lo tanto, $\lim_{n\to\infty} |\|A_n\|_F - \|A_0\|_F| = 0$ y entonces $\lim_{n\to\infty} \|A_n\|_F = \|A_0\|_F$.

Ejercicio 4.8. Sea $A = \begin{bmatrix} 0{,}7 & 0{,}2 & 0{,}1 \\ 0{,}3 & 0{,}5 & 0{,}2 \\ 0{,}2 & 0{,}4 & 0{,}4 \end{bmatrix}$.

a) Hallar el espectro de A y comprobar que es diagonalizable.

b) Usando la descomposición espectral de A comprobar que existe $M > 0$ tal que

$$\|A^n - G_1\|_F \leq M(\frac{3+\sqrt{2}}{10})^n,$$

donde G_1 es la matriz de la proyección sobre $\text{nul}(A - I)$ en la dirección $\text{col}(A - I)$.

Dem. $a)$: Calculamos el polinomio característico de A :

$$\chi_A(x) = \det(A - xI) = \det \begin{bmatrix} 0{,}7 - x & 0{,}2 & 0{,}1 \\ 0{,}3 & 0{,}5 - x & 0{,}2 \\ 0{,}2 & 0{,}4 & 0{,}4 - x \end{bmatrix} = -x^3 + 1{,}6x^2 - 0{,}67x + 0{,}07.$$

Por lo tanto, los autovalores de A son $\lambda_1 = 1, \lambda_2 = \frac{3+\sqrt{2}}{10}$ y $\lambda_3 = \frac{3-\sqrt{2}}{10}$. Entonces $\sigma(A) = \{1, \frac{3+\sqrt{2}}{10}, \frac{3-\sqrt{2}}{10}\}$. Como A tiene 3 autovalores distintos, por el Teorema 4.2.2, A resulta diagonalizable.

$b)$: Por el Teorema 4.4.1, existen matrices G_1, G_2, G_3 que cumplen 1., 2. y 3. de dicho teorema, tales que

$$A = \lambda_1 G_1 + \lambda_2 G_2 + \lambda_3 G_3 = G_1 + \frac{3+\sqrt{2}}{10} G_2 + \frac{3-\sqrt{2}}{10} G_3.$$

Entonces, por el Corolario 4.4.2,

$$A^n = G_1 + (\frac{3+\sqrt{2}}{10})^n G_2 + (\frac{3-\sqrt{2}}{10})^n G_3.$$

Por lo tanto, usando que $(\frac{3-\sqrt{2}}{10})^n < (\frac{3+\sqrt{2}}{10})^n$ y las propiedades de la norma, se sigue que, para todo $n \in \mathbb{N}$,

$$\begin{aligned}0 \leq \|A^n - G_1\|_F = \|(\frac{3+\sqrt{2}}{10})^n G_2 + (\frac{3-\sqrt{2}}{10})^n G_3\|_F &\leq (\frac{3+\sqrt{2}}{10})^n \|G_2\|_F + (\frac{3-\sqrt{2}}{10})^n \|G_3\|_F \\ &\leq (\frac{3+\sqrt{2}}{10})^n \|G_2\|_F + (\frac{3+\sqrt{2}}{10})^n \|G_3\|_F = (\|G_2\|_F + \|G_3\|_F)(\frac{3+\sqrt{2}}{10})^n \\ &= M(\frac{3+\sqrt{2}}{10})^n,\end{aligned}$$

Donde $M := (\|G_2\|_F + \|G_3\|_F) > 0$. Entonces, como $\frac{3+\sqrt{2}}{10} < 1$, se sigue que

$$\lim_{n\to\infty} M(\frac{3+\sqrt{2}}{10})^n = M \lim_{n\to\infty} (\frac{3+\sqrt{2}}{10})^n = 0.$$

Entonces, como para todo $n \in \mathbb{N}$, vale que

$$0 \leq \|A^n - G_1\|_F \leq M(\frac{3+\sqrt{2}}{10})^n$$

y $\lim_{n\to\infty} M(\frac{3+\sqrt{2}}{10})^n = 0$, se sigue que $\lim_{n\to\infty} \|A^n - G_1\|_F = 0$ y por lo tanto $\lim_{n\to\infty} A^n = G_1$. ■

Ejercicio 4.9. Sea $A = \frac{1}{2}\begin{bmatrix} 7 & 6 & 6 \\ -2 & 0 & -2 \\ -1 & -2 & 0 \end{bmatrix} = \begin{bmatrix} \frac{7}{2} & 3 & 3 \\ -1 & 0 & -1 \\ -\frac{1}{2} & -1 & 0 \end{bmatrix}$.

a) Probar que $\lim_{n\to\infty} \frac{\|A^n\|_F}{2^n} > 0$ y concluir que no existe $\lim_{n\to\infty} A^n$.

b) Probar que el conjunto $\{x \in \mathbb{R}^3 : \lim_{n\to\infty} A^n x = 0\}$ es un subespacio de $\mathbb{R}^3$ y hallar una base del mismo.

c) Probar que el conjunto $\mathcal{S} := \{x \in \mathbb{R}^3 :$ existe $\lim_{n\to\infty} A^n x\}$ es un subespacio de $\mathbb{R}^3$ y hallar una base del mismo.

d) Comprobar que $[0\ -2\ 2]^T \in \{\lim_{n\to\infty} A^n x : x \in \mathcal{S}\}$ y hallar todas las soluciones de la ecuación $\lim_{n\to\infty} A^n x = [0\ -2\ 2]^T$.

Dem. $a)$: Calculamos el polinomio característico de A :

$$\chi_A(x) = \det(A - xI) = \det(\frac{1}{2}\begin{bmatrix} 7-x & 6 & 6 \\ -2 & -x & -2 \\ -1 & -2 & -x \end{bmatrix}) = -x^3 + \frac{7}{2}x^2 - \frac{7}{2}x + 1.$$

Por lo tanto, los autovalores de A son $\lambda_1 = 2, \lambda_2 = 1$ y $\lambda_3 = \frac{1}{2}$. Como A tiene 3 autovalores distintos por el Teorema 4.2.2, A resulta diagonalizable. Por el Teorema 4.4.1, existen matrices G_1, G_2, G_3 que cumplen 1., 2. y 3. de dicho teorema, tales que

$$A = \lambda_1 G_1 + \lambda_2 G_2 + \lambda_3 G_3 = 2G_1 + 1G_2 + \frac{1}{2}G_3.$$

Entonces, por el Corolario 4.4.2,

$$A^n = 2^n G_1 + G_2 + (\frac{1}{2})^n G_3.$$

Entonces
$$\frac{A^n}{2^n} = \frac{2^nG_1 + G_2 + (\frac{1}{2})^nG_3}{2^n} = G_1 + \frac{1}{2^n}G_2 + \frac{1}{2^{2n}}G_3.$$
Observar que, para todo $n \in \mathbb{N}$, tenemos que
$$0 \leq \|\frac{A^n}{2^n} - G_1\|_F = \|\frac{1}{2^n}G_2 + \frac{1}{2^{2n}}G_3\|_F \leq \frac{1}{2^n}\|G_2\|_F + \frac{1}{2^{2n}}\|G_3\|_F.$$
Entonces como
$$\lim_{n\to\infty}(\frac{1}{2^n}\|G_2\|_F + \frac{1}{2^{2n}}\|G_3\|_F) = 0,$$
se sigue que $\lim\limits_{n\to\infty}\|\frac{A^n}{2^n} - G_1\|_F = 0$ y por lo tanto $\lim\limits_{n\to\infty}\frac{A^n}{2^n} = G_1$. Entonces, por (4.11),
$$\lim_{n\to\infty}\frac{\|A^n\|_F}{2^n} = \lim_{n\to\infty}\|\frac{A^n}{2^n}\|_F = \|\lim_{n\to\infty}\frac{A^n}{2^n}\|_F = \|G_1\|_F > 0.$$
Esto implica que la sucesión $\|A^n\|_F$ no está acotada en $\mathbb{R}$ y, por lo tanto, no existe $\lim\limits_{n\to\infty} A^n$.

b) : Llamemos $\mathcal{T} := \{x \in \mathbb{R}^3 : \lim\limits_{n\to\infty} A^nx = 0\}$. Veamos que $\mathcal{T}$ es un subespacio de $\mathbb{R}^3$. Claramente $0 \in \mathcal{T}$, pues $A^n0 = 0$ para todo n y por lo tanto $\lim\limits_{n\to\infty} A^n0 = 0$. Por otra parte, supongamos que $x, y \in \mathcal{T}$ y sea $\alpha \in \mathbb{R}$. Entonces, como $x \in \mathcal{T}$ tenemos que $\lim\limits_{n\to\infty} A^nx = 0$ y como $y \in \mathcal{T}$ tenemos que $\lim\limits_{n\to\infty} A^ny = 0$. Por lo tanto, como $A^n(\alpha x + y) = \alpha A^nx + A^ny$, por el álgebra de límites, se sigue que $\lim\limits_{n\to\infty} A^n(\alpha x + y) = \alpha \lim\limits_{n\to\infty} A^nx + \lim\limits_{n\to\infty} A^ny = \alpha 0 + 0 = 0$. Entonces $\alpha x + y \in \mathcal{T}$ y concluimos que $\mathcal{T}$ es un subespacio.

Calculemos los autoespacios asociados a los autovalores de A :
$$\mathcal{S}_{\lambda_1} = \mathrm{nul}(A - 2I) = \mathrm{nul}(\begin{bmatrix} \frac{3}{2} & 3 & 3 \\ -1 & -2 & -1 \\ -\frac{1}{2} & -1 & -2 \end{bmatrix}) = gen\{\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}\}.$$
$$\mathcal{S}_{\lambda_2} = \mathrm{nul}(A - I) = \mathrm{nul}(\begin{bmatrix} \frac{5}{2} & 3 & 3 \\ -1 & -1 & -1 \\ -\frac{1}{2} & -1 & -1 \end{bmatrix}) = gen\{\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\}.$$
$$\mathcal{S}_{\lambda_3} = \mathrm{nul}(A - \frac{1}{2}I) = \mathrm{nul}(\begin{bmatrix} 3 & 3 & 3 \\ -1 & -\frac{1}{2} & -1 \\ -\frac{1}{2} & -1 & -\frac{1}{2} \end{bmatrix}) = gen\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\}.$$

Como A es diagonalizable $B := \{\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\}$ es una base de autovectores $\mathbb{R}^3$.

Afirmamos que $\mathcal{T} = gen\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\}$. Es claro que $gen\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\} \subseteq \mathcal{T}$, porque si $x \in gen\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\}$ entonces $x = \alpha \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ con $\alpha \in \mathbb{R}$. Entonces $A^nx = A^n\alpha \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \alpha A^n \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \alpha\frac{1}{2^n}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ y vale que
$$\lim_{n\to\infty} A^nx = \lim_{n\to\infty} \alpha\frac{1}{2^n}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Por otra parte, si $x \in \mathcal{T}$ entonces como $x \in \mathbb{R}^3$, tenemos que $x = \alpha \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, con $\alpha, \beta, \gamma \in \mathbb{R}$. Entonces

$$A^n x = \alpha A^n \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + \beta A^n \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \gamma A^n \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \alpha 2^n \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \gamma \frac{1}{2^n} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

que converge al vector $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ si y sólo si $\alpha = \beta = 0$. Por lo tanto $\mathcal{T} \subseteq gen\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \}$. Entonces $\mathcal{T} = gen\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \}$ y una base de $\mathcal{T}$ puede ser $B_{\mathcal{T}} = \{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \}$.

c) : Claramente $0 \in \mathcal{S}$, pues $A^n 0 = 0$ para todo n y como $\lim_{n\to\infty} A^n 0 = 0$, entonces existe $\lim_{n\to\infty} A^n 0$. Por otra parte, supongamos que $x, y \in \mathcal{S}$ y sea $\alpha \in \mathbb{R}$. Entonces, como $x \in \mathcal{S}$ existe $\lim_{n\to\infty} A^n x$ y como $y \in \mathcal{S}$, existe $\lim_{n\to\infty} A^n y$. Por lo tanto, como $A^n(\alpha x + y) = \alpha A^n x + A^n y$, por el álgebra de límites, se sigue que existe $\lim_{n\to\infty} A^n(\alpha x + y)$. Más aún,

$$\lim_{n\to\infty} A^n(\alpha x + y) = \alpha \lim_{n\to\infty} A^n x + \lim_{n\to\infty} A^n y.$$

Entonces $\alpha x + y \in \mathcal{S}$ y concluimos que $\mathcal{S}$ es un subespacio.

Afirmamos que $\mathcal{S} = gen\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \}$. Es claro que $gen\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \} \subseteq \mathcal{S}$, porque si $x \in gen\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \}$ entonces $x = \alpha \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ con $\alpha, \beta \in \mathbb{R}$. Entonces

$$\begin{aligned} A^n x &= A^n(\alpha \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}) = \alpha A^n \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \beta A^n \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \\ &= \alpha \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \beta \frac{1}{2^n} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \underset{n\to\infty}{\to} \alpha \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \end{aligned}$$

y concluimos que $A^n x$ converge. Por otra parte, si $x \in \mathcal{S}$ entonces como $x \in \mathbb{R}^3$, tenemos que $x = \alpha \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, con $\alpha, \beta, \gamma \in \mathbb{R}$. Entonces

$$A^n x = \alpha A^n \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + \beta A^n \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \gamma A^n \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \alpha 2^n \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \gamma \frac{1}{2^n} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

que converge (el límite existe) si y sólo si $\alpha = 0$. Por lo tanto $\mathcal{S} \subseteq gen\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \}$.

Entonces $\mathcal{S} = gen\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \}$ y una base de $\mathcal{S}$ puede ser $B_{\mathcal{S}} = \{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \}$.

d) : Vamos a describir el conjunto $\{\lim_{n\to\infty} A^n x : x \in \mathcal{S}\}$. Sea $x \in \mathcal{S}$ (es decir x es tal que existe $\lim_{n\to\infty} A^n x$) entonces $x = \alpha \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ con $\alpha, \beta \in \mathbb{R}$. Entonces

$$\begin{aligned} A^n x = A^n(\alpha \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}) &= \alpha A^n \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \beta A^n \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \\ &= \alpha \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \beta \frac{1}{2^n} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}. \end{aligned}$$

Por lo tanto

$$\lim_{n\to\infty} A^n x = \lim_{n\to\infty} (\alpha \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \beta \frac{1}{2^n} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}) = \alpha \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

Entonces $\{\lim_{n\to\infty} A^n x : x \in \mathcal{S}\} = gen\{\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\}$ y es claro que $\begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix} \in \{\lim_{n\to\infty} A^n x : x \in \mathcal{S}\}$.

Ahora busquemos las soluciones de la ecuación $\lim_{n\to\infty} A^n x = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix}$. Observar que la ecuación anterior es lineal, por lo tanto (si existe solución) todas las soluciones son de la forma $x_s = x_p + x_h$, donde x_p es una solución particular, es decir $\lim_{n\to\infty} A^n x_p = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix}$ y x_h son las soluciones del sistema homogéneo, es decir $\lim_{n\to\infty} A^n x_h = 0$. Por ejemplo, si $x_p := \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$ entonces $\lim_{n\to\infty} A^n x_p = 2 \lim_{n\to\infty} A^n \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = 2 \lim_{n\to\infty} 1 \cdot \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix}$, entonces x_p es una solución particular. Tal como vimos en *b*) el conjunto $\mathcal{T} = \{x \in \mathbb{R}^3 : \lim_{n\to\infty} A^n x = 0\} = gen\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\}$, entonces $x_h \in gen\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\}$.

Conclusión, todas las soluciones de la ecuación $\lim_{n\to\infty} A^n x = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix}$ son

$$x_s = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix} + a \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \text{con } a \in \mathbb{R}.$$

■

4.5. Formas de Jordan

Cuando una matriz $A \in \mathbb{K}^{n\times n}$ no es diagonalizable, es necesario recurrir a lo que se denominan *formas de Jordan.*

A continuación, vamos a demostrar cómo son los bloques de Jordan J para el caso $A \in \mathbb{K}^{3\times 3}$ y cómo obtener una matriz $P \in \mathbb{K}^{3\times 3}$ inversible tal que

$$A = PJP^{-1}.$$

Pensando de manera similar, podemos deducir cómo sería la forma de Jordan para $A \in \mathbb{K}^{2\times 2}$.

Teorema 4.5.1. *Sea $A \in \mathbb{R}^{3\times 3}$ una matriz no diagonalizable. Entonces existe una matriz inversible P tal que $P^{-1}AP = J$, donde J tiene alguna de las siguientes formas:*

$$\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{bmatrix},$$

con $\lambda, \mu \in \mathbb{R}$ y $\lambda \neq \mu$.

Dem. Si $A \in \mathbb{R}^{3\times 3}$ es no diagonalizable, es porque tenemos alguno de los siguientes 3 casos:

1. λ es un autovalor triple de A con multiplicidad geométrica 1.
2. λ es un autovalor triple de A con multiplicidad geométrica 2.
3. λ es un autovalor doble de A con multiplicidad geométrica 1 y μ es un autovalor simple de A.

Caso 1: λ es un autovalor triple con multiplicidad geométrica 1. En este caso, vamos a tomar

$$J := \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}.$$

Respecto de la matriz $P = [v_1\ v_2\ v_3]$, (donde v_1, v_2, v_3 denotan sus columnas) P debe ser inversible y como queremos que $A = PJP^{-1}$, multiplicando a izquierda por P a ambos lados de la ecuación anterior, tenemos que se debe cumplir que

$$AP = PJ.$$

Por un lado, P es inversible si y sólo si $\{v_1, v_2, v_3\}$ es linealmente independiente. Por otra parte, como

$$AP = A[v_1\ v_2\ v_3] = [Av_1\ Av_2\ Av_3] \text{ y}$$

$$PJ = [v_1\ v_2\ v_3] \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} = [\lambda v_1\ v_1 + \lambda v_2\ v_2 + \lambda v_3].$$

Tenemos que $AP = PJ$ si y sólo si $Av_1 = \lambda v_1$, $Av_2 = v_1 + \lambda v_2$ y $Av_3 = v_2 + \lambda v_3$. Es decir,

$$(A - \lambda I)v_1 = 0,\ (A - \lambda I)v_2 = v_1 \text{ y } (A - \lambda I)v_3 = v_2.$$

En resumen, la matriz P se obtiene buscando 3 vectores $v_1, v_2, v_3 \in \mathbb{R}^3$ linealmente independientes tales que $v_1 \in \mathcal{S}_\lambda \cap \operatorname{col}(A - \lambda I)$, $(A - \lambda I)v_2 = v_1$ y $v_2 \in \operatorname{col}(A - \lambda I)$ y $(A - \lambda I)v_3 = v_2$.

Caso 2: λ es un autovalor triple con multiplicidad geométrica 2. En este caso, vamos a tomar

$$J := \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}.$$

Respecto de la matriz $P = [v_1\ v_2\ v_3]$, (donde v_1, v_2, v_3 denotan sus columnas) P debe ser inversible y se debe cumplir que $AP = PJ$. Por un lado, P es inversible si y sólo si $\{v_1, v_2, v_3\}$ es linealmente independiente. Por otra parte, como

$$AP = A[v_1\ v_2\ v_3] = [Av_1\ Av_2\ Av_3] \text{ y}$$

$$PJ = [v_1\ v_2\ v_3] \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} = [\lambda v_1\ v_1 + \lambda v_2\ \lambda v_3].$$

Tenemos que $AP = PJ$ si y sólo si $Av_1 = \lambda v_1$, $Av_2 = v_1 + \lambda v_2$ y $Av_3 = \lambda v_3$. Es decir,

$$(A - \lambda I)v_1 = 0, \; (A - \lambda I)v_2 = v_1 \text{ y } (A - \lambda I)v_3 = 0.$$

En resumen, la matriz P se obtiene buscando 3 vectores $v_1, v_2, v_3 \in \mathbb{R}^3$ linealmente independientes tales que $v_1 \in \mathcal{S}_\lambda \cap \operatorname{col}(A - \lambda I)$, v_3 es un autovector asociado a λ y $(A - \lambda I)v_2 = v_1$.

Caso 3: λ es un autovalor doble con multiplicidad geométrica 1 y μ es un autovalor simple.

En este caso, vamos a tomar

$$J := \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{bmatrix}.$$

Respecto de la matriz $P = [v_1 \; v_2 \; v_3]$, (donde v_1, v_2, v_3 denotan sus columnas) P debe ser inversible y se debe cumplir que $AP = PJ$. Por un lado, P es inversible si y sólo si $\{v_1, v_2, v_3\}$ es linealmente independiente. Por otra parte, como

$$AP = A[v_1 \; v_2 \; v_3] = [Av_1 \; Av_2 \; Av_3] \text{ y}$$

$$PJ = [v_1 \; v_2 \; v_3] \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{bmatrix} = [\lambda v_1 \; v_1 + \lambda v_2 \; \mu v_3].$$

Tenemos que $AP = PJ$ si y sólo si $Av_1 = \lambda v_1$, $Av_2 = v_1 + \lambda v_2$ y $Av_3 = \mu v_3$. Es decir,

$$(A - \lambda I)v_1 = 0, \; (A - \lambda I)v_2 = v_1 \text{ y } (A - \mu I)v_3 = 0.$$

En resumen, la matriz P se obtiene buscando 3 vectores $v_1, v_2, v_3 \in \mathbb{R}^3$ linealmente independientes tales que $v_1 \in \mathcal{S}_\lambda \cap \operatorname{col}(A - \lambda I)$, v_3 es un autovector asociado a μ y $(A - \lambda I)v_2 = v_1$. ■

A partir del Teorema 4.5.1 podemos probar que la traza de una matriz cuadrada es la suma de sus autovalores:

Sea $A \in \mathbb{K}^{n\times n}$ y sean $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$ los autovalores de A. Entonces

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n,$$

es decir $\operatorname{tr}(A)$ es la suma de los autovalores de A.

Vamos a probarlo por casos:

- Si A es diagonalizable, entonces $A = P\Lambda P^{-1}$ para ciertos $P \in \mathbb{K}^{n\times n}$ inversible y $\Lambda = \operatorname{diag}(\lambda_1, \cdots, \lambda_n) \in \mathbb{K}^{n\times n}$ (matriz diagonal que tiene los autovalores de A es su diagonal). Entonces
$$\operatorname{tr}(A) = \operatorname{tr}(P\Lambda P^{-1}) = \operatorname{tr}(P^{-1}P\Lambda) = \operatorname{tr}(\Lambda) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$
y probamos lo que queríamos.

- Si $A \in \mathbb{K}^{3\times 3}$ y A no es diagonalizable, por el **Teorema 4.5.1**, existe una matriz inversible P tal que $P^{-1}AP = J$, donde J tiene alguna de las siguientes formas:
$$J_1 = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad J_2 = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}, \quad J_3 = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{bmatrix},$$
con $\lambda, \mu \in \mathbb{R}$ y $\lambda \neq \mu$. Entonces
$$\operatorname{tr}(A) = \operatorname{tr}(PJ_iP^{-1}) = \operatorname{tr}(P^{-1}PJ_i) = \operatorname{tr}(J_i) = \lambda + \lambda + \lambda,$$
para el caso $i = 1, 2$ ó
$$\operatorname{tr}(A) = \operatorname{tr}(PJ_3P^{-1}) = \operatorname{tr}(P^{-1}PJ_3) = \operatorname{tr}(J_3) = \lambda + \lambda + \mu,$$
y concluimos que $\operatorname{tr}(A)$ es la suma de los autovalores de A.

- Si $A \in \mathbb{K}^{n\times n}$ y A no es diagonalizable, de manera similar a como hicimos en el **Teorema 4.5.1**, se puede demostrar que existe una matriz inversible P tal que $P^{-1}AP = J$, y J es una matriz que no es diagonal pero tiene los autovalores de A en su diagonal. En este caso, tal como hicimos en el caso 3×3, también se puede demostrar que $\mathrm{tr}(A)$ es la suma de los autovalores de A.

El siguiente ejercicio similar al **Ejercicio 4.10** es un ejemplo de cálculo de la forma de Jordan de una matriz de 3×3.

Ejercicio 4.H. Encontrar los autovalores de la siguiente matriz, sus multiplicidades algebraicas y geométricas. En caso de que A no sea diagonalizable, hallar P inversible tal que $P^{-1}AP = J$, donde J tiene alguna de las formas del Teorema 4.5.1.

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

Dem. Calculamos el polinomio característico de A :

$$\begin{aligned} \chi_A(x) = \det(A - xI) = \det(&\begin{bmatrix} 3-x & 1 & -1 \\ 0 & 2-x & 0 \\ 1 & 1 & 1-x \end{bmatrix}) \\ = (2-x)&[(3-x)(1-x)+1] \\ = (2-x)&(x^2-4x+4) = (2-x)^3. \end{aligned}$$

Los autovalores de A son las raíces de su polinomio característico. En este caso, la única raíz del polinomio característico es $\lambda_{1,2,3} = 2$. La multiplicidad algebraica del autovalor $\lambda = 2$ es entonces 3. Veamos cuál es su multiplicidad geométrica calculando el autoespacio asociado a $\lambda = 2$:

$$\mathrm{nul}(A-2I) = \mathrm{nul}(\begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & -1 \end{bmatrix}) = \mathrm{nul}(\begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\}.$$

Por lo tanto la multiplicidad geométrica de $\lambda = 2$ es 2 y A no es diagonalizable. Estamos en el **Caso 2** del **Teorema 4.5.1**. En este caso,

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

Si $P := [v_1 \; v_2 \; v_3]$ es inversible, entonces $v_1, v_2, v_3 \in \mathbb{R}^3$ (las columnas de P) son vectores linealmente independientes tales que $v_1 \in \mathcal{S}_{\lambda=2} \cap \mathrm{col}(A-2I)$, v_3 es un autovector asociado a $\lambda = 2$ y $(A-2I)v_2 = v_1$. Como $A - 2I = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & -1 \end{bmatrix}$, tomamos $v_1 := \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \in \mathcal{S}_{\lambda=2} \cap \mathrm{col}(A-2I)$ y $v_3 := \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \in \mathcal{S}_{\lambda=2}$ (que son vectores claramente linealmente independientes). Entonces, buscamos $v_2 \in \mathbb{R}^3$ linealmente independiente con v_1 y v_3 tal que

$$\begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & -1 \end{bmatrix} v_2 = (A-2I)v_2 = v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Si $v_2 = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$, nos queda el siguiente sistema no homogéneo a resolver:

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

La solución de dicho sistema es:

$$v_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix},$$

con $\alpha, \beta \in \mathbb{R}$. Como sólo necesitamos un vector $v_2 \in \mathbb{R}^3$, tomamos $\alpha = \beta = 0$ y $v_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ que es li con v_1 y v_3 (verificarlo). Entonces $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}$ y tenemos que $A = PJP^{-1}$. ■

4.6. Matrices semejantes

Definición. Sean $A, B \in \mathbb{K}^{n \times n}$. Diremos que A y B son *semejantes*, y lo notamos $A \sim B$, si existe una matriz $P \in \mathbb{K}^{n \times n}$ inversible tal que

$$A = PBP^{-1}.$$

Sean $A, B \in \mathbb{K}^{n \times n}$. Como $A = IAI$ e I es inversible, se sigue que

$$A \sim A.$$

Si $A \sim B$, entonces $A = PBP^{-1}$, para cierta P inversible. Entonces

$$P^{-1}AP = P^{-1}PBP^{-1}P = B,$$

y entonces, como P^{-1} también es una matriz inversible, se sigue que $B \sim A$.

Finalmente, si $A \sim B$ y $B \sim C$, tenemos que $A = PBP^{-1}$, para cierta P inversible y $B = QCQ^{-1}$ para cierta Q inversible. Entoces $A = PBP^{-1} = PQCQ^{-1}P^{-1}$. Sea $W := PQ$, entonces W es inversible pues $\det(W) = \det(P)\det(Q) \neq 0$ y además, $W^{-1} = Q^{-1}P^{-1}$. Entonces, tenemos que

$$A = WCW^{-1},$$

por lo tanto $A \sim C$.

Como se cumplen las tres propiedades que vimos, se dice que la relación $\sim$ es *reflexiva*, *simétrica* y *transitiva*.

La siguiente observación es inmediata:

Sea $A \in \mathbb{K}^{n \times n}$, entonces A es **diagonalizable** si y sólo si existe $\Lambda \in \mathbb{K}^{n \times n}$ diagonal, tal que

$$A \sim \Lambda.$$

Proposición 4.6.1. *Sean $A, B \in \mathbb{K}^{n \times n}$ tales que $A \sim B$. Entonces:*

1. $\det(A) = \det(B)$,
2. $\operatorname{tr}(A) = \operatorname{tr}(B)$,
3. $\chi_A = \chi_B$,

4. *A y B tienen los mismos autovalores y además*

$$m^A(\lambda) = m^B(\lambda) \text{ y } \mu^A(\lambda) = \mu^B(\lambda),$$

para todo λ autovalor de A (y de B).

Dem. Si $A \sim B$, entonces $A = PBP^{-1}$ para cierta $P \in \mathbb{K}^{n \times n}$ inversible. Entonces:

1. : $\det(A) = \det(PBP^{-1}) = \det(P)\det(B)\det(P^{-1}) = \det(P)\det(B)\det(P)^{-1} = \det(B)$.

2. : $\operatorname{tr}(A) = \operatorname{tr}(PBP^{-1}) = \operatorname{tr}(P^{-1}PB) = \operatorname{tr}(IB) = \operatorname{tr}(B)$.

3. : $\chi_A(x) = \det(A - xI) = \det(PBP^{-1} - xI) = \det(P(B - xI)P^{-1}) = \det(P)\det(P^{-1})\det(B - xI) = \det(B - xI) = \chi_B(x)$, para todo $x \in \mathbb{K}$.

4. : Por el item 3., como los polinomios característicos de A y B coinciden, entonces tienen las mismas raíces con las mismas multiplicidades. Entonces, λ es un autovalor de A con cierta multiplicidad algebraica si y sólo si λ es un autovalor de B con la misma multiplicidad algebraica.

Por otra parte, observar que como P es una matriz inversible, vale que (hacer la cuenta)

$$\operatorname{col}(PBP^{-1} - \lambda I) = \operatorname{col}(P(B - \lambda I)P^{-1}) = \operatorname{col}(P(B - \lambda I)) \text{ y}$$

$$\operatorname{nul}(P(B - \lambda I)) = \operatorname{nul}(B - \lambda I).$$

Supongamos que λ es un autovalor de A. Entonces, por un lado λ es también un autovalor de B y además, usando el Teorema de la dimensión, tenemos que

$$\begin{aligned}\mu^A(\lambda) &= \dim(\operatorname{nul}(A - \lambda I)) = \dim(\operatorname{nul}(PBP^{-1} - \lambda I)) = n - \dim(\operatorname{col}(PBP^{-1} - \lambda I)) \\ &= n - \dim(\operatorname{col}(P(B - \lambda I)) = \dim(\operatorname{nul}(P(B - \lambda I)) = \dim(\operatorname{nul}(B - \lambda I)) = \mu^B(\lambda).\end{aligned}$$

■

El siguiente ejemplo muestra que **la vuelta de la Proposición 4.6.1, en general es falsa**.

Ejemplo 4.a. Sean

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ y } B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Observar que, $\det(A) = 0 = \det(B)$, $\operatorname{tr}(A) = 0 = \operatorname{tr}(B)$, $\chi_A(x) = x^4 = \chi_B(x)$. Entonces, $\lambda = 0$ es el único autovalor de A y B y $m^A(\lambda = 0) = 4 = m^B(\lambda = 0)$. Por otra parte,

$$\mu^A(\lambda = 0) = \dim(\operatorname{nul}(A)) = \dim(\operatorname{nul}(\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix})) = 2 \text{ y}$$

$$\mu^B(\lambda = 0) = \dim(\operatorname{nul}(B)) = \dim(\operatorname{nul}(\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix})) = 2.$$

Por lo tanto, tenemos que los ítems $1., 2., 3.$ y 4. de la Proposición 4.6.1 se cumplen.

Observar que $A^2 = 0$. Pero $B^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Entonces, $B \not\sim A$, porque si existiera P inversible tal que $B = PAP^{-1}$, tendríamos que

$$0 \neq B^2 = (PAP^{-1})PAP^{-1} = PA^2P^{-1} = P0P^{-1} = 0,$$

lo cual es absurdo. Por lo tanto **la vuelta de la Proposición anterior, en general es falsa**.

Ejercicio 4.I. Determinar cuáles de las siguientes parejas de matrices son semejantes:

a) $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

b) $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$.

c) $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

Dem. *a)* : Observar que $\text{tr}(A) = 2 \neq 0 = \text{tr}(B)$. Entonces, por la Proposición 4.6.1, A no puede ser semejante a B.

b) : Observar que $\det(A) = -1 \neq -\frac{1}{4} = \det(B)$. Entonces, por la Proposición 4.6.1, A no puede ser semejante a B.

c) : Observar que $\chi_A(x) = \det(A - xI) = (1-x)(2-x)(3-x)$. Entonces, los autovalores de A son $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$. Por lo tanto, como los tres autovalores de A son distintos, A es diagonalizable. Si $\Lambda := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = B$, tenemos que $A \sim \Lambda = B$. ■

El siguiente resultado que enunciaremos para $\mathbb{R}^{3\times 3}$ y $\mathbb{R}^{2\times 2}$ vale en general para $\mathbb{R}^{n\times n}$.

Teorema 4.6.2. *Sean $A, B \in \mathbb{R}^{n\times n}$ con $n \leq 3$. Entonces $A \sim B$ si y sólo si tienen los mismos bloques de Jordan.*

Dem. Como A y B tienen los mismos bloques de Jordan, por el Teorema 4.5.1, existen $P, Q \in \mathbb{R}^{n\times n}$ inversibles tales que $A = PJP^{-1}$ y $B = QJQ^{-1}$. Entonces $A \sim J$ y $B \sim J$. Entonces, como la relación $\sim$ es transitiva y simétrica, vale que $A \sim B$. ■

Veamos una aplicación del Teorema 4.6.2:

Ejercicio 4.11. Determinar cuáles de las siguientes parejas de matrices A_1 y A_2 son semejantes, y en caso de serlo hallar S inversible tal que $A_1 = SA_2S^{-1}$.

a) $A_1 = \frac{1}{4}\begin{bmatrix} 6 & 8 & 6 \\ -5 & 22 & 9 \\ 4 & -12 & 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & 2 & \frac{3}{2} \\ -\frac{5}{4} & \frac{11}{2} & \frac{9}{4} \\ 1 & -3 & 0 \end{bmatrix}$, $A_2 = \frac{1}{4}\begin{bmatrix} 18 & -13 & -57 \\ -2 & 13 & 9 \\ 2 & -3 & -3 \end{bmatrix} = \begin{bmatrix} \frac{9}{2} & -\frac{13}{4} & -\frac{57}{4} \\ -\frac{1}{2} & \frac{13}{4} & \frac{9}{4} \\ \frac{1}{2} & -\frac{3}{4} & -\frac{3}{4} \end{bmatrix}$.

b) $A_1 = \frac{1}{4}\begin{bmatrix} 5 & 6 & 3 \\ -6 & 10 & 6 \\ 5 & -10 & 3 \end{bmatrix}$, $A_2 = \frac{1}{4}\begin{bmatrix} 6 & -5 & 4 \\ 8 & 22 & -12 \\ 6 & 9 & 0 \end{bmatrix}$.

Dem. *a)* : Calculemos el polinomio característico de A_1 :

$$\begin{aligned}\chi_{A_1}(x) = \det(A_1 - xI) &= \det \begin{bmatrix} \frac{3}{2} - x & 2 & \frac{3}{2} \\ -\frac{5}{4} & \frac{11}{2} - x & \frac{9}{4} \\ 1 & -3 & -x \end{bmatrix} \\ &= -x^3 + 7x^2 - 16x + 12.\end{aligned}$$

Los autovalores de A_1 son las raíces de su polinomio característico. En este caso, las raíces son $\lambda_1 = 3$ y $\lambda_{2,3} = 2$. La multiplicidad algebraica del autovalor $\lambda = 2$ es entonces 2. Veamos cuál es su multiplicidad geométrica calculando el autoespacio asociado a $\lambda = 2$:

$$\mathrm{nul}(A_1 - 2I) = \mathrm{nul}(\begin{bmatrix} -\frac{1}{2} & 2 & \frac{3}{2} \\ -\frac{5}{4} & \frac{7}{2} & \frac{9}{4} \\ 1 & -3 & -2 \end{bmatrix}) = gen\{\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}\}.$$

Por lo tanto la multiplicidad geométrica de $\lambda = 2$ es 1 y A_1 no es diagonalizable. Estamos en el **Caso 3** del **Teorema 4.5.1**. En este caso, $J = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ y, por el Teorema 4.5.1, $A_1 \sim \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

Ahora, calculemos el polinomio característico de A_2 :

$$\begin{aligned} \chi_{A_2}(x) = \det(A_2 - xI) = \det &\begin{bmatrix} \frac{9}{2} - x & -\frac{13}{4} & -\frac{57}{4} \\ -\frac{1}{2} & \frac{13}{4} - x & \frac{9}{4} \\ \frac{1}{2} & -\frac{3}{4} & -\frac{3}{4} - x \end{bmatrix}. \\ = -x^3 + 7x^2 - 16x + 12. \end{aligned}$$

Los autovalores de A_2 son las raíces de su polinomio característico. En este caso, las raíces son $\lambda_1 = 3$ y $\lambda_{2,3} = 2$. La multiplicidad algebraica del autovalor $\lambda = 2$ es entonces 2. Veamos cuál es su multiplicidad geométrica calculando el autoespacio asociado a $\lambda = 2$:

$$\mathrm{nul}(A_2 - 2I) = \mathrm{nul}(\begin{bmatrix} \frac{5}{2} & -\frac{13}{4} & -\frac{57}{4} \\ -\frac{1}{2} & \frac{5}{4} & \frac{9}{4} \\ \frac{1}{2} & -\frac{3}{4} & -\frac{11}{4} \end{bmatrix}) = gen\{\begin{bmatrix} 7 \\ 1 \\ 1 \end{bmatrix}\}.$$

Por lo tanto la multiplicidad geométrica de $\lambda = 2$ es 1 y A_2 no es diagonalizable. Estamos en el **Caso 3** del **Teorema 4.5.1**. En este caso, $J = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ y, por el Teorema 4.5.1, $A_2 \sim \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. Entonces, $A_1 \sim \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \sim A_2$. Por lo tanto A_1 y A_2 son semejantes.

Para hallar S inversible tal que $A_1 = SA_2S^{-1}$. Busquemos P y Q inversibles tales que $A_1 = PJP^{-1}$ y $A_2 = QJQ^{-1}$.

Si $P := [v_1\ v_2\ v_3]$ es inversible y cumple que $A_1 = PJP^{-1}$, entonces $v_1, v_2, v_3 \in \mathbb{R}^3$ (las columnas de P) son vectores linealmente independientes tales que $v_1 \in \mathcal{S}_{\lambda=2} \cap \mathrm{col}(A_1 - 2I)$, v_3 es un autovector asociado a $\lambda = 3$ y $(A_1 - 2I)v_2 = v_1$.

Como $A_1 - 2I = \begin{bmatrix} -\frac{1}{2} & 2 & \frac{3}{2} \\ -\frac{5}{4} & \frac{7}{2} & \frac{9}{4} \\ 1 & -3 & -2 \end{bmatrix}$, tomamos $v_1 := \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \in \mathcal{S}_{\lambda=2} \cap \mathrm{col}(A_1 - 2I)$ y calculamos el autoespacio asociado a $\lambda = 3$:

$$\mathrm{nul}(A_1 - 3I) = \mathrm{nul}(\begin{bmatrix} -\frac{3}{2} & 2 & \frac{3}{2} \\ -\frac{5}{4} & \frac{5}{2} & \frac{9}{4} \\ 1 & -3 & -3 \end{bmatrix}) = gen\{\begin{bmatrix} -3 \\ -6 \\ 5 \end{bmatrix}\}.$$

Entonces, tomamos $v_3 := \begin{bmatrix} -3 \\ -6 \\ 5 \end{bmatrix} \in \mathcal{S}_{\lambda=3}$. Nos resta encontrar un vector $v_2 \in \mathbb{R}^3$ linealmente independiente con v_1 y v_3 tal que

$$\begin{bmatrix} -\frac{1}{2} & 2 & \frac{3}{2} \\ -\frac{5}{4} & \frac{7}{2} & \frac{9}{4} \\ 1 & -3 & -2 \end{bmatrix} v_2 = (A_1 - 2I)v_2 = v_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}.$$

Si $v_2 = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$, nos queda el siguiente sistema no homogéneo a resolver:

$$\begin{bmatrix} -\frac{1}{2} & 2 & \frac{3}{2} \\ -\frac{5}{4} & \frac{7}{2} & \frac{9}{4} \\ 1 & -3 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}.$$

La solución de dicho sistema es:

$$v_2 = \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

con $\alpha \in \mathbb{R}$. Como sólo necesitamos un vector $v_2 \in \mathbb{R}^3$, tomamos $\alpha = 0$ y $v_2 = \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix}$ que es li con v_1 y v_3 (verificarlo). Entonces $P = \begin{bmatrix} -1 & -2 & -3 \\ -1 & -1 & -6 \\ 1 & 0 & 5 \end{bmatrix}$ y tenemos que $A_1 = PJP^{-1}$.

Por otra parte, si $Q := [w_1\ w_2\ w_3]$ es inversible y cumple que $A_2 = QJQ^{-1}$, entonces $w_1, w_2, w_3 \in \mathbb{R}^3$ (las columnas de Q) son vectores linealmente independientes tales que tales que $w_1 \in \mathcal{S}_{\lambda=2} \cap \operatorname{col}(A_2 - 2I)$, w_3 es un autovector asociado a $\lambda = 3$ y $(A_2 - 2I)w_2 = w_1$.

Como $A_2 - 2I = \begin{bmatrix} \frac{5}{2} & -\frac{13}{4} & -\frac{57}{4} \\ -\frac{1}{2} & \frac{5}{4} & \frac{9}{4} \\ \frac{1}{2} & -\frac{3}{4} & -\frac{11}{4} \end{bmatrix}$, tomamos $w_1 := \begin{bmatrix} 7 \\ 1 \\ 1 \end{bmatrix} \in \mathcal{S}_{\lambda=2} \cap \operatorname{col}(A_2 - 2I)$ y calculamos el autoespacio asociado a $\lambda = 3$:

$$\operatorname{nul}(A_2 - 3I) = \operatorname{nul}(\begin{bmatrix} \frac{3}{2} & -\frac{13}{4} & -\frac{57}{4} \\ -\frac{1}{2} & \frac{1}{4} & \frac{9}{4} \\ \frac{1}{2} & -\frac{3}{4} & -\frac{15}{4} \end{bmatrix}) = gen\{\begin{bmatrix} 3 \\ -3 \\ 1 \end{bmatrix}\}.$$

Entonces, tomamos $w_3 := \begin{bmatrix} 3 \\ -3 \\ 1 \end{bmatrix} \in \mathcal{S}_{\lambda=3}$. Nos resta encontrar un vector $w_2 \in \mathbb{R}^3$ linealmente independiente con w_1 y w_3 tal que

$$\begin{bmatrix} \frac{5}{2} & -\frac{13}{4} & -\frac{57}{4} \\ -\frac{1}{2} & \frac{5}{4} & \frac{9}{4} \\ \frac{1}{2} & -\frac{3}{4} & -\frac{11}{4} \end{bmatrix} w_2 = (A_2 - 2I)w_2 = w_1 = \begin{bmatrix} 7 \\ 1 \\ 1 \end{bmatrix}.$$

Si $w_2 = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$, nos queda el siguiente sistema no homogéneo a resolver:

$$\begin{bmatrix} \frac{5}{2} & -\frac{13}{4} & -\frac{57}{4} \\ -\frac{1}{2} & \frac{5}{4} & \frac{9}{4} \\ \frac{1}{2} & -\frac{3}{4} & -\frac{11}{4} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \\ 1 \end{bmatrix}.$$

La solución de dicho sistema es:

$$w_2 = \begin{bmatrix} 8 \\ 4 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} 7 \\ 1 \\ 1 \end{bmatrix}$$

con $\alpha \in \mathbb{R}$. Como sólo necesitamos un vector $w_2 \in \mathbb{R}^3$, tomamos $\alpha = 0$ y $w_2 = \begin{bmatrix} 8 \\ 4 \\ 0 \end{bmatrix}$ que es li con w_1 y w_3 (verificarlo). Entonces $Q = \begin{bmatrix} 7 & 8 & 3 \\ 1 & 4 & -3 \\ 1 & 0 & 1 \end{bmatrix}$ y tenemos que $A_2 = QJQ^{-1}$.

Finalmente, notar que como $A_1 = PJP^{-1}$ y $J = Q^{-1}A_2Q$ tenemos que

$$A_1 = PJP^{-1} = PQ^{-1}A_2QP^{-1} = PQ^{-1}A_2(PQ^{-1})^{-1}.$$

Por lo tanto, tomando

$$S := PQ^{-1} = \begin{bmatrix} -1 & -2 & -3 \\ -1 & -1 & -6 \\ 1 & 0 & 5 \end{bmatrix} \frac{1}{4} \begin{bmatrix} -1 & 2 & 9 \\ 1 & -1 & -6 \\ 1 & -2 & -5 \end{bmatrix} = \begin{bmatrix} -1 & \frac{3}{2} & \frac{9}{2} \\ -\frac{3}{2} & \frac{11}{4} & \frac{27}{4} \\ 1 & -2 & -4 \end{bmatrix},$$

se sigue que $A_1 = SA_2S^{-1}$.

b) : Notar que $\operatorname{tr}(A_1) = \frac{18}{4}$ y $\operatorname{tr}(A_2) = 7$. Entonces como $\operatorname{tr}(A_1) \neq \operatorname{tr}(A_2)$, por la Proposición 4.6.1, A_1 no puede ser semejante a A_2. ■

4.7. Sistemas de ecuaciones diferenciales

En este sección vamos a resolver sistemas de ecuaciones diferenciales a coeficientes constantes. Empecemos con un ejemplo:

Ejercicio 4.J. Resolver el siguiente sistema de ecuaciones diferenciales a coeficientes constantes:

$$\begin{cases} y_1' = -4y_1 - 6y_2 + 3y_3 \\ y_2' = 2y_1 + 4y_2 - 2y_3 \\ y_3' = -2y_1 - 2y_2 + y_3 \end{cases}.$$

Dem. Para resolver el sistema de ecuaciones diferenciales, primero lo escribimos en forma matricial. Sea $A := \begin{bmatrix} -4 & -6 & 3 \\ 2 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix}$ e $Y(t) := \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix}$. Entonces $Y'(t) = \begin{bmatrix} y_1'(t) \\ y_2'(t) \\ y_3'(t) \end{bmatrix}$ y resolver el sistema de arriba, es equivalente a resolver

$$Y' = AY.$$

A continuación, buscamos diagonalizar la matriz A (si es posible). En caso de que A no sea diagonalizable buscamos la forma de Jordan de A tal como hicimos en el Teorema 4.5.1.

Calculamos el polinomio característico de A:

$$\chi_A(x) = \det(A - xI) = \det(\begin{bmatrix} -4-x & -6 & 3 \\ 2 & 4-x & -2 \\ -2 & -2 & 1-x \end{bmatrix}) = -x^3 + x^2 + 2x.$$

Las raíces de χ_A son $\lambda_1 = 2,\ \lambda_2 = -1,\ \lambda_3 = 0$ todas raíces distintas, por lo tanto A es diagonalizable.

Calculamos los autoespacios asociados a cada autovalor.

$$\mathcal{S}_{\lambda=2} = \operatorname{nul}(A - 2I) = \operatorname{nul}(\begin{bmatrix} -6 & -6 & 3 \\ 2 & 2 & -2 \\ -2 & -2 & -1 \end{bmatrix}) = \operatorname{nul}(\begin{bmatrix} -6 & -6 & 3 \\ 0 & 0 & -3 \\ 0 & 0 & -3 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\}.$$

$$\mathcal{S}_{\lambda=-1} = \operatorname{nul}(A + I) = \operatorname{nul}(\begin{bmatrix} -3 & -6 & 3 \\ 2 & 5 & -2 \\ -2 & -2 & 2 \end{bmatrix}) = \operatorname{nul}(\begin{bmatrix} -3 & -6 & 3 \\ 0 & 3 & 0 \\ 0 & 6 & 0 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\}.$$

$$\mathcal{S}_{\lambda=0} = \operatorname{nul}(A) = \operatorname{nul}(\begin{bmatrix} -4 & -6 & 3 \\ 2 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix}) = \operatorname{nul}(\begin{bmatrix} -4 & -6 & 3 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{bmatrix}) = gen\{\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}\}.$$

Construimos la matriz P colocando en sus columnas tres autovectores linealmente independientes, por ejemplo:

$$P := \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

Si $\Lambda := \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, entonces

$$A = P\Lambda P^{-1}.$$

Para resolver el sistema de ecuaciones diferenciales, proponemos el siguiente cambio de variables:

$$Y(t) := PZ(t).$$

Entonces, intuitivamente podemos ver que $Y'(t) = PZ'(t)$ y el sistema a resolver nos queda:

$$Y'(t) = PZ'(t) = AY(t) = APZ(t).$$

Multiplicando a izquierda por P^{-1} la ecuación anterior, nos queda:

$$Z'(t) = P^{-1}PZ'(t) = P^{-1}APZ(t) = \Lambda Z(t).$$

Entonces, $\begin{bmatrix} z_1'(t) \\ z_2'(t) \\ z_3'(t) \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \end{bmatrix}$ y nos queda el siguiente sistema desacoplado a resolver:

$$\begin{cases} z_1' = 2z_1 \\ z_2' = -z_2 \\ z_3' = 0 \end{cases}.$$

Recordando el **Ejercicio 2.27**, tenemos que la solución de cada ecuación diferencial es:

$$z_1(t) = c_1e^{2t}, \; z_2(t) = c_2e^{-t}, \; z_3(t) = c_3,$$

con $c_1, c_2, c_3 \in \mathbb{R}$. Volviendo a la variable original, nos queda que:

$$Y(t) = PZ(t) = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} c_1e^{2t} \\ c_2e^{-t} \\ c_3 \end{bmatrix} = c_1e^{2t} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + c_2e^{-t} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} =$$

$$= \begin{bmatrix} c_1e^{2t} + c_2e^{-t} \\ -c_1e^{2t} + c_3 \\ c_2e^{-t} + 2c_3 \end{bmatrix},$$

$c_1, c_2, c_3 \in \mathbb{R}$. ■

En general, supongamos que $A \in \mathbb{C}^{n\times n}$ y queremos resolver el sistema de ecuaciones diferenciales

$$Y' = AY. \tag{4.12}$$

Si A es diagonalizable, entonces $A = P\Lambda P^{-1}$ con $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ una matriz diagonal que en su diagonal tiene a los autovalores de A (no necesariamente distintos) y $P = [v_1 \; v_2 \; \cdots \; v_n] \in \mathbb{C}^{n\times n}$ una matriz inversible cuyas columnas $v_1, v_2, \cdots, v_n$ forman una base de autovectores de A y cada v_i es un autovector asociado al autovalor λ_i.

A partir del cambio de variable $Y(t) = PZ(t)$, operando de la misma manera que hicimos en el ejercicio anterior, podemos ver que una base de soluciones (o un conjunto fundamental de soluciones) de (4.12) es

$$\{e^{\lambda_1 t}v_1, e^{\lambda_2 t}v_2, \cdots, e^{\lambda_n t}v_n\}$$

y todas las soluciones del sistema de ecuaciones diferenciales (4.12) son:

$$Y(t) = c_1 e^{\lambda_1 t}v_1 + c_2 e^{\lambda_2 t}v_2 + \cdots + c_n e^{\lambda_n t}v_n,$$

con $c_1, c_2, \cdots, c_n \in \mathbb{C}$.

Ejercicio 4.13 b). Hallar la solución del problema a valores inciales y analizar el comportamiento asintótico de:

$$Y' = AY, \; Y(0) = Y_0, \tag{4.13}$$

con $A = \begin{bmatrix} -1 & -2 \\ 0 & -3 \end{bmatrix}$.

Dem. Tal como hicimos en el ejercicio anterior, veamos si A es diagonalizable. Calculamos el polinomio característico de A:

$$\chi_A(x) = \det(A - xI) = \det(\begin{bmatrix} -1-x & -2 \\ 0 & -3-x \end{bmatrix}) = (-1-x)(-3-x).$$

Observar como χ_A nos quedó factorizado. Por lo tanto, las raíces del χ_A son $\lambda_1 = -1$ y $\lambda_2 = -3$ todas raíces distintas, por lo tanto A es diagonalizable.

Fácilmente vemos que los autoespacios asociados son:

$$\mathcal{S}_{\lambda=-1} = \text{nul}(A + I) = \text{nul}(\begin{bmatrix} 0 & -2 \\ 0 & -2 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\}.$$

$$\mathcal{S}_{\lambda=-3} = \text{nul}(A + 3I) = \text{nul}(\begin{bmatrix} 2 & -2 \\ 0 & 0 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}.$$

Por lo tanto, como vimos arriba, todas las soluciones del sistema de ecuaciones diferenciales son:

$$Y(t) = c_1 e^{-t} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

con $c_1, c_2 \in \mathbb{R}$.

Usando la condición inicial podemos hallar los valores de c_1 y c_2. De hecho,

$$Y(0) = c_1 e^0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 e^0 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 + c_2 \\ c_2 \end{bmatrix} = \begin{bmatrix} Y_{01} \\ Y_{02} \end{bmatrix},$$

donde Y_{01} e Y_{02} son la primera y segunda coordenada del vector Y_0 respectivamente. Despejando, tenemos que $c_2 = Y_{02}$ y $c_1 = Y_{01} - c_2 = Y_{01} - Y_{02}$. Por lo tanto, la solución del sistema de ecuaciones diferenciales a valores iniciales es:

$$Y(t) = (Y_{01} - Y_{02})e^{-t} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + Y_{02}e^{-3t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} (Y_{01} - Y_{02})e^{-t} + Y_{02}e^{-3t} \\ Y_{02}e^{-3t} \end{bmatrix}.$$

Por lo tanto,

$$\begin{aligned} \lim_{t\to\infty} Y(t) &= \lim_{t\to\infty} \begin{bmatrix} (Y_{01} - Y_{02})e^{-t} + Y_{02}e^{-3t} \\ Y_{02}e^{-3t} \end{bmatrix} \\ &:= \begin{bmatrix} \lim_{t\to\infty} [(Y_{01} - Y_{02})e^{-t} + Y_{02}e^{-3t}] \\ \lim_{t\to\infty} Y_{02}e^{-3t} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned}$$

■

Vimos que para matrices $A \in \mathbb{C}^{2\times 2}$ (con coeficientes reales), el polinomio característico de A es

$$\chi_A(x) = x^2 - \operatorname{tr}(A)x + \det(A),$$

y en consecuencia las raíces de χ_A son

$$\lambda_{1,2} = \frac{1}{2}[\operatorname{tr}(A) \pm \sqrt{\operatorname{tr}(A)^2 - 4\det(A)}].$$

Si A es diagonalizable, vimos que las soluciones del sistema de ecuaciones (4.13) son

$$Y(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2,$$

donde $c_1, c_2 \in \mathbb{C}$, $\lambda_1, \lambda_2 \in \mathbb{C}$ son los autovalores de A y $v_1, v_2 \in \mathbb{C}^2$ los autovectores asociados a dichos autovalores.

Por lo tanto, $\lim_{t\to\infty} Y(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (para cualquier valor inicial) si y sólo si $\lambda_1 < 0$ y $\lambda_2 < 0$. Si $4\det(A) \leq \operatorname{tr}(A)^2$ entonces, por la fórmula de los autovalores de A que escribimos arriba, $\lambda_1 \in \mathbb{R}$ y $\lambda_2 \in \mathbb{R}$. En ese caso, si $\operatorname{tr}(A) < 0$ y $\det(A) > 0$, tendremos que

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 < 0 \text{ y } \det(A) = \lambda_1\ \lambda_2 > 0.$$

Por lo tanto, no queda otra que $\lambda_1 < 0$ y $\lambda_2 < 0$ (¿por qué?) y en ese caso, siempre tendremos que $\lim_{t\to\infty} Y(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

El siguiente ejercicio es (un poco) similar al **Ejercicio 4.18**.

Ejercicio 4.K. Sea $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Hallar todos los $Y_0 \in \mathbb{R}^2$ tales que la solución del problema a valores inciales

$$Y' = AY,\ Y(0) = Y_0,$$

tiene norma acotada. Es decir, existe $C > 0$ tal que

$$\|Y(t)\| \leq C$$

para todo $t \geq 0$. La norma que usamos es la del producto interno canónico de $\mathbb{R}^2$.

Dem. Operando de manera similar al ejercicio anterior, tenemos que $\lambda = 3$ y $\lambda_2 = -1$ son los autovalores de A con autoespacios asociados $\mathcal{S}_{\lambda=3} = gen\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$ y $\mathcal{S}_{\lambda=-1} = gen\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$. Por lo tanto A es diagonalizable y todas las soluciones del sistema de ecuaciones diferenciales son:

$$Y(t) = c_1 e^{3t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{-t} \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

con $c_1, c_2 \in \mathbb{R}$.

Usando la condición inicial podemos hallar los valores de c_1 y c_2. De hecho,

$$Y(0) = c_1 e^0 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^0 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 - c_2 \\ c_1 + c_2 \end{bmatrix} = \begin{bmatrix} Y_{01} \\ Y_{02} \end{bmatrix},$$

donde Y_{01} e Y_{02} son la primera y segunda coordenada del vector Y_0 respectivamente. Despejando, tenemos que $c_1 = \frac{Y_{01}+Y_{02}}{2}$ y $c_2 = \frac{Y_{02}-Y_{01}}{2}$. Por lo tanto, la solución del sistema de ecuaciones diferenciales a valores iniciales es:

$$Y(t) = (\frac{Y_{01}+Y_{02}}{2})e^{3t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (\frac{Y_{02}-Y_{01}}{2})e^{-t} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Considerando el producto interno canónico de $\mathbb{R}^2$ y usando Pitágoras (o haciendo la cuenta) tenemos que

$$\|Y(t)\|^2 = \langle Y(t), Y(t) \rangle = (\frac{Y_{01}+Y_{02}}{2})^2 e^{6t} 2 + (\frac{Y_{02}-Y_{01}}{2})^2 e^{-2t} 2.$$

Buscamos los valores de Y_0 para que $\|Y(t)\|^2$ sea acotada. Es decir, buscamos para qué valores de Y_0 existe $C > 0$ tal que

$$\|Y(t)\| \leq C$$

para todo $t \geq 0$.

Observar que si $Y_{01} + Y_{02} \neq 0$ entonces $\|Y(t)\|^2$ nunca es acotada pues como $\lim_{t\to\infty} e^{6t} = \infty$ y $\lim_{t\to\infty}(\frac{Y_{02}-Y_{01}}{2})^2 e^{-2t} 2 = 0$, tendríamos que $\lim_{t\to\infty} \|Y(t)\|^2 = \infty$, eso significa que dado cualquier $M > 0$ siempre existe $t_0 > 0$ tal que si $t \geq t_0$ entonces $\|Y(t)\|^2 > M$, entonces $\|Y(t)\|^2$ no es acotada y por lo tanto $\|Y(t)\|$ tampoco.

Afirmamos que $\|Y(t)\|$ es acotada si y sólo si $Y_{01} = -Y_{02}$. De hecho, en ese caso,

$$\|Y(t)\|^2 = (\frac{Y_{02}-Y_{01}}{2})^2 e^{-2t} 2 = 2Y_{02}^2 e^{-2t} \leq 2Y_{02}^2,$$

para todo $t \geq 0$. Por lo tanto $\|Y(t)\| \leq \sqrt{2}\,|Y_{02}|$, para todo $t \geq 0$ e $Y(t)$ resulta acotada. Conclusión, la norma de $Y(t)$ resulta acotada si y sólo si $Y_0 \in gen\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$. ■

4.8. Sistemas de ecuaciones diferenciales: caso no diagonalizable

A continuación veremos un ejemplo, similar al **Ejercicio 4.16**, de cómo resolver un sistema de ecuaciones diferenciales a coficientes constantes cuando la matriz A no es diagonalizable:

Ejercicio 4.L. Sea $A = \begin{bmatrix} 3 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. Hallar la solución del problema a valores inciales y analizar el comportamiento asintótico de:

$$Y' = AY, \; Y(0) = Y_0,$$

Dem. Tal como vimos en el **Ejercicio 4.H**, A no es diagonalizable.

En este caso, calculamos la forma de Jordan de A. Recordar que en la resolución del **Ejercicio 4.H**, probamos que si

$$J = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \text{ y}$$

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

entonces la forma de Jordan de A es

$$A = PJP^{-1}.$$

En este caso, también haremos el cambio de variables $Y(t) := PZ(t)$. Entonces, $Y'(t) = PZ'(t)$ y el sistema a resolver nos queda:

$$Y'(t) = PZ'(t) = AY(t) = APZ(t).$$

Multiplicando a izquierda por P^{-1} la ecuación anterior, nos queda:

$$Z'(t) = P^{-1}PZ'(t) = P^{-1}APZ(t) = JZ(t).$$

Entonces,

$$\begin{bmatrix} z_1'(t) \\ z_2'(t) \\ z_3'(t) \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \end{bmatrix}$$

y nos queda el siguiente sistema a resolver,

$$\begin{cases} z_1' = 2z_1 + z_2 \\ z_2' = 2z_2 \\ z_3' = 2z_3 \end{cases}.$$

Recordando (por ejemplo) el **Ejercicio 2.27**, tenemos que la solución de las ecuaciones diferenciales homogéneas $z_2' = 2z_2$ y $z_3' = 2z_3$ es

$$z_2 = c_2e^{2t}, \ z_3(t) = c_3e^{2t},$$

con $c_2, c_3 \in \mathbb{R}$. Nos queda a resolver la siguiente ecuación diferencial no homogénea

$$z_1'(t) = 2z_1(t) + z_2(t) = 2z_1(t) + c_2e^{2t}. \tag{4.14}$$

Recordando (por ejemplo) el **Ejercicio 2.27**, tenemos que todas las soluciones de (4.14) son

$$z_1(t) = z_1^h(t) + z_1^p(t),$$

donde z_1^h son las soluciones del sistema homogéneo asociado y z_1^p es una solución particular. En este caso $z_1^h \in gen\{e^{2t}\}$ y $z_1^p(t) = f(t)e^{2t}$, donde $f'(t) = (c_2e^{2t})e^{-2t} = c_2$. Entonces, $f(t) = c_2t$ y tenemos que $z_1^p(t) = c_2te^{2t}$. Por lo tanto,

$$z_1(t) = c_1e^{2t} + c_2te^{2t},$$

con $c_1 \in \mathbb{R}$.

Volviendo a la variable original, nos queda que:

$$\begin{aligned} Y(t) = PZ(t) &= \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} c_1e^{2t} + c_2te^{2t} \\ c_2e^{2t} \\ c_3e^{2t} \end{bmatrix} \\ &= c_1e^{2t} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_2e^{2t}(t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}) + c_3e^{2t} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} c_1e^{2t} + c_2e^{2t}(t+1) + c_3e^{2t} \\ -c_3e^{2t} \\ c_1e^{2t} + c_2e^{2t}t \end{bmatrix}. \end{aligned}$$

Usando la condición inicial podemos hallar los valores de c_1, c_2 y c_3. De hecho,

$$Y(0) = \begin{bmatrix} c_1e^0 + c_2e^0(0+1) + c_3e^0 \\ -c_3e^0 \\ c_1e^0 + c_2e^00 \end{bmatrix} = \begin{bmatrix} c_1 + c_2 + c_3 \\ -c_3 \\ c_1 \end{bmatrix} = \begin{bmatrix} Y_{01} \\ Y_{02} \\ Y_{03} \end{bmatrix},$$

donde Y_{01}, Y_{02} e Y_{03} son la primera, segunda y tercera coordenada del vector Y_0 respectivamente. Despejando, tenemos que $c_1 = Y_{03}$, $c_3 = -Y_{02}$ y $c_2 = Y_{01} - c_3 - c_1 = Y_{01} + Y_{02} - Y_{03}$. Por lo tanto, la solución del sistema de ecuaciones diferenciales a valores iniciales es:

$$Y(t) = Y_{03}e^{2t} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (Y_{01} + Y_{02} - Y_{03})e^{2t}(t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}) - Y_{02}e^{2t} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

■

En general, supongamos que $A \in \mathbb{C}^{3\times 3}$ y queremos resolver el sistema de ecuaciones diferenciales

$$Y' = AY. \tag{4.15}$$

Supongamos que A no es diagonalizable y que la forma de Jordan de A es $A = PJP^{-1}$ con $P = [v_1\ v_2\ v_3] \in \mathbb{C}^{3\times 3}$ una matriz inversible. Teníamos 3 casos posibles:

Caso 1: λ es un autovalor triple de A con multiplicidad geométrica 1. En ese caso

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

Caso 2: λ es un autovalor triple de A con multiplicidad geométrica 2. En ese caso

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

Caso 3: λ es un autovalor doble de A con multiplicidad geométrica 1 y μ es un autovalor simple de A. En ese caso

$$J = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{bmatrix}.$$

A partir del cambio de variable $Y(t) = PZ(t)$ y operando de la misma manera que hicimos en el ejercicio anterior, podemos obtener todas las soluciones del sistema de ecuaciones diferenciales (4.15) según el caso en el que estemos:

Caso 1: λ es un autovalor triple con multiplicidad geométrica 1. En ese caso, todas las soluciones del sistema de ecuaciones diferenciales (4.15) son:

$$Y(t) = c_1 e^{\lambda t} v_1 + c_2 e^{\lambda t}(tv_1 + v_2) + c_3 e^{\lambda t}(\frac{t^2}{2} v_1 + tv_2 + v_3),$$

con $c_1, c_2, c_3 \in \mathbb{C}$.

Caso 2: λ es un autovalor triple de A con multiplicidad geométrica 2. En ese caso, todas las soluciones del sistema de ecuaciones diferenciales (4.15) son:

$$Y(t) = c_1 e^{\lambda t} v_1 + c_2 e^{\lambda t}(tv_1 + v_2) + c_3 e^{\lambda t} v_3,$$

con $c_1, c_2, c_3 \in \mathbb{C}$.

Caso 3: λ es un autovalor doble de A con multiplicidad geométrica 1 y μ es un autovalor simple de A. En ese caso, todas las soluciones del sistema de ecuaciones diferenciales (4.15) son:

$$Y(t) = c_1 e^{\lambda t} v_1 + c_2 e^{\lambda t}(tv_1 + v_2) + c_3 e^{\mu t} v_3,$$

con $c_1, c_2, c_3 \in \mathbb{C}$.

Pensando de manera similar, podemos deducir cómo sería la forma de Jordan para $A \in \mathbb{C}^{2\times 2}$ y todas las soluciones del sistema de ecuaciones diferenciales (4.15).

Ejercicio 4.17. Hallar todas las soluciones $Y \in C(\mathbb{R}, \mathbb{R}^2)$ del sistema

$$Y' = \begin{bmatrix} 6 & -5 \\ 5 & -2 \end{bmatrix} Y.$$

Dem. Primero, veamos si $A := \begin{bmatrix} 6 & -5 \\ 5 & -2 \end{bmatrix}$ es diagonalizable. Calculamos el polinomio característico de A:

$$\chi_A(x) = \det(A - xI) = \det(\begin{bmatrix} 6-x & -5 \\ 5 & -2-x \end{bmatrix}) = x^2 + 4x + 13.$$

Por lo tanto, las raíces del χ_A son $\lambda_1 = 2 + 3i$ y $\lambda_2 = 2 - 3i$ ambas raíces distintas, por lo tanto, si consideramos que $A \in \mathbb{C}^{n\times n}$, se sigue que A es diagonalizable.

Fácilmente vemos que los autoespacios asociados son:

$$\mathcal{S}_{\lambda=2+3i} = \mathrm{nul}(A - (2+3i)I) = \mathrm{nul}(\begin{bmatrix} 4-3i & -5 \\ 5 & -4-3i \end{bmatrix}) = gen\{\begin{bmatrix} 4+3i \\ 5 \end{bmatrix}\}.$$

$$\mathcal{S}_{\lambda=2-3i} = \mathrm{nul}(A - (2-3i)I) = \mathrm{nul}(\begin{bmatrix} 4+3i & -5 \\ 5 & -4+3i \end{bmatrix}) = gen\{\begin{bmatrix} 4-3i \\ 5 \end{bmatrix}\}.$$

Por lo tanto, como vimos arriba, todas las soluciones del sistema de ecuaciones diferenciales son:

$$Y(t) = c_1 e^{(2+3i)t} \begin{bmatrix} 4+3i \\ 5 \end{bmatrix} + c_2 e^{(2-3i)t} \begin{bmatrix} 4-3i \\ 5 \end{bmatrix},$$

con $c_1, c_2 \in \mathbb{R}$.

Notar que las funciones $\begin{bmatrix} 4+3i \\ 5 \end{bmatrix} \in C(\mathbb{R}, \mathbb{C}^2)$ y $\begin{bmatrix} 4-3i \\ 5 \end{bmatrix} \in C(\mathbb{R}, \mathbb{C}^2)$, como el ejercicio pide todas las soluciones $Y \in C(\mathbb{R}, \mathbb{R}^2)$, tendremos que maniobrar un poco.

Recordar que $e^{(2+3i)t} = e^{2t}(\cos(3t) + i\sin(3t))$ y $e^{(2-3i)t} = e^{2t}(\cos(3t) - i\sin(3t))$. Entonces

$$\begin{aligned}
Y(t) = c_1 & \begin{bmatrix} e^{2t}(\cos(3t) + i\sin(3t))(4+3i) \\ 5e^{2t}(\cos(3t) + i\sin(3t)) \end{bmatrix} + c_2 \begin{bmatrix} e^{2t}(\cos(3t) - i\sin(3t))(4-3i) \\ 5e^{2t}(\cos(3t) - i\sin(3t)) \end{bmatrix} \\
= c_1 & \begin{bmatrix} e^{2t}(4\cos(3t) - 3\sin(3t)) \\ 5e^{2t}\cos(3t) \end{bmatrix} + ic_1 \begin{bmatrix} e^{2t}(3\cos(3t) + 4\sin(3t)) \\ 5e^{2t}\sin(3t) \end{bmatrix} + \\
+ c_2 & \begin{bmatrix} e^{2t}(4\cos(3t) - 3\sin(3t)) \\ 5e^{2t}(\cos(3t)) \end{bmatrix} - ic_2 \begin{bmatrix} e^{2t}(3\cos(3t) + 4\sin(3t)) \\ 5e^{2t}\sin(3t) \end{bmatrix} \\
= (c_1 + c_2) & \begin{bmatrix} e^{2t}(4\cos(3t) - 3\sin(3t)) \\ 5e^{2t}\cos(3t) \end{bmatrix} + i(c_1 - c_2) \begin{bmatrix} e^{2t}(3\cos(3t) + 4\sin(3t)) \\ 5e^{2t}\sin(3t) \end{bmatrix}.
\end{aligned}$$

Tomando $\alpha := c_1 + c_2 \in \mathbb{C}$ y $\beta := i(c_1 - c_2) \in \mathbb{C}$, tenemos que todas las soluciones de la ecuación diferencial $Y' = AY$, son de la forma

$$Y(t) = \alpha Y_1(t) + \beta Y_2(T),$$

con $Y_1(t) := \begin{bmatrix} e^{2t}(4\cos(3t) - 3\sin(3t)) \\ 5e^{2t}(\cos(3t)) \end{bmatrix} \in C(\mathbb{R}, \mathbb{R}^2)$ e $Y_2(t) := \begin{bmatrix} e^{2t}(3\cos(3t) + 4\sin(3t)) \\ 5e^{2t}(\sin(3t)) \end{bmatrix} \in C(\mathbb{R}, \mathbb{R}^2)$. ■

El siguiente ejercicio demuestra que toda simetría y todo proyector es diagonalizable, esto ya lo habíamos notado en Proposición 2.5.1. Vamos a reescribir lo que hicimos para adaptarlo a la notación de este capítulo.

Ejercicio 4.M. Sea $\mathbb{V}$ un $\mathbb{C}$-espacio vectorial de dimensión $n > 1$. Demostrar que

a) Si $S \in \mathcal{L}(\mathbb{V}) \setminus \{I_{\mathbb{V}}\}$ es una simetría (es decir, $S^2 = I_{\mathbb{V}}$), existen $k \in \mathbb{N}$ y una base $\mathcal{B}$ de $\mathbb{V}$ tales que

$$[S]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{bmatrix},$$

con I_k la matriz identidad de $k \times k$ e I_{n-k} la matriz identidad de $n - k \times n - k$.

b) Si $T \in \mathcal{L}(\mathbb{V}) \setminus \{0_{\mathbb{V}}, I_{\mathbb{V}}\}$ es una proyección (es decir, $T^2 = T$), existen $k \in \mathbb{N}$ y una base $\mathcal{B}$ de $\mathbb{V}$ tales que
$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix},$$
con I_k la matriz identidad de $k \times k$.

Dem. $a)$: Como S cumple que $S^2 = I_{\mathbb{V}}$, en el **Ejercicio 2.23 e)** vimos que S es una simetría y probamos que
$$\mathbb{V} = \mathrm{Nu}(S - I_{\mathbb{V}}) \oplus \mathrm{Nu}(S + I_{\mathbb{V}})$$
(en este caso el símbolo $\oplus$ denota suma directa, no necesariamente ortogonal).

Sea $k := \dim(\mathrm{Nu}(S - I_{\mathbb{V}}))$ (como $S \neq I_{\mathbb{V}}$ tenemos que $k < n$).

Si $k = 0$, entonces $\mathrm{Nu}(S - I_{\mathbb{V}}) = \{0_{\mathbb{V}}\}$ y tenemos que $\mathbb{V} = \mathrm{Nu}(S - I_{\mathbb{V}}) \oplus \mathrm{Nu}(S + I_{\mathbb{V}}) = \mathrm{Nu}(S + I_{\mathbb{V}})$. Por lo tanto $(S + I_{\mathbb{V}})v = 0_{\mathbb{V}}$, para todo $v \in \mathbb{V}$ y entonces se sigue $S = -I_{\mathbb{V}}$. En ese caso, tomando cualquier base $\mathcal{B}$ de $\mathbb{V}$ se sigue que
$$[S]_{\mathcal{B}}^{\mathcal{B}} = -I_n.$$

Si $k \geq 1$, sea $\{v_1, v_2, \cdots, v_k\}$ una base de $\mathrm{Nu}(S - I_{\mathbb{V}})$. Entonces,
$$S(v_i) = v_i,$$
para todo $i = 1, 2, \cdots, k$. Por otra parte, como $\mathbb{V} = \mathrm{Nu}(S - I_{\mathbb{V}}) \oplus \mathrm{Nu}(S + I_{\mathbb{V}})$, tenemos que
$$\dim(\mathrm{Nu}(S + I_{\mathbb{V}})) = \dim(\mathbb{V}) - \dim(\mathrm{Nu}(S - I_{\mathbb{V}})) = n - k \geq 1.$$
Sea $\{v_{k+1}, \cdots, v_n\}$ una base de $\mathrm{Nu}(S + I_{\mathbb{V}})$. Entonces,
$$S(v_i) = -v_i$$
para todo $i = k + 1, \cdots, n$.

Sea $\mathcal{B} := \{v_1, v_2, \cdots, v_k, v_{k+1}, \cdots, v_n\}$ entonces $\mathcal{B}$ es una base de $\mathbb{C}^n$ (¿por qué?). Recordemos además que $[v_i]^{\mathcal{B}} = e_i$, donde e_i es el vector i-ésimo de la base canónica de $\mathbb{C}^n$, es decir el vector de $\mathbb{C}^n$ con todas componentes nulas excepto en el lugar i donde vale 1. Por ejemplo,
$$[v_1]^{\mathcal{B}} = [1\ 0 \cdots\ 0]^T = e_1.$$

Entonces, por definición de $[S]_{\mathcal{B}}^{\mathcal{B}}$, tenemos que
$$\begin{aligned}
[S]_{\mathcal{B}}^{\mathcal{B}} &= [\ [S(v_1)]^{\mathcal{B}}\ [S(v_2)]^{\mathcal{B}} \cdots\ [S(v_k)]^{\mathcal{B}}\ [S(v_{k+1})]^{\mathcal{B}}\ \cdots\ [S(v_n)]^{\mathcal{B}}\ ] \\
&= [\ [v_1]^{\mathcal{B}}\ [v_2]^{\mathcal{B}}\ \cdots\ [v_k]^{\mathcal{B}}\ [-v_{k+1}]^{\mathcal{B}}\ \cdots\ [-v_n]^{\mathcal{B}}\ ] \\
&= [\ e_1\ e_2\ \cdots\ e_k\ -e_{k+1}\ \cdots\ -e_n] \\
&= \begin{bmatrix} I_{k \times k} & 0_{k \times n-k} \\ 0_{n-k \times k} & -I_{n-k \times n-k} \end{bmatrix}.
\end{aligned}$$

$b)$: Como T cumple que $T^2 = T$, en el **Ejercicio 2.23** vimos que T es la proyección sobre $\mathrm{Im}(T)$ en la dirección de $\mathrm{Nu}(T)$. Por otra parte, en el **Ejercicio 2.19 a)**, probamos que
$$\mathbb{V} = \mathrm{Im}(T) \oplus \mathrm{Nu}(T)$$
(en este caso el símbolo $\oplus$ denota suma directa, no necesariamente ortogonal).

Sea $k := \dim(\mathrm{Im}(T))$ (como $T \neq I_{\mathbb{V}}$ y $T \neq 0_{\mathbb{V}}$ entonces $1 \leq k < n$) y sea $\{v_1, v_2, \cdots, v_k\}$ una base de $\mathrm{Im}(T)$. Entonces, como T es un proyector y $v_i \in \mathrm{Im}(T)$ para todo $i = 1, 2, \cdots, k$, tenemos que
$$T(v_i) = v_i,$$

para todo $i = 1, 2, \cdots, k$. Por otra parte, como $\mathbb{V} = \mathrm{Im}(T) \oplus \mathrm{Nu}(T)$, tenemos que

$$\dim(\mathrm{Nu}(T)) = \dim(\mathbb{V}) - \dim(\mathrm{Im}(T)) = n - k \geq 1.$$

Sea $\{v_{k+1}, \cdots, v_n\}$ una base de $\mathrm{Nu}(T)$. Entonces,

$$T(v_i) = 0_{\mathbb{V}}$$

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

Sea $\mathcal{B} := \{v_1, v_2, \cdots, v_k, v_{k+1}, \cdots, v_n\}$ entonces $\mathcal{B}$ es una base de $\mathbb{C}^n$ (¿por qué?).

Por lo tanto, por definición de $[T]_{\mathcal{B}}^{\mathcal{B}}$, tenemos que

$$\begin{aligned}
[T]_{\mathcal{B}}^{\mathcal{B}} &= [\ [T(v_1)]^{\mathcal{B}}\ [T(v_2)]^{\mathcal{B}} \cdots\ [T(v_k)]^{\mathcal{B}}\ [T(v_{k+1})]^{\mathcal{B}}\ \cdots\ [T(v_n)]^{\mathcal{B}}\ ] \\
&= [\ [v_1]^{\mathcal{B}}\ [v_2]^{\mathcal{B}}\ \cdots\ [v_k]^{\mathcal{B}}\ [0_{\mathbb{V}}]^{\mathcal{B}}\ \cdots\ [0_{\mathbb{V}}]^{\mathcal{B}}\ ] \\
&= [\ e_1\ e_2\ \cdots\ e_k\ 0_{\mathbb{C}^n}\ \cdots\ 0_{\mathbb{C}^n}] \\
&= \begin{bmatrix} I_{k\times k} & 0_{k\times n-k} \\ 0_{n-k\times k} & 0_{n-k\times n-k} \end{bmatrix}.
\end{aligned}$$

■

Conclusión: Si $S^2 = I_{\mathbb{V}}$, entonces S es una simetría con respecto a $\mathrm{Nu}(S - I_{\mathbb{V}})$ en la dirección $\mathrm{Im}(S - I_{\mathbb{V}}) = \mathrm{Nu}(S + I_{\mathbb{V}})$ (ver **Ejercicio 2.23 d)**). En ese caso, S es diagonalizable con autovalores 1 y -1. La multiplicidad algebraica y geométrica del autovalor 1 es igual a la dimensión del autoespacio $\mathrm{Nu}(S - I_{\mathbb{V}})$ y la multiplicidad algebraica y geométrica del autovalor -1 es igual a la dimensión del autoespacio $\mathrm{Im}(S - I_{\mathbb{V}}) = \mathrm{Nu}(S + I_{\mathbb{V}})$.

Si $T^2 = T$, entonces T es un proyector con respecto a $\mathrm{Im}(T)$ en la dirección $\mathrm{Nu}(T)$ (ver **Ejercicio 2.23 a)**). En ese caso, T es diagonalizable, con autovalores 1 y 0. La multiplicidad algebraica y geométrica del autovalor 1 es igual a la dimensión del autoespacio $\mathrm{Im}(T) = \mathrm{Nu}(T - I_{\mathbb{V}})$ y la multiplicidad algebraica y geométrica del autovalor 0 es igual a la dimensión del autoespacio $\mathrm{Nu}(T)$.

Ahora sí, resolvamos el siguiente ejercicio de examen:

Ejercicio 4.N (De Parcial). Sea $\mathcal{S} = gen\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}\}$. Sean $[P_{\mathcal{S}}]_{\mathcal{E}}^{\mathcal{E}}$ y $[P_{\mathcal{S}^\perp}]_{\mathcal{E}}^{\mathcal{E}}$ las matrices en la base canónica de $\mathbb{R}^3$ de las proyecciones sobre $\mathcal{S}$ y $\mathcal{S}^\perp$ respectivamente considerando el producto interno canónico de $\mathbb{R}^3$. Si $A := 2[P_{\mathcal{S}}]_{\mathcal{E}}^{\mathcal{E}} + 3[P_{\mathcal{S}^\perp}]_{\mathcal{E}}^{\mathcal{E}}$, encontrar todas las soluciones del problema a valores iniciales:

$$Y' = A^{21}Y, \quad Y(0) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Dem. En primer lugar, recordar que como $P_{\mathcal{S}}$ es la proyección ortogonal sobre $\mathcal{S}$, tenemos que $P_{\mathcal{S}^\perp} = I - P_{\mathcal{S}}$ y por lo tanto, $[P_{\mathcal{S}^\perp}]_{\mathcal{E}}^{\mathcal{E}} = [I - P_{\mathcal{S}}]_{\mathcal{E}}^{\mathcal{E}} = I_3 - [P_{\mathcal{S}}]_{\mathcal{E}}^{\mathcal{E}}$, donde I_3 es la matriz identidad de 3×3. Entonces,

$$A = 2[P_{\mathcal{S}}]_{\mathcal{E}}^{\mathcal{E}} + 3[P_{\mathcal{S}^\perp}]_{\mathcal{E}}^{\mathcal{E}} = 2[P_{\mathcal{S}}]_{\mathcal{E}}^{\mathcal{E}} + 3(I_3 - [P_{\mathcal{S}}]_{\mathcal{E}}^{\mathcal{E}}) = 3I_3 - [P_{\mathcal{S}}]_{\mathcal{E}}^{\mathcal{E}}.$$

Veamos que $P_{\mathcal{S}}$ siempre es diagonalizable. De hecho, tenemos que $P_{\mathcal{S}}(v) = v$ para todo $v \in \mathrm{Im}(P_{\mathcal{S}}) = \mathcal{S}$ y $P_{\mathcal{S}}(v) = 0$ para todo $v \in \mathrm{Nu}(P_{\mathcal{S}}) = \mathcal{S}^\perp$. Por lo tanto, los autovalores de la transformación lineal $P_{\mathcal{S}}$ son $\lambda_{1,2} = 1$ y $\lambda_3 = 0$ con autoespacios asociados $\mathcal{S}_{\lambda=1} = \mathrm{Im}(P_{\mathcal{S}}) = \mathcal{S}$ y $\mathcal{S}_{\lambda=0} = \mathrm{Nu}(P_{\mathcal{S}}) = \mathcal{S}^\perp$.

Calculemos $\mathcal{S}^{\perp}$, recordemos que $\mathcal{S}^{\perp} = \{x \in \mathbb{R}^3 : \left\langle x, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\rangle = \left\langle x, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \right\rangle = 0\}$. Entonces, $x \in \mathcal{S}^{\perp}$ si y sólo si $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ es solución del sistema

$$0 = \left\langle x, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\rangle = x_1 + x_2, \quad 0 = \left\langle x, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \right\rangle = 2x_1 + x_2 + x_3.$$

Resolviendo, nos queda que $x_1 = -x_2$ y $x_3 = -2x_1 - x_2 = 2x_2 - x_2 = x_2$. Entonces $x = \begin{bmatrix} -x_2 \\ x_2 \\ x_2 \end{bmatrix}$, con $x_2 \in \mathbb{R}$ y por lo tanto $\mathcal{S}^{\perp} = gen\{ \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \}$.

Sea $\mathcal{B} := \{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \}$, entonces $\mathcal{B}$ es una base (formada por autovectores) de $\mathbb{R}^3$. Sea $P := M_{\mathcal{B}}^{\mathcal{E}} = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ y $\Lambda := [P_{\mathcal{S}}]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, tenemos que

$$[P_{\mathcal{S}}]_{\mathcal{E}}^{\mathcal{E}} = P\Lambda P^{-1}.$$

Por lo tanto,

$$\begin{aligned} A = 3I_3 - [P_{\mathcal{S}}]_{\mathcal{E}}^{\mathcal{E}} &= 3PP^{-1} - P\Lambda P^{-1} = P(3I_3 - \Lambda)P^{-1} \\ &= P \begin{bmatrix} 3-1 & 0 & 0 \\ 0 & 3-1 & 0 \\ 0 & 0 & 3-0 \end{bmatrix} P^{-1} = P \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} P^{-1}. \end{aligned}$$

Entonces A es diagonalizable y entonces

$$A^{21} = P(\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix})^{21} P^{-1} = P \begin{bmatrix} 2^{21} & 0 & 0 \\ 0 & 2^{21} & 0 \\ 0 & 0 & 3^{21} \end{bmatrix} P^{-1}. \tag{4.16}$$

Además, es claro que (4.16) es la diagonalización de la matriz A^{21}.

Entonces, tal como vimos arriba, todas las soluciones del sistema de ecuaciones diferenciales $Y' = A^{21}Y$, son

$$Y(t) = c_1 e^{2^{21}t} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2 e^{2^{21}t} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + c_3 e^{3^{21}t} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix},$$

con $c_1, c_2, c_3 \in \mathbb{R}$. Usando la condición inicial podemos hallar los valores de c_1, c_2 y c_3. De hecho,

$$Y(0) = c_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Resolviendo el sistema, tenemos que $c_1 = 1$, $c_2 = 0$ y $c_3 = 0$. Por lo tanto, la solución del sistema de ecuaciones diferenciales a valores iniciales es:

$$Y(t) = e^{2^{21}t} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

■

4.9. Sistema de ecuaciones diferenciales no homogéneo

Finalmente, veamos un ejemplo de resolución de un sistema de ecuaciones diferenciales no homogéneo.

Ejercicio 4.Ñ. Resolver el problema no homogéneo

$$Y'(t) = AY(t) + F(t),$$

con $A = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}$ y $F(t) = \begin{bmatrix} e^{3t} \\ e^{3t} \end{bmatrix}$.

Dem. Para resolver el sistema de ecuaciones diferenciales no homogéneo, primero veamos si A es diagonalizable. En caso de que A no sea diagonalizable, calcularemos la forma de Jordan de A y prodeceremos de manera similar.

Calculamos el polinomio característico de A:

$$\chi_A(x) = \det(A - xI) = \det(\begin{bmatrix} 2-x & -1 \\ 3 & -2-x \end{bmatrix}) = x^2 - 1.$$

Las raíces del χ_A son $\lambda_1 = -1$ y $\lambda_2 = 1$ todas raíces distintas, por lo tanto A es diagonalizable. Fácilmente vemos que los autoespacios asociados son:

$$\mathcal{S}_{\lambda=-1} = \text{nul}(A + I) = \text{nul}(\begin{bmatrix} 3 & -1 \\ 3 & -1 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 3 \end{bmatrix}\}.$$

$$\mathcal{S}_{\lambda=1} = \text{nul}(A - I) = \text{nul}(\begin{bmatrix} 1 & -1 \\ 3 & -3 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}.$$

Construimos la matriz P colocando en sus columnas dos autovectores linealmente independientes, por ejemplo: $P := \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix}$ y, si $\Lambda := \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, entonces

$$A = P\Lambda P^{-1},$$

con $P^{-1} = \frac{1}{2}\begin{bmatrix} -1 & 1 \\ 3 & -1 \end{bmatrix}$.

Tal como hicimos en los ejercicios anteriores, proponemos el siguiente cambio de variables, $Y(t) := PZ(t)$. Entonces, $Y'(t) = PZ'(t)$ y el sistema a resolver nos queda:

$$Y'(t) = PZ'(t) = AY(t) + F(t) = APZ(t) + F(t).$$

Multiplicando a izquierda por P^{-1} la ecuación anterior, nos queda:

$$\begin{aligned} Z'(t) &= P^{-1}APZ(t) + P^{-1}F(t) = \Lambda Z(t) + P^{-1}F(t) \\ &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} + \frac{1}{2}\begin{bmatrix} -1 & 1 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} e^{3t} \\ e^{3t} \end{bmatrix} \\ &= \begin{bmatrix} -z_1(t) \\ z_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ e^{3t} \end{bmatrix} = \begin{bmatrix} -z_1(t) \\ z_2(t) + e^{3t} \end{bmatrix}. \end{aligned}$$

Entonces, nos queda el siguiente sistema desacoplado no homogéneo a resolver,

$$\begin{cases} z_1' = -z_1 \\ z_2' = z_2 + e^{3t} \end{cases}.$$

Recordando (por ejemplo) el **Ejercicio 2.27**, tenemos que todas las soluciones de la ecuación diferencial homogénea $z_1' + z_1 = 0$ son

$$z_1(t) = c_1 e^{-t},$$

con $c_1 \in \mathbb{R}$.

Por otro lado, recordando también (por ejemplo) el **Ejercicio 2.27**, tenemos que todas las soluciones de de la ecuación diferencial no homogénea $z_2' - z_2 = g(t)$ con $g(t) = e^{3t}$, son

$$z_2(t) = z_2^h(t) + z_2^p(t),$$

donde z_2^h son las soluciones del sistema homogéneo asociado y z_2^p es una solución particular.

En este caso, $z_2^h \in gen\{e^t\}$ y $z_2^p(t) = f(t)e^t$, donde $f'(t) = g(t)e^{-t} = e^{3t}e^{-t} = e^{2t}$. Entonces, $f(t) = \frac{e^{2t}}{2}$ y tenemos que $z_2^p(t) = \frac{e^{2t}}{2}e^t = \frac{e^{3t}}{2}$. Por lo tanto,

$$z_2(t) = c_2e^t + \frac{e^{3t}}{2},$$

con $c_2 \in \mathbb{R}$.

Volviendo a la variable original, nos queda que todas las soluciones del sistema de ecuaciones diferenciales no homogéneo son:

$$Y(t) = PZ(t) = \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} c_1e^{-t} \\ c_2e^t + \frac{e^{3t}}{2} \end{bmatrix} = c_1e^{-t} \begin{bmatrix} 1 \\ 3 \end{bmatrix} + c_2e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{e^{3t}}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

con $c_1, c_2 \in \mathbb{R}$.

Observar que si llamamos $Y_h(t) := c_1e^{-t} \begin{bmatrix} 1 \\ 3 \end{bmatrix} + c_2e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, con $c_1, c_2 \in \mathbb{R}$ e $Y_p(t) := \frac{e^{3t}}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, entonces,

$$Y(t) = Y_h(t) + Y_p(t),$$

donde Y_h son las soluciones del sistema homogéneo asociado $Y' = AY$ e Y_p es una solución particular, es decir $Y_p'(t) = AY_p(t) + F(t)$. ■

CAPÍTULO 5

Matrices Unitarias

En este capítulo, vamos a tomar como cuerpo $\mathbb{K} = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$ y vamos a considerar el producto interno canónico, definido por $\langle x, y \rangle = y^* x$, para $x, y \in \mathbb{K}^n$.

Sea $A \in \mathbb{K}^{m\times n}$ entonces

$$\langle Ax, y \rangle = y^*(Ax) = (A^*y)^*(x) = \langle x, A^*y \rangle \text{ para todo } x \in \mathbb{K}^n,\ y \in \mathbb{K}^m. \tag{5.1}$$

5.1. Matrices unitarias y ortogonales

Definición. Diremos que $U \in \mathbb{C}^{n\times n}$ es *unitaria* si $UU^* = U^*U = I$. Si $U \in \mathbb{R}^{n\times n}$ y $UU^T = U^TU = I$, diremos que U es *ortogonal*.

Observar que $U \in \mathbb{C}^{n\times n}$ es unitaria, si y sólo si $U^{-1} = U^*$ y $U \in \mathbb{R}^{n\times n}$ es ortogonal, si y sólo si $U^{-1} = U^T$.

Por ejemplo, la matriz $U := \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix}$ es unitaria, pues

$$\begin{aligned} UU^* &= \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} -i & 0 \\ 0 & -i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \\ &= \begin{bmatrix} -i & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix} = U^*U. \end{aligned}$$

La matriz, $P := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ es ortogonal. Pues $P \in \mathbb{R}^{2\times 2}$ y

$$\begin{aligned} PP^T &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = P^TP. \end{aligned}$$

La siguiente propiedad demuestra que si una matriz cuadrada tiene inversa a izquierda o a derecha entonces es inversible:

Sea $A \in \mathbb{K}^{n\times n}$ tal que existe $B \in \mathbb{K}^{n\times n}$ tal que

$$AB = I,$$

entonces

$$BA = I \text{ y } A^{-1} = B.$$

De hecho, si existe $B \in \mathbb{K}^{n\times n}$ tal que $AB = I$, entonces,

$$\det(AB) = \det(A)\det(B) = \det(I) = 1 \neq 0.$$

Por lo tanto, $\det(A) \neq 0$. Entonces, A es de rango máximo. Por lo tanto, existe A^{-1}. Multiplicando a izquierda por A^{-1} tenemos que

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

Entonces $A^{-1} = B$, por lo tanto, también vale que $BA = A^{-1}A = I$.

Eso significa que (por ejemplo) dada $U \in \mathbb{C}^{n\times n}$, basta ver que $UU^* = I$ ó que $U^*U = I$ para concluir que U es unitaria.

En la siguiente proposición veremos varias propiedades que caracterizan a las matrices unitarias (ortogonales).

Proposición 5.1.1. *Sea $U \in \mathbb{C}^{n\times n}$. Entonces, son equivalentes:*

i) *U es unitaria,*

ii) *U^* es unitaria,*

iii) *U es inversible y $U^{-1} = U^*$,*

iv) *U preserva el producto interno (canónico), es decir*

$$\langle Ux, Uy \rangle = \langle x, y \rangle$$

para todo $x, y \in \mathbb{C}^n$.

v) *Si $\{v_1, v_2, \cdots, v_n\}$ es una bon de $\mathbb{C}^n$, entonces $\{Uv_1, Uv_2, \cdots, Uv_n\}$ también es una bon de $\mathbb{C}^n$.*

vi) *Las columnas de U forman una bon de $\mathbb{C}^n$.*

vii) *Las filas de U forman una bon de $\mathbb{C}^n$.*

viii) *U es una* isometría*, es decir*

$$\|Ux\| = \|x\|$$

para todo $x \in \mathbb{C}^n$.

Dem. $i) \Rightarrow ii)$: Si U es unitaria, entonces $UU^* = I$ y $U^*U = I$. Entonces, es claro que $U^*U = I$ y $UU^* = I$, por lo tanto U^* es unitaria.

$ii) \Rightarrow iii)$: Si U^* es unitaria, entonces $UU^* = I$ y $U^*U = I$, eso implica como vimos arriba que $U^{-1} = U^*$.

$iii) \Rightarrow iv)$: Si $U^{-1} = U^*$. Usando (5.1) tenemos que, para todo $x, y \in \mathbb{C}^n$,

$$\langle Ux, Uy \rangle = \langle x, U^*Uy \rangle = \left\langle x, U^{-1}Uy \right\rangle = \langle x, Iy \rangle = \langle x, y \rangle .$$

$iv) \Rightarrow v)$: Si $\{v_1, v_2, \cdots, v_n\}$ es una bon de $\mathbb{C}^n$, entonces $\langle v_i, v_j \rangle = \delta_{ij}$, donde $\delta_{ij} = 1$ si $i = j$ y $\delta_{ij} = 0$ si $i \neq j$ (la *delta de Kronecker*). Usando $iv)$, tenemos que

$$\langle Uv_i, Uv_j \rangle = \langle v_i, v_j \rangle = \delta_{ij}.$$

Por lo tanto, $\{Uv_1, Uv_2, \cdots, Uv_n\}$ también es una bon de $\mathbb{C}^n$.

$v) \Rightarrow vi)$: Supongamos que $U := [u_1\ u_2\ \cdots\ u_n]$, donde $u_i \in \mathbb{C}^n$ es la i-ésima columna de U. Sea $\{e_1, e_2, \cdots, e_n\}$ la base canónica de $\mathbb{C}^n$, que es una base ortonormal de $\mathbb{C}^n$. Entonces, recordemos que

$$Ue_1 = u_1,\ Ue_2 = u_2,\ \cdots,\ Ue_n = u_n.$$

Usando v), tenemos que el conjunto

$$\{Ue_1, Ue_2, \cdots, Ue_n\} = \{u_1, u_2, \cdots, u_n\}$$

es una bon de $\mathbb{C}^n$. Por lo tanto, las columnas de U forman una bon de $\mathbb{C}^n$.

$vi) \Rightarrow vii)$: Supongamos que $U = [u_1\ u_2\ \cdots\ u_n]$, donde $u_i \in \mathbb{C}^n$ es la i-ésima columna de U. Por hipótesis, $\{u_1, u_2, \cdots, u_n\}$ es una bon de $\mathbb{C}^n$, entonces

$$U^*U = \begin{bmatrix} u_1^* \\ u_2^* \\ \vdots \\ u_n^* \end{bmatrix} [u_1\ u_2\ \cdots\ u_n] = \begin{bmatrix} u_1^*u_1 & u_1^*u_2 & \cdots & u_1^*u_n \\ u_2^*u_1 & u_2^*u_2 & \cdots & u_2^*u_n \\ \vdots & \vdots & \vdots & \vdots \\ u_n^*u_1 & u_n^*u_2 & \cdots & u_n^*u_n \end{bmatrix} = I.$$

Entonces, U es unitaria. Tomando conjugado en la ecuación anterior, tenemos que

$$I = \overline{I} = \overline{U^*U} = \overline{U^*}\ \overline{U} = U^T(U^T)^*.$$

Entonces U^T es unitaria. Como ya probamos que $i) \Rightarrow vi)$ (porque probamos $i) \Rightarrow ii) \Rightarrow iii) \Rightarrow iv) \Rightarrow v) \Rightarrow vi)$), tenemos que las columnas de U^T forman una bon de $\mathbb{C}^n$. Como las columnas de U^T son las filas de U, concluimos entonces que las filas de U forman una bon de $\mathbb{C}^n$.

$vii) \Rightarrow viii)$: Si las filas de U forman una bon de $\mathbb{C}^n$ entonces las columnas de U^T (que son las filas de U) forman una bon de $\mathbb{C}^n$. Entonces, operando de manera similar que en $vi) \Rightarrow vii)$, tendremos que vale que $(U^T)^*U^T = \overline{U}U^T = I$. Entonces, $I = I^T = (\overline{U}U^T)^T = UU^*$, entonces U es unitaria. Como ya probamos que $i) \Rightarrow iv)$ (porque probamos $i) \Rightarrow ii) \Rightarrow iii) \Rightarrow iv)$), se sigue en particular que, para todo $x \in \mathbb{C}^n$

$$\|Ux\|^2 = \langle\, Ux, Ux \,\rangle = \langle\, x, x \,\rangle = \|x\|^2.$$

$viii) \Rightarrow i)$: Si $\|Ux\| = \|x\|$ para todo $x \in \mathbb{C}^n$. Entonces,

$$\langle\, Ux, Ux \,\rangle = \langle\, U^*Ux, x \,\rangle = \langle\, Ix, x \,\rangle,$$

para todo $x \in \mathbb{C}^n$. Entonces

$$\langle\, (U^*U - I)x, x \,\rangle = \langle\, U^*Ux - x, x \,\rangle = \langle\, U^*Ux, x \,\rangle - \langle\, x, x \,\rangle = 0,$$

para todo $x \in \mathbb{C}^n$.

Llamemos $A := U^*U - I$. Observar que, si $z, w \in \mathbb{C}^n$ (hacer la cuenta) vale la *fórmula de polarización*:

$$\begin{aligned} \langle\, Az, w \,\rangle &= \langle\, A(z+w), z+w \,\rangle + i\,\langle\, A(z+iw), z+iw \,\rangle \\ &\quad + (-1)\,\langle\, A(z+(-1)w), z+(-1)w \,\rangle + (-i)\,\langle\, A(z+(-i)w), z+(-i)w \,\rangle. \end{aligned}$$

Entonces, como $\langle\, Ax, x \,\rangle = 0$, para todo $x \in \mathbb{C}^n$, vale que $\langle\, Az, w \,\rangle = \langle\, A(z+w), z+w \,\rangle + i\,\langle\, A(z+iw), z+iw \,\rangle + (-1)\,\langle\, A(z+(-1)w), z+(-1)w \,\rangle + (-i)\,\langle\, A(z+(-i)w), z+(-i)w \,\rangle = 0$, para todo $z, w \in \mathbb{C}^n$. Entonces

$$\langle\, (U^*U - I)z, w \,\rangle = 0,$$

para todo $z, w \in \mathbb{C}^n$. En particular, si tomamos $w := (U^*U - I)z$, tenemos que

$$\langle\, (U^*U - I)z, (U^*U - I)z \,\rangle = \|(U^*U - I)z\|^2 = 0,$$

para todo $z \in \mathbb{C}^n$. Entonces

$$(U^*U - I)z = 0,$$

para todo $z \in \mathbb{C}^n$. Por lo tanto $U^*U - I = 0$ ó, equivalentemente, $U^*U = I$ y U resulta unitaria. ■

Ejercicio 5.A. Pensar cómo se modifica la Propoposición 5.1.1, si tenemos U ortogonal en vez de U unitaria.

5.2. Matrices hermíticas y diagonalización unitaria

Definición. Sea $A \in \mathbb{C}^{n\times n}$. Diremos que A es *hermítica* si $A^* = A$ y diremos que A es *simétrica*, si $A \in \mathbb{R}^{n\times n}$ y $A^T = A$. Finalmente, diremos que A es *anti simétrica*, si $A \in \mathbb{R}^{n\times n}$ y $A^T = -A$.

Ejemplo 5.a. La matriz $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ es anti simétrica, pues $A^T = -A$.

La matriz $\begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix}$ es hermítica, pues $A^* = A$.

Proposición 5.2.1. *Sea $A \in \mathbb{C}^{n\times n}$ tal que A es hermítica, es decir $A^* = A$. Entonces:*

1. *Los autovalores de A son reales.*

2. *Autovectores de A correspondientes a distintos autovalores son ortogonales (con el producto interno canónico).*

Dem. 1. : Supongamos que $\lambda \in \mathbb{C}$ es un autovalor de A. Entonces, existe $v \in \mathbb{C}^n \setminus \{0\}$ tal que $Av = \lambda v$. Entonces, usando (5.1),

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Av, v \rangle = \langle v, A^* v \rangle = \langle v, Av \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \langle v, v \rangle .$$

Entonces $(\lambda - \overline{\lambda}) \langle v, v \rangle = 0$. Como $\langle v, v \rangle \neq 0$, pues $v \neq 0$, se sigue que $\lambda = \overline{\lambda}$ y por lo tanto $\lambda \in \mathbb{R}$.

2. : Sean $\lambda, \mu \in \mathbb{R}$ (ya probamos que son reales) autovalores distintos de A y sean $v, w \in \mathbb{C}^n \setminus \{0\}$ los autovectores asociados correspondientes. Entonces, usando (5.1),

$$\lambda \langle v, w \rangle = \langle \lambda v, w \rangle = \langle Av, w \rangle = \langle v, A^* w \rangle = \langle v, Aw \rangle = \langle v, \mu w \rangle = \mu \langle v, w \rangle .$$

Entonces

$$(\lambda - \mu) \langle v, w \rangle = 0.$$

Como $\lambda \neq \mu$, concluimos que $\langle v, w \rangle = 0$ y entonces v y w son ortogonales. ■

Definición. Sea $A \in \mathbb{C}^{n\times n}$. Diremos que A es *diagonalizable unitariamente* si existen $U \in \mathbb{C}^{n\times n}$ unitaria y $\Lambda \in \mathbb{C}^{n\times n}$ diagonal (Λ no tiene por qué ser real) tales que

$$A = U \Lambda U^*.$$

Sea $A \in \mathbb{R}^{n\times n}$. Diremos que A es *diagonalizable ortogonalmente* si existen $U \in \mathbb{R}^{n\times n}$ ortogonal y $\Lambda \in \mathbb{R}^{n\times n}$ diagonal tales que

$$A = U \Lambda U^T.$$

Enunciaremos uno de los teoremas más importantes que veremos en la materia:

Teorema 5.2.2. *Sea $A \in \mathbb{C}^{n\times n}$ hermítica, es decir $A^* = A$ entonces A es diagonalizable unitariamente.*

Dem. Lo demostraremos por inducción en n.

Si $n = 1$, entonces $A \in \mathbb{C}$ (es un número) y no hay nada que probar.

Supongamos que $n > 1$ y que para toda matriz $A \in \mathbb{C}^{n-1\times n-1}$ hermítica vale que A es diagonalizable unitariamente. Veamos que eso mismo vale para $A \in \mathbb{C}^{n\times n}$ hermítica.

Si $A \in \mathbb{C}^{n\times n}$ es hermítica, por la Proposición 5.2.1, existe $\lambda \in \mathbb{R}$ autovalor de A. Sea $v \in \mathbb{C}^n \setminus \{0\}$ un autovector asociado a λ y tomemos $w := \frac{v}{\|v\|}$, entonces w es un vector de norma 1 asociado al mismo autovalor λ. Entonces, observar que

$$w^*(Aw) = w^*(\lambda w) = \lambda w^* w = \lambda \|w\|^2 = \lambda.$$

Ahora, definamos $\mathcal{S} := gen\{w\}^{\perp}$. Entonces, recordemos que

$$\dim(\mathcal{S}) = \dim(gen\{w\}^{\perp}) = \dim(\mathbb{C}^n) - \dim(gen\{w\}) = n - 1.$$

Además, si $s \in \mathcal{S}$, vale que $As \in \mathcal{S}$. De hecho, si $s \in \mathcal{S} = gen\{w\}^{\perp}$, entonces $\langle s, w \rangle = 0$. Por lo tanto, como $A^* = A$,

$$\langle As, w \rangle = \langle s, A^*w \rangle = \langle s, Aw \rangle = \langle s, \lambda w \rangle = \lambda \langle s, w \rangle = 0.$$

Entonces $As \in gen\{w\}^{\perp} = \mathcal{S}$, que era lo que queríamos ver.

A continuación, completemos el conjunto $\{w\}$ de manera tal que $\mathcal{B} := \{w, v_1, \cdots, v_{n-1}\}$ sea una base ortonormal de $\mathbb{C}^n$. Entonces, como $v_1, v_2, \cdots, v_{n-1} \in gen\{w\}^{\perp}$ (son parte de una base ortonormal) por lo que acabamos de probar, vale que $Av_i \in gen\{w\}^{\perp}$ para $i = 1, 2, \cdots, n-1$ y entonces,

$$0 = \langle Av_i, w \rangle = w^*Av_i = \langle w, Av_i \rangle = (Av_i)^*w = v_i^*A^*w = v_i^*Aw$$

para $i = 1, 2, \cdots, n-1$. Sea

$$V := [w\ v_1\ \cdots\ v_{n-1}],$$

(donde $w, v_1, \cdots, v_{n-1}$ son las columnas de V). Entonces, como las columnas de V son una bon de $\mathbb{C}^n$, por la Proposición 5.1.1, V es unitaria. Además, usando lo que vimos arriba, se sigue que

$$\begin{aligned} V^*AV &= \begin{bmatrix} w^* \\ v_1^* \\ \vdots \\ v_{n-1}^* \end{bmatrix} A[\ w\ v_1\ \cdots\ v_{n-1}] = \begin{bmatrix} w^*Aw & w^*Av_1 & \cdots & w^*Av_{n-1} \\ v_1^*Aw & v_1^*Av_1 & \cdots & v_1^*Av_{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ v_{n-1}^*Aw & v_{n-1}^*Av_1 & \cdots & v_{n-1}^*Av_{n-1} \end{bmatrix} \\ &= \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & v_1^*Av_1 & \cdots & v_1^*Av_{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & v_{n-1}^*Av_1 & \cdots & v_{n-1}^*Av_{n-1} \end{bmatrix}. \end{aligned}$$

Llamemos

$$A_1 := \begin{bmatrix} v_1^*Av_1 & \cdots & v_1^*Av_{n-1} \\ \vdots & \vdots & \vdots \\ v_{n-1}^*Av_1 & \cdots & v_{n-1}^*Av_{n-1} \end{bmatrix}.$$

Entonces, por la cuenta anterior, nos quedó la siguiente matriz en bloques:

$$V^*AV = \begin{bmatrix} \lambda & 0 \\ 0 & A_1 \end{bmatrix}.$$

Observar que $A_1 \in \mathbb{C}^{n-1 \times n-1}$ y $A_1^* = A_1$ (hacer la cuenta usando que $A^* = A$). Entonces, por hipótesis inductiva, A_1 es diagonalizable unitariamente. Es decir, existe $U_1 \in \mathbb{C}^{n-1 \times n-1}$ unitaria y $\Lambda_1 \in \mathbb{R}^{n-1 \times n-1}$ (como A_1 es hermítica sus autovalores son reales) diagonal tales que

$$A_1 = U_1 \Lambda_1 U_1^*.$$

Entonces (verificarlo haciendo la cuenta), tenemos que

$$A = V \begin{bmatrix} \lambda & 0 \\ 0 & A_1 \end{bmatrix} V^* = V \begin{bmatrix} 1 & 0 \\ 0 & U_1 \end{bmatrix} \begin{bmatrix} \lambda & 0 \\ 0 & \Lambda_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & U_1^* \end{bmatrix} V^*.$$

Llamemos $U := V \begin{bmatrix} 1 & 0 \\ 0 & U_1 \end{bmatrix} \in \mathbb{C}^{n \times n}$. Entonces, notar que

$$\begin{aligned} U^*U &= (V \begin{bmatrix} 1 & 0 \\ 0 & U_1 \end{bmatrix})^*\ V \begin{bmatrix} 1 & 0 \\ 0 & U_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & U_1^* \end{bmatrix} V^*V \begin{bmatrix} 1 & 0 \\ 0 & U_1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & U_1^* \end{bmatrix} I_n \begin{bmatrix} 1 & 0 \\ 0 & U_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & U_1^*U_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & I_{n-1} \end{bmatrix} = I_n. \end{aligned}$$

Entonces U es unitaria y además $U^* = \begin{bmatrix} 1 & 0 \\ 0 & U_1^* \end{bmatrix} V^*$. Por lo tanto, si llamamos $\Lambda := \begin{bmatrix} \lambda & 0 \\ 0 & \Lambda_1 \end{bmatrix} \in \mathbb{R}^{n\times n}$, tenemos que Λ es diagonal y

$$A = U\Lambda U^*.$$

Entonces, A resulta diagonalizable unitariamente. ■

De manera similar a como probamos el Teorema 5.2.2, podemos probar que si $A \in \mathbb{R}^{n\times n}$ es simétrica, entonces A es diagonalizable ortogonalmente.

Corolario 5.2.3. *Sea $A \in \mathbb{R}^{n\times n}$ (en el sentido de que consideramos $A \in \mathbb{C}^{n\times n}$ con coeficientes reales) antisimétrica, es decir $A = -A^T$. Entonces A es diagonalizable unitariamente.*

Dem. De hecho, observar que

$$(iA)^* = i^* \, (\overline{A})^T = -iA^T = i(-A^T) = iA.$$

Entonces $iA \in \mathbb{C}^{n\times n}$ es hermítica. Entonces, por el Teorema 5.2.2, existe $U \in \mathbb{C}^{n\times n}$ unitaria y $\Lambda \in \mathbb{R}^{n\times n}$ diagonal, tales que

$$iA = U\Lambda U^*.$$

Entonces,

$$A = -i(iA) = (-i)U\Lambda U^* = U(-i\Lambda)U^*$$

y por lo tanto A es diagonalizable unitariamente. Observar que como $(-i\Lambda) \in \mathbb{C}^{n\times n}$ los autovalores de A son números complejos puros. ■

Ejercicio 5.B. Explicar por qué las siguientes matrices son diagonalizables unitariamente y hallar una diagonalización unitaria de las mismas.

- $A_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$,
- $A_2 = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$,
- $A_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,
- $A_4 = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Dem. Observar que

$$A_1^T = (\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix})^T = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = -A_1.$$

Entonces, por lo que acabamos de ver A_1 es diagonalizable unitariamente. De hecho, tenemos que el polinomio característico de A_1 es $\chi_{A_1}(x) = x^2+1$. Por lo tanto $\lambda_1 = i$ y $\lambda_1 = -i$ son los autovalores de A_1 y como $\text{nul}(A_1 - iI) = gen\{\begin{bmatrix} i \\ 1 \end{bmatrix}\}$ y $\text{nul}(A_1 + iI) = gen\{\begin{bmatrix} -i \\ 1 \end{bmatrix}\}$. Entonces $\{\frac{1}{\sqrt{2}}\begin{bmatrix} i \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} -i \\ 1 \end{bmatrix}\}$ es una bon de $\mathbb{C}^2$. Tomemos

$$U := \frac{1}{\sqrt{2}}\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix},$$

entonces U es unitaria y

$$A_1 = U \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} U^*.$$

Observar que

$$A_2 = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 \\ 0 & \cos\theta \end{bmatrix} + \sin\theta \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \cos\theta\ I_2 + \sin\theta\ A_1.$$

Acabamos de ver, que $A_1 = U \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} U^*$, con $U = \frac{1}{\sqrt{2}} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$ unitaria. Entonces,

$$\begin{aligned} A_2 &= \cos\theta\ I_2 + \sin\theta\ A_1 = \cos\theta\ (UU^*) + \sin\theta\ (U \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} U^*) \\ &= U \begin{bmatrix} \cos\theta + i\sin\theta & 0 \\ 0 & \cos\theta - i\sin\theta \end{bmatrix} U^* = U \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix} U^* \end{aligned}$$

y entonces A_2 es diagonalizable unitariamente.

Observar que A_3 se puede escribir como la siguiente matriz en bloques:

$$A_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Acabamos de ver, que $A_1 = U\Lambda U^*$, con $\Lambda = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$ y $U = \frac{1}{\sqrt{2}} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$ unitaria. Entonces, definamos la siguiente matriz en bloques

$$V := \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{C}^{3\times 3}.$$

Observar que $V^* = \begin{bmatrix} U^* & 0 \\ 0 & 1 \end{bmatrix}$ y además

$$V^*V = \begin{bmatrix} U^* & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} U^*U & 0 \\ 0 & 1\cdot 1 \end{bmatrix} = \begin{bmatrix} I_2 & 0 \\ 0 & 1 \end{bmatrix} = I_3.$$

Por lo tanto, V es unitaria. Además,

$$A_3 = \begin{bmatrix} A_1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} U\Lambda U^* & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Lambda & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & 1 \end{bmatrix} = V \begin{bmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 1 \end{bmatrix} V^*$$

y entonces A_3 es diagonalizable unitariamente.

Por último, de manera similar, observar que A_4 se puede escribir como la siguiente matriz en bloques:

$$A_4 = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} A_2 & 0 \\ 0 & 1 \end{bmatrix}.$$

Acabamos de ver, que $A_2 = U\tilde{\Lambda}U^*$, con $\tilde{\Lambda} := \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}$ y $U = \frac{1}{\sqrt{2}} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$ unitaria. Entonces, tomando nuevamente la matriz en bloques unitaria $V = \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix}$. Tenemos que

$$\begin{aligned} A_4 &= \begin{bmatrix} A_2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} U\tilde{\Lambda}U^* & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{\Lambda} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & 1 \end{bmatrix} \\ &= V \begin{bmatrix} e^{i\theta} & 0 & 0 \\ 0 & e^{-i\theta} & 0 \\ 0 & 0 & 1 \end{bmatrix} V^* \end{aligned}$$

y entonces A_4 es diagonalizable unitariamente. ∎

5.3. Clasificación de transformaciones ortogonales

Recordemos que $U \in \mathbb{R}^{n\times n}$ es ortogonal si $UU^T = U^TU = I$.

Si $U \in \mathbb{R}^{n\times n}$ es una matriz ortogonal entonces:

- $|\det(U)| = 1$; luego $\det(U) = \pm 1$.
- Si λ es autovalor de U entonces $|\lambda| = 1$; si además $\lambda \in \mathbb{R}$, $\lambda = \pm 1$.

De hecho, $\det(U^TU) = \det(U^T)\det(U) = \det(U)^2 = \det(I) = 1$. Entonces $|\det(U)| = 1$. Como $U \in \mathbb{R}^{n\times n}$, se sigue que $\det(U) \in \mathbb{R}$ y entonces $\det(U) = \pm 1$.

Tener en cuenta que con $U \in \mathbb{R}^{n\times n}$ nos referimos a $U \in \mathbb{C}^{n\times n}$ con todos sus coeficientes reales. Sea λ un autovalor de U, entonces existe $v \neq 0$, tal que $Uv = \lambda v$. Por la Proposición 5.2.1, $\|Uv\| = \|v\|$. Entonces

$$\|v\| = \|Uv\| = \|\lambda v\| = |\lambda|\|v\|.$$

Entonces $|\lambda| = 1$ pues $v \neq 0$. Si además $\lambda \in \mathbb{R}$, $\lambda = \pm 1$.

Definición. Sea $U \in \mathbb{R}^{n\times n}$ con $n = 2$ ó $n = 3$ y consideremos la transformación lineal $T(x) := Ux$. Entonces:

- Diremos que T es una *rotación* si $\det(U) = 1$.
- Diremos que T es una *simetría ortogonal respecto del subespacio* $\mathcal{S}$ si $T(v) = v$ para todo $v \in \mathcal{S}$ y $T(v) = -v$ para todo $v \in \mathcal{S}^{\perp}$.

Proposición 5.3.1. *Sea $U \in \mathbb{R}^{2\times 2}$ ortogonal y consideremos la transformación lineal $T(x) := Ux$. Entonces, T es una rotación o una simetría.*

Más aún, existe una bon $B = \{v_1, v_2\}$ de $\mathbb{R}^2$, tal que

$$[T]_B^B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

con $\theta \in [0, \pi]$. En ese caso $\det(U) = 1$, $\operatorname{tr}(U) = 2\cos\theta$ y T es una rotación de ángulo θ. Ó, existe una bon $B = \{v_1, v_2\}$ de $\mathbb{R}^2$, tal que

$$[T]_B^B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

En ese caso $\det(U) = -1$, $\operatorname{tr}(U) = 0$ y T es una simetría ortogonal respecto del subespacio $\mathcal{S} = \mathcal{S}_{\lambda=1}$, el autoespacio asociado al autovalor $\lambda = 1$.

Tener en cuenta que con $U \in \mathbb{R}^{2\times 2}$ nos referimos a $U \in \mathbb{C}^{2\times 2}$ con todos sus coeficientes reales. Además, recordar que si E es la base canónica de $\mathbb{R}^2$, vale que $[T]_E^E = U$.

Dem. Sea $B = \{v_1, v_2\}$ una bon de $\mathbb{R}^2$. Entonces, existen $a, b, a', b' \in \mathbb{R}$ tales que

$$Uv_1 = av_1 + bv_2 \text{ y } Uv_2 = a'v_1 + b'v_2.$$

Por la Proposición 5.2.1, $\{Uv_1, Uv_2\}$ es una bon de $\mathbb{R}^2$. Entonces

$$a^2 + b^2 = \langle Uv_1, Uv_1 \rangle = 1, \quad a'^2 + b'^2 = \langle Uv_2, Uv_2 \rangle = 1, \quad aa' + bb' = \langle Uv_1, Uv_2 \rangle = 0.$$

Despejando, nos queda que $a^2 + b^2 = 1$ y que $(a', b') = (-b, a)$ ó $(a', b') = (b, -a)$. Entonces,

1. $[T]_B^B = [[T(v_1)]^B \ [T(v_2)]^B] = [[Uv_1]^B \ [Uv_2]^B] = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ ó
2. $[T]_B^B = [[T(v_1)]^B \ [T(v_2)]^B] = [[Uv_1]^B \ [Uv_2]^B] = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$.

Recordar que
$$U = [T]_E^E = [M]_B^E [T]_B^B [M]_E^B = [M]_B^E [T]_B^B ([M]_B^E)^{-1},$$
donde $[M]_B^E$ es la matriz de cambio de base de B a E. Entonces

$$\chi_U(x) = \det(U - xI) = \det([M]_B^E([T]_B^B - xI)([M]_B^E)^{-1}) = \det([T]_B^B - xI).$$

Caso 1: En este caso $\chi_U(x) = \det([T]_B^B - xI) = (a-x)^2 + b^2 = x^2 - 2ax + a^2 + b^2 = x^2 - 2ax + 1$, con discriminante $\Delta = 4a^2 - 4 = 4(a^2-1) = -4b^2$. Si $a = \pm 1$ entonces, $b^2 = 1 - a^2 = 0$ y se sigue que $[T]_B^B = \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Si no, χ_U no tiene raíces reales.

Por otro lado, como $a^2 + b^2 = 1$, existe $\theta \in [0, 2\pi)$ tal que $a = \cos\theta, b = \sin\theta$. Luego, $[T]_B^B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Eventualmente, cambiando la base $\{v_1, v_2\}$ por $\{v_1, -v_2\}$ (que también es una bon de $\mathbb{R}^2$) podemos tomar $\theta \in [0, \pi]$.

En este caso, $U = [M]_B^E \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} ([M]_B^E)^{-1}$. Entonces

$$\begin{aligned} \det(U) &= \det([M]_B^E [T]_B^B ([M]_B^E)^{-1}) = \det([M]_B^E) \det([M]_E^B)^{-1} \det([T]_B^B) \\ &= \det \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \cos\theta^2 + \sin\theta^2 = 1. \end{aligned}$$

Además,

$$\begin{aligned} \operatorname{tr}(U) &= \operatorname{tr}([M]_B^E [T]_B^B ([M]_B^E)^{-1}) = \operatorname{tr}([T]_B^B ([M]_B^E)^{-1} [M]_B^E) \\ &= \operatorname{tr}([T]_B^B) = \operatorname{tr}(\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}) = 2\cos\theta. \end{aligned}$$

Caso 2: En este caso $\chi_U(x) = \det([T]_B^B - xI) = (a-x)(-a+x) - b^2 = x^2 - a^2 - b^2 = x^2 - 1$, entonces los autovalores de U son $\lambda_1 = 1$ y $\lambda_2 = -1$. Sea $w_1 \in \mathcal{S}_{\lambda=1}$ tal que $\|w_1\| = 1$ y sea $w_2 \in \mathcal{S}_{\lambda=-1}$ tal que $\|w_2\| = 1$. Entonces $B' := \{w_1, w_2\}$ es una bon de $\mathbb{R}^2$ tal que $[T]_{B'}^{B'} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Operando como en el caso anterior, nos queda que $\det(U) = \det([T]_B^B) = -1$ y $\operatorname{tr}(U) = \operatorname{tr}([T]_B^B) = 1 - 1 = 0$. ■

Orientación del ángulo de rotación.
Sea $U \in \mathbb{R}^{2\times 2}$ una matriz ortogonal y sea $T(x) = Ux$ la isometría asociada. Como vimos en la Proposición 5.3.1, si tenemos una rotación, sabiendo $\operatorname{tr}(U)$ podemos determinar el ángulo de rotación θ de manera que $\theta \in [0, \pi]$. Sólo nos resta determinar si el ángulo se toma con signo positivo o negativo, dependiendo de la orientación que estemos tomando. Para determinar el ángulo de rotación y su signo (respecto de la base canónica de $\mathbb{R}^2$), tomamos la base $B' := \{e_1, Ue1\}$ donde $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Calculamos $\det([M]_{B'}^E)$, donde $[M]_{B'}^E$ es la matriz de cambio de base de B' a E (la base canónica de $\mathbb{R}^2$). Si $\det([M]_{B'}^E) > 0$ entonces $\theta = +\arccos(\frac{\operatorname{tr}(U)}{2})$. Si $\det([M]_{B'}^E) < 0$ entonces $\theta = -\arccos(\frac{\operatorname{tr}(U)}{2})$.

Ejercicio 5.1. Comprobar que las siguientes matrices $U \in \mathbb{R}^{2\times 2}$ son ortogonales

- $U_1 = \frac{1}{5} \begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix}$;
- $U_2 = \frac{1}{5} \begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix}$.

En cada caso caracterizar la isometría $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T_i(x) := U_i x$ con $i = 1, 2$.

Dem. Observar que en todos los casos las columnas de cada matriz U_1 y U_2 forman una base ortonormal de $\mathbb{R}^2$. Por lo tanto, por la Propoposición 5.1.1, las matrices son ortogonales.

Consideremos la isometría definida por $T_1(x) := U_1 x$. Como $\det(U_1) = \det(\frac{1}{5}\begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix}) = \frac{1}{25}(9+16) = 1$, por la Proposición 5.3.1, T_1 define una rotación de ángulo $\theta = \arccos(\frac{\operatorname{tr}(U_1)}{2}) = \arccos(\frac{3}{5})$ ó $\theta = -\arccos(\frac{3}{5})$.

Para determinar el signo de θ (respecto de la base canónica de $\mathbb{R}^2$), tomamos la base $B' := \{e_1, U_1 e1\}$ donde $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $U_1 e_1 = \frac{1}{5}\begin{bmatrix} 3 \\ -4 \end{bmatrix}$. Calculamos

$$\det([M]_{B'}^{E}) = \det \begin{bmatrix} 1 & \frac{3}{5} \\ 0 & -\frac{4}{5} \end{bmatrix} = -\frac{4}{5} < 0.$$

Entonces $\theta = -\arccos(\frac{3}{5})$.

Consideremos la isometría definida por $T_2(x) := U_2 x$. Como $\det(U_2) = -1$, por la Proposición 5.3.1, T_2 es una simetría ortogonal respecto del subespacio

$$\mathcal{S}_{\lambda=1} = \operatorname{nul}(U_2 - I) = \operatorname{nul}(\begin{bmatrix} -\frac{1}{5} & \frac{3}{5} \\ \frac{3}{5} & \frac{-9}{5} \end{bmatrix}) = gen\{\begin{bmatrix} 3 \\ 1 \end{bmatrix}\}.$$

■

Proposición 5.3.2. *Sea $U \in \mathbb{R}^{3\times 3}$ ortogonal y consideremos la transformación lineal $T(x) := Ux$.*

Entonces, $\lambda = 1$ ó $\lambda = -1$ es una autovalor de U y T es una rotación, una simetría o una composición de ambas.

Más aún, si $\lambda = 1$ es un autovalor de U, existe una bon $B = \{v_1, v_2, v_3\}$ de $\mathbb{R}^3$, tal que

$$[T]_B^B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix},$$

con $\theta \in [0, \pi]$. En ese caso $\det(U) = 1$, $\operatorname{tr}(U) = 1 + 2\cos\theta$ y T es una rotación de ángulo θ con eje de rotación dado por el subespacio $\mathcal{S}_{\lambda=1}$.

Ó, existe una bon $B = \{v_1, v_2, v_3\}$ de $\mathbb{R}^3$, tal que

$$[T]_B^B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

En ese caso $\det(U) = -1$, $\operatorname{tr}(U) = 1$ y T es una simetría ortogonal respecto del subespacio $\mathcal{S} = \mathcal{S}_{\lambda=1}$ el autoespacio asociado al autovalor $\lambda = 1$.

Si $\lambda = 1$ no es autovalor de U entonces $\lambda = -1$ es un autovalor de U y existe una bon $B = \{v_1, v_2, v_3\}$ de $\mathbb{R}^3$, tal que

$$[T]_B^B = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix},$$

con $\theta \in [0, \pi]$. En ese caso $\det(U) = -1$, $\operatorname{tr}(U) = -1 + 2\cos\theta$ y T es la composición de una rotación de ángulo θ con eje de rotación $\mathcal{S}_{\lambda=-1}$ y una simetría respecto de $\mathcal{S}_{\lambda=-1}^{\perp}$.

Dem. El polinomio característico de U es un polinomio de grado 3 con coeficientes reales, es decir $\chi_U \in \mathbb{R}_3[x]$. Por el Teorema Fundamental del Álgebra, χ_U tiene 3 raíces y como los coeficientes de χ_U son reales hay dos opciones: las 3 raíces de χ_U son reales ó el polinomio característico χ_U tiene

1 raíz real y dos raíces complejas conjugadas. Conclusión, χ_U al menos tiene una raíz $\lambda_1 \in \mathbb{R}$. Por lo tanto, U tiene al menos un autovalor $\lambda_1 \in \mathbb{R}$ y, como vimos que $|\lambda_1| = 1$, se sigue que $\lambda_1 = 1$ ó $\lambda_1 = -1$.

Supongamos que $\lambda_1 = 1$. Sea $B = \{v_1, v_2, v_3\}$ una bon de $\mathbb{R}^3$ tal que $Uv_1 = v_1$, es decir con v_1 un autovector asociado a $\lambda_1 = 1$ (que ya vimos que es un autovalor posible). Como $\{v_1, v_2, v_3\}$ es una bon de $\mathbb{R}^3$, tenemos que $gen\{v_2, v_3\} = gen\{v_1\}^{\perp}$. Entonces, usando que $v_1 = Uv_1$, tenemos que

$$\langle U(v_2), v_1 \rangle = \langle Uv_2, Uv_1 \rangle = \left\langle v_2, U^T Uv_1 \right\rangle = \langle v_2, v_1 \rangle = 0.$$

Entonces $U(v_2) \in gen\{v_1\}^{\perp} = gen\{v_2, v_3\}$. De la misma manera,

$$\langle U(v_3), v_1 \rangle = \langle Uv_3, Uv_1 \rangle = \left\langle v_3, U^T Uv_1 \right\rangle = \langle v_3, v_1 \rangle = 0.$$

Entonces $U(v_3) \in gen\{v_1\}^{\perp} = gen\{v_2, v_3\}$.

Por lo tanto, existen $a, b, c, d \in \mathbb{R}$ tales que $U(v_2) = av_2 + bv_3$ y $U(v_3) = a'v_2 + b'v_3$. Entonces, como U es ortogonal, por la Proposición 5.1.1 y usando que $\langle v_i, v_j \rangle = \delta_{ij}$, tenemos que

$$1 = \langle v_2, v_2 \rangle = \langle Uv_2, Uv_2 \rangle = \langle av_2 + bv_3, av_2 + bv_3 \rangle = a^2 + b^2,$$

$$1 = \langle v_3, v_3 \rangle = \langle Uv_3, Uv_3 \rangle = \langle a'v_2 + b'v_3, a'v_2 + b'v_3 \rangle = a'^2 + b'^2 \text{ y}$$

$$0 = \langle v_2, v_3 \rangle = \langle Uv_2, Uv_3 \rangle = \langle av_2 + bv_3, a'v_2 + b'v_3 \rangle = ab' + bb'.$$

Entonces, despejando nos queda que $a = b'$ y $a' = -b$ ó $a = -b'$ y $a' = b$. Entonces, reemplazando, nos queda que $U(v_3) = a'v_2 + b'v_3 = -bv_2 + av_3$ ó $U(v_3) = a'v_2 + b'v_3 = bv_2 - av_3$.

Por lo tanto,

$$[T]_B^B = [[U(v_1)]^B \ [U(v_2)]^B \ [U(v_3)]^B] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{bmatrix}$$

ó

$$[T]_B^B = [[U(v_1)]^B \ [U(v_2)]^B \ [U(v_3)]^B] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & b & -a \end{bmatrix}.$$

Entonces, de la misma manera que hicimos en la Proposición 5.3.1, tenemos uno de los siguientes casos

- $[T]_B^B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$, para cierto $\theta \in [0, \pi]$. En este caso, $\det(U) = 1$, $\operatorname{tr}(U) = 1 + 2\cos\theta$ y T es una rotación de ángulo θ con eje de rotación dado por el subespacio $\mathcal{S}_{\lambda=1}$.
- $[T]_B^B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. En este caso, $\det(U) = -1$, $\operatorname{tr}(U) = 1$ y T es una simetría respecto del subespacio $\mathcal{S} = \mathcal{S}_{\lambda=1}$ el autoespacio asociado al autovalor $\lambda = 1$.

Finalmente, supongamos que $\lambda_1 = -1$ y 1 no es autovalor de U. Sea $B = \{v_1, v_2, v_3\}$ una bon de $\mathbb{R}^3$ tal que $Uv_1 = -v_1$, es decir con v_1 un autovector asociado a $\lambda_1 = -1$ (que ya vimos que es un autovalor posible). Como $\{v_1, v_2, v_3\}$ es una bon de $\mathbb{R}^3$, de manera anáologa a lo hecho en el caso anterior, se sigue que, para cierto $\theta \in [0, \pi]$,

$$[T]_B^B = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

y tenemos que T es una rotación compuesta con una simetría. En este caso, $\det(U) = -1$, $\operatorname{tr}(U) = -1 + 2\cos\theta$ y T es la composición de una rotación de ángulo θ con eje de rotación $\mathcal{S}_{\lambda=-1}$ y una simetría respecto de $\mathcal{S}_{\lambda=-1}^{\perp}$. ■

Orientación del ángulo de rotación.
Sea $U \in \mathbb{R}^{3\times 3}$ una matriz ortogonal y sea $T(x) = Ux$ la isometría asociada. Como vimos en la Proposición 5.3.2, si tenemos una rotación o una rotación compuesta con una simetría, sabiendo $\text{tr}(U)$ podemos determinar el ángulo de rotación θ de manera que $\theta \in [0, \pi]$. Sólo nos resta determinar si el ángulo se toma con signo positivo o negativo, dependiendo de la orientación que estemos tomando. Si T es una rotación ($\det(U) = 1$), para determinar el ángulo de rotación y su signo (respecto de la base canónica de $\mathbb{R}^3$), tomamos la base $B' := \{v_1, v_2, Uv_2\}$ donde v_1 es un autovector asociado a $\lambda = 1$ y v_2 es un vector ortogonal a v_1. Calculamos $\det([M]^E_{B'})$, donde $[M]^E_{B'}$ es la matriz de cambio de base de B' a E (la base canónica de $\mathbb{R}^3$). Si $\det([M]^E_{B'}) > 0$ entonces $\theta = +\arccos(\frac{\text{tr}(U)-1}{2})$. Si $\det([M]^E_{B'}) < 0$ entonces $\theta = -\arccos(\frac{\text{tr}(U)-1}{2})$.
Por otro lado, si T es una rotación compuesta con una simetría ($\det(U) = -1$), para determinar el ángulo de rotación y su signo (respecto de la base canónica de $\mathbb{R}^3$), tomamos la base $B' := \{v_1, v_2, Uv_2\}$ donde v_1 es un autovector asociado a $\lambda = -1$, y v_2 es un vector ortogonal a v_1. Calculamos $\det([M]^E_{B'})$, donde $[M]^E_{B'}$ es la matriz de cambio de base de B' a E (la base canónica de $\mathbb{R}^3$). Si $\det([M]^E_{B'}) > 0$ entonces $\theta = +\arccos(\frac{\text{tr}(U)+1}{2})$. Si $\det([M]^E_{B'}) < 0$ entonces $\theta = -\arccos(\frac{\text{tr}(U)+1}{2})$.

Ejercicio 5.2. Comprobar que las siguientes matrices $U \in \mathbb{R}^{3\times 3}$ son ortogonales

- $U_1 = \frac{1}{7}\begin{bmatrix} 2 & -6 & 3 \\ -6 & -3 & -2 \\ 3 & -2 & -6 \end{bmatrix}$.
- $U_2 = \frac{1}{7}\begin{bmatrix} -2 & 6 & 3 \\ 6 & 3 & -2 \\ -3 & 2 & -6 \end{bmatrix}$.
- $U_3 = \frac{1}{7}\begin{bmatrix} 2 & 3 & -6 \\ -6 & -2 & -3 \\ 3 & -6 & -2 \end{bmatrix}$.

En cada caso caracterizar la isometría $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T_i(x) := U_i x$ con $i = 1, 2, 3$.

Dem. Observar que en todos los casos las columnas de cada matriz U forman una bon de $\mathbb{R}^3$. Por lo tanto, por la Propoposición 5.1.1, las matrices son ortogonales.

Consideremos $T_1(x) := U_1 x$. Calculamos el polinomio característico de U_1 :

$$\chi_{U_1}(x) = \det(U_1 - xI) = \det \begin{bmatrix} \frac{2}{7} - x & \frac{-6}{7} & \frac{3}{7} \\ \frac{-6}{7} & \frac{-3}{7} - x & \frac{-2}{7} \\ \frac{3}{7} & \frac{-2}{7} & \frac{-6}{7} - x \end{bmatrix} = -x^3 - x^2 + x + 1.$$

Los autovalores de A son las raíces de χ_{U_1} y son $\lambda_1 = 1, \lambda_{2,3} = -1$. Entonces, $\det(U_1) = \lambda_1\lambda_2\lambda_3 = 1$. Por lo tanto, por la Propoposición 5.3.2, T_1 es una rotación. Para obtener el eje de rotación, calculamos:

$$\mathcal{S}_{\lambda=1} = \text{nul}(U_1 - I) = \text{nul}(\begin{bmatrix} \frac{-5}{7} & \frac{-6}{7} & \frac{3}{7} \\ \frac{-6}{7} & \frac{-10}{7} & \frac{-2}{7} \\ \frac{3}{7} & \frac{-2}{7} & \frac{-13}{7} \end{bmatrix}) = gen\{\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}\}.$$

Entonces el eje de rotación es el subespacio $gen\{\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}\}$. Finalmente, como $\mathrm{tr}(U_1) = -1$, el ángulo de rotación es $\theta = \arccos(\frac{\mathrm{tr}(U_1)-1}{2}) = \arccos(-1) = \pi$. En este caso, también podemos decir que T_1 es la simetría ortogonal respecto de la recta $\mathcal{S}_{\lambda=1} = gen\{\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}\}$.

Consideremos $T_2(x) := U_2 x$. Calculamos el polinomio característico de U_2 :

$$\chi_{U_2}(x) = \det(U_2 - xI) = \det \begin{bmatrix} \frac{-2}{7} - x & \frac{6}{7} & \frac{3}{7} \\ \frac{6}{7} & \frac{3}{7} - x & \frac{-2}{7} \\ \frac{-3}{7} & \frac{2}{7} & \frac{-6}{7} - x \end{bmatrix} = -x^3 - \frac{5}{7}x^2 + \frac{5}{7}x + 1.$$

Los autovalores de A son las raíces de χ_{U_2} y son $\lambda_1 = 1, \lambda_{2,3} = \frac{1}{7}(-6 \pm i\sqrt{13})$. Entonces, $\det(U_2) = \lambda_1\lambda_2\lambda_3 = 1$. Por lo tanto, por la Propoposición 5.3.2, T_2 es una rotación. Para obtener el eje de rotación, calculamos:

$$\mathcal{S}_{\lambda=1} = \mathrm{nul}(U_2 - I) = \mathrm{nul}(\begin{bmatrix} \frac{-9}{7} & \frac{6}{7} & \frac{3}{7} \\ \frac{6}{7} & \frac{-4}{7} & \frac{-2}{7} \\ \frac{-3}{7} & \frac{2}{7} & \frac{-13}{7} \end{bmatrix}) = gen\{\begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}\}.$$

Entonces el eje de rotación es el subespacio $gen\{\begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}\}$. Finalmente, como $\mathrm{tr}(U_2) = -\frac{5}{7}$, el ángulo de rotación es $\theta = \arccos(\frac{\mathrm{tr}(U_2)-1}{2}) = +\arccos(-\frac{6}{7})$ ó $\theta = -\arccos(-\frac{6}{7})$. Para determinar el signo de θ (respecto de la base canónica de $\mathbb{R}^3$), tomamos la base $B' := \{v_1, v_2, U_2v_2\}$ donde $v_1 = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$ es un autovector asociado a $\lambda = 1$, $v_2 = \begin{bmatrix} 3 \\ -2 \\ 0 \end{bmatrix}$ es un vector ortogonal a v_1 y $U_2v_2 = \frac{1}{7}\begin{bmatrix} -18 \\ 12 \\ -13 \end{bmatrix}$. Calculamos $\det([M]_{B'}^E)$, donde $[M]_{B'}^E$ es la matriz de cambio de base de B' a E (la base canónica de $\mathbb{R}^3$). Luego

$$\det[M]_{B'}^E = \det \begin{bmatrix} 2 & 3 & -\frac{18}{7} \\ 3 & -2 & \frac{12}{7} \\ 0 & 0 & -\frac{13}{7} \end{bmatrix} = \frac{169}{7} > 0.$$

Entonces $\theta = \arccos(-\frac{6}{7})$.

Consideremos $T_3(x) := U_3 x$. Calculamos el polinomio característico de U_3 :

$$\chi_{U_3}(x) = \det(U_3 - xI) = \det \begin{bmatrix} \frac{2}{7} - x & \frac{3}{7} & \frac{-6}{7} \\ \frac{-6}{7} & \frac{-2}{7} - x & \frac{-3}{7} \\ \frac{3}{7} & \frac{-6}{7} & \frac{-2}{7} - x \end{bmatrix} = -x^3 - \frac{2}{7}x^2 - \frac{2}{7}x - 1.$$

Los autovalores de A son las raíces de χ_{U_3} y son $\lambda_1 = -1, \lambda_{2,3} = \frac{1}{14}(5 \pm i3\sqrt{19})$. Entonces, $\det(U_3) = \lambda_1\lambda_2\lambda_3 = -1$. Por lo tanto, por la Propoposición 5.3.2, T_3 es una rotación compuesta por una simetría. Para obtener el eje de rotación, calculamos:

$$\mathcal{S}_{\lambda=-1} = \mathrm{nul}(U_3 + I) = \mathrm{nul}(\begin{bmatrix} \frac{9}{7} & \frac{3}{7} & \frac{-6}{7} \\ \frac{-6}{7} & \frac{5}{7} & \frac{-3}{7} \\ \frac{3}{7} & \frac{-6}{7} & \frac{5}{7} \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}\}.$$

Entonces T_3 es una rotación compuesta por una simetría: el eje de rotación es el subespacio $gen\{\begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}\}$ y, como $\mathrm{tr}(U_3) = -\frac{2}{7}$, el ángulo de rotación es $\theta = \arccos(\frac{\mathrm{tr}(U_3)+1}{2}) = \arccos(\frac{5}{14})$

ó $\theta = -\arccos(\frac{5}{14})$. Por otra parte, la simetría es respecto del subespacio $gen\{\begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}\}^{\perp} = gen\{\begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\}$. Para determinar el signo de θ (respecto de la base canónica de $\mathbb{R}^3$), tomamos la base $B' := \{v_1, v_2, U_3 v_2\}$ donde $v_1 = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}$ es un autovector asociado a $\lambda = -1$, $v_2 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}$ es un vector ortogonal a v_1 y $U_3 v_2 = \frac{1}{7}\begin{bmatrix} -3 \\ 16 \\ -15 \end{bmatrix}$. Calculamos $\det([M]_{B'}^{E})$, donde $[M]_{B'}^{E}$ es la matriz de cambio de base de B' a E (la base canónica de $\mathbb{R}^3$). Luego

$$\det[M]_{B'}^{E} = \det \begin{bmatrix} 1 & -3 & -\frac{3}{7} \\ 3 & 1 & \frac{16}{7} \\ 3 & 0 & -\frac{15}{7} \end{bmatrix} = -\frac{285}{7} < 0.$$

Entonces $\theta = -\arccos(\frac{5}{14})$. ■

Ejercicio 5.3.

a) Hallar la matriz de rotación de ángulo $\frac{\pi}{3}$ alrededor del eje generado por $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

b) Hallar la simetría ortogonal respecto del subespacio $\mathcal{S} = gen\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\}$.

Dem. *a*): Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la rotación de ángulo $\frac{\pi}{3}$ alrededor del eje generado por $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Sea E la base canónica de $\mathbb{R}^3$, el ejercicio nos pide $U := [T]_E^E$. Una forma de obtener la matriz U es primero obtener una bon de $\mathbb{R}^3$ que contenga al eje de rotación. Para eso, si $\mathcal{S} := gen\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\}$ entonces $\mathcal{S}^{\perp} = gen\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\}$. Ahora, obtenemos una base ortogonal de $\mathcal{S}^{\perp}$ usando Gram-Schmidt. En ese caso

$$w_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - \frac{\left\langle \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\rangle}{\| \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \|^2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -1 \end{bmatrix}.$$

Por lo tanto $\{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}\}$ es una bon de $\mathcal{S}^{\perp}$ y

$$B := \{\frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}\}$$

es una bon de $\mathbb{R}^3$ que contiene al eje de rotación. Ahora veamos que la base B está bien orientada respecto de la base canónica, para eso calculamos el determinante de la matriz de cambio de base $[M]_B^E$:

$$\det([M]_B^E) = \det \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{bmatrix} = 1 > 0.$$

Como $\det([M]_B^E) = 1$ (es positivo) la base B está bien orientada, si nos hubiera dado que $\det([M]_B^E) = -1$ entonces basta con cambiarle el signo a uno de los vectores de la base para obtener una base bien orientada. Por la Proposición 5.3.2, tenemos que como $\theta = \frac{\pi}{3}$, entonces

$$[T]_B^B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}.$$

Para obtener $U = [T]_E^E$ solo nos resta realizar un cambio de base. Observar que como B es una bon de $\mathbb{R}^3$ entonces la matriz de cambio de base $[M]_B^E$ es ortogonal, por lo tanto $[M]_E^B = ([M]_B^E)^{-1} = ([M]_E^B)^T$. Entonces

$$\begin{aligned} U = [T]_E^E &= [M]_B^E[T]_B^B[M]_E^B = [M]_B^E[T]_B^B([M]_B^E)^T \\ &= \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ 1 & 2 & 2 \end{bmatrix}. \end{aligned}$$

b): Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la simetría ortogonal respecto del subespacio $\mathcal{S} = gen\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \}$.

Sea E la base canónica de $\mathbb{R}^3$, el ejercicio nos pide $U := [T]_E^E$. Una forma de obtener la matriz U es primero obtener una bon de $\mathbb{R}^3$ que contenga los generadores de $\mathcal{S}$. En la parte *a*) vimos que

$$B := \{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \}$$

es una bon de $\mathbb{R}^3$ que contiene a los generadores de $\mathcal{S}$. Como T es la simetría ortogonal respecto de $\mathcal{S}$, se sigue que

$$\begin{aligned} T(\frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}) &= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \\ T(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}) &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \\ T(\frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}) &= -\frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}. \end{aligned}$$

Entonces, es claro que

$$[T]_B^E = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix}.$$

Para obtener $U = [T]_E^E$ solo nos resta realizar un cambio de base. Observar que como B es una bon de $\mathbb{R}^3$, la matriz de cambio de base $[M]_B^E$ es ortogonal, por lo tanto $[M]_E^B = ([M]_B^E)^{-1} = ([M]_E^B)^T$. Entonces

$$U = [T]_E^E = [T]_B^E [M]_E^B = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & -1 & 2 \\ -1 & 2 & 2 \\ 2 & 2 & -1 \end{bmatrix}.$$

■

Ejercicio 5.5. Sea $A \in \mathbb{R}^{3\times 3}$ tal que $v_1 = [1\ 1\ 1]^T, v_2 = [1\ -1\ 0]^T,\ v_3 = [0\ 1\ -1]^T$, es una base de autovectores de A asociados a los autovalores $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, respectivamente. Probar que A es simétrica si y sólo si $\lambda_2 = \lambda_3$.

Dem. Supongamos que A es simétrica. Entonces $A = A^T$. Entonces, usando que $Av_2 = \lambda_2 v_2$, $Av_3 = \lambda_3 v_3$ y que $\lambda_2, \lambda_3 \in \mathbb{R}$ se sigue que

$$\lambda_2 \langle v_2, v_3 \rangle = \langle \lambda_2 v_2, v_3 \rangle = \langle Av_2, v_3 \rangle = \langle v_2, A^T v_3 \rangle = \langle v_2, Av_3 \rangle = \langle v_2, \lambda_3 v_3 \rangle = \lambda_3 \langle v_2, v_3 \rangle .$$

Entonces, como $\langle v_2, v_3 \rangle = v_3^T v_2 = [0\ 1\ -1] \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = -1$, tenemos que

$$0 = (\lambda_2 - \lambda_3) \langle v_2, v_3 \rangle = (\lambda_2 - \lambda_3)(-1) = \lambda_3 - \lambda_2.$$

Por lo tanto $\lambda_2 = \lambda_3$.

Recíprocamente, supongamos que $\lambda_2 = \lambda_3$. Entonces, como $\{v_1, v_2, v_3\}$ es una base $\mathbb{R}^3$ y además, para todo $a, b, c \in \mathbb{R}$,

$$\begin{aligned} \langle av_1, bv_2 + cv_3 \rangle &= ab \langle v_1, v_2 \rangle + ac \langle v_1, v_3 \rangle = ab(v_1^T v_2) + ac(v_1^T v_2) \\ &= ab\ [1\ 1\ 1] \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + ac\ [1\ 1\ 1] \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = 0. \end{aligned}$$

Tenemos que

$$gen\{v_1\} = gen\{v_2, v_3\}^\perp$$

y además, $\mathcal{S}_{\lambda=\lambda_1} = gen\{v_1\}$ y $\mathcal{S}_{\lambda=\lambda_2} = gen\{v_2, v_3\}$. Busquemos bases ortonormales de ambos autoespacios. Por un lado, $\{\frac{1}{\sqrt{3}}[1\ 1\ 1]^T\}$ es una bon de $\mathcal{S}_{\lambda=\lambda_1}$ y, usando el algoritmo de Gram-Schmid, podemos obtener una bon del autoespacio $\mathcal{S}_{\lambda=\lambda_2}$. De hecho, tomamos $u_2 = v_2 = [1\ -1\ 0]^T$ y

$$u_3 = v_3 - \frac{\langle u_2, v_3 \rangle}{\|u_2\|^2} u_2 = [0\ 1\ -1]^T - \frac{-1}{2}[1\ -1\ 0]^T = [\frac{1}{2}\ \frac{1}{2} - 1]^T.$$

Entonces $\{\frac{1}{\sqrt{2}}[1\ -1\ 0]^T, \frac{1}{\sqrt{6}}[1\ 1 - 2]^T\}$ es una bon de $\mathcal{S}_{\lambda=\lambda_2}$.

Tomemos $U := \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{bmatrix}$. Entonces, como las columnas de U son una bon de $\mathbb{R}^3$, por la Proposición 5.1.1, U es una matriz ortogonal y además, tenemos que

$$A = U \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} U^T.$$

Entonces

$$\begin{aligned} A^T &= (U \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} U^T)^T = (U^T)^T (\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix})^T U^T \\ &= U \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} U^T = A \end{aligned}$$

y A resulta simétrica. ■

Antes de resolver el próximo ejercicio, recordemos que si $(\mathbb{V}, \|\cdot\|)$ es un espacio normado y $(v_n)_{n\in\mathbb{N}}$ una sucesión de elementos de $\mathbb{V}$, decimos que la sucesión $(v_n)_{n\in\mathbb{N}}$ es *convergente* si existe $v_0 \in \mathbb{V}$ tal que

$$\lim_{n\to\infty} \|v_n - v_0\| = 0,$$

ó, equivalentemente, si para cada $\varepsilon > 0$ existe $N_0 \in \mathbb{N}$ tal que

$$\|v_n - v_0\| < \varepsilon,$$

para todo $n \geq N_0$. En este caso, escribimos

$$\lim_{n\to\infty} v_n = v_0.$$

Además, sea $A \in \mathbb{K}^{n\times n}$ recordemos que si $\lambda \in \mathbb{K}$ es un autovalor de A con autovector asociado $v \in \mathbb{K}^n$ entonces

$$A^k v = \lambda^k v \tag{5.2}$$

Ejercicio 5.8. Para $A = \frac{1}{18} \begin{bmatrix} 13 & -2 & -4 \\ -2 & 10 & 2 \\ -4 & 2 & 13 \end{bmatrix}$,

a) Hallar $\{x \in \mathbb{R}^3 : \lim_{k\to\infty} A^k x = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\}$.

b) Comprobar que $\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} \in \{\lim_{k\to\infty} A^k x : x \in \mathbb{R}^3\}$ y hallar todas las soluciones de la ecuación $\lim_{k\to\infty} A^k x = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$.

Dem. *a*) : Primero, calculamos el polinomio característico de A :

$$\begin{aligned} \chi_A(x) = \det(A - xI) &= \det(\frac{1}{18} \begin{bmatrix} 13-x & -2 & -4 \\ -2 & 10-x & 2 \\ -4 & 2 & 13-x \end{bmatrix}) = \\ &= -x^3 + 2x^2 - \frac{5x}{4} + \frac{1}{4}. \end{aligned}$$

Los autovalores de A son las raíces de χ_A y son $\lambda_1 = 1, \lambda_{2,3} = \frac{1}{2}$. Para obtener los autoespacios asociados, calculamos:

$$\mathcal{S}_{\lambda=1} = \text{nul}(A - I) = \text{nul}(\begin{bmatrix} \frac{-5}{18} & \frac{-1}{9} & \frac{-2}{9} \\ \frac{-1}{9} & \frac{-4}{9} & \frac{1}{9} \\ \frac{-2}{9} & \frac{1}{9} & \frac{-5}{18} \end{bmatrix}) = gen\{ \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} \}.$$

$$\mathcal{S}_{\lambda=\frac{1}{2}} = \mathrm{nul}(A - \frac{1}{2}I) = \mathrm{nul}(\begin{bmatrix} \frac{2}{9} & \frac{-1}{9} & \frac{-2}{9} \\ \frac{-1}{9} & \frac{1}{18} & \frac{1}{9} \\ \frac{-2}{9} & \frac{1}{9} & \frac{2}{9} \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\}.$$

Si tomamos $\mathcal{B} := \{\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\}$ entonces $\mathcal{B}$ es una base (de autovectores) de $\mathbb{R}^3$. Calculemos entonces $\lim_{k\to\infty} A^k x$ para cada $x \in \mathbb{R}^3$. Sea $x \in \mathbb{R}^3$, como $\mathcal{B}$ es una base (de autovectores) de $\mathbb{R}^3$ existen (únicos) $a, b, c \in \mathbb{R}$ tales que $x = a\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} + b\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$. Entonces, usando (5.2),

$$\begin{aligned} A^k x &= aA^k \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} + bA^k \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + cA^k \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \\ &= a(1)^k \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} + b(\frac{1}{2})^k \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c(\frac{1}{2})^k \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \\ &= a \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} + \frac{1}{2^k}(b \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}). \end{aligned}$$

Usando que $\lim_{k\to\infty} \frac{1}{2^k} = 0$, tenemos que $\lim_{k\to\infty} A^k x = a\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$. Por lo tanto, $\lim_{k\to\infty} A^k x = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ si y sólo si $a = 0$. Entonces

$$\begin{aligned} \{x \in \mathbb{R}^3 : \lim_{k\to\infty} A^k x = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\} &= \{x \in \mathbb{R}^3 : x = b\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \ b,\ c \in \mathbb{R}\} \\ &= gen\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\} = \mathcal{S}_{\lambda=\frac{1}{2}}. \end{aligned}$$

$b)$: Tal como vimos en el item $a)$, si $x \in \mathbb{R}^3$, existen $a, b, c \in \mathbb{R}$ tales que $x = a\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} + b\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$. Entonces

$$\lim_{k\to\infty} A^k x = \lim_{k\to\infty} (a\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} + \frac{1}{2^k}(b\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix})) = a\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}.$$

Por lo tanto, es claro que $\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} \in \{\lim_{k\to\infty} A^k x : x \in \mathbb{R}^3\}$ (tomando $a = 1$). Por otra parte, tal como vimos arriba, $\lim_{k\to\infty} A^k x = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$ si y sólo si $a = 1$ y $b, c \in \mathbb{R}$. Por lo tanto, todas las soluciones de la ecuación $\lim_{k\to\infty} A^k x = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$ son

$$x_{sol} = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} + b \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \quad b,\ c \in \mathbb{R}.$$

■

5.4. Matrices definidas y semidefinidas positivas

En esta sección veremos las definiciones de matrices definidas (semidefinidas) positivas, negativas e indefinidas y demostraremos varias propiedades útiles.

Definición. Sea $A \in \mathbb{C}^{n\times n}$ hermítica, es decir $A^* = A$, decimos que:

- A es *definida positiva* si $\langle Ax, x \rangle > 0$, para todo $x \in \mathbb{C}^n \setminus \{0\}$.
- A es *semidefinida positiva* si $\langle Ax, x \rangle \geq 0$, para todo $x \in \mathbb{C}^n$.
- A es *definida negativa* si $\langle Ax, x \rangle < 0$, para todo $x \in \mathbb{C}^n \setminus \{0\}$.
- A es *semidefinida negativa* si $\langle Ax, x \rangle \leq 0$, para todo $x \in \mathbb{C}^n$.
- A es *indefinida* si no es ninguna de las anteriores.

Observar que para que tenga sentido la definición anterior es necesario que A sea hermítica, es decir $A^* = A$. En ese caso, recordemos que $\langle Ax, x \rangle = \langle x, A^*x \rangle = \langle x, Ax \rangle = \overline{\langle Ax, x \rangle}$ y entonces $\langle Ax, x \rangle \in \mathbb{R}$.

Proposición 5.4.1. *Sea $A \in \mathbb{C}^{n\times n}$ hermítica, es decir $A^* = A$ y sean $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{R}$ sus autovalores. Entonces:*

- *A es definida positiva si y sólo si $\lambda_i > 0$ para todo $i = 1, 2, \cdots, n$.*
- *A es semidefinida positiva si y sólo si $\lambda_i \geq 0$ para todo $i = 1, 2, \cdots, n$.*
- *A es definida negativa si y sólo si $\lambda_i < 0$ para todo $i = 1, 2, \cdots, n$.*
- *A es semidefinida negativa si y sólo si $\lambda_i \leq 0$ para todo $i = 1, 2, \cdots, n$.*
- *A es indefinida si A tiene algún autovalor negativo y algún autovalor positivo.*

Dem. Supongamos que A es definida positiva, entonces $\langle Ax, x \rangle > 0$ para todo $x \in \mathbb{C}^n \setminus \{0\}$. Sea $\lambda_i \in \mathbb{R}$ un autovalor de A y $v_i \neq 0$ algún autovector asociado tal que $Av_i = \lambda_i v_i$, entonces:

$$0 < \langle Av_i, v_i \rangle = \langle \lambda_i v_i, v_i \rangle = \lambda_i \langle v_i, v_i \rangle,$$

entonces como $\langle v_i, v_i \rangle > 0$, se sigue que $\lambda_i > 0$. Concluimos que si A es definida positiva todo autovalor de A es positivo (mayor estricto que 0).

Recíprocamente, supongamos que todos los autovalores de A son positivos. Es decir $\lambda_i > 0$ para $i = 1, 2, \cdots, n$. Sea $\{v_1, v_2, \cdots, v_n\}$ una bon de autovectores de $\mathbb{C}^n$ (existe pues A es hermítica) tal que $Av_i = \lambda_i v_i$ para $i = 1, 2, \cdots, n$. Sea $x \in \mathbb{C}^n \setminus \{0\}$, entonces $x = a_1 v_2 + a_2 v_2 + \cdots + a_n v_n$ con $a_1, a_2, \cdots, a_n \in \mathbb{C}$ no todos nulos. Entonces

$$\begin{aligned} \langle Ax, x \rangle &= \langle A(a_1 v_2 + a_2 v_2 + \cdots + a_n v_n), a_1 v_2 + a_2 v_2 + \cdots + a_n v_n \rangle \\ &= \langle a_1 A v_1 + a_2 A v_2 + \cdots + a_n A v_n, a_1 v_2 + a_2 v_2 + \cdots + a_n v_n \rangle \\ &= |a_1|^2 \lambda_1 \langle v_1, v_1 \rangle + |a_2|^2 \lambda_2 \langle v_2, v_2 \rangle + \cdots + |a_n|^2 \lambda_n \langle v_n, v_n \rangle \\ &= |a_1|^2 \lambda_1 + |a_2|^2 \lambda_2 + \cdots + |a_n|^2 \lambda_n > 0, \end{aligned}$$

y A resulta definida positiva.

De manera similar se prueban los demás ítems. ■

Submatrices principales y un criterio para determinar si una matriz es definida positiva

A continuación veremos otro criterio para determinar si una matriz es definida positiva, definida negativa o indefinida. Sea $A \in \mathbb{R}^{n\times n}$ de la forma

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Denotaremos $A^{(k)}$ a la *submatriz principal* de $\mathbb{R}^{k\times k}$ que se obtiene a partir de la esquina superior izquierda de A, es decir

$$A^{(k)} := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix},$$

para $k = 1, 2, \cdots, n$.

Por ejemplo $A^{(1)} = [a_{11}]$, $A^{(2)} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ y $A^{(n)} = A$.

Recordemos que si $\lambda_1, \lambda_2, \cdots, \lambda_n$ son los autovalores de A, tenemos que

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n.$$

Por lo tanto,

- Si A es definida positiva, como todos sus autovalores son positivos, tendremos que $\det(A) > 0$.
- Si A es definida negativa, como todos sus autovalores son negativos, tendremos que todos los autovalores de $-A$ son positivos y entonces $-A$ es definida positiva y por lo tanto, $\det(-A) > 0$.
- Si A es semidefinida, como algún autovalor es nulo, tendremos que $\det(A) = 0$.
- Si A es indefinida, no podemos decir nada del $\det(A)$.

El siguiente teorema nos provee de otro criterio para determinar si una matriz es definida positiva, definida negativa o indefinida:

Teorema 5.4.2. *Sea $A \in \mathbb{R}^{n\times n}$ simétrica. Entonces:*

1. *A es definida positiva si y sólo si $\det(A^{(k)}) > 0$ para todo $k = 1, 2, \cdots, n$. Es decir, el determinante de todas las submatrices principales de A es positivo.*
2. *A es definida negativa si y sólo si $(-1)^k \det(A^{(k)}) > 0$ para todo $k = 1, 2, \cdots, n$. Es decir, el determinante de todas las submatrices principales de A van alternándose entre negativos y positivos:* $\det(A^{(1)}) < 0$, $\det(A^{(2)}) > 0$, *etc.*
3. *A es indefinida si existe $k \in \{1, 2, \cdots, n\}$ tal que $\det(A^{(k)})$ rompe el patrón. Es decir, si por ejemplo,* $\det(A^{(1)}) > 0, \cdots, \det(A^{(k-1)}) > 0$ *y luego tenemos que* $\det(A^{(k)}) < 0$.

Dem. 1. : Lo vamos a probar por inducción en n.

Si $n = 1$, entonces $A = [a_{11}]$ es definida positiva si y sólo si $a_{11} > 0$ si y sólo si $\det([a_{11}]) = a_{11} > 0$ y en este caso es válido el teorema.

Hipótesis inductiva (HI): supongamos que el teorema vale para $n-1$, es decir: $A \in \mathbb{R}^{(n-1)\times(n-1)}$ es definida positiva si y sólo si $\det(A^{(k)}) > 0$ para todo $k = 1, 2, \cdots, n-1$. Veamos que entonces, el teorema vale para todo n.

Primero veamos que, por HI, A sólo puede tener como mucho un autovalor negativo. Supongamos que no, es decir, supogamos que A tiene dos autovalores negativos. Entoces, como A es simétrica, existen dos autovectores $u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n \setminus \{0\}, v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n \setminus \{0\}$ asociados a dichos autovalores negativos, que son ortogonales entre sí, es decir $\langle\, u, v \,\rangle = u^T v = v^T u = 0$.

Sea
$$w := v_n u - u_n v \neq 0.$$

Entonces
$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_n u_1 - u_n v_1 \\ \vdots \\ v_n u_n - u_n v_n \end{bmatrix} = \begin{bmatrix} v_n u_1 - u_n v_1 \\ \vdots \\ 0 \end{bmatrix},$$

entonces $w_n = 0$. Por lo tanto (comprobarlo haciendo la cuenta):

$$\langle\, Aw, w \,\rangle = \left\langle A^{(n-1)} \begin{bmatrix} w_1 \\ \vdots \\ w_{n-1} \end{bmatrix}, \begin{bmatrix} w_1 \\ \vdots \\ w_{n-1} \end{bmatrix} \right\rangle > 0$$

donde usamos que, por HI, $A^{(n-1)} \in \mathbb{R}^{(n-1)\times(n-1)}$ es definida positiva. Por otra parte (verificarlo haciendo la cuenta y recordando que $u^T v = v^T u = 0$):

$$\langle\, Aw, w \,\rangle = (v_n u - u_n v)^T A (v_n u - u_n v) = v_n^2 (u^T A u) + u_n^2 (v^T A v) = v_n^2 \langle\, Au, u \,\rangle + u_n^2 \langle\, Av, v \,\rangle < 0,$$

pues u y v eran autovectores asociados a autovalores negativos.

Entonces $\langle\, Aw, w \,\rangle > 0$ y $\langle\, Aw, w \,\rangle < 0$, lo cual es absurdo. Entonces, A tiene a lo sumo 1 único autovalor negativo. Pero eso tampoco puede pasar. Supongamos que A tiene un autovalor negativo, entoces $\det(A) = \lambda_1 \cdots \lambda_n < 0$ pues tendríamos el producto de $n-1$ autovalores positivos y uno negativo. Pero estamos suponiendo que $\det(A) = \det(A^{(n)}) > 0$, entonces tenemos otra contradicción. Tampoco A puede tener algún autovalor nulo porque en ese caso, obtendríamos que $\det(A) = 0$ y estamos suponiendo que $\det(A) > 0$. Por lo tanto, A tiene todos sus autovalores positivos. Entonces, probamos por inducción que A es definida positiva.

Recíprocamente, si A es definida positiva, tenemos que $\langle\, Ax, x \,\rangle > 0$ para todo $x \in \mathbb{R}^n \setminus \{0\}$. Sea $x \in \mathbb{R}^n \setminus \{0\}$ y tomemos $x^{(k)} := [x_1\ x_2\ \cdots\ x_k]^T$, es decir $x^{(k)}$ es el vector de $\mathbb{R}^k$ igual a x hasta la componente k. Entonces

$$\left\langle A^{(k)} x^{(k)}, x^{(k)} \right\rangle = \langle\, A[x_1\ x_2\ \cdots\ x_k\ 0\ \cdots\ 0]^T, [x_1\ x_2\ \cdots\ x_k\ 0\ \cdots\ 0]^T \,\rangle > 0$$

para todo $x^{(k)} \in \mathbb{R}^k \setminus \{0\}$ y todo $k = 1, 2, \cdots, n$. Entonces, $A^{(k)}$ es definida positiva para todo $k = 1, 2, \cdots, n$. Por lo tanto, $\det(A^{(k)}) > 0$ para todo $k = 1, 2, \cdots, n$ y probamos lo que queríamos.

2. : Esta condición se prueba aplicando el item 1 a la matriz $-A$ (A es definida negativa si y sólo si $-A$ es definida positiva) y recordando que

$$\det((-A)^{(k)}) = \det(-A^{(k)}) = (-1)^k \det(A^{(k)}),$$

pues $A^{(k)} \in \mathbb{R}^{k\times k}$.

3. : Recordando el paso inductivo en el item 1 vimos que, si por ejemplo,

$$\det(A^{(1)}) > 0, \cdots, \det(A^{(k-1)}) > 0$$

entonces $A^{(k)}$ tiene sólo un autovalor negativo. Por ende, $A^{(k)}$ es indefinida (tiene $k-1$ autovalores positivos y uno negativo) y entonces A también es indefinida. ■

Observar que si existe algún $k \in \{1, 2, \cdots, n\}$ tal que $\det(A^{(k)}) = 0$ el Teorema 5.4.2 es inconcluso.

5.5. Descomposición en valores singulares

Sea $A \in \mathbb{C}^{m\times n}$ entonces la matriz $A^*A \in \mathbb{C}^{n\times n}$ es hermítica y semidefinida positiva. De hecho

$$(A^*A)^* = A^*(A^*)^* = A^*A,$$

entonces A^*A es hermítica. Por otro lado, tenemos que

$$\langle A^*Ax, x\rangle = \langle Ax, Ax\rangle = \|Ax\|^2 \geq 0, \text{ para todo } x \in \mathbb{C}^n.$$

Entonces A es semidefinida positiva y, por la Proposición 5.4.1, si $\lambda \in \mathbb{R}$ es un autovalor de A entonces $\lambda \geq 0$.

Definimos, el $i-$ésimo *valor singular* de A como

$$\sigma_i := \lambda_i^{1/2},$$

donde λ_i es el i-ésimo autovalor de A^TA.

Definición. Sea $A \in \mathbb{C}^{m\times n}$, una *descomposición en valores singulares (DVS)* de A es una factorización

$$A = U\Sigma V^*,$$

con $U \in \mathbb{C}^{m\times m}$ y $V \in \mathbb{C}^{n\times n}$ unitarias y $\Sigma \in \mathbb{R}^{m\times n}$ con $\Sigma = \begin{bmatrix} D & 0_{r\times n-r} \\ 0_{m-r\times r} & 0_{m-r\times n-r} \end{bmatrix}$ donde $D = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix} \in \mathbb{R}^{r\times r}$ con $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, los *valores singulares* no nulos de A.

Si $A \in \mathbb{R}^{m\times n}$ entonces en la DVS de A tendremos que $U \in \mathbb{R}^{m\times m}$ y $V \in \mathbb{R}^{n\times n}$ son ortogonales y $A = U\Sigma V^T$.

El próximo teorema nos brinda un algoritmo para obtener una DVS de cualquier matriz.

Teorema 5.5.1. *Sea $A \in \mathbb{C}^{m\times n}$ entonces existe una DVS de A.*

Dem. Sean $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ los n autovalores de A^*A (contados con multiplicidad) y $\{v_1, v_2, \cdots, v_n\}$ una bon de $\mathbb{C}^n$ tal que $A^*Av_i = \lambda_i v_i$ para todo $i = 1, 2, \cdots, n$ (existe pues A^*A es hermítica y por ende diagonalizable unitariamente). Entonces, definimos V y Σ de la siguiente manera:

$$V := [v_1\ v_2 \cdots\ v_n] \in \mathbb{C}^{n\times n}$$

que es unitaria (pues sus columnas son una bon de $\mathbb{C}^n$),

$$\Sigma = \begin{bmatrix} D & 0_{r\times n-r} \\ 0_{m-r\times r} & 0_{m-r\times n-r} \end{bmatrix} \text{ donde } D = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix} \in \mathbb{R}^{r\times r},$$

con $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, los *valores singulares* no nulos de A.

Para definir U, consideremos los vectores

$$u_1 := \frac{Av_1}{\sigma_1},\ u_2 := \frac{Av_2}{\sigma_2},\ \cdots, u_r := \frac{Av_r}{\sigma_r}.$$

Observar que,

$$\langle Av_i, Av_j \rangle = \langle v_i, A^*Av_j \rangle = \langle v_i, \lambda_j v_j \rangle = \lambda_j \langle v_i, v_j \rangle = \begin{cases} \lambda_j \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}.$$

Entonces,

$$\langle u_i, u_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle v_i, A^*Av_j \rangle = \begin{cases} \frac{1}{\sigma_j\sigma_j}\lambda_j = \frac{\lambda_j}{\sigma_j^2} = 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}.$$

Por lo tanto, $\{u_1, u_2, \cdots, u_r\}$ es un conjunto ortonormal de $\mathbb{C}^m$. Si $r < m$, buscamos vectores $u_{r+1}, \cdots, u_m$, tales que $\{u_1, u_2, \cdots, u_m\}$ sea una bon de $\mathbb{C}^m$. Definimos U como

$$U := [u_1\ u_2\ \cdots\ u_r\ \cdots\ u_m] \in \mathbb{C}^{m\times m}$$

que es una matriz unitaria (pues sus columnas forman una bon de $\mathbb{C}^m$).

Además, observar que $\{u_1, u_2, \cdots, u_r\}$ es una bon de $\mathrm{col}(A)$. Para ver eso, como ya vimos que $\{u_1, u_2, \cdots, u_r\}$ es un conjunto ortonormal, sólo nos basta ver que

$$gen\{u_1, u_2, \cdots, u_r\} = gen\{Av_1, Av_2, \cdots, Av_r\} = \mathrm{col}(A).$$

Para eso, consideremos la transformación lineal $T : \mathbb{C}^n \to \mathbb{C}^m$, $T(x) = Ax$. Entonces, $\mathrm{Im}(T) = \mathrm{col}(A)$ y como $\{v_1, v_2, \cdots, v_n\}$ es una base (de autovectores) de $\mathbb{C}^n$, tenemos que

$$\mathrm{col}(A) = \mathrm{Im}(T) = gen\{T(v_1), \cdots, T(v_n)\} = gen\{Av_1, \cdots, Av_n\}.$$

Por otra parte, si $i \geq r+1$, entonces $\lambda_i = 0$ y tenemos que

$$\|Av_i\|^2 = \langle Av_i, Av_i \rangle = \langle v_i, A^*Av_i \rangle = \lambda_i \langle v_i, v_i \rangle = 0,$$

entonces $Av_i = 0$ y tenemos que

$$\mathrm{col}(A) = \mathrm{Im}(T) = gen\{Av_1, \cdots, Av_n\} = gen\{Av_1, \cdots, Av_r\}$$

y probamos lo que queríamos.

Finalmente, veamos que $A = U\Sigma V^*$. Por un lado,

$$AV = [Av_1\ Av_2\ \cdots\ Av_r\ Av_{r+1} \cdots\ Av_n] = [\sigma_1 u_1\ \sigma_2 u_2\ \cdots\ \sigma_r u_r\ 0\ \cdots\ 0],$$

donde usamos que $Av_i = 0$, si $i \geq r+1$.

Por otro lado,

$$U\Sigma = [u_1\ u_2\ \cdots\ u_r\ \cdots\ u_m] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} = [\sigma_1 u_1\ \cdots\ \sigma_r u_r\ 0 \cdots\ 0].$$

Por lo tanto, $AV = U\Sigma$. Como $V^{-1} = V^*$, concluimos que $A = U\Sigma V^*$. ■

A partir del Teorema 5.5.1, podemos sacar algunas conclusiones extras:

- Si $A = U\Sigma V^*$ es una DVS de A. Entonces:

$$\{u_2, u_2, \cdots, u_r\} = \{\frac{Av_1}{\sigma_1}, \frac{Av_2}{\sigma_2}, \cdots, \frac{Av_r}{\sigma_r}\}$$

es una bon de col(A). Por lo tanto,

$$\boxed{r = \dim(\mathrm{col}(A)) = \mathrm{rg}(A) = \text{número de valores singulares no nulos.}}$$

- Como $\{v_{r+1}, \cdots, v_n\}$ es un conjunto ortonormal tal que $Av_i = 0$ si $i = r+1, \cdots, n$ y, por el Teorema de la Dimensión, $\dim(\mathrm{nul}(A)) = n - \mathrm{rg}(A) = n - r$, se sigue que

$$\{v_{r+1}, \cdots, v_n\}$$

es una bon de nul(A).

Ejercicio 5.10 b). Hallar una DVS de la matriz $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$.

Dem. Para resolver este ejercicio, simplemente seguimos los pasos que hicimos en la demostración del Teorema 5.5.1:

1. Calcular los autovalores de $A^T A$ y ordenarlos de mayor a menor. De hecho,

$$A^T A = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}.$$

Entonces, el polinomio caraceterístico de $A^T A$ es

$$\chi_{A^T A}(x) = -x^3 + 34x^2 - 225x.$$

Por lo tanto, los autovalores de $A^T A$ (ordenados de mayor a menor) son: $\lambda_1 = 25$, $\lambda_2 = 9, \lambda_3 = 0$.

2. Hallar una bon $\{v_1, v_2, v_3\}$ de $\mathbb{R}^3$ de autovectores de $A^T A$ tal que $A^T A v_i = \lambda_i v_i$ para $i = 1, 2, 3$.

De hecho, como $\mathrm{nul}(A^T A - 25I) = gen\{[1\ 1\ 0]^T\}$, $\mathrm{nul}(A^T A - 9I) = gen\{[1\ -1\ 4]^T\}$ y $\mathrm{nul}(A^T A) = gen\{[-2\ 2\ 1]^T\}$. Podemos tomar

$$v_1 = \frac{1}{\sqrt{2}}[1\ 1\ 0]^T,\ v_2 = \frac{1}{3\sqrt{2}}[1\ -1\ 4]^T,\ v_3 = \frac{1}{3}[-2\ 2\ 1]^T.$$

3. Calcular los valores singulares de A.

$$\sigma_1 = \lambda_1^{1/2} = 25^{1/2} = 5,\ \sigma_2 = \lambda_2^{1/2} = 9^{1/2} = 3,\ \sigma_3 = \lambda_3^{1/2} = 0^{/12} = 0.$$

4. Determinar r el número de valores singulares no nulos de A (que es igual a rg(A)) y para $i = 1, \cdots, r$ definir

$$u_1 = \frac{Av_i}{\sigma_i}.$$

De hecho, $r = 2$ (el número de valores singulares no nulos). Entonces, como $Av_1 = \frac{1}{\sqrt{2}}[5\ 5]^T$ y $Av_2 = \frac{1}{3\sqrt{2}}[9\ -9]^T$, tenemos que

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{1}{\sqrt{2}}[5\ 5]^T \frac{1}{5} = \frac{1}{\sqrt{2}}[1\ 1]^T,\ \ u_2 = \frac{Av_2}{\sigma_2} = \frac{1}{3\sqrt{2}}[9\ -9]^T \frac{1}{3} = \frac{1}{\sqrt{2}}[1\ -1]^T.$$

5. Si $r < m$, hallar $u_{r+1}, \cdots, u_m$ tales que $\{u_1, \cdots, u_m\}$ es una bon de $\mathbb{R}^m$.

   Como en este caso $r = 2 = m$, este paso no es necesario.

6. Definir las matrices

$$V := [v_1\ v_2\ v_3] = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & \frac{1}{3} \end{bmatrix},\ U := [u_1\ u_2] = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}.$$

Entonces,

$$A = U\Sigma V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}.$$

■

5.6. DVS y los subespacios fundamentales de una matriz

Sea $A \in \mathbb{R}^{m\times n}$ tal que $\text{rg}(A) = r$ y $A = U\Sigma V^T$ es una DVS de A, donde

$$V = [v_1\ \cdots\ v_r\ v_{r+1}\ \cdots\ v_n] \text{ y } U = [u_1\ \cdots\ u_r\ u_{r+1}\ \cdots\ u_m].$$

Entonces, definimos las matrices:

$$V_r := [v_1\ \cdots\ v_r],\quad V_{n-r} := [v_{r+1}\ \cdots\ v_n],\quad U_r := [u_1\ \cdots\ u_r],\quad U_{m-r} := [u_{r+1}\ \cdots\ u_m].$$

Entonces:

1. $\{v_1, \cdots, v_r\}$ es una bon de $\text{col}(A^T) = \text{fil}(A)$ y $\{v_{r+1}, \cdots, v_n\}$ es una bon de $\text{nul}(A)$. Además, si $\mathcal{E}$ es la base canónica de $\mathbb{R}^n$, tenemos que:

$$V_rV_r^T = [P_{col(A^T)}]_{\mathcal{E}}^{\mathcal{E}},\quad V_r^TV_r = I_r,\quad V_{n-r}V_{n-r}^T = [P_{nul(A)}]_{\mathcal{E}}^{\mathcal{E}},\quad V_{n-r}^TV_{n-r} = I_{n-r}.$$

   De hecho, en el Teorema 5.5.1, vimos que $\{v_{r+1}, \cdots, v_n\}$ es una bon de $\text{nul}(A)$. Por lo tanto, como $\{v_1, \cdots, v_n\}$ es una bon de $\mathbb{R}^n$, no queda otra que $\{v_1, \cdots, v_r\}$ sea una bon de $\text{nul}(A)^\perp = \text{col}(A^T) = \text{fil}(A)$, donde usamos la Proposición 3.7.1.

   Además, si $\mathcal{E}$ es la base canónica de $\mathbb{R}^n$, recordando la fórmula 3.16, se deduce que $V_rV_r^T = [P_{\text{col}(A^T)}]_{\mathcal{E}}^{\mathcal{E}}$, $V_r^TV_r = I_r$, $V_{n-r}V_{n-r}^T = [P_{\text{nul}(A)}]_{\mathcal{E}}^{\mathcal{E}}$, $V_{n-r}^TV_{n-r} = I_{n-r}$.

2. $\{u_1, \cdots, u_r\}$ es una bon de $\text{col}(A)$ y $\{u_{r+1}, \cdots, u_m\}$ es una bon de $\text{col}(A)^\perp = \text{nul}(A^T)$. Además, si $\mathcal{E}$ es la base canónica de $\mathbb{R}^m$, tenemos que:

$$U_rU_r^T = [P_{col(A)}]_{\mathcal{E}}^{\mathcal{E}},\quad U_r^TU_r = I_r,\quad U_{m-r}U_{m-r}^T = [P_{\text{nul}(A^T)}]_{\mathcal{E}}^{\mathcal{E}},\quad U_{m-r}^TU_{m-r} = I_{m-r}.$$

   De hecho, en el Teorema 5.5.1, vimos que $\{u_1, \cdots, u_m\}$ es una bon de $\text{col}(A)$. Por lo tanto, como $\{u_1, \cdots, u_m\}$ es una bon de $\mathbb{R}^m$, no queda otra que $\{u_{r+1}, \cdots, u_m\}$ sea una bon de $\text{col}(A)^\perp = \text{nul}(A^T)$, donde usamos la Proposición 3.7.1.

   Además, si $\mathcal{E}$ es la base canónica de $\mathbb{R}^m$, recordando la fórmula 3.16, se deduce que $U_rU_r^T = [P_{\text{col}(A)}]_{\mathcal{E}}^{\mathcal{E}}$, $U_r^TU_r = I_r$, $U_{m-r}U_{m-r}^T = [P_{\text{nul}(A^T)}]_{\mathcal{E}}^{\mathcal{E}}$, $U_{m-r}^TU_{m-r} = I_{m-r}$.

Finalmente, observar que $V = [V_r \ V_{n-r}]$ y $U = [U_r \ U_{m-r}]$. Entonces,

$$A = U\Sigma V^T = [U_r \ U_{m-r}] \begin{bmatrix} D & 0_{r\times n-r} \\ 0_{m-r\times r} & 0_{m-r\times n-r} \end{bmatrix} \begin{bmatrix} V_r^T \\ V_{n-r}^T \end{bmatrix} = U_r D V_r^T.$$

A la factorización $A = U_r D V_r^T$, la llamaremos *DVS reducida* de A.

Notar que como $D \in \mathbb{R}^{r\times r}$ es una matriz diagonal con elementos positivos en su diagonal (los valores singulares no nulos de A), entonces D es inversible.

Ejercicio 5.11. Sea

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \\ 1 & -1 & 4 \\ 2\sqrt{2} & -2\sqrt{2} & -\sqrt{2} \end{bmatrix}.$$

a) Hallar los valores singulares de A, bon de sus cuatro subespacios fundamentales y sus respectivas matrices de proyección.

b) Hallar una DVS reducida de A.

Dem. a): La factorización que tenemos de A es “casi” una DVS de A. La matriz $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ tiene columnas ortogonales pero no ortonormales y lo mismo sucede con $\begin{bmatrix} 3 & 3 & 0 \\ 1 & -1 & 4 \\ 2\sqrt{2} & -2\sqrt{2} & -\sqrt{2} \end{bmatrix}$.

Procedemos a normalizar las columnas de $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ dividiendo cada una de ellas por su norma que es $\sqrt{2}$. Para que el producto siga siendo A, es necesario multiplicar también por $\sqrt{2}$. Es decir,

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \sqrt{2} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \\ 1 & -1 & 4 \\ 2\sqrt{2} & -2\sqrt{2} & -\sqrt{2} \end{bmatrix}.$$

Finalmente, como la norma de las 3 columnas de la matriz $\begin{bmatrix} 3 & 3 & 0 \\ 1 & -1 & 4 \\ 2\sqrt{2} & -2\sqrt{2} & -\sqrt{2} \end{bmatrix}$ es $3\sqrt{2}$ para normalizarlas, procedemos a multiplicar y dividir por $3\sqrt{2}$. Entonces, nos queda

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \sqrt{2} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} 3\sqrt{2} \frac{1}{3\sqrt{2}} \begin{bmatrix} 3 & 3 & 0 \\ 1 & -1 & 4 \\ 2\sqrt{2} & -2\sqrt{2} & -\sqrt{2} \end{bmatrix}.$$

Entonces obtenemos

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3\sqrt{2}\,3\sqrt{2} & 0 & 0 \\ 0 & 2\sqrt{2}\,3\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 18 & 0 & 0 \\ 0 & 12 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix},$$

que ahora sí es una DVS de A, ya que $A = U\Sigma V^T$, con

$$U := \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 18 & 0 & 0 \\ 0 & 12 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & -\frac{1}{3} \end{bmatrix}.$$

Entonces, $\sigma_1 = 18, \sigma_2 = 12$ y $r = 2 = \text{rg}(A)$ es el número de valores singulares no nulos.

Una bon de $\text{col}(A) = \mathbb{R}^2$, puede ser $\{[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^T, [\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}]^T\}$. Por otra parte, $\text{col}(A)^\perp = \text{nul}(A^T) = \{0\}$. Entonces,

$$U_2 U_2^T = I_2 = [P_{\text{col}(A)}]_{\mathcal{E}}^{\mathcal{E}}.$$

Finalmente, tenemos que $\{[\frac{1}{\sqrt{2}} \ \frac{1}{\sqrt{2}} \ 0]^T, [\frac{1}{3\sqrt{2}} \ -\frac{1}{3\sqrt{2}} \ \frac{4}{3\sqrt{2}}]^T\}$ es una bon de $\text{col}(A^T) = \text{fil}(A) = \text{nul}(A)^\perp$ y $\{[\frac{2}{3} \ -\frac{2}{3} \ -\frac{1}{3}]^T\}$ es una bon de $\text{nul}(A)$. Entonces,

$$V_2 V_2^T = [P_{\text{col}(A^T)}]_{\mathcal{E}}^{\mathcal{E}} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ 0 & \frac{4}{3\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{5}{9} & \frac{4}{9} & \frac{2}{9} \\ \frac{4}{9} & \frac{5}{9} & -\frac{2}{9} \\ \frac{2}{9} & -\frac{2}{9} & \frac{8}{9} \end{bmatrix}.$$

$$V_1 V_1^T = [P_{\text{nul}(A)}]_{\mathcal{E}}^{\mathcal{E}} = \begin{bmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ -\frac{1}{3} \end{bmatrix} [\frac{2}{3} \ -\frac{2}{3} \ -\frac{1}{3}] = \begin{bmatrix} \frac{4}{9} & -\frac{4}{9} & -\frac{2}{9} \\ -\frac{4}{9} & \frac{4}{9} & \frac{2}{9} \\ -\frac{2}{9} & \frac{2}{9} & \frac{1}{9} \end{bmatrix}$$

$b)$: Una DVS reducida de A puede ser

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 18 & 0 \\ 0 & 12 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \end{bmatrix}.$$

■

5.7. Solución por mínimos cuadrados de norma mínima. Pseudo-inversa de Moore-Penrose

Recordemos que en la Sección 3.7, vimos que, si $A \in \mathbb{R}^{m\times n}$ y $b \in \mathbb{R}^n$, la ecuación $Ax = b$ siempre admite solución por mínimos cuadrados y dicha solución es única si y sólo si $\text{nul}(A) = \{0\}$, ó, equivalentemente $\text{rg}(A) = n$. Cuando existan infinitas soluciones por mínimos cuadrados, nos interesa disponer de un criterio para seleccionar una de ellas.

Definición. Diremos que $\tilde{x}$ es una *solución por mínimos cuadrados de norma mínima* de la ecuación $Ax = b$, si $\tilde{x}$ es solución por mínimos cuadrados de $Ax = b$ y

$$\|\tilde{x}\| \leq \|\hat{x}\|$$

para toda $\hat{x}$ solución por mínimos cuadrados de $Ax = b$.

Sea $\hat{x}$ una solución por mínimos cuadrados de $Ax = b$. Entonces, por un lado, recordemos que $\hat{x}$ cumple que

$$A\hat{x} = P_{\text{col}(A)} b.$$

Por el otro lado, tenemos que

$$\hat{x} = P_{\text{nul}(A)}\hat{x} + P_{\text{nul}(A)^\perp}\hat{x}.$$

Sea $\tilde{x} := P_{\text{nul}(A)^\perp}\hat{x} \in \text{nul}(A)^\perp$. Entonces

$$P_{\text{col}(A)} b = A\hat{x} = A(P_{\text{nul}(A)}\hat{x} + P_{\text{nul}(A)^\perp}\hat{x}) = AP_{\text{nul}(A)}\hat{x} + AP_{\text{nul}(A)^\perp}\hat{x} = A\tilde{x}.$$

Entonces $\tilde{x}$ es una solución por mínimos cuadrados que pertenece a $\text{nul}(A)^{\perp}$.

Supongamos que x' es otra solución por mínimos cuadrados que pertenece a $\text{nul}(A)^{\perp}$. Entonces, como $\tilde{x}$ y x' son soluciones por mínimos cuadrados, tenemos que

$$Ax' = P_{\text{col}(A)}b = A\tilde{x}.$$

Entonces $x' - \tilde{x} \in \text{nul}(A)$. Además, como $\tilde{x}$ y x' pertenecen al subespacio $\text{nul}(A)^{\perp}$, se sigue que $x' - \tilde{x} \in \text{nul}(A)^{\perp}$. Por lo tanto, $x' - \tilde{x} \in \text{nul}(A) \cap \text{nul}(A)^{\perp} = \{0\}$. Entonces $\tilde{x} = x'$. Concluimos que:

> Existe una única solución por mínimos cuadrados de la ecuación $Ax = b$ que pertenece a $\text{nul}(A)^{\perp}$.

Ahora veamos que $\tilde{x}$ es una solución por mínimos cuadrados de norma mínima. De hecho, sea $\hat{x}$ una solución por mínimos cuadrados de $Ax = b$. Entonces, vimos que $\hat{x} - \tilde{x} \in \text{nul}(A)$. Entonces, como $\hat{x} = \hat{x} - \tilde{x} + \tilde{x}$, con $\hat{x} - \tilde{x} \in \text{nul}(A)$ y $\tilde{x} \in \text{nul}(A)^{\perp}$, por el Teorema de Pitágoras (o haciendo la cuenta), tenemos que

$$\|\hat{x}\|^2 = \|\hat{x} - \tilde{x}\|^2 + \|\tilde{x}\|^2 \geq \|\tilde{x}\|^2.$$

Tomando raíz (que es una función creciente), tenemos que

$$\|\hat{x}\| \geq \|\tilde{x}\|$$

para toda $\hat{x}$ solución por mínimos cuadrados de $Ax = b$. Por lo tanto,

> $\tilde{x}$ es una solución por mínimos cuadrados de norma mínima de la ecuación $Ax = b$.

Además, $\tilde{x}$ es la única solución por mínimos cuadrados de norma mínima. De hecho, si z también es una solución por mínimos cuadrados de norma mínima, entonces $\|z\| = \|\tilde{x}\|$. Además, como z y $\tilde{x}$ son solución por mínimos cuadrados, tenemos que $z - \tilde{x} \in \text{nul}(A)$ y tenemos que $\tilde{x} \in \text{nul}(A)^{\perp}$. Por lo tanto, aplicando Pitágoras nuevamente, se sigue que

$$\|z\|^2 = \|z - \tilde{x} + \tilde{x}\|^2 = \|z - \tilde{x}\|^2 + \|\tilde{x}\|^2.$$

Entonces

$$\|z - \tilde{x}\|^2 = \|z\|^2 - \|\tilde{x}\|^2 = 0.$$

Por lo tanto $z = \tilde{x}$. Entonces,

> $\tilde{x}$ es la única solución por mínimos cuadrados de norma mínima de la ecuación $Ax = b$.

Veamos una manera de encontrar $\tilde{x}$. Sea $A \in \mathbb{R}^{m \times n}$ tal que $\text{rg}(A) = r$ y $A = U_r D V_r$ es una DVS reducida de A. Recordemos que $D \in \mathbb{R}^{r \times r}$ es inversible, $U_r U_r^T = P_{\text{col}(A)}, V_r^T V_r = I_r$ y $col(V_r) = \text{nul}(A)^{\perp}$. Entonces, afirmamos que

> $$\tilde{x} = (V_r D^{-1} U_r^T)b.$$

De hecho, $(V_r D^{-1} U_r^T)b \in \text{col}(V_r) = \text{nul}(A)^{\perp}$ y además

$$\begin{aligned} A[(V_r D^{-1} U_r^T)b] &= U_r D V_r^T (V_r D^{-1} U_r^T)b = U_r D (V_r^T V_r) D^{-1} U_r^T b = U_r D I_r D^{-1} U_r^T b = U_r I_r U_r^T b \\ &= U_r U_r^T b = P_{\text{col}(A)} b. \end{aligned}$$

Entonces, $(V_r D^{-1} U_r^T)b$ es una solución por mínimos cuadrados que pertenece a $\text{nul}(A)^{\perp}$. Acabamos de ver $\tilde{x}$ es la única solución por mínimos cuadrados que pertence a $\text{nul}(A)^{\perp}$. Por lo tanto, tenemos que $\tilde{x} = (V_r D^{-1} U_r^T)b$ como queríamos ver.

La matriz $A^\dagger := U_r D V_r^T$ es la *matriz pseudo inversa de Moore-Penrose* de A. Por lo tanto

$$\tilde{x} = A^\dagger b$$

es la única solución por mínimos cuadrados de norma mínima de la ecuación $Ax = b$.

Ejercicio 134 b). Sean $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$ y $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$. Hallar $A^\dagger$ y determinar la solución por mínimos cuadrados de norma mínima de la ecuación $Ax = b$.

Dem. En el **Ejercicio 5.10 b)**, probamos que $A = U\Sigma V^T$ con

$$V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & \frac{1}{3} \end{bmatrix}, \; U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \; \Sigma = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}.$$

Entonces, $\operatorname{rg}(A) = 2$,

$$V_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ 0 & \frac{4}{3\sqrt{2}} \end{bmatrix}, \quad U_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \; D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix},$$

y tenemos que $A = U_2 D V_2^T$ es una DVS reducida de A.

Aplicando la fórmula de la definición de la pseudo inversa Moore-Penrose de A tenemos que

$$A^\dagger = V_2 D^{-1} U_2^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ 0 & \frac{4}{3\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{5} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{7}{45} & \frac{2}{45} \\ \frac{2}{45} & \frac{7}{45} \\ \frac{2}{9} & -\frac{2}{9} \end{bmatrix}.$$

Entonces, la solución por mínimos cuadrados de norma mínima de la ecuación $Ax = b$ es

$$\tilde{x} = A^\dagger b = \begin{bmatrix} \frac{7}{45} & \frac{2}{45} \\ \frac{2}{45} & \frac{7}{45} \\ \frac{2}{9} & -\frac{2}{9} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \frac{7b_1+2b_2}{45} \\ \frac{2b_1+7b_2}{45} \\ \frac{2b_1-2b_2}{9} \end{bmatrix}.$$

■

Antes de ver el **Ejercicio 5.14**, vamos a probar algunas propiedades que serán necesarias para su resolución.

Teorema 5.7.1. *Sea $A \in \mathbb{C}^{m\times n}$ y $\sigma_1, \cdots, \sigma_n$, los valores singulares de A (ordenados de mayor a menor). Entonces*

$$\max_{\|x\|=1} \|Ax\| = \sigma_1,$$

y el máximo se alcanza en los vectores $x \in \mathcal{S}_{\lambda_1=\sigma_1^2}$ (el autoespacio asociado al mayor autovalor $\lambda_1 = \sigma_1^2$ de A^*A) tal que $\|x\| = 1$.

$$\min_{\|x\|=1} \|Ax\| = \sigma_n,$$

y el mínimo se alcanza en los vectores $x \in \mathcal{S}_{\lambda_n=\sigma_n^2}$ (el autoespacio asociado al menor autovalor $\lambda_n = \sigma_n^2$ de A^*A) tal que $\|x\| = 1$.

Dem. Sea $A = U\Sigma V^*$ una DVS de A, donde $U \in \mathbb{C}^{m\times m}$ y $V \in \mathbb{C}^{n\times n}$ son unitarias y $\sigma_1, \cdots, \sigma_n$ son los valores singulares de A (ordenados de mayor a menor). Sea $x \in \mathbb{C}^m$ tal que $\|x\| = 1$, entonces,

$$\|Ax\|^2 = \|U\Sigma V^* x\|^2.$$

Como U es unitaria, por la Proposición 5.1.1 tenemos que

$$\|U\Sigma V^*x\|^2 = \|\Sigma V^*x\|^2.$$

Sea $y := V^*x$, entonces como V^* es unitaria, por la Proposición 5.1.1 tenemos que $\|y\| = \|V^*x\| = \|x\|$ y entonces,

$$\|Ax\|^2 = \|\Sigma V^*x\|^2 = \|\Sigma y\|^2 = \sigma_1^2|y_1|^2 + \cdots + \sigma_n^2|y_n|^2.$$

Veamos que

$$\max_{\|y\|=1} (\sigma_1^2|y_1|^2 + \cdots + \sigma_n^2|y_n|^2) = \sigma_1^2.$$

De hecho, si $y \in \mathbb{C}^n$ es tal que $\|y\|^2 = 1$, entonces

$$\begin{aligned}\sigma_1^2 &= \sigma_1^2\|y\|^2 = \sigma_1^2(|y_1|^2 + \cdots + |y_n|^2) = \sigma_1^2|y_1|^2 + \sigma_1^2|y_2|^2 + \cdots + \sigma_1^2|y_n|^2\\ &\geq \sigma_1^2|y_1|^2 + \sigma_2^2|y_2|^2 + \cdots + \sigma_n^2|y_n|^2,\end{aligned}$$

donde usamos que los valores singulares de A están ordenados de mayor a menor, entonces

$$\sigma_1^2|y_i|^2 \geq \sigma_i^2|y_i|^2,$$

para $i = 1, 2, \cdots, n$. Por lo tanto, σ_1^2 es una cota superior del conjunto en cuestión, es decir

$$\sigma_1^2 \geq \sigma_1^2|y_1|^2 + \sigma_2^2|y_2|^2 + \cdots + \sigma_n^2|y_n|^2,$$

para todo $y \in \mathbb{C}^n$ tal que $\|y\|^2 = 1$.

Además el vector $y := [1\ 0\ \cdots\ 0]^T \in \mathbb{C}^n$, cumple $\|y\| = 1$ y

$$\sigma_1^2|y_1|^2 + \sigma_2^2|y_2|^2 + \cdots + \sigma_n^2|y_n|^2 = \sigma_1^2.$$

Entonces, encontramos un elemento del conjunto en cuestión que alcanza la cota superior. Por lo tanto σ_1^2 es el máximo, es decir

$$\max_{\|y\|=1} (\sigma_1^2|y_1|^2 + \cdots + \sigma_n^2|y_n|^2) = \sigma_1^2.$$

Como

$$\max_{\|x\|=1} \|Ax\|^2 = \max_{\|y\|=1} \|\Sigma y\|^2 = \max_{\|y\|=1} (\sigma_1^2|y_1|^2 + \cdots + \sigma_n^2|y_n|^2) = \sigma_1^2.$$

Tomando raíz (que es una función creciente), se sigue que:

$$\boxed{\max_{\|x\|=1} \|Ax\| = \sigma_1}$$

Veamos que el máximo se alcanza en los vectores $x \in \mathcal{S}_{\lambda_1=\sigma_1^2}$ (el autoespacio asociado al mayor autovalor $\lambda_1 = \sigma_1^2$ de A^*A) tales que $\|x\| = 1$. Supongamos que $\{v_1, v_2, \cdots, v_t\}$ es una bon de $\mathcal{S}_{\lambda_1=\sigma_1^2}$, donde $v_1, \cdots, v_t$ son tales que

$$A^*Av_i = \sigma_1^2 v_i$$

para $i = 1, 2 \cdots, t$. Entonces, recordando cómo construimos la DVS de A, se sigue que $v_1, \cdots, v_t$ son las primeras t columnas de la matriz V y tenemos que, para $i = 1, \cdots, t$,

$$V^*v_i = \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_n^* \end{bmatrix} v_i = \begin{bmatrix} \langle v_i, v_1 \rangle \\ \langle v_i, v_2 \rangle \\ \vdots \\ \langle v_i, v_n \rangle \end{bmatrix} = e_i,$$

donde e_i es el i-ésimo vector de la base canónica de $\mathbb{C}^n$. Sea $x \in \mathcal{S}_{\lambda_1=\sigma_1^2} = gen\{v_1, v_2, \cdots, v_t\}$, entonces $x = a_1v_1 + \cdots + a_tv_t$, para ciertos $a_1, \cdots, a_t \in \mathbb{C}$. Entonces, por un lado, como $\|x\| = 1$, tenemos que (usando el Teorema de Pitágoras)

$$1 = \|x\| = \|a_1v_1 + \cdots + a_tv_t\| = (|a_1|^2 + \cdots + |a_t|^2)^{1/2}$$

y, por el otro lado,

$$\begin{aligned}\|Ax\| &= \|U\Sigma V^*x\| = \|\Sigma V^*(a_1v_1 + \cdots + a_tv_t)\| = \|\Sigma(a_1V^*v_1 + a_2V^*v_2 + \cdots + a_tV^*v_t)\| \\ &= \|\Sigma(a_1e_1 + a_2e_2 + \cdots a_te_t)\| = \|\sigma_1a_1e_1 + \sigma_1a_2e_2 + \cdots + \sigma_1a_te_t\| \\ &= \sigma_1\|a_1e_1 + \cdots a_te_t\| = \sigma_1(|a_1|^2 + \cdots + |a_t|^2)^{1/2} = \sigma_1.\end{aligned}$$

Entonces, $\max_{\|x\|=1} \|Ax\| = \sigma_1$ y el máximo se alcanza en los vectores $x \in \mathcal{S}_{\lambda_1=\sigma_1^2}$ (el autoespacio asociado al mayor autovalor $\lambda_1 = \sigma_1^2$ de A^*A) tal que $\|x\| = 1$.

De manera similar, se prueba que:

$$\min_{\|x\|=1} \|Ax\| = \sigma_n$$

y el mínimo se alcanza en los vectores $x \in \mathcal{S}_{\lambda_n=\sigma_n^2}$ (el autoespacio asociado al menor autovalor $\lambda_n = \sigma_n^2$ de A^*A) tal que $\|x\| = 1$.

■

Ejercicio 5.14. Sea $T : \mathbb{R}^2 \to \mathbb{R}^3$ la transformación lineal definida por $T(x) := Ax$ con

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ -3 & 1 \end{bmatrix}.$$

Hallar entre todos los $x \in \mathbb{R}^2$ de norma 1 aquellos que maximizan y minimizan $\|T(x)\|$ y calcular dicho máximo y mínimo.

Dem. Tal como vimos en la prueba del Teorema 5.7.1, para resolver el ejercicio, primero hallemos una DVS de A. Busquemos los autovalores y autovectores de

$$A^TA = \begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix}.$$

De hecho, $\chi_{A^TA}(t) = (11-t)^2 - 1$. Entonces $\lambda_1 = 12$ y $\lambda_2 = 10$ son los autovalores de A ordenados de mayor a menor y $v_1 = \frac{1}{\sqrt{2}}[1\ 1]^T$ y $v_2 = \frac{1}{\sqrt{2}}[-1\ 1]^T$ los autovectores asociados correspondientes. Entonces, $\sigma_1 = \lambda_1^{1/2} = 2\sqrt{3}$, $\sigma_2 = \lambda_2^{1/2} = \sqrt{10}$ y $V = [v_1\ v_2] = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$.

Por lo tanto, por el Teorema 5.7.1,

$$\max_{\|x\|=1} \|T(x)\| = \max_{\|x\|=1} \|Ax\| = \sigma_1 = 2\sqrt{3}$$

y el máximo se alcanza en $x \in \mathcal{S}_{\lambda_1=\sigma_1^2} = \mathcal{S}_{\lambda_1=12} = gen\{v_1\} = gen\{\frac{1}{\sqrt{2}}[1\ 1]^T\}$ tal que $\|x\| = 1$. Entonces, el vector x realiza el máximo si $x = \alpha\ \frac{1}{\sqrt{2}}[1\ 1]^T$ con $\alpha \in \mathbb{R}$ y

$$1 = \|x\| = \|\alpha\ \frac{1}{\sqrt{2}}[1\ 1]^T\| = |\alpha|\|\frac{1}{\sqrt{2}}[1\ 1]^T\| = |\alpha|.$$

Entonces $\alpha = \pm 1$ y

$$x_{\text{máx}} = \pm\frac{1}{\sqrt{2}}[1\ 1]^T.$$

De manera similar, tenemos que

$$\min_{\|x\|=1} \|T(x)\| = \min_{\|x\|=1} \|Ax\| = \sigma_2 = \sqrt{10}$$

y el mínimo se alcanza en $x \in \mathcal{S}_{\lambda_2=\sigma_2^2} = \mathcal{S}_{\lambda_2=10} = gen\{v_2\} = gen\{\frac{1}{\sqrt{2}}[-1\ 1]^T\}$ tal que $\|x\| = 1$. Entonces, el vector x realiza el mínimo si $x = \alpha\ \frac{1}{\sqrt{2}}[-1\ 1]^T$ con $\alpha \in \mathbb{R}$ y

$$1 = \|x\| = \|\alpha\ \frac{1}{\sqrt{2}}[-1\ 1]^T\| = |\alpha|\|\frac{1}{\sqrt{2}}[-1\ 1]^T\| = |\alpha|.$$

Entonces $\alpha = \pm 1$ y

$$x_{\text{mín}} = \pm\frac{1}{\sqrt{2}}[-1\ 1]^T.$$

■

5.8. Imagen por una transformación lineal de la circunferencia unitaria

Los siguientes ejercicios tienen como objetivo interpretar geométricamente la DVS de una matriz estudiando la imagen por una transformación lineal de la circunferencia unitaria

$$S_{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\}.$$

Ejercicio 5.16 a). Sea $T \in \mathcal{L}(\mathbb{R}^2)$ la transformación lineal definida por $T(x) := Ax$. Hallar la imagen por T de la circunferencia unitaria $S_1 := \{x \in \mathbb{R}^2 : \|x\| = 1\}$. Donde $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

Dem. Recordar que

$$T(S_1) = \{T(x) : x \in S_1\} = \{T(x) : \|x\| = 1\}.$$

Entonces

$$z \in T(S_1) \text{ si y sólo si } z = T(x) = Ax$$

para cierto $x \in \mathbb{R}^2$ con $\|x\| = 1$.

Supongamos que $A = U\Sigma V^T$ es una DVS de A y consideremos el cambio de variable $x = Vy$. Entonces, como V es ortogonal, por la Proposición 5.1.1 vale que $\|x\| = \|y\|$. Entonces

$$z \in T(S_1) \text{ si y sólo si } z = T(x) = Ax = AVy = U\Sigma V^T Vy = U\Sigma y$$

con $\|y\| = 1$. Entonces, como $A \in \mathbb{R}^{2\times 2}$,

$$z \in T(S_1) \text{ si y sólo si } z = U\Sigma y = \sigma_1 u_1 y_1 + \sigma_2 u_2 y_2$$

con $y = [y_1\ y_2]^T$ tal que $\|y\|^2 = y_1^2 + y_2^2 = 1$.

Si consideramos $\mathcal{B} := \{u_1, u_2\}$ que es una bon de $\mathbb{R}^2$. Tenemos que

$$[z]^{\mathcal{B}} := [w_1\ w_2]^T = [\sigma_1 y_1\ \sigma_2 y_2]^T,$$

con $y = [y_1\ y_2]^T$ tal que $\|y\|^2 = y_1^2 + y_2^2 = 1$. Por lo tanto, $z \in T(S_1)$ si y sólo si $z = w_1 u_1 + w_2 u_2$ con $w_1 = \sigma_1 y_1, w_2 = \sigma_2 y_2$ tal que $\|y\|^2 = y_1^2 + y_2^2 = 1$.

Ahora calculemos los valores singulares de A y las columnas u_1, u_2 de U. Para eso, busquemos los autovalores y autovectores de

$$A^T A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

De hecho $\chi_{A^T A}(t) = (2-t)^2 - 4$. Entonces, $\lambda_1 = 4$, $\lambda_2 = 0$ son los autovalores de $A^T A$ ordenados de mayor a menor y $v_1 = \frac{1}{\sqrt{2}}[1\ 1]^T, v_2 = \frac{1}{\sqrt{2}}[1\ -1]^T$ son los autovectores asociados correspondientes.

Por lo tanto, $\sigma_1 = \lambda_1^{1/2} = 2$, $\sigma_2 = 0$ y además, $u_1 = \frac{Av_1}{\sigma_1} = \frac{1}{\sqrt{2}}\frac{[2\ 2]^T}{2} = \frac{1}{\sqrt{2}}[1\ 1]^T$ y $u_2 = \frac{1}{\sqrt{2}}[-1\ 1]^T$ (cualquier vector para completar a una bon de $\mathbb{R}^2$).

Entonces, como $[z]^{\mathcal{B}} = [w_1\ w_2]^T = [\sigma_1 y_1\ \sigma_2 y_2]^T$, despejando tenemos que

$$y_1 = \frac{w_1}{\sigma_1} = \frac{w_1}{2},\ w_2 = 0.$$

Entonces,

$$z \in T(S_1) \text{ si y sólo si } z = w_1 u_1 \text{ con } \frac{w_1^2}{\sigma_1^2} = y_1^2 \leq y_1^2 + y_2^2 = 1.$$

Entonces, $w_1^2 \leq \sigma_1^2$, por lo tanto $-2 = -\sigma_1 \leq w_1 \leq \sigma_1 = 2$ y $w_2 = 0$. Entonces

$$\begin{aligned} T(S_1) &= \{z \in \mathbb{R}^2 : z = w_1 u_1 + w_2 u_2, -2 \leq w_1 \leq 2,\ w_2 = 0\} = \{z \in \mathbb{R}^2 : z = w_1 u_1, -2 \leq w_1 \leq 2\} \\ &= \{z \in \mathbb{R}^2 : z = \frac{w_1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, -2 \leq w_1 \leq 2\}. \end{aligned}$$

Entonces la imagen por T de la circunferencia unitaria de $\mathbb{R}^2$ es el segmento en $\mathbb{R}^2$ de extremos $\frac{-2}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\frac{2}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. ■

Ejercicio 5.18 b). Sea $T \in \mathcal{L}(\mathbb{R}^3)$ la transformación lineal definida por $T(x) := Ax$. Hallar la imagen por T de la esfera unitaria $S_2 := \{x \in \mathbb{R}^3 : \|x\| = 1\}$. Donde $A = \begin{bmatrix} 2 & 4 & -2 \\ 4 & 10 & 0 \\ -2 & 0 & 10 \end{bmatrix}$.

Dem. Recordar que

$$T(S_2) = \{T(x) : x \in S_2\} = \{T(x) : \|x\| = 1\}.$$

De la misma manera que hicimos en el **Ejercicio 5.16 a)**, si $A = U\Sigma V^T$ es una DVS de A y consideramos el cambio de variable $x = Vy$. Entonces, como V es ortogonal, vale que $\|x\| = \|y\|$. Entonces,

$$z \in T(S_2) \text{ si y sólo si } z = U\Sigma V^T x = U\Sigma y = \sigma_1 u_1 y_1 + \sigma_2 u_2 y_2 + \sigma_3 u_3 y_3$$

con $y = [y_1\ y_2\ y_3]^T$ tal que $\|y\|^2 = y_1^2 + y_2^2 + y_3^2 = 1$. Sea $\mathcal{B} := \{u_1, u_2, u_3\}$ que es una bon de $\mathbb{R}^3$, entonces

$$[z]^{\mathcal{B}} = [w_1\ w_2\ w_3]^T = [\sigma_1 y_1\ \sigma_2 y_2\ \sigma_3 y_3]^T,$$

con $y = [y_1\ y_2\ y_3]^T$ tal que $\|y\|^2 = y_1^2 + y_2^2 + y_3^2 = 1$.

Calculemos los valores singulares de A y las columnas u_1, u_2, u_3 de la matriz U. Para eso, busquemos los autovalores y autovectores de

$$A^T A = \begin{bmatrix} 24 & 48 & -24 \\ 48 & 116 & -8 \\ -24 & -8 & 104 \end{bmatrix}.$$

De hecho $\chi_{A^T A}(t) = -t^3 + 244t^2 - 14400t$. Entonces, $\lambda_1 = 144$, $\lambda_2 = 100$, $\lambda_3 = 0$ son los autovalores de $A^T A$ ordenados de mayor a menor y $v_1 = \frac{1}{\sqrt{6}}[-1\ -2\ 1]^T$, $v_2 = \frac{1}{\sqrt{5}}[0\ 1\ 2]^T$ y $v_3 = \frac{1}{\sqrt{30}}[5\ -2\ 1]^T$ los autovectores asociados correspondientes. Por lo tanto, $\sigma_1 = \lambda_1^{1/2} = 12$, $\sigma_2 = 100^{1/2} = 10$, $\sigma_3 = 0$ y $u_1 = \frac{Av_1}{\sigma_1} = \frac{[-12\ -24\ 12]^T}{\sqrt{6}\ 12} = \frac{1}{\sqrt{6}}[-1\ -2\ -1]^T$, $u_2 = \frac{Av_2}{\sigma_2} = \frac{[0\ 10\ 20]^T}{\sqrt{5}\ 10} = \frac{1}{\sqrt{5}}[0\ 1\ 2]^T$ y u_3 cualquier vector para completar a una bon de $\mathbb{R}^3$.

Por lo tanto, como $[z]^{\mathcal{B}} = [w_1\ w_2\ w_3]^T = [\sigma_1 y_1\ \sigma_2 y_2\ \sigma_3 y_3]^T$, tenemos que

$$y_1 = \frac{w_1}{\sigma_1} = \frac{w_1}{12},\ y_2 = \frac{w_2}{\sigma_2} = \frac{w_2}{10},\ w_3 = 0.$$

Entonces, $z \in T(S_1)$ si y sólo si $z = w_1u_1 + w_2u_2$ con $\frac{w_1^2}{144} + \frac{w_2^2}{100} = \frac{w_1^2}{\sigma_1^2} + \frac{w_2^2}{\sigma_2^2} = y_1^2 + y_2^2 \leq y_1^2 + y_2^2 + y_3^2 = 1$ y $w_3 = 0$. Entonces,

$$\begin{aligned} T(S_1) &= \{z \in \mathbb{R}^3 : z = w_1u_1 + w_2u_2 + w_3u_3,\ \frac{w_1^2}{144} + \frac{w_2^2}{100} \leq 1, w_3 = 0\} \\ &= \{z \in \mathbb{R}^3 : z = w_1u_1 + w_2u_2,\ \frac{w_1^2}{144} + \frac{w_2^2}{100} \leq 1\} \\ &= \{z \in \mathbb{R}^3 : z = w_1\frac{1}{\sqrt{6}}[-1\ \ -2\ \ -1]^T + w_2\frac{1}{\sqrt{5}}[0\ 1\ 2]^T,\ \frac{w_1^2}{144} + \frac{w_2^2}{100} \leq 1\}. \end{aligned}$$

Entonces la imagen por T de la circunferencia unitaria de $\mathbb{R}^3$ es la superficie en $\mathbb{R}^3$ contenida en el plano generado por $\mathrm{col}(A) = gen\{u_1, u_2\} = gen\{\frac{1}{\sqrt{6}}[-1\ \ -2\ \ -1]^T, \frac{1}{\sqrt{5}}[0\ 1\ 2]^T\}$. Dicha superficie contiene al origen y está limitada por la elipse centrada en el origen, cuyo eje mayor tiene longitud $2\sigma_1 = 24$, y está contenido en la recta generada por $u_1 = \frac{1}{\sqrt{6}}[-1\ \ -2\ \ -1]^T$ y su eje menor tiene longitud $2\sigma_2 = 20$, y está contenido en la recta generada por $u_2 = \frac{1}{\sqrt{5}}[0\ 1\ 2]^T$. ■

5.9. Descomposición Polar

A continuación veremos que toda matriz $A \in \mathbb{C}^{n\times n}$ (cuadrada) se puede factorizar como $A = W|A|$, donde $U \in \mathbb{C}^{n\times n}$ es unitaria y $|A| \in \mathbb{C}^{n\times n}$ es semidefinida positiva.

Teorema 5.9.1. *Sea $A \in \mathbb{C}^{n\times n}$ (cuadrada) entonces existen matrices $U \in \mathbb{C}^{n\times n}$ unitaria y $|A| \in \mathbb{C}^{n\times n}$ semidefinida positiva tales que*

$$A = W|A|.$$

A esta factorización se la conoce como *descomposición polar* de A.

Dem. Sea $= U\Sigma V^*$ una DVS de A. Como $A \in \mathbb{C}^{n\times n}$ (es decir es cuadrada), tenemos que $U, V \in \mathbb{C}^{n\times n}$ unitarias y

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

Definimos

$$|A| := V\Sigma V^* = V \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_1 \end{bmatrix} V^*,$$

entonces como

$$|A|^* = (V\Sigma V^*)^* = V\Sigma^* V^* = V\Sigma V^* = |A|,$$

es decir $|A|$ es hermítica y los autovalores de $|A|$ son $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ (observar que definimos $|A|$ a partir de su digonalización unitaria) entonces, por la Proposición 3.5.1 5.4.1, $|A|$ es semidefinida positiva. Por otra parte, usando que $VV^* = V^*V = I_n$, tenemos que

$$A = U\Sigma V^* = UV^*V\Sigma V^* = (UV^*)|A|.$$

Tomando $W := UV^*$ tenemos que W es unitaria, pues $W^*W = (UV^*)^*(UV^*) = VU^*UV^* = VI_nV^* = I_n$ y

$$A = W|A|,$$

como queríamos ver. ■

Notar que como $A \in \mathbb{C}^{n\times n}$ podemos calcular su determinante. Además, dado que $|A|$ es semidefinida positiva, tenemos que todos sus autovalores son no negativos y entonces $\det(|A|) \geq 0$. Por otra parte, recordemos que como W es unitaria, vale que $|\det(W)| = 1$. Entonces, tenemos una fórmula para el módulo del determinante de A :

$$|\det(A)| = |\det(W|A|)| = |\det(W)||\det(|A|)| = 1 \cdot \det(V\Sigma V^*) = \det(\Sigma) = \sigma_1\sigma_2\cdots\sigma_n.$$

Ejercicio 5.C. Hallar una descomposición polar para $A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$.

Dem. Siguiendo las ideas de la demostración del Teorema 5.9.1, obtengamos primero una DVS de A. Busquemos los autovalores y autovectores de

$$A^*A = \begin{bmatrix} 1 & -i \\ -i & -1 \end{bmatrix} \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix} = \begin{bmatrix} 2 & 2i \\ -2i & 2 \end{bmatrix}.$$

De hecho, $\chi_{A^*A}(x) = x^2 - 4x$. Entonces $\lambda_1 = 4$ y $\lambda_2 = 0$ son los autovalores de A ordenados de mayor a menor y $v_1 = \frac{1}{\sqrt{2}}[i\ 1]^T$ y $v_2 = \frac{1}{\sqrt{2}}[-i\ 1]^T$ los autovectores asociados correspondientes. Entonces, $\sigma_1 = \lambda_1^{1/2} = 2$ y $\sigma_2 = \lambda_2^{1/2} = 0$ y $V = \begin{bmatrix} \frac{i}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$. Para obtener U, tenemos que $u_1 = \frac{Av_1}{\sigma_1} = \frac{[2i\ -2]}{\sqrt{2}\ 2} = \frac{1}{\sqrt{2}}[i-1]^T$ y $u_2 = \frac{1}{\sqrt{2}}[-1\ i]^T$ (cualquier vector para completar a una bon de $\mathbb{C}^2$). Entonces $U = \begin{bmatrix} \frac{i}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{bmatrix}$.

Por lo tanto, tal como vimos arriba tenemos que

$$W = UV^* = \begin{bmatrix} \frac{i}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1-i}{2} & \frac{i-1}{2} \\ \frac{i-1}{2} & \frac{-1+i}{2} \end{bmatrix} = \frac{i-1}{2} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}.$$

y

$$|A| = V\Sigma V^* = \begin{bmatrix} \frac{i}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}.$$

Entonces, una descomposición polar de A es

$$A = W|A| = \left(\frac{i-1}{2} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}.$$

■

CAPÍTULO 6

Formas Cuadraticas

6.1. Definición de formas cuadráticas

Definición. Un forma cuadrática es una función $Q : \mathbb{R}^n \to \mathbb{R}$ que se puede expresar de la forma

$$Q(x) = x^T A x \text{ con } A = A^T \in \mathbb{R}^{n\times n} \text{ (simétrica).}$$

Observar que si Q es una forma cuadrática y $\alpha \in \mathbb{R}$ entonces

$$Q(\alpha x) = (\alpha x)^T A(\alpha x) = \alpha\alpha x^T A x = \alpha^2 Q(x). \tag{6.1}$$

Sea $B \in \mathbb{R}^{n\times n}$ no necesariamente simétrica y $Q : \mathbb{R}^n \to \mathbb{R}$ definida como

$$Q(x) = x^T B x.$$

Entonces, como $Q(x) \in \mathbb{R}$, tenemos que $[Q(x)]^T = Q(x)$. Entonces

$$2Q(x) = Q(x) + Q(x) = Q(x) + [Q(x)]^T = x^T Bx + (x^T Bx)^T = x^T Bx + x^T B^T x = x^T (B + B^T)x.$$

Por lo tanto

$$Q(x) = \frac{x^T (B + B^T)x}{2} = x^T \frac{(B + B^T)}{2} x.$$

En este caso, la matriz $\frac{B+B^T}{2}$ es simétrica, pues $(\frac{B+B^T}{2})^T = \frac{B^T+B}{2} = \frac{B+B^T}{2}$. Entonces, a la función $Q(x) = x^T Bx$ con $B \in \mathbb{R}^{n\times n}$ la pudimos expresar de la forma $Q(x) = x^T Ax$ con $A := \frac{B+B^T}{2}$ simétrica.

Ejercicio 6.A. Sea $Q : \mathbb{R}^3 \to \mathbb{R}$ definida por

$$Q(x) = x_1^2 + 2x_2^2 + 3x_3^3 + 10x_1x_2 + 16x_2x_3 + 26x_1x_2.$$

Comprobar que Q es una forma cuadrática y expresarla de la forma $Q(x) = x^T Ax$ con $A \in \mathbb{R}^{3\times 3}$ simétrica.

Dem. Observar que $Q(x) = x_1^2 + 2x_2^2 + 3x_3^3 + 10x_1x_2 + 16x_2x_3 + 26x_1x_2 = x_1^2 + 2x_2^2 + 3x_3^2 + 36x_1x_2 + 16x_2x_3 = x_1^2 + 2x_2^2 + 3x_3^2 + 2(18x_1x_2 + 8x_2x_3)$.

Si $A \in \mathbb{R}^{3\times 3}$ es tal que $A = A^T$. Entonces A tiene la siguiente forma

$$A = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix},$$

con $a, b, c, d, e, f \in \mathbb{R}$. Si $Q(x) = x^T Ax$, entonces

$$\begin{aligned} Q(x) = x^T Ax = [x_1 \ x_2 \ x_3] \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ = ax_1^2 + dx_2^2 + fx_3^2 + 2bx_1x_2 + 2cx_1x_3 + 2ex_2x_3. \end{aligned}$$

Entonces $Q(x) = x_1^2 + 2x_2^2 + 3x_3^2 + 2(18x_1x_2 + 8x_2x_3) = ax_1^2 + dx_2^2 + fx_3^2 + 2bx_1x_2 + 2cx_1x_3 + 2ex_2x_3$. Comparando, nos queda que $a = 1, d = 2, f = 3, b = 18, c = 0, e = 8$. Entonces,

$$A = \begin{bmatrix} 1 & 18 & 0 \\ 18 & 2 & 8 \\ 0 & 8 & 3 \end{bmatrix}$$

y tenemos que $Q(x) = x^T Ax$. ■

Ejercicio 6.B. Sea $Q : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$Q(x) = x^T \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} x.$$

Comprobar que Q es una forma cuadrática y expresarla de la forma $Q(x) = x^T Ax$ con $A \in \mathbb{R}^{2\times 2}$ simétrica.

Dem. La matriz $B := \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ no es simétrica. Sin embargo, arriba vimos que si tomamos

$$A = \frac{B + B^T}{2} = \frac{\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}}{2} = \begin{bmatrix} 1 & \frac{5}{2} \\ \frac{5}{2} & 4 \end{bmatrix}$$

entonces $Q(x) = x^T Ax$ y en este caso A resulta simétrica y Q es una forma cuadrática. ■

6.2. Eliminación de productos cruzados

Sea $Q : \mathbb{R}^n \to \mathbb{R}$ la forma cuadrática definida por $Q(x) = x^T Ax$ con $A = A^T$. Como A es simétrica, vimos que existe $P \in \mathbb{R}^{n\times n}$ ortogonal (cuyas columnas son autovectores de A que forman una bon de $\mathbb{R}^n$) y $\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$ con $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$ los autovalores de A tales que:

$$A = P\Lambda P^T.$$

Con el objetivo de eliminar los términos cruzados de la forma $x_i x_j$ con $i \neq j$, vamos a considerar el cambio de variables

$$x = Py.$$

Entonces

$$Q(x) = x^T Ax = (Py)^T A(Py) = y^T P^T APy = y^T \Lambda y =: \tilde{Q}(y).$$

Por lo tanto

$$Q(x) = \tilde{Q}(y) = y^T \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} y = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

Ejercicio 6.3 d). Expresar la forma cuadrática $Q : \mathbb{R}^3 \to \mathbb{R}$ definida por

$$Q(x) = x_3^2 + x_1x_2 + x_1x_3 + x_2x_3$$

como $x^T Ax$ con $A \in \mathbb{R}^{3\times 3}$ simétrica, diagonalizar ortogonalmente $A = P\Lambda P^T$ y mediante el cambio de variables $x = Py$ expresar Q sin productos cruzados.

Dem. Tal como hicimos en el **Ejercicio 6.A**, podemos ver que

$$Q(x) = x^T \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix} x$$

con $A := \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$ simétrica. Para diagonalizar ortogonalmente A, calculemos su polinomio característico, sus autovalores y obtengamos una bon de autovectores de $\mathbb{R}^3$. De hecho

$$\chi_A(t) = \det(A - tI) = -t^3 + t^2 + \frac{3t}{4}.$$

Entonces, los autovalores de A son $\lambda_1 = -\frac{1}{2}$, $\lambda_2 = \frac{3}{2}$ y $\lambda_3 = 0$. Como

$$\mathrm{nul}(A + \frac{1}{2}I) = \mathrm{nul}(\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{3}{2} \end{bmatrix} = gen\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \},$$

$$\mathrm{nul}(A - \frac{3}{2}I) = \mathrm{nul}(\begin{bmatrix} -\frac{3}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} = gen\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \},$$

$$\mathrm{nul}(A) = \mathrm{nul}(\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix} = gen\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \}.$$

Tenemos que $\{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}\}$ es una bon de autovectores de $\mathbb{R}^3$. Tomemos

$$P := \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Entonces

$$A = P \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} P^T.$$

Como vimos arriba, si hacemos el cambio de variables $x = Py$, tenemos que

$$Q(x) = \tilde{Q}(y) = y^T \Lambda y = -\frac{1}{2}y_1^2 + \frac{3}{2}y_2^2.$$

■

6.3. Clasificación de formas cuadráticas

Definición. Dada una forma cuadrática $Q : \mathbb{R}^n \to \mathbb{R}$ decimos que:

- Q es *definida positiva* si $Q(x) > 0$ para todo $x \in \mathbb{R}^n \setminus \{0\}$.
- Q es *semidefinida positiva* si $Q(x) \geq 0$ para todo $x \in \mathbb{R}^n$.
- Q es *definida negativa* si $Q(x) < 0$ para todo $x \in \mathbb{R}^n \setminus \{0\}$.
- Q es *semidefinida negativa* si $Q(x) \leq 0$ para todo $x \in \mathbb{R}^n$.
- Q es *indefinida* si no es ninguna de las anteriores. Es decir, existen $x_1, x_2 \in \mathbb{R}^n$ tales que $Q(x_1) > 0$ y $Q(x_2) < 0$.

Si expresamos la forma cuadrática

$$Q(x) = x^T Ax$$

con A simétrica, entonces recordando lo que vimos en la Proposición 5.4.1, tenemos que:

Proposición 6.3.1. *Sea Q una forma cuadrática en $\mathbb{R}^n$ donde $Q(x) = x^T Ax$ con $A \in \mathbb{R}^{n\times n}$ simétrica y sean $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{R}$ los autovalores de A. Entonces:*

- *Q es definida positiva si y sólo si A es definida positiva si y sólo si $\lambda_i > 0$ para todo $i = 1, 2, \cdots, n$.*
- *Q es semidefinida positiva si y sólo si A es semidefinida positiva si y sólo si $\lambda_i \geq 0$ para todo $i = 1, 2, \cdots, n$.*
- *Q es definida negativa si y sólo si A es definida negativa si y sólo si $\lambda_i < 0$ para todo $i = 1, 2, \cdots, n$.*
- *Q es semidefinida negativa si y sólo si A es semidefinida negativa si y sólo si $\lambda_i \leq 0$ para todo $i = 1, 2, \cdots, n$.*
- *Q es indefinida si y sólo si A es indefinida si y sólo A tiene algún autovalor negativo y algún autovalor positivo.*

6.4. Conjuntos de nivel

Dada una forma cuadrática $Q : \mathbb{R}^n \to \mathbb{R}$, para $c \in \mathbb{R}$ se define el *conjunto de nivel* c como

$$\mathcal{N}_c(Q) = \{x \in \mathbb{R}^n : Q(x) = c\}.$$

Si $Q(x) = x^T Ax$ con $A = A^T$ y diagonalizamos ortgonalmente a A de la forma $A = P\Lambda P^T$, resulta más simple determinar qué tipo de conjunto es $\mathcal{N}_c(Q)$ empleando el cambio de variables $x = Py$. En ese caso, tendremos que

$$Q(x) = \tilde{Q}(y) = y^T \Lambda y$$

con Λ diagonal. Para $c \in \mathbb{R}$ dado, tenemos que

$$Q(x) = c \text{ si y sólo si } \tilde{Q}(y) = c.$$

Veremos a continuación que el conjunto de nivel $\mathcal{N}_c(Q)$ se obtiene rotando el conjunto de nivel $\mathcal{N}_c(\tilde{Q})$ en los ejes principales de A (determinados por las columnas ortonormales de la matriz P). En particular, ambos conjuntos de nivel tienen la misma forma geométrica.

Ejercicio 6.2 a). Sea $Q(x) = 9x_1^2 + 3x_2^2 - 8x_1x_2$ en $\mathbb{R}^2$. Clasificar dicha forma cuadrática y graficar sus conjuntos de nivel $\mathcal{N}_c(Q)$.

Dem. Tal como hicimos en los ejercicios anteriores, tenemos que

$$Q(x) = x^T \begin{bmatrix} 9 & -4 \\ -4 & 3 \end{bmatrix} x.$$

Entonces si $A = \begin{bmatrix} 9 & -4 \\ -4 & 3 \end{bmatrix}$, tenemos que $Q(x) = x^T Ax$ con A simétrica.

Ahora, diagonalizemos ortogonalmente A. Para eso, calculemos su polinomio característico, sus autovalores y obtengamos una bon de autovectores de $\mathbb{R}^2$. De hecho

$$\chi_A(t) = \det(A - tI) = (9 - t)(3 - t) - 16 = t^2 - 12t + 11.$$

Entonces, los autovalores de A son $\lambda_1 = 1$, $\lambda_2 = 11$. Los autoespacios asociados son:

$$\text{nul}(A - 1I) = \text{nul}(\begin{bmatrix} 8 & -4 \\ -4 & 2 \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\},$$

$$\text{nul}(A - 11I) = \text{nul}(\begin{bmatrix} -2 & -4 \\ -4 & -8 \end{bmatrix}) = gen\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\}.$$

Entonces $\{\frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}\}$ es una bon de $\mathbb{R}^2$. Si tomamos

$$P = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

tenemos que

$$A = P \begin{bmatrix} 1 & 0 \\ 0 & 11 \end{bmatrix} P^T.$$

Entonces, como los autovalores de A son ambos positivos, tenemos que A es **definida positiva** y por lo tanto Q es **definida positiva**.

Por otra parte, con el cambio de variables $x = Py$, tenemos que

$$Q(x) = \tilde{Q}(y) = 1y_1^2 + 11y_2^2.$$

Por lo tanto, dado $c \in \mathbb{R}$ tenemos que las conjuntos de nivel $\mathcal{N}_c(Q)$ son todos los $x \in \mathbb{R}^2$ tales que

$$Q(x) = 9x_1^2 + 3x_2^2 - 8x_1x_2 = c.$$

Veamos 3 casos posibles:

- Si $c < 0$, como Q es definida positiva, entonces
$$\mathcal{N}_c(Q) = \emptyset.$$
- Si $c = 0$, como Q es definida positiva, entonces
$$\mathcal{N}_c(Q) = \{0, 0\}.$$
- Si $c > 0$. Como $Q(x) = c$ si y sólo si $\tilde{Q}(y) = c$, se sigue que $x \in \mathcal{N}_c(Q)$ si y sólo si $y = [y_1 \ y_2]^T \in \mathbb{R}^2$ cumple que
$$\tilde{Q}(y) = y_1^2 + 11y_2^2 = c,$$
ó equivalentemente $(c > 0)$,
$$\frac{y_1^2}{(\sqrt{c})^2} + \frac{y_2^2}{(\sqrt{\frac{c}{11}})^2} = 1.$$

Entonces, los conjuntos de nivel $\mathcal{N}_c(Q) = \{x \in \mathbb{R}^2 : 9x_1^2 + 3x_2^2 - 8x_1x_2 = c\}$ $(c > 0)$ son elipses en $\mathbb{R}^2$ en los nuevos ejes y_1, y_2 determinados por los versores $v_1 := \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ y $v_2 := \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ que son las columnas de P y que son una rotación de un ángulo (positivo) de los ejes originales x_1, x_2 (ver Figura 6.1). ■

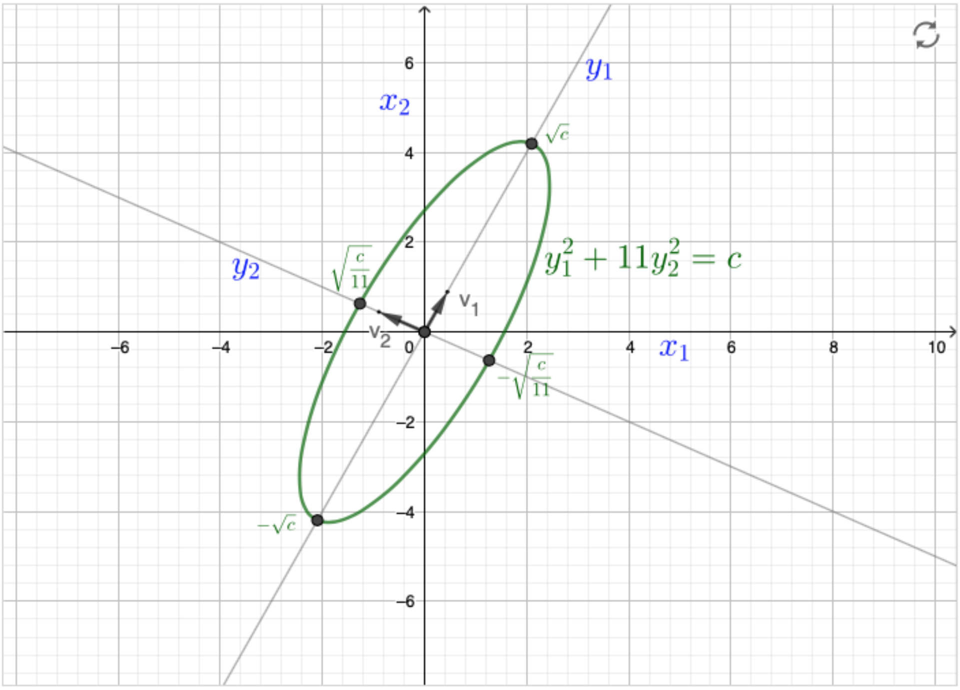

Figura 6.1: Ejercicio 6.2 a)

Ejercicio 6.C (De Final)**.** Sea $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ el producto interno canónico en $\mathbb{R}^n$ y sea $A \in \mathbb{R}^{n\times n}$. Demostrar que

$$\langle x, y \rangle_A := \langle Ax, y \rangle$$

define un producto interno en $\mathbb{R}^n$ si y sólo si A es definida positiva.

Recordemos que A es definida positiva si A es simétrica, es decir $A = A^T$ y $\langle Ax, x \rangle > 0$ para todo $x \in \mathbb{R}^n \setminus \{0\}$.

Dem. Como el producto interno canónico es lineal en la primera variable, para cualquier matriz $A \in \mathbb{R}^{n\times n}$, la función

$$\langle x, y \rangle_A := \langle Ax, y \rangle$$

es lineal en la primera variable. De hecho,

$$\begin{aligned} \langle x+z, y \rangle_A &= \langle A(x+z), y \rangle = \langle Ax + Az, y \rangle = \langle Ax, y \rangle + \langle Az, y \rangle \\ &= \langle x, y \rangle_A + \langle z, y \rangle_A \text{ para todo } x, y, z \in \mathbb{R}^n. \end{aligned}$$

Y también tenemos que

$$\langle \alpha x, y \rangle_A = \langle A(\alpha x), y \rangle = \langle \alpha Ax, y \rangle = \alpha \langle Ax, y \rangle = \alpha \langle x, y \rangle_A \text{ para todo } x, y \in \mathbb{R}^n, \alpha \in \mathbb{R}.$$

Veamos entonces, que $\langle \cdot, \cdot \rangle_A$ define un producto interno si y sólo si A es definida positiva.

Supongamos que A es definida positiva, entonces por un lado tenemos que $A = A^T$ y usando que el producto interno canónico es simétrico, tenemos que

$$\langle x, y \rangle_A = \langle Ax, y \rangle = \left\langle x, A^T y \right\rangle = \langle x, Ay \rangle = \langle Ay, x \rangle = \langle y, x \rangle_A ,$$

donde usamos que siempre vale que $\langle Ax, y \rangle = \left\langle x, A^T y \right\rangle$.

Entonces, $\langle x, y \rangle_A = \langle y, x \rangle_A$ para todo $x, y \in \mathbb{R}^n$. Finalmente, como A es definida positiva tenemos que

$$\langle x, x \rangle_A = \langle Ax, x \rangle > 0 \text{ para todo } x \neq 0.$$

Entonces, como la función $\langle \cdot, \cdot \rangle_A$ es lineal en la primera variable, simétrica y cumple que $\langle x, x \rangle_A > 0$ para todo $x \neq 0$, concluimos que $\langle \cdot, \cdot \rangle_A$ define un producto interno en $\mathbb{R}^n$.

Recíprocamente, supongamos que la función $\langle \cdot, \cdot \rangle_A$ define un producto interno en $\mathbb{R}^n$. Entonces, en particular, tenemos que para todo $x, y \in \mathbb{R}^n$, vale que $\langle x, y \rangle_A = \langle y, x \rangle_A$. Por lo tanto, usando que el producto interno canónico es simétrico, se sigue que

$$\langle Ax, y \rangle = \langle x, y \rangle_A = \langle y, x \rangle_A = \langle Ay, x \rangle = \langle x, Ay \rangle \text{ para todo } x, y \in \mathbb{R}^n. \tag{6.2}$$

Recordemos que siempre vale que $\langle Ax, y \rangle = \left\langle x, A^T y \right\rangle$. Entoces, usando la ecuación (6.2), tenemos que

$$\left\langle x, A^T y \right\rangle = \langle Ax, y \rangle = \langle x, Ay \rangle ,$$

para todo $x, y \in \mathbb{R}^n$. Entonces, usando la linealidad del producto interno canónico, tenemos que

$$\left\langle x, (A^T - A)y \right\rangle = \left\langle x, A^T y \right\rangle - \langle x, Ay \rangle = 0 \text{ para todo } x, y \in \mathbb{R}^n.$$

Entonces, tomando $x := (A^T - A)y$, tenemos que

$$\left\langle (A^T - A)y, (A^T - A)y \right\rangle = \|(A^T - A)y\|^2 = 0,$$

para todo $y \in \mathbb{R}^n$. Como $\| \cdot \|$ es la norma inducida del producto interno canónico, se sigue que

$$(A^T - A)y = 0 \text{ para todo } y \in \mathbb{R}^n.$$

Es decir, $A^T y = Ay$ para todo $y \in \mathbb{R}^n$. Entonces concluimos que $A^T = A$ y A es simétrica.

Finalmente, como $\langle \cdot, \cdot \rangle_A$ define un producto interno en $\mathbb{R}^n$, vale que

$$\langle Ax, x \rangle = \langle x, x \rangle_A > 0$$

para todo $x \neq 0$. Entonces, como A es simétrica y $\langle Ax, x \rangle > 0$ para todo $x \neq 0$, concluimos que A es definida positiva. ■

Ejercicio 6.D (De Final). Determinar los valores de $k \in \mathbb{R}$ (si existen) para los cuales la forma cuadrática definida en $\mathbb{R}^2$ por

$$Q(x) = x_1^2 + kx_1x_2 + x_2^2,$$

es definida positiva.

Dem. Observar que

$$Q(x) = x^T \begin{bmatrix} 1 & \frac{k}{2} \\ \frac{k}{2} & 1 \end{bmatrix} x.$$

Sea $A := \begin{bmatrix} 1 & \frac{k}{2} \\ \frac{k}{2} & 1 \end{bmatrix}$. Entonces, claramente $A^T = A$. Por la Proposición 6.3.1, Q es definida positiva si y sólo si A es definida positiva, si y sólo si los autovalores de A son positivos.

El polinomio característico de A es

$$\chi_A(t) = (1 - t)^2 - \frac{k^2}{4}.$$

Entonces, los autovalores de A son

$$\lambda_1 = 1 + \frac{|k|}{2} > 0 \text{ y } \lambda_2 = 1 - \frac{|k|}{2}.$$

Claramente λ_1 es positivo para cualquier valor de k y queremos también que $\lambda_2 > 0$. Entonces, $|k| < 2$. Por lo tanto, A es definida positiva si y sólo $|k| < 2$ ó, equivalentemente, si $-2 < k < 2$.

También podemos resolver este ejercicio usando el Teorema 5.4.2. Dicho teorema afirma que A es definida positiva si y sólo si el determinate de todas sus submatrices principales es positivo.

En este caso, tenemos dos submatrices principales: $A^{(1)} = [1]$ cuyo determinante es $\det(A^{(1)}) = 1 > 0$ y $A^{(2)} = A$ cuyo determinante es $\det(A^{(2)}) = \det(A) = 1 - \frac{k^2}{4}$. Si imponemos que $1 - \frac{k^2}{4} > 0$, entonces, $k^2 < 4$ y, tomando raíz, nos queda que $|k| < 2$, ó equivalentemente, $-2 < k < 2$ y llegamos al mismo resultado. ■

6.5. Optimización de formas cuadráticas

El siguiente ejercicio nos va a permitir entender la demostración del Teorema de Rayleigh que veremos luego.

Ejercicio 6.E. Sea $Q : \mathbb{R}^4 \to \mathbb{R}$ la forma cuadrática definida por

$$Q(x) = 5x_1^2 - 3x_2^2 + 2x_3^2 + 5x_4^2.$$

a) Hallar los valores máximos y mínimos de los *cocientes de Rayleigh*

$$\frac{Q(x)}{\|x\|^2}, \quad x \neq 0.$$

b) Hallar el conjunto de todos los x_M y $x_m \in \mathbb{R}^4 \setminus \{0\}$ tales que

$$\frac{Q(x_M)}{\|x_M\|^2} = \max_{x\neq 0} \frac{Q(x)}{\|x\|^2} \text{ y } \frac{Q(x_m)}{\|x_m\|^2} = \min_{x\neq 0} \frac{Q(x)}{\|x\|^2}.$$

c) Hallar el conjunto de todos los $\hat{x}_M \in S_3$ y $\hat{x}_m \in S_3$ tales que

$$Q(\hat{x}_M) = \max_{x\in S_3} Q(x) = \max_{\|x\|=1} Q(x) \text{ y } Q(\hat{x}_m) = \min_{x\in S_3} Q(x) = \min_{\|x\|=1} Q(x).$$

Como Q no tiene productos cruzados nos va a resultar sencillo resolver el ejercicio. Luego veremos cómo vamos a proceder en el caso en que Q tenga productos cruzados.

Dem. $a)$: Observar que, como $-3x_2^2 \leq 5x_2^2$ y $2x_3^2 \leq 5x_3^2$, tenemos que, para todo $x \in \mathbb{R}^4$,

$$Q(x) = 5x_1^2 - 3x_2^2 + 2x_3^2 + 5x_4^2 \leq 5x_1^2 + 5x_2^2 + 5x_3^2 + 5x_4^2 = 5\|x\|^2.$$

Por otro lado, como $-3x_1^2 \leq 5x_1^2$, $-3x_3^2 \leq 2x_3^2$ y $-3x_4^2 \leq 5x_4^2$, tenemos que, para todo $x \in \mathbb{R}^4$,

$$-3\|x\|^2 = -3x_1^2 - 3x_2^2 - 3x_3^2 - 3x_4^2 \leq Q(x) = 5x_1^2 - 3x_2^2 + 2x_3^2 + 5x_4^2.$$

Por lo tanto,

$$-3\|x\|^2 \leq Q(x) \leq 5\|x\|^2 \text{ para todo } x \in \mathbb{R}^4.$$

Con lo cual, si $x \neq 0$, tenemos que

$$-3 \leq \frac{Q(x)}{\|x\|^2} \leq 5 \quad \text{para todo } x \in \mathbb{R}^4 \setminus \{0\}.$$

Veamos entonces, que el máximo del cociente de Rayleigh es 5 y el mínimo del cociente de Rayleigh es -3. Como ya vimos que 5 es una cota superior y -3 es una cota inferior del cociente de

Rayleigh, sólo nos resta encontrar algún $x_M \in \mathbb{R}^4 \setminus \{0\}$ tal que $\frac{Q(x_M)}{\|x_M\|^2} = 5$ y algún $x_m \in \mathbb{R}^4 \setminus \{0\}$ tal que $\frac{Q(x_m)}{\|x_m\|^2} = -3$.

Observar que $x_M := [1\ 0\ 0\ 0]^T \in \mathbb{R}^4 \setminus \{0\}$, $Q(x_M) = 5$ y $\|x_M\| = 1$. Entonces $\frac{Q(x_M)}{\|x_M\|^2} = 5$ y concluimos que

$$\max_{x \neq 0} \frac{Q(x)}{\|x\|^2} = 5.$$

Por otra parte, $x_m := [0\ 1\ 0\ 0]^T \in \mathbb{R}^4 \setminus \{0\}$, $Q(x_m) = -3$ y $\|x_m\| = 1$. Entonces $\frac{Q(x_m)}{\|x_m\|^2} = -3$ y concluimos que

$$\min_{x \neq 0} \frac{Q(x)}{\|x\|^2} = -3.$$

b) : Queremos hallar el conjunto de todos los $x_M \in \mathbb{R}^4 \setminus \{0\}$ tales que $\frac{Q(x_M)}{\|x_M\|^2} = \max_{x \neq 0} \frac{Q(x)}{\|x\|^2} = 5$. Planteamos para qué valores de $x \in \mathbb{R}^4$ tenemos que $Q(x) = 5\|x\|^2$ y eso vale, si y sólo si,

$$Q(x) = 5x_1^2 - 3x_2^2 + 2x_3^2 + 5x_4^2 = 5\|x\|^2 = 5x_1^2 + 5x_2^2 + 5x_3^2 + 5x_4^2.$$

Entonces,

$$8x_2^2 + 3x_3^2 = 0.$$

Por lo tanto, $x_2 = x_3 = 0$. Entonces,

$$Q(x) = 5\|x\|^2 \text{ si y sólo si } x = [x_1\ 0\ 0\ x_4] \text{ con } x_1, x_4 \in \mathbb{R}$$

$$\text{ó, equivalentemente, } Q(x) = 5\|x\|^2 \text{ si y sólo si } x \in gen\{[1\ 0\ 0\ 0]^T, [0\ 0\ 0\ 1]^T\}.$$

Por lo tanto, si $x_M \in gen\{[1\ 0\ 0\ 0]^T, [0\ 0\ 0\ 1]^T\} \setminus \{0\}$ tenemos que

$$\frac{Q(x_M)}{\|x_M\|^2} = \max_{x \neq 0} \frac{Q(x)}{\|x\|^2} = 5.$$

Para hallar el conjunto de todos los $x_m \in \mathbb{R}^4 \setminus \{0\}$ tales que $\frac{Q(x_m)}{\|x_m\|^2} = \min_{x \neq 0} \frac{Q(x)}{\|x\|^2} = -3$, planteamos para qué valores de $x \in \mathbb{R}^4$ tenemos que $Q(x) = -3\|x\|^2$ y eso vale, si y sólo si,

$$Q(x) = 5x_1^2 - 3x_2^2 + 2x_3^2 + 5x_4^2 = -3\|x\|^2 = -3x_1^2 - 3x_2^2 - 3x_3^2 - 3x_4^2.$$

Entonces,

$$8x_1^2 + 5x_3^2 + 8x_4^2 = 0.$$

Por lo tanto, $x_2 = x_3 = x_4 = 0$. Entonces,

$$Q(x) = -3\|x\|^2 \text{ si y sólo si } x = [0\ x_2\ 0\ 0] \text{ con } x_2 \in \mathbb{R}$$

$$\text{ó, equivalentemente } Q(x) = -3\|x\|^2 \text{ si y sólo si } x \in gen\{[0\ 1\ 0\ 0]^T\}.$$

Por lo tanto, si $x_m \in gen\{[0\ 1\ 0\ 0]^T\} \setminus \{0\}$ tenemos que

$$\frac{Q(x_m)}{\|x_m\|^2} = \min_{x \neq 0} \frac{Q(x)}{\|x\|^2} = -3.$$

c) : En el item *b*), vimos que $Q(x) = 5\|x\|^2$ si y sólo si $x \in gen\{[1\ 0\ 0\ 0]^T, [0\ 0\ 0\ 1]^T\}$. Si ahora imponemos la restricción $x \in S_3 = \{x \in \mathbb{R}^4 : \|x\| = 1\}$, tenemos que

$$\max_{x \in S_3} Q(x) = \max_{\|x\|=1} Q(x) = 5$$

y los $\hat{x}_M \in S_3$ tales que

$$Q(\hat{x}_M) = \max_{x \in S_3} Q(x) = \max_{\|x\|=1} Q(x) = 5$$

son los $x \in gen\{[1\ 0\ 0\ 0]^T, [0\ 0\ 0\ 1]^T\}$ tales que $\|x\| = 1$. Entonces, el máximo se alcanza en los vectores $\hat{x}_M = \alpha[1\ 0\ 0\ 0]^T + \beta[0\ 0\ 0\ 1]^T$ con $\alpha, \beta \in \mathbb{R}$ tales que (aplicando Pitágoras o haciendo la cuenta)

$$1 = \|x\|^2 = \|\alpha[1\ 0\ 0\ 0]^T + \beta[0\ 0\ 0\ 1]^T\|^2 = \alpha^2\|[1\ 0\ 0\ 0]^T\|^2 + \beta^2\|[0\ 0\ 0\ 1]^T\|^2 = \alpha^2 + \beta^2.$$

Entonces, el máximo se alcanza en los vectores $\hat{x}_M = \alpha[1\ 0\ 0\ 0]^T + \beta[0\ 0\ 0\ 1]^T$ tales que $\alpha^2 + \beta^2 = 1$.

En el item c), vimos que $Q(x) = -3\|x\|^2$ si y sólo si $x \in gen\{[0\ 1\ 0\ 0]^T\}$. Si ahora imponemos la restricción $x \in S_3 = \{x \in \mathbb{R}^4 : \|x\| = 1\}$, tenemos que

$$\min_{x \in S_3} Q(x) = \min_{\|x\|=1} Q(x) = -3$$

y los $\hat{x}_m \in S_3$ tales que

$$Q(\hat{x}_m) = \min_{x \in S_3} Q(x) = \min_{\|x\|=1} Q(x) = -3$$

son los $x \in gen\{[0\ 1\ 0\ 0]^T\}$ tales que $\|x\| = 1$. Entonces, el mínimo se alcanza en los vectores $\hat{x}_m = \alpha[0\ 1\ 0\ 0]^T$ con $\alpha \in \mathbb{R}$ tales que $1 = \|x\| = \|\alpha[0\ 1\ 0\ 0]^T\| = |\alpha|\|[0\ 1\ 0\ 0]^T\| = |\alpha|$. Entonces, $\alpha = \pm 1$ y, el mínimo se alcanza en los vectores $\hat{x}_m = \pm[0\ 1\ 0\ 0]^T$. ■

Teorema 6.5.1 (Teorema de Rayleigh). *Sea $A \in \mathbb{R}^{n \times n}$ simétrica y $Q(x) = x^T A x$. Sean λ_M y λ_m los autovalores máximo y mínimo de A respectivamente y sean $\mathcal{S}_{\lambda_M}$ y $\mathcal{S}_{\lambda_m}$ los autespacios asociados respectivos. Entonces:*

$$\lambda_m\|x\|^2 \leq Q(x) \leq \lambda_M\|x\|^2 \text{ para todo } x \in \mathbb{R}^n. \text{ Además,}$$

$$Q(x) = \lambda_m\|x\|^2 \text{ si y sólo si } x \in \mathcal{S}_{\lambda_m} \text{ y } Q(x) = \lambda_M\|x\|^2 \text{ si y sólo si } x \in \mathcal{S}_{\lambda_M}.$$

Lo vamos a demostrar de manera sencilla usando las ideas del **Ejercicio 6.E**.

Dem. Sea $A = P\Lambda P^T$ una descomposición ortogonal de A, que existe pues A es simétrica. Supongamos que

$$\Lambda = \begin{bmatrix} \lambda_M & \cdots & 0 \\ \dots & \ddots & \dots \\ 0 & \cdots & \lambda_m \end{bmatrix}.$$

Donde $\lambda_M = \lambda_1 = \cdots = \lambda_r > \lambda_{r+1} \geq \cdots > \lambda_{n-k+1} = \cdots = \lambda_n = \lambda_m$. Es decir, ordenamos los autovalores de A de mayor a menor y suponemos que la multiplicidad algebraica (y geométrica) de λ_M es r y la multiplicidad algebraica (y geométrica) de λ_m es k.

Entonces, si $P = [v_1\ v_2\ \cdots\ v_n]$, tendremos que

$$\mathcal{S}_{\lambda_M} = gen\{v_1, \cdots, v_r\} \text{ y } \mathcal{S}_{\lambda_m} = gen\{v_{n-k+1}, \cdots, v_n\}.$$

Si hacemos el cambio de variables $x = Py$, por un lado, como P es ortogonal, vale que

$$\|x\| = \|Py\| = \|y\|$$

y, por el otro lado,

$$\begin{aligned} Q(x) = (Py)^T A(Py) = y^T \Lambda y := \tilde{Q}(y) &= \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \\ &= \lambda_M y_1^2 + \cdots + \lambda_M y_r^2 + \cdots + \lambda_m y_{n-k+1}^2 + \cdots + \lambda_m y_n^2. \end{aligned}$$

Entonces, tenemos una expresión sin productos cruzados y podemos usar ideas similares a las del **Ejercicio 6.E**. Entonces, para todo $y \in \mathbb{R}^n$,

$$\lambda_m\|y\|^2 \leq \tilde{Q}(y) \leq \lambda_M\|y\|^2.$$

Usado que $\|x\| = \|y\|$, nos queda que, para todo $x \in \mathbb{R}^n$,

$$\lambda_m\|x\|^2 = \lambda_m\|y\|^2 \leq \tilde{Q}(y) = Q(x) \leq \lambda_M\|y\|^2 = \lambda_M\|x\|^2.$$

Entonces, probamos que

$$\lambda_m\|x\|^2 \leq Q(x) \leq \lambda_M\|x\|^2 \text{ para todo } x \in \mathbb{R}^n.$$

Por otra parte, también usando ideas similares a las del **Ejercicio 6.E**, tenemos que

$$Q(x) = \tilde{Q}(y) = \lambda_M\|y\|^2 = \lambda_M\|x\|^2,$$

si y sólo si $y \in gen\{e_1, \cdots, e_r\}$ donde e_i es el i-ésimo vector de la base canónica de $\mathbb{R}^n$. Entonces $y = [a_1\ a_2\ \cdots\ a_r\ 0\ \cdots\ 0]^T$ para ciertos $a_1, \cdots, a_r \in \mathbb{R}$ y, como $x = Py = [v_1\ \cdots\ v_r\ \cdots\ v_{n-k+1}\ \cdots\ v_n]y = a_1v_1 + \cdots + a_rv_r$, nos queda que

$$Q(x) = \lambda_M\|x\|^2 \text{ si y sólo si } x \in gen\{v_1, \cdots, v_r\} = \mathcal{S}_{\lambda_M}.$$

De la misma manera, tenemos que

$$Q(x) = \tilde{Q}(y) = \lambda_m\|y\|^2 = \lambda_m\|x\|^2,$$

si y sólo si $y \in gen\{e_{n-k+1}, \cdots, e_n\}$. Entonces $y = [0\ \cdots\ 0\ a_{n-k+1}\ \cdots\ a_n\ ]^T$ para ciertos $a_{n-k+1}, \cdots, a_n \in \mathbb{R}$ y, como $x = Py = [v_1\ \cdots\ v_r\ \cdots\ v_{n-k+1}\ \cdots\ v_n]y = a_{n-k+1}v_{n-k+1} + \cdots + a_nv_n$, nos queda que

$$Q(x) = \lambda_m\|x\|^2 \text{ si y sólo si } x \in gen\{v_{n-k+1}, \cdots, v_n\} = \mathcal{S}_{\lambda_m}.$$

■

A partir del Teorema de Rayleigh que acabamos de probar, es inmediato ver que:

Dada $A \in \mathbb{R}^{n\times n}$ simétrica y $Q(x) = x^TAx$. Si λ_M y λ_m son los autovalores máximo y mínimo de A respectivamente y $\mathcal{S}_{\lambda_M}$ y $\mathcal{S}_{\lambda_m}$ los autespacios asociados respectivos. Entonces:

$$\max_{x\neq 0} \frac{Q(x)}{\|x\|^2} = \lambda_M \text{ y el máximo se alcanza en los } x \in \mathcal{S}_{\lambda_M} \setminus \{0\}.$$

$$\min_{x\neq 0} \frac{Q(x)}{\|x\|^2} = \lambda_m \text{ y el máximo se alcanza en los } x \in \mathcal{S}_{\lambda_m} \setminus \{0\}.$$

En particular, $\max\limits_{\|x\|=1} Q(x) = \lambda_M$ y el máximo se alcanza en los $x \in \mathcal{S}_{\lambda_M}$ tales que $\|x\| = 1$ y $\min\limits_{\|x\|=1} Q(x) = \lambda_m$ y el mínimo se alcanza en los $x \in \mathcal{S}_{\lambda_m}$ tales que $\|x\| = 1$.

Ejercicio 6.7. Sea $Q : \mathbb{R}^3 \to \mathbb{R}$ la forma cuadrática definida por

$$Q(x) = 5x_1^2 + 6x_2^2 + 7x_3^2 - 4x_1x_2 - 4x_2x_3.$$

a) Hallar $\max_{\|x\|=1} Q(x)$ y $\min_{\|x\|=1} Q(x)$.

b) Verificar que para todo $x \in \mathbb{R}^3$ vale que $3\|x\|^2 \leq Q(x) \leq 9\|x\|^2$.

c) Caracterizar el conjunto de nivel $\mathcal{N}_1(Q)$ y, utilizando el resultado del inciso anterior, hallar los puntos de $\mathcal{N}_1(Q)$ cuya distancia al origen sea mínima y aquellos cuya distancia al origen sea máxima. ¿Qué valores tienen esas distancias?

Observar que $Q(x) = x^T Ax$, con $A = \begin{bmatrix} 5 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 7 \end{bmatrix}$ simétrica. El polinomio característico de A es $\chi_A(t) = -t^3 + 18t^2 - 99t + 162$ y sus raíces son $\lambda_1 = 9, \lambda_2 = 6$ y $\lambda_3 = 3$. Por otra parte, $\mathcal{S}_{\lambda_1=9} = \text{nul}(A - 9I) = \text{nul}(\begin{bmatrix} -4 & -2 & 0 \\ -2 & -3 & -2 \\ 0 & -2 & -2 \end{bmatrix}) = gen\{\begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix}\}$, $\mathcal{S}_{\lambda_2=6} = \text{nul}(A - 6I) = \text{nul}(\begin{bmatrix} -1 & -2 & 0 \\ -2 & 0 & -2 \\ 0 & -2 & 1 \end{bmatrix}) + gen\{\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}\}$ y $\mathcal{S}_{\lambda_2=3} = \text{nul}(A - 3I) = \text{nul}(\begin{bmatrix} 2 & -2 & 0 \\ -2 & 3 & -2 \\ 0 & -2 & 4 \end{bmatrix}) = gen\{\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}\}$. Entonces, $\{\frac{1}{3}\begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix}, \frac{1}{3}\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}, \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}\}$ es una bon de $\mathbb{R}^3$. Tomamos $P := \frac{1}{3}\begin{bmatrix} -1 & -2 & 2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$ entonces $\det(P) = 1$ y una diagonalización ortogonal de A es

$$A = P \begin{bmatrix} 9 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{bmatrix} P^T.$$

Dem. $a)$: Usando la diagonalización ortogonal de A y el Teorema de Rayleigh, tenemos que

$$\max_{\|x\|=1} Q(x) = \lambda_M = 9 \text{ y } \min_{\|x\|=1} Q(x) = \lambda_m = 3.$$

$b)$: Por el ítem $a)$ tenemos que $3 \leq Q(x) \leq 9$ para todo $x \in \mathbb{R}^3$ tal que $\|x\| = 1$. Es claro que si $x = 0$ entonces $3\|x\|^2 \leq Q(x) \leq 9\|x\|^2$ (pues todos los términos dan 0). Por otra parte, si $x \in \mathbb{R}^3 \setminus \{0\}$ y tomamos $\hat{x} := \frac{x}{\|x\|}$ vale que $\|\hat{x}\| = \|\frac{x}{\|x\|}\| = 1$. Entonces, $3 \leq Q(\hat{x}) \leq 9$ y usando la propiedad (6.1), tenemos que $Q(x) = Q(\hat{x}\|x\|) = Q(\hat{x})\|x\|^2$. Por lo tanto, para todo $x \in \mathbb{R}^3$,

$$3\|x\|^2 \leq Q(x) = Q(\hat{x})\|x\|^2 \leq 9\|x\|^2.$$

$c)$: El conjunto de nivel $\mathcal{N}_1(Q)$ es el conjunto $\mathcal{N}_1(Q) = \{x \in \mathbb{R}^3 : Q(x) = 1\}$. Si hacemos el cambio de variables $x = Py$, con $P := \frac{1}{3}\begin{bmatrix} -1 & -2 & 2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$ tenemos que $Q(x) = 1$ si y sólo si $\tilde{Q}(y) = 9y_1^2 + 6y_2^2 + 3y_3^2 = 1$. Podemos reescribir la ecuación anterior como $\frac{y_1^2}{(\frac{1}{3})^2} + \frac{y_2^2}{(\frac{1}{\sqrt{6}})^2} + \frac{y_3^2}{(\frac{1}{\sqrt{3}})^2} = 1$.

Si tomamos los nuevos ejes rotados y_1, y_2, y_3 determinados por los versores de $\mathbb{R}^3$: $v_1 := \frac{1}{3}\begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix}$, $v_2 := \frac{1}{3}\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$ y $v_3 := \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$ el conjunto de nivel $\mathcal{N}_1(Q)$ es un elipsoide con centro en el origen de coordenadas y ejes coincidentes con los ejes y_1, y_2, y_3. Los valores $\frac{1}{3}$, $\frac{1}{\sqrt{6}}$ y $\frac{1}{\sqrt{3}}$ son las longitudes de los semiejes del elipsoide respecto de los ejes y_1, y_2, y_3, respectivamente. Por lo tanto, en los nuevos ejes, los puntos $y_m := \pm \begin{bmatrix} \frac{1}{3} \\ 0 \\ 0 \end{bmatrix}$ son los puntos de $\mathcal{N}_1(Q)$ más cercanos al origen. En la variable original, estos puntos son $x_m = Py_m = \pm\frac{1}{9} \begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix}$ y el valor de la distancia mínima es $\|y_m\| = \|x_m\| = \| \pm \frac{1}{9} \begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix} \| = \dfrac{1}{3}$. De manera similar, los puntos $y_M := \pm \begin{bmatrix} 0 \\ 0 \\ \frac{1}{\sqrt{3}} \end{bmatrix}$ son los puntos de $\mathcal{N}_1(Q)$ más alejados al origen. En la variable original, estos puntos son $x_M = Py_M = \pm\frac{1}{\sqrt{3}}\frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$ y el valor de la distancia máxima es $\|y_M\| = \|x_M\| = \| \pm \frac{1}{3\sqrt{3}} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \| = \dfrac{1}{\sqrt{3}}$. ■

Ejercicio 6.8. Sea Q la forma cuadrática en $\mathbb{R}^2$ definida por $Q(x) := x^T(aI + bA)x$, con $a, b \in \mathbb{R}$ donde A es la matriz en base canónica de una proyección ortogonal de $\mathbb{R}^2$ sobre un subespacio unidimensional.

a) Hallar los valores de a y b para que $\max\limits_{\|x\|=1} Q(x) = 5$ y $\min\limits_{\|x\|=1} Q(x) = 2$.

b) Para los valores hallados en el inciso anterior, y sabiendo que $\mathrm{col}(A) = \{x \in \mathbb{R}^2 : 3x_1 + 4x_2 = 0\}$, graficar el conjunto $\{x \in \mathbb{R}^2 : Q(x) \leq 3\}$.

Dem. $a)$: Vamos a llamar $\mathcal{S}$ al subespacio unidimensional de $\mathbb{R}^3$ sobre el que proyecta A. Recordemos que si A es una proyección ortogonal (considerando el producto interno canónico) entonces $A = A^2 = A^T$, ver ejemplos (3.a) y (3.b).

Más aún, vimos en el **Ejercicio 4.M** que los autovalores de A son $\lambda_1 = 1$ con autoespacio asociado $\mathcal{S}$ y $\lambda_2 = 0$ con autoespacio asociado el núcleo de A que es $\mathcal{S}^\perp$.

Si llamamos $B := aI + bA$ tenemos que $Q(x) = x^T Bx$. Como A es simétrica y $a, b \in \mathbb{R}$ es claro que $B^T = B$. Entonces, podemos aplicar el Teorema de Rayleigh para concluir que

$$\max_{\|x\|=1} Q(x) = \lambda_M(B) = 5 \text{ y } \min_{\|x\|=1} Q(x) = \lambda_m(B) = 2$$

donde $\lambda_M(B)$ y $\lambda_m(B)$ son los autovalores máximo y mínimo de B, respectivamente.

Aplicando las propiedades de autovalores y autovectores que vimos en **4.1**, tenemos que los autovalores de B son $\lambda_1 = a + b$ y $\lambda_2 = a$ (acá NO estamos ordenando los autovalores de mayor a menor). Hagamos la cuenta de todas maneras: si $v \in \mathcal{S}$ tenemos $Av = 1v$, entonces $Bv = (aI + bA)v = av + bAv = av + bv = (a + b)v$ y entonces $a + b$ es un autovalor de B con autovector asociado v. De la misma manera, si $w \in \mathcal{S}^\perp$ tenemos $Aw = 0w$, entonces $Bw = (aI + bA)w = aw + bAw = aw + b0 = aw$ y a es el otro autovalor de B con autovector asociado w. Como B es una matriz de simétrica de 2×2, concluimos que los autovalores de B son $a + b$ con autoespacio asociado $\mathcal{S}$ y a con autoespacio asociado $\mathcal{S}^\perp$. Como no sabemos si a y b son negativos, positivos o nulos, no podemos deducir si $a + b \geq a$ o si $a \geq a + b$. Consideramos los dos casos: **Caso 1:** supongamos que $\lambda_M(B) = a + b = 5$ y $\lambda_m(B) = a = 2$. Entonces, despejando nos queda que $a = 2$ y $b = 3$. Por otra parte, **Caso 2:** supongamos que $\lambda_M(B) = a = 5$ y $\lambda_m(B) = a + b = 2$. Entonces, despejando nos queda que $a = 5$ y $b = -3$.

$b)$: Recordemos que $\operatorname{col}(A)$ es el subespacio sobre el que proyecta A que llamamos $\mathcal{S}$. Es decir, $\mathcal{S} = \operatorname{col}(A) = \{x \in \mathbb{R}^2 : 3x_1 + 4x_2 = 0\} = gen\{\begin{bmatrix} 4 \\ -3 \end{bmatrix}\}$. Como vimos en el ítem $a)$ los autovalores de B son $\lambda_1 = a + b$ con autoespacio asociado $\mathcal{S} = gen\{\begin{bmatrix} 4 \\ -3 \end{bmatrix}\}$ y $\lambda_2 = a$ con autoespacio asociado $\mathcal{S}^\perp = gen\{\begin{bmatrix} 3 \\ 4 \end{bmatrix}\}$. Entonces, $\{\frac{1}{5}\begin{bmatrix} 4 \\ -3 \end{bmatrix}, \frac{1}{5}\begin{bmatrix} 3 \\ 4 \end{bmatrix}\}$ es una bon de $\mathbb{R}^2$. Tomamos $P := \frac{1}{5}\begin{bmatrix} 4 & 3 \\ -3 & 4 \end{bmatrix}$ entonces $\det(P) = 1$ y una diagonalización ortogonal de B es $B = P\begin{bmatrix} a+b & 0 \\ 0 & a \end{bmatrix}P^T$. Si hacemos el cambio de variables $x = Py$ tenemos que $Q(x) \leq 3$ si y sólo si

$$\tilde{Q}(y) = (a+b)y_1^2 + ay_2^2 \leq 3. \tag{6.3}$$

- **Caso 1:** $a = 2$ y $b = 3$. En este caso (6.3), nos queda $\tilde{Q}(y) = 5y_1^2 + 2y_2^2 \leq 3$. El conjunto de nivel$\{x \in \mathbb{R}^2 : Q(x) \leq 3\}$ es el contorno y el interior de una elipse en $\mathbb{R}^2$ en los nuevos ejes y_1, y_2 determinados por los versores $v_1 := \frac{1}{5}\begin{bmatrix} 4 \\ -3 \end{bmatrix}$ y $v_2 := \frac{1}{5}\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ que son las columnas de P y que son una rotación de los ejes originales x_1, x_2 (ver Figura 6.2).

- **Caso 2:** $a = 5$ y $b = -3$. En este caso (6.3), nos queda $\tilde{Q}(y) = 2y_1^2 + 5y_2^2 \leq 3$. El conjunto de nivel$\{x \in \mathbb{R}^2 : Q(x) \leq 3\}$ es el contorno y el interior de una elipse en $\mathbb{R}^2$ en los nuevos ejes y_1, y_2 determinados por los versores $v_1 := \frac{1}{5}\begin{bmatrix} 4 \\ -3 \end{bmatrix}$ y $v_2 := \frac{1}{5}\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ que son las columnas de P y que son una rotación de los ejes originales x_1, x_2 (ver Figura 6.3).

■

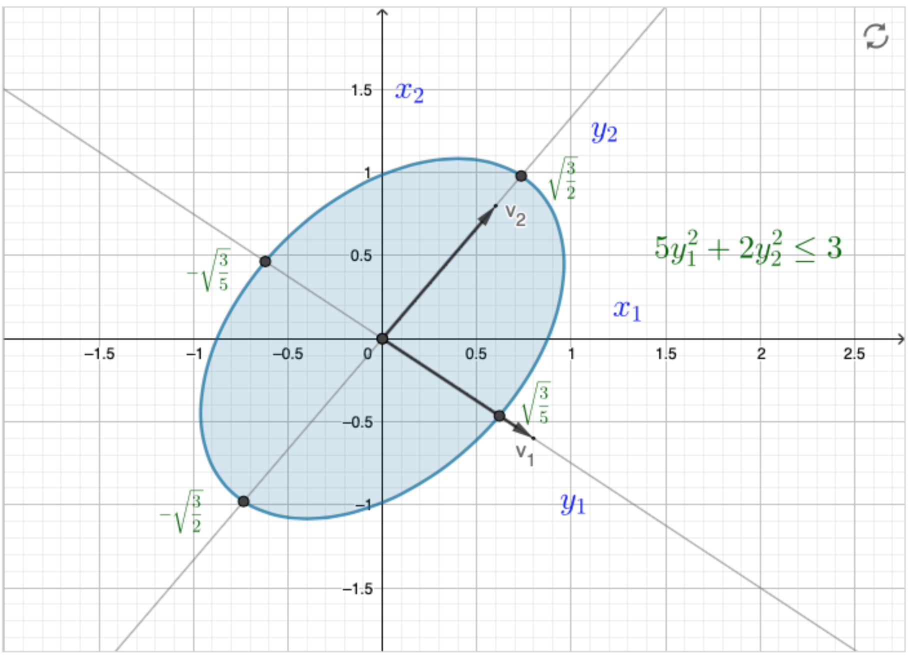

Figura 6.2: Ejercicio 6.8) Caso 1

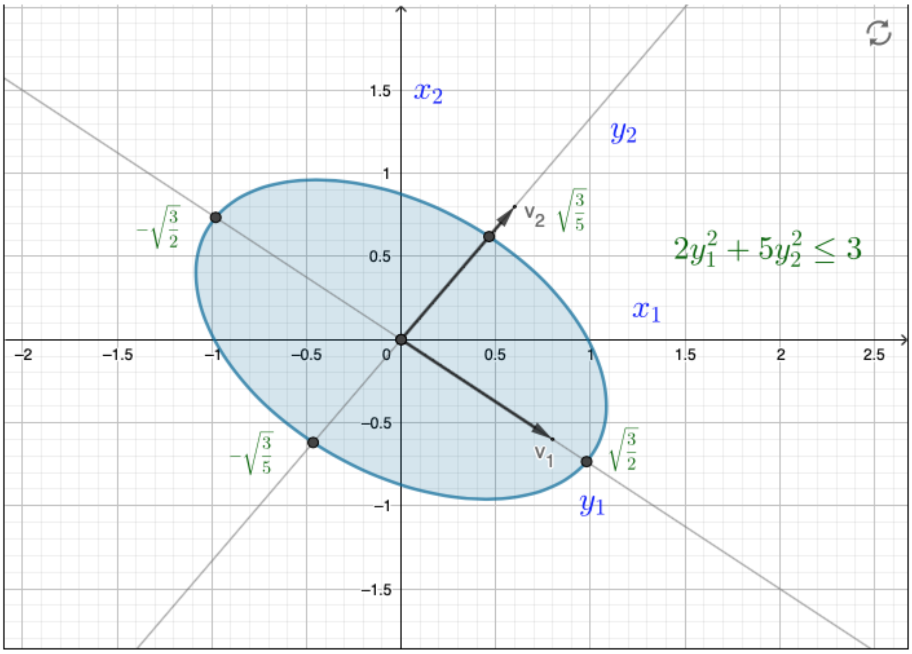

Figura 6.3: Ejercicio 6.8) Caso 2

Ejercicio 6.F (De Final)**.** Sea $\langle \cdot, \cdot \rangle_A : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ el producto interno definido por $\langle x, y \rangle_A := y^T Ax$ con

$$A = \begin{bmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{bmatrix},$$

donde $\theta \in (0, \pi)$.

a) Observar que $Q_A(x) = \langle x, x \rangle_A$ es una forma cuadrática en $\mathbb{R}^2$.

b) Hallar $P \in \mathbb{R}^{2\times 2}$ ortogonal con $\det(P) = 1$ tal que el cambio de variables $x = Py$ transforme Q_A en una forma cuadrática sin productos cruzados.

c) Determinar los ejes principales de la forma cuadrática Q_A.

d) Graficar el conjunto de nivel $\mathcal{N}_1(Q_A) = \{x \in \mathbb{R}^2 : Q_A(x) = 1\}$.

e) Determinar cómo son los conjuntos de nivel $\mathcal{N}_c(Q_A)$ con $c > 0$.

Dem. *a)* : Observar en primer lugar que $A^T = A$. Además, observar que $\operatorname{tr}(A) = 2 > 0$ y $\det(A) = 1 - \cos\theta^2 = \sin\theta^2 > 0$. Entonces, si $\lambda_1, \lambda_2 \in \mathbb{R}$ son los autovalores de A,

$$\lambda_1 + \lambda_2 = \operatorname{tr}(A) > 0 \text{ y } \lambda_1\lambda_2 = \det(A) > 0$$

y deducimos que $\lambda_1 > 0$ y $\lambda_2 > 0$. Por lo tanto, A es simétrica y sus autovalores son positivos. Entonces, por la Proposición 6.3.1, A es definida positiva. Conclusión: por el **Ejercicio 6.C**, $\langle \cdot, \cdot \rangle_A$ efectivamente resulta un producto interno y $Q_A := \langle x, x \rangle_A$ es una forma cuadrática definida positiva.

b) : Usando que $\operatorname{tr}(A) = 2$ y $\det(A) = \sin\theta^2$, haciendo la cuenta, tenemos que el polinomio característico de A es

$$\chi_A(t) = \det(A - tI) = t^2 - \operatorname{tr}(A)t + \det(A) = t^2 - 2t + \sin\theta^2.$$

Entonces, como

$$\Delta := \operatorname{tr}(A)^2 - 4\det(A) = 4 - 4\sin\theta^2 = 4(1 - \sin\theta^2) = 4\cos\theta^2 > 0.$$

Tenemos que los autovalores de A son

$$\lambda_{1,2} = \frac{1}{2}\big(\operatorname{tr}(A) \pm \sqrt{\operatorname{tr}(A)^2 - 4\det(A)}\big) = \frac{1}{2}\big(\operatorname{tr}(A) \pm \sqrt{\Delta}\big).$$

Es decir:

$$\lambda_1 = \frac{2+\sqrt{\Delta}}{2} = \frac{2+2\cos\theta}{2} = 1+\cos\theta \text{ y } \lambda_2 = \frac{2-\sqrt{\Delta}}{2} = 1-\cos\theta.$$

Por lo tanto, el autoespacio asociado a λ_1 es

$$\operatorname{nul}(A-\lambda_1 I) = \operatorname{nul}(A-(1+\cos\theta)I) = \operatorname{nul}(\begin{bmatrix} -\cos\theta & \cos\theta \\ \cos\theta & -\cos\theta \end{bmatrix}) = gen\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}.$$

De la misma manera, el autoespacio asociado a λ_2 es

$$\operatorname{nul}(A-\lambda_2 I) = \operatorname{nul}(A-(1-\cos\theta)I) = \operatorname{nul}(\begin{bmatrix} \cos\theta & \cos\theta \\ \cos\theta & \cos\theta \end{bmatrix}) = gen\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}.$$

Entonces

$$\{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$$

es una bon de $\mathbb{R}^2$. Tomamos $P := \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$. Claramente $\det(P) = 1$ y $A = P\begin{bmatrix} 1+\cos\theta & 0 \\ 0 & 1-\cos\theta \end{bmatrix}P^T$.

$c)$: Los ejes principales de la forma cuadrática Q_A son los vectores unitarios de la bon de $\mathbb{R}^2$ formada por los autovectores de A: $\{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$.

$d)$: Si hacemos el cambio de variables $x = Py$, con P definida en el item $b)$, usando que $A = P\begin{bmatrix} 1+\cos\theta & 0 \\ 0 & 1-\cos\theta \end{bmatrix}P^T$, tenemos que $Q(x) = 1$, si y sólo si

$$\tilde{Q}(y) = (1+\cos\theta)y_1^2 + (1-\cos\theta)y_2^2 = 1.$$

Podemos reescribir la ecuación anterior como

$$\tilde{Q}(y) = \frac{y_1^2}{[\frac{1}{\sqrt{(1+\cos\theta)}}]^2} + \frac{y_2^2}{[\frac{1}{\sqrt{(1-\cos\theta)}}]^2} = 1.$$

Como $\theta \in (0,\pi)$ se ve que $1+\cos\theta > 0$ y $1-\cos\theta > 0$ (y ambos autovalores son distintos excepto en el caso $\theta = \frac{\pi}{2}$).

Entonces, si $\theta \in (0,\pi)$ y $\theta \neq \frac{\pi}{2}$, el conjunto de nivel $\mathcal{N}_1(Q) = \{x \in \mathbb{R}^2 : x_1^2+x_2^2+2\cos\theta x_1x_2 = 1\}$ es una elipse en $\mathbb{R}^2$ en los nuevos ejes y_1, y_2 determinados por los versores $v_1 := \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $v_2 := \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$ que son las columnas de P y que son una rotación de un ángulo (positivo) de los ejes originales x_1, x_2 (ver Figura 6.4). En el caso $\theta = \frac{\pi}{2}$, el conjunto de nivel $\mathcal{N}_1(Q)$ es una circunferencia de radio 1.

Vamos a graficar el conjunto de nivel $\mathcal{N}_1(Q)$ para 3 valores distintos de θ :

- Gráfico 1: $\theta = \frac{\pi}{3}$ (color verde). En este caso, nos queda el conjunto de nivel $x_1^2+x_2^2+x_1x_2 = 1$ (en los ejes x_1, x_2) ó $\frac{y_1^2}{[\sqrt{\frac{2}{3}}]^2} + \frac{y_2^2}{[\sqrt{2}]^2} = 1$ (en los ejes y_1, y_2).

- Gráfico 2: $\theta = \frac{\pi}{2}$ (color negro). En este caso, nos queda el conjunto de nivel $x_1^2+x_2^2 = 1$ (en los ejes x_1, x_2) ó $y_1^2+y_2^2 = 1$ (en los ejes y_1, y_2). Observar que en realidad obtenemos una circunferencia de radio 1.

- Gráfico 3: $\theta = \frac{2\pi}{3}$ (color azul). En este caso, nos queda el conjunto de nivel $x_1^2 + x_2^2 - x_1x_2 = 1$ (en los ejes x_1, x_2) ó $\frac{y_1^2}{[\sqrt{2}]^2} + \frac{y_2^2}{[\sqrt{\frac{2}{3}}]^2} = 1$ (en los ejes y_1, y_2).

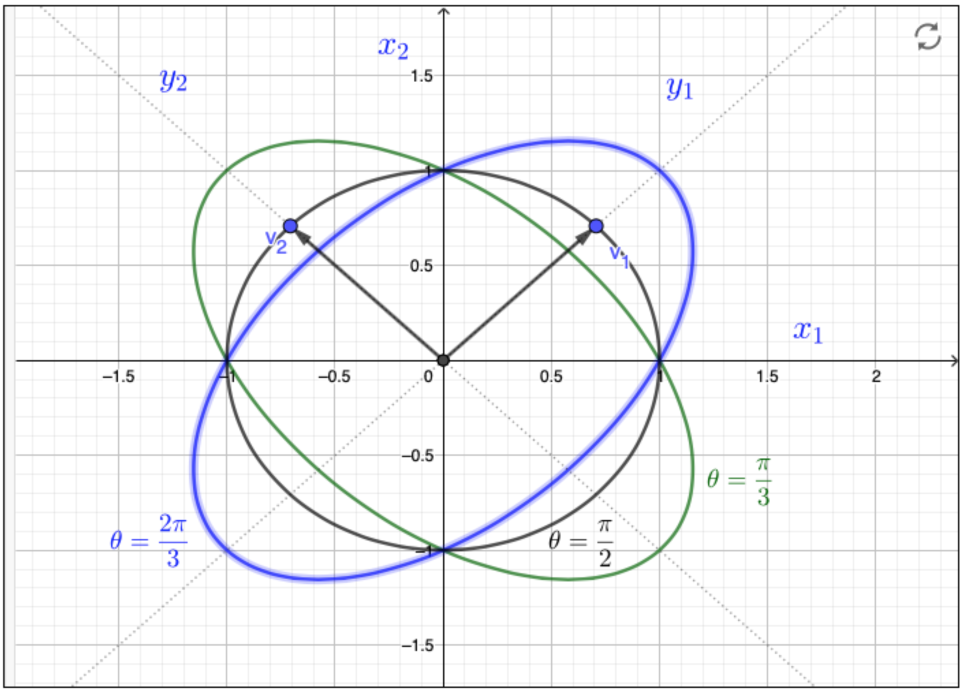

Figura 6.4: Ejercicio 6.F

$e)$: De la misma manera que hicimos en el item d, tenemos que $Q(x) = c$, si y sólo si

$$\tilde{Q}(y) = (1 + \cos\theta)y_1^2 + (1 - \cos\theta)y_2^2 = c.$$

Usando que $c > 0$ podemos reescribir la ecuación anterior como

$$\tilde{Q}(y) = \frac{y_1^2}{[\sqrt{\frac{c}{(1+\cos\theta)}}]^2} + \frac{y_2^2}{[\sqrt{\frac{c}{(1-\cos\theta)}}]^2} = 1.$$

Entonces, el conjunto de nivel $\mathcal{N}_1(Q) = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 + 2\cos\theta x_1x_2 = c\}$ es una elipse en $\mathbb{R}^2$ en los nuevos ejes y_1, y_2 determinados por los versores $v_1 := \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $v_2 := \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$ que son las columnas de P y que son una rotación de un ángulo (positivo) de los ejes originales x_1, x_2. Como vimos, en el caso $\theta = \frac{\pi}{2}$, obtenemos una circunferencia de radio c. ■

Ejercicio 6.9. Sea $A \in \mathbb{R}^{3\times3}$ la matriz simétrica de traza nula tal que

$$\text{nul}(A - 2I) = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\},$$

y sea $Q : \mathbb{R}^3 \to \mathbb{R}$ la forma cuadrática definida por $Q(x) := x^TAx$. Si $x_0 \in \mathbb{R}^3$ es un vector cuya distancia al subespacio $gen\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\}$ es 5, ¿qué valor debe tener la distancia de x_0 al $\text{nul}(A - 2I)$ para que $Q(x_0) = 14$?

Dem. Notar que $\text{nul}(A-2I) = gen\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}\} := \mathcal{S}_{\lambda=2}$ por lo tanto $\lambda = 2$ es un autovalor de A con multiplicidad geométrica 2. Más aún, como A es simétrica entonces es diagonalizable y por lo tanto, la multiplicidad algebraica de $\lambda = 2$ también es 2. Por otra parte, como

$$0 = \text{tr}(A) = \lambda_1 + \lambda_2 + \lambda_3 = 2 + 2 + \lambda_3,$$

tenemos que $\lambda = -4$ es un autovalor de A de multiplicidad 1 con autoespacio asociado el complemento ortogonal de $\text{nul}(A-2I)$ que (afortunadamente) es $\mathcal{S}_{\lambda=-4} := \mathcal{S}_{\lambda=2}^{\perp} = gen\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\}$.

Por lo tanto,

$$\{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\}$$

es una bon de autovectores de $\mathbb{R}^3$. Vamos a llamar $v_1 := \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ $v_2 := \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$ y $v_3 := \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Como $x_0 \in \mathbb{R}^3$, existen $\alpha, \beta, \gamma \in \mathbb{R}$ tales que

$$x_0 = \alpha v_1 + \beta v_2 + \gamma v_3.$$

Además, es claro que $P_{\mathcal{S}_{\lambda=-4}}(x_0) = \gamma v_3$ y $P_{\mathcal{S}_{\lambda=2}}(x_0) = \alpha v_1 + \beta v_2$. Entonces

$$\begin{aligned} 25 &= d(x_0, gen\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\})^2 = d(x_0, \mathcal{S}_{\lambda=-4})^2 = \|x_0 - P_{\mathcal{S}_{\lambda=-4}}(x_0)\|^2 \\ &= \|\alpha v_1 + \beta v_2\|^2 = \alpha^2 + \beta^2, \end{aligned}$$

donde usamos el Teorema de Pitágoras para la última igualdad.

Por otra parte, buscamos que $Q(x_0) = x_0^T A x_0 = 14$. Entonces, como $Av_1 = 2v_1, Av_2 = 2v_2$, $Av_3 = -4v_3$ y $\{v_1, v_2, v_3\}$ es una bon de $\mathbb{R}^3$, tenemos que

$$\begin{aligned} 14 = Q(x_0) &= x_0^T A x_0 = (\alpha v_1^T + \beta v_2^T + \gamma v_3^T)A(\alpha v_1 + \beta v_2 + \gamma v_3) \\ &= (\alpha v_1^T + \beta v_2^T + \gamma v_3^T)(2\alpha v_1 + 2\beta v_2 - 4\gamma v_3) = 2\alpha^2 + 2\beta^2 - 4\gamma^2. \end{aligned}$$

Entonces, $\gamma^2 = \frac{2(\alpha^2+\beta^2)-14}{4} = \frac{2\times 25-14}{4} = 9$. Finalmente, notar que

$$d(x_0, \text{nul}(A-2I)) = d(x_0, \mathcal{S}_{\lambda=2}) = \|x_0 - P_{\mathcal{S}_{\lambda=2}}(x_0)\| = \|\gamma v_3\| = |\gamma|.$$

Entonces, la distancia de x_0 al $\text{nul}(A-2I)$ debe valer

$$d(x_0, \text{nul}(A-2I)) = |\gamma| = \sqrt{9} = 3.$$

■

Ejercicio 6.G (De Final). Decidir si la siguiente afirmación es verdadera o falsa: sean

$$Q(x) = x^T \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} x \text{ y } R(x) = x^T \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} x.$$

Entonces

$$\max_{\|x\|=1} Q(x) = \max_{\|x\|=1} R(x).$$

Dem. La afirmación es FALSA. Por el Teorema de Rayleigh, claramente

$$\max_{\|x\|=1} R(x) = 2,$$

esto es debido a que el máximo autovalor de la matriz simétrica (en particular diagonal) $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ es 2. Por otra parte, si llamamos $C := \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$ entonces $Q(x) = x^T Cx$. El polinomio característico de C es

$$\chi_C(t) = \det(A - tI) = (2-t)(1-t).$$

Por lo tanto, los autovalores de C son $\lambda_1 = 2$ y $\lambda_2 = 1$. Sin embargo, como C claramente NO es simétrica NO podemos aplicar el Teorema de Rayleigh. Para aplicar el Teorema de Rayleigh necesitamos escribir a Q como $Q(x) = x^T Ax$ con A simétrica. Si recordamos lo que hicimos arriba, vimos que tomando

$$A := \frac{C + C^T}{2} = \begin{bmatrix} 2 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}$$

entonces $Q(x) = x^T Ax$ y ahora A es simétrica. Ahora sí podemos aplicar el Teorema de Rayleigh. El polinomio característico de A es

$$\chi_A(t) = (2-t)(1-t) - \frac{1}{4} = t^2 - 3t + \frac{7}{4}$$

y sus raíces son $\lambda_1 = \frac{3+\sqrt{2}}{2}$ y $\lambda_2 = \frac{3-\sqrt{2}}{2}$. Por lo tanto, por el Teorema de Rayleigh,

$$\max_{\|x\|=1} Q(x) = \frac{3+\sqrt{2}}{2} \neq 2 = \max_{\|x\|=1} R(x)$$

y concluimos que la afirmación es FALSA. ∎

Ejercicio 6.H (De Final)**.** Sea $B \in \mathbb{R}^{3\times 3}$ simétrica con autovalores $-2, -1, 1$. Hallar todos los $r \in \mathbb{R}$ tales que

$$-2 \leq x^T(B + rI)x \leq 16 \text{ para todo } x \in \mathbb{R}^3 \text{ tal que } \|x\| = 2.$$

Dem. Sea $Q(x) := x^T Ax$ donde $A := B + rI$. Como B es simétrica y $r \in \mathbb{R}$, entonces A también es simétrica. De hecho

$$A^T = (B + rI)^T = B^T + (rI)^T = B + rI = A.$$

Por otra parte, si $\lambda \in \mathbb{R}$ es un autovalor de B, entonces $Bv = \lambda v$, para cierto $v \in \mathbb{R}^3 \setminus \{0\}$. Entonces, como

$$Av = (B + rI)v = Bv + rv = \lambda v + rv = (\lambda + r)v,$$

tenemos que $\lambda + r$ es un autovalor de A con el mismo autovector asociado.

Por lo tanto, los autovalores de A (ordenados de mayor a menor) son: $\lambda_1 = 1 + r, \lambda_2 = -1 + r$ y $\lambda_3 = -2 + r$. Sean $\lambda_M = 1 + r$ y $\lambda_m = -2 + r$ el mayor y menor autovalor de A respectivamente y sean $\mathcal{S}_{\lambda_M}$ y $\mathcal{S}_{\lambda_m}$ los autoespacios asociados correspondientes. Entonces, por el Teorema de Rayleigh,

$$\lambda_m \|x\|^2 \leq Q(x) \leq \lambda_M \|x\|^2 \text{ para todo } x \in \mathbb{R}^3$$

y $Q(x) = \lambda_m \|x\|^2$ si y sólo si $x \in \mathcal{S}_{\lambda_m}$ y $Q(x) = \lambda_M \|x\|^2$ si y sólo si $x \in \mathcal{S}_{\lambda_M}$. Entonces, es inmediato ver que

$$4\lambda_m \leq Q(x) \leq 4\lambda_M \text{ para todo } x \in \mathbb{R}^3 \text{ tal que } \|x\| = 2$$

y $Q(x) = 4\lambda_m$ si y sólo si $x \in \mathcal{S}_{\lambda_m}$ y $\|x\| = 2$ y $Q(x) = 4\lambda_M$ si y sólo si $x \in \mathcal{S}_{\lambda_M}$ y $\|x\| = 2$. Por lo tanto,

$$\min_{\|x\|=2} Q(x) = 4\lambda_m = 4(-2 + r) \quad \text{y} \quad \max_{\|x\|=2} Q(x) = 4\lambda_M = 4(1 + r).$$

Si $Q(x) = x^T Ax = x^T(B + rI)x \leq 16$ para todo $x \in \mathbb{R}^3$ tal que $\|x\| = 2$, entonces 16 es una cota superior del conjunto $\{Q(x) : x \in \mathbb{R}^3 \text{ con } \|x\| = 2\}$. Entonces, el máximo de dicho conjunto es menor o igual a esa cota superior y por lo tanto,

$$\max_{\|x\|=2} Q(x) = 4\lambda_M = 4(1 + r) \leq 16$$

y entonces, tenemos que $r \leq \frac{16}{4} - 1 = 3$.

De la misma manera, si $-2 \leq x^T(B + rI)x = x^T Ax = Q(x)$ para todo $x \in \mathbb{R}^3$ tal que $\|x\| = 2$, entonces -2 es una cota inferior del conjunto $\{Q(x) : x \in \mathbb{R}^3 \text{ con } \|x\| = 2\}$. Entonces, el mínimo de dicho conjunto es mayor o igual a esa cota inferior y por lo tanto,

$$-2 \leq \min_{\|x\|=2} Q(x) = 4\lambda_m = 4(-2 + r)$$

y entonces, tenemos que $-\frac{1}{2} + 2 = \frac{3}{2} \leq r$. Por lo tanto, $\frac{3}{2} \leq r \leq 3$. ■

6.6. Optimización de formas cuadráticas con restricciones definidas positivas

En lo que sigue, buscamos maximizar o minimizar una forma cudrática $Q(x) = x^T Ax$ (con A simétrica) sujeta a la restricción $R(x) = 1$, donde $R(x) = x^T Bx$ es una forma cuadrática con B simétrica y **definida positiva**.

Queremos resolver

$$\max_{R(x)=1} Q(x) \quad \text{y} \quad \min_{R(x)=1} Q(x).$$

Si $R(x) = x^T x = \|x\|^2$, entonces estamos en los casos que ya analizamos y que sabemos resolver usando el Teorema de Rayleigh. La idea entonces es, a partir de algún cambio de variables conveniente, trasformar la restricción $R(x) = 1$ en una restricción de la forma $z^T z = 1$ y luego usar lo que ya sabemos en las nuevas variables.

Para aplicar la idea anterior B debe ser necesariamente **definida positiva**. De hecho, supongamos que existe una matriz inversible $F \in \mathbb{R}^{n\times n}$ tal que con el cambio de variables $z = Fx$, transformamos la restricción $R(x) = 1$ en una restricción de la forma $z^T z = 1$. Entonces, por un lado $x = F^{-1}z$ y por el otro lado,

$$R(x) = x^T Bx = z^T(F^{-1})^T BF^{-1}z.$$

Si queremos que $z^T(F^{-1})^T BF^{-1}z = z^T z$ para todo $z \in \mathbb{R}^n$. Entonces,

$$z^T((F^{-1})^T BF^{-1} - I)z = z^T(F^{-1})^T BF^{-1}z - z^T z = 0 \text{ para todo } z \in \mathbb{R}^n.$$

Entonces, eso implica (meditar por qué) que

$$(F^{-1})^T BF^{-1} = I.$$

Si ahora multiplicamos la ecuación anterior a izquierda por F^T y a derecha por F, recordando que $(F^{-1})^T = (F^T)^{-1}$, tenemos que

$$B = F^T[(F^{-1})^T BF^{-1}]F = F^T IF = F^T F.$$

Entonces, para cada $x \in \mathbb{R}^n$ tenemos que

$$\langle Bx, x \rangle = \langle F^T Fx, x \rangle = \langle Fx, Fx \rangle = \|Fx\|^2 \geq 0,$$

y entonces B tiene que ser semidefinida positiva. Y además, como queremos que F sea inversible, tenemos que $\text{nul}(F) = \{0\}$. Entonces,

$$\langle Bx, x \rangle = 0 = \|Fx\|^2$$

si y sólo si $Fx = 0$ si y sólo si $x \in \mathrm{nul}(F) = \{0\}$. Por lo tanto, B tiene que ser necesariamente **definida positiva**. Si B no cumple esa condición, no tendremos esperanzas de encontrar un cambio de variables que transforme la restricción $R(x) = 1$ en una restricción de la forma $z^T z = 1$.

Ahora sí, supongamos que $B \in \mathbb{R}^{n\times n}$ es definida positiva. Entonces, B es simétrica, sus autovalores son todos positivos y existe $Q \in \mathbb{R}^{n\times n}$ ortogonal, tal que $B = Q\Lambda Q^T$, donde $\Lambda = \begin{bmatrix} \mu_1 & \cdots & 0 \\ \ldots & \ddots & \ldots \\ 0 & \cdots & \mu_n \end{bmatrix}$ es la matriz con todos los autovalores positivos de B.

Sea

$$F := Q \begin{bmatrix} \mu_1^{1/2} & \cdots & 0 \\ \ldots & \ddots & \ldots \\ 0 & \cdots & \mu_n^{1/2} \end{bmatrix} Q^T.$$

Entonces, esa es una diagonalización ortogonal de F por cómo la definimos. Es más, por definición, los autovalores de F son la raíz cuadrada de los autovalores de B. Por lo tanto, F es simétrica y tiene todos autovalores positivos (no nulos). Es decir, F es inversible y definida positiva y además

$$\begin{aligned} F^2 &= FF = FF^T = F^T F \\ &= Q \begin{bmatrix} \mu_1^{1/2} & \cdots & 0 \\ \ldots & \ddots & \ldots \\ 0 & \cdots & \mu_n^{1/2} \end{bmatrix} Q^T Q \begin{bmatrix} \mu_1^{1/2} & \cdots & 0 \\ \ldots & \ddots & \ldots \\ 0 & \cdots & \mu_n^{1/2} \end{bmatrix} Q^T = Q\Lambda Q^T = B. \end{aligned}$$

La matriz F es un buen candidato para el cambio de variables que buscamos. Entonces, hagamos el cambio de variables $z = Fx$, $(x = F^{-1}z)$. Entonces, como $B = F^2 = F^T F = FF^T$, tenemos que

$$R(x) = x^T Bx = x^T F^T Fx = (Fx)^T (Fx) = z^T z \ \text{ y}$$

$$Q(x) = x^T Ax = (F^{-1}z)^T A(F^{-1}z) = z^T (F^{-1})^T AF^{-1}z = z^T (F^{-1}AF^{-1})z.$$

Entonces, usando el Teorema de Rayleigh, tenemos que

$$\max_{R(x)=1} Q(x) = \max_{z^T z=1} z^T (F^{-1}AF^{-1})z = \lambda_M(F^{-1}AF^{-1}),$$

donde $\lambda_M(F^{-1}AF^{-1})$ es el autovalor máximo de la matriz simétrica $F^{-1}AF^{-1}$. Por el mismo teorema, tenemos que el máximo se alcanza en los $x = F^{-1}z$ con $z \in \mathcal{S}_{\lambda_M(F^{-1}AF^{-1})}$ tales que $\|z\| = 1$.

De manera análoga, tenemos que

$$\min_{R(x)=1} Q(x) = \min_{z^T z=1} z^T (F^{-1}AF^{-1})z = \lambda_m(F^{-1}AF^{-1}),$$

donde $\lambda_m(F^{-1}AF^{-1})$ es el autovalor mínimo de la matriz simétrica $F^{-1}AF^{-1}$. Por el mismo teorema, tenemos que el mínimo se alcanza en los $x = F^{-1}z$ con $z \in \mathcal{S}_{\lambda_m(F^{-1}AF^{-1})}$ tales que $\|z\| = 1$.

Para evitar hacer todas las cuentas que implicaría el cálculo de los autovalores y autoespacios de la matriz $F^{-1}AF^{-1}$, en lo que sigue veremos una manera simple de hallar $\lambda_M(F^{-1}AF^{-1})$, $\lambda_m(F^{-1}AF^{-1})$, $\mathcal{S}_{\lambda_M(F^{-1}AF^{-1})}$ y $\mathcal{S}_{\lambda_m(F^{-1}AF^{-1})}$.

Con la notación que usamos arriba, valen las siguientes observaciones:

1. λ es un autovalor de $F^{-1}AF^{-1}$ si y sólo si λ es una raíz de $\det(A - tB)$.
2. Sea λ un autovalor de $F^{-1}AF^{-1}$. Entonces $x = F^{-1}z$ con $z \in \mathcal{S}_\lambda$ si y sólo si $x \in \mathrm{nul}(A - \lambda B)$.

Probemos estas observaciones.

1. : Recordemos que $B = F^2$, entonces

$$F^{-1}BF^{-1} = F^{-1}FFF^{-1} = I.$$

Por lo tanto,

$$F^{-1}AF^{-1} - \lambda I = F^{-1}AF^{-1} - \lambda F^{-1}BF^{-1} = F^{-1}(A - \lambda B)F^{-1}.$$

Entonces, λ es un autovalor de $F^{-1}AF^{-1}$ si y sólo si

$$\begin{aligned} 0 &= \det(F^{-1}AF^{-1} - \lambda I) = \det(F^{-1}(A - \lambda B)F^{-1}) \\ &= \det(F)\det(A - \lambda B)\det(F^{-1}) = \det(A - \lambda B) \end{aligned}$$

si y sólo si λ es una raíz de $\det(A - tB)$.

2. : Sea λ un autovalor de $F^{-1}AF^{-1}$ y supongamos que $x = F^{-1}z$ con $z \in \mathcal{S}_\lambda$, entonces $(F^{-1}AF^{-1})z = \lambda z$. Por lo tanto, $AF^{-1}z = F(F^{-1}AF^{-1})z = F\lambda z = \lambda Fz$ y

$$(A - \lambda B)x = (A - \lambda B)F^{-1}z = AF^{-1}z - \lambda BF^{-1}z = \lambda Fz - \lambda FFF^{-1} = \lambda Fz - \lambda Fz = 0,$$

entonces $x \in \text{nul}(A - \lambda B)$.

Recíprocamente, supongamos que $x \in \text{nul}(A - \lambda B)$ entonces $(A - \lambda B)x = 0$, ó equivalentemente, $Ax = \lambda Bx = \lambda FFx$. Entonces

$$F^{-1}Ax = \lambda F^{-1}FFx = \lambda Fx.$$

Sea $z := Fx$, entonces $x = F^{-1}z$ y

$$(F^{-1}AF^{-1})z = F^{-1}Ax = \lambda Fx = \lambda z.$$

Por lo tanto λ es un autovalor de $F^{-1}AF^{-1}$, $z \in \mathcal{S}_\lambda$ y $x = F^{-1}z$.

Vamos a aplicar estas herramientas en el siguiente ejercicio:

Ejercicio 6.12. Sea $Q : \mathbb{R}^2 \to \mathbb{R}$ las forma cuadrática en $\mathbb{R}^2$ definida por

$$Q(x) := x_1^2 + x_2^2 + x_1x_2.$$

a) Observar que Q es definida positiva y hallar un cambio de variables $x = My$ tal que $\tilde{Q}(y) = Q(My) = \|y\|^2$.

b) Hallar, si existen, el máximo y el mínimo de $Q(x)$ sujetos a la restricción $R(x) := 9x_1^2 + 3x_2^2 - 8x_1x_2 = 1$ y determinar los vectores que los realizan.

Dem. $a)$: Observar que $Q(x) = x^TAx$, con $A = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$. El polinomio característico de A es $\chi_A(t) = (1-t)^2 - \frac{1}{4}$, cuyas raíces son $\frac{3}{2}$ y $\frac{1}{2}$ (ambas positivas). Por lo tanto, A es definida positiva (también se puede probar que A es definida positiva viendo que el determinante de las submatrices principales de A son ambos positivos). Por otra parte, los autoespacios correspondientes son $\mathcal{S}_{\lambda=\frac{3}{2}} = gen\{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$ y $\mathcal{S}_{\lambda=\frac{1}{2}} = gen\{\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$. Entonces, si $P := \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, tal como vimos arriba, podemos definir

$$G := P\begin{bmatrix} \sqrt{\frac{3}{2}} & 0 \\ 0 & \sqrt{\frac{1}{2}} \end{bmatrix}P^T.$$

Por definición G es simétrica e inversible. Más aún, $G^TG = GG = A$. Tomemos

$$M := G^{-1} = P \begin{bmatrix} \sqrt{\frac{2}{3}} & 0 \\ 0 & \sqrt{2} \end{bmatrix} P^T,$$

notar que $GM = MG = I$. Entonces, si consideramos el cambio de variables $x = My$, tenemos que

$$\tilde{Q}(y) = Q(My) = y^TM^TAMy = y^TM^TG^TGMy = y^T(GM)^T(GM)y = y^Ty = \|y\|^2.$$

$b)$: Observar que $R(x) = x^TBx$ con $B = \begin{bmatrix} 9 & -4 \\ -4 & 3 \end{bmatrix}$. El polinomio característico de B es $\chi_B(t) = (9-t)(3-t) - 16$, cuyas raíces son $\lambda_1 = 11$ y $\lambda_2 = 1$ (ambas positivas). Por lo tanto, B es definida positiva, entonces la restricción R también es una forma cuadrática definida positiva y podemos aplicar lo que vimos arriba. Recordar que probamos que si $B = FF$, con $F \in \mathbb{R}^{2\times 2}$ definida positiva, entonces

$$\max_{R(x)=1} Q(x) = \lambda_M(F^{-1}AF^{-1}) \text{ y } \min_{R(x)=1} Q(x) = \lambda_m(F^{-1}AF^{-1})$$

donde $\lambda_M(F^{-1}AF^{-1})$ y $\lambda_m(F^{-1}AF^{-1})$ son el máximo y mínimo autovalor de $F^{-1}AF^{-1}$ respectivamente. También vimos que λ es una autovalor de $F^{-1}AF^{-1}$ si y sólo si λ es una raíz de $\det(A - tB)$. Busquemos entonces las raíces de

$$\begin{aligned}\det(A - tB) &= \det(\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} - t \begin{bmatrix} 9 & -4 \\ -4 & 3 \end{bmatrix}) = \det(\begin{bmatrix} 1-9t & \frac{1}{2}+4t \\ \frac{1}{2}+4t & 1-3t \end{bmatrix}) \\ &= 27t^2 - 12t + 1 - (\frac{1}{2} + 4t)^2 = 27t^2 - 12t + 1 - 16t^2 - \frac{1}{4} - 4t \\ &= 11t^2 - 16t + \frac{3}{4}.\end{aligned}$$

Las ráices de $\det(A - tB)$ son $\lambda_1 = \frac{16+\sqrt{223}}{22} = \lambda_M(F^{-1}AF^{-1})$ y $\lambda_2 = \frac{16-\sqrt{223}}{22} = \lambda_m(F^{-1}AF^{-1})$. Entonces,

$$\max_{R(x)=1} Q(x) = \frac{16+\sqrt{223}}{22} \text{ y } \min_{R(x)=1} Q(x) = \frac{16-\sqrt{223}}{22}.$$

$b)$: Vimos que los $x \in \mathbb{R}^2$ tales que $R(x) = 1$ y que maximizan Q son los $x = F^{-1}z$ con $z \in \mathcal{S}_{\lambda_M(F^{-1}AF^{-1})}$ tales que $\|z\| = 1$. También vimos que $x = F^{-1}z$ con $z \in \mathcal{S}_{\lambda_M(F^{-1}AF^{-1})}$ si y sólo si $x \in \mathrm{nul}(A - \lambda_M(F^{-1}AF^{-1})B)$. Calculemos

$$\begin{aligned}\mathrm{nul}(A - \lambda_M(F^{-1}AF^{-1})B) &= \mathrm{nul}(A - \frac{16+\sqrt{223}}{22}B) \\ &= \begin{bmatrix} \frac{-122-9\sqrt{223}}{22} & \frac{75+4\sqrt{223}}{22} \\ \frac{75+4\sqrt{223}}{22} & \frac{-26-3\sqrt{223}}{22} \end{bmatrix} = gen\{\begin{bmatrix} 75+4\sqrt{223} \\ 122+9\sqrt{223} \end{bmatrix}\}.\end{aligned}$$

Entonces, el conjunto de todos los $x \in \mathbb{R}^2$ tales que $R(x) = 1$ y que maximizan Q son los x de la forma $x = \alpha \begin{bmatrix} 75+4\sqrt{223} \\ 122+9\sqrt{223} \end{bmatrix}$ con $\alpha \in \mathbb{R}$ tales que $R(x) = 1$. Llamemos $v := \begin{bmatrix} 75+4\sqrt{223} \\ 122+9\sqrt{223} \end{bmatrix}$, como

$$1 = R(x) = R(\alpha v) = \alpha^2 R(v),$$

tenemos que $\alpha = \pm\frac{1}{R(v)^{1/2}}$ y el conjunto de todos los $x \in \mathbb{R}^2$ tales que $R(x) = 1$ y que maximizan Q, es el conjunto de dos puntos:

$$\{\frac{1}{R(v)^{1/2}}v, -\frac{1}{R(v)^{1/2}}v\}.$$

c) : Vimos que los $x \in \mathbb{R}^2$ tales que $R(x) = 1$ y que minimizan Q son los $x = F^{-1}z$ con $z \in \mathcal{S}_{\lambda_m(F^{-1}AF^{-1})}$ tales que $\|z\| = 1$. También vimos que $x = F^{-1}z$ con $z \in \mathcal{S}_{\lambda_m(F^{-1}AF^{-1})}$ si y sólo si $x \in \text{nul}(A - \lambda_m(F^{-1}AF^{-1})B)$. Calculemos

$$\begin{aligned}\text{nul}(A - \lambda_m(F^{-1}AF^{-1})B) &= \text{nul}(A - \frac{16-\sqrt{223}}{22}B) \\ &= \begin{bmatrix} \frac{-122+9\sqrt{223}}{22} & \frac{75-4\sqrt{223}}{22} \\ \frac{75-4\sqrt{223}}{22} & \frac{-26+3\sqrt{223}}{22} \end{bmatrix} = gen\{\begin{bmatrix} 75-4\sqrt{223} \\ 122-9\sqrt{223} \end{bmatrix}\}.\end{aligned}$$

Entonces el conjunto de todos los $x \in \mathbb{R}^2$ tales que $R(x) = 1$ y que minimizan Q son los x de la forma $x = \alpha \begin{bmatrix} 75-4\sqrt{223} \\ 122-9\sqrt{223} \end{bmatrix}$ con $\alpha \in \mathbb{R}$ tales que $R(x) = 1$. Llamemos $w := \begin{bmatrix} 75-4\sqrt{223} \\ 122-9\sqrt{223} \end{bmatrix}$, como

$$1 = R(x) = R(\alpha w) = \alpha^2 R(w),$$

tenemos que $\alpha = \pm\frac{1}{R(w)^{1/2}}$ y el conjunto de todos los $x \in \mathbb{R}^2$ tales que $R(x) = 1$ y que minimizan Q, es el conjunto de dos puntos:

$$\{\frac{1}{R(w)^{1/2}}w, -\frac{1}{R(w)^{1/2}}w\}.$$

■

En el próximo ejercicio, vamos a resolver un problema de minimización con una restricción que no es definida positiva. Por lo tanto, no podremos aplicar las herramientas que vimos hasta ahora. Sin embargo, como la función a minimizar es $Q_1(x) = \|x\|^2$ (que veremos es una función relativamente sencilla) podremos resolver el ejercicio geométricamente.

El siguiente ejercicio es similar al **Ejercicio 6.11**.

Ejercicio 6.I. Hallar si existen, el máximo y el mínimo y (los puntos donde se alcanzan) de la forma cuadrática $Q_1 : \mathbb{R}^2 \to \mathbb{R}$, $Q_1(x) = \|x\|^2$, sujeto a la restricción

$$Q_2(x) = 2x_1^2 - 6x_2^2 + 6x_1x_2 = 1.$$

Dem. Observar que $Q_2(x) = x^T Bx$, con $B := \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$. El polinomio característico de B es

$$\chi_B(t) = (2-t)(-6-t) - 6 = t^2 + 4t - 21.$$

Entonces, los autovalores de B son $\lambda_1 = 3$ y $\lambda_2 = -7$. Por lo tanto B es indefinida y NO podemos aplicar la técnica que usamos en el **Ejercicio 6.12**.

Sin embargo, como la forma cuadrática a optimizar es $Q_1(x) = \|x\|^2$ no todo está perdido. Imaginemos que graficamos en $\mathbb{R}^2$ la restricción $Q_2(x) = 1$ que es el gráfico del conjunto de nivel $c = 1$ para la forma cuadrática indefinida Q_2. Si podemos observar cuáles puntos de dicho gráfico se encuentran más cercanos al origen es decir al punto $[0\ 0]^T \in \mathbb{R}^2$, entonces es claro que esos puntos van a minimizar la función $\|x\|$, luego (si no se entiende la próxima igualdad demostrarla) tenemos que

$$(\min_{Q_2(x)=1} \|x\|)^2 = \min_{Q_2(x)=1} \|x\|^2 = \min_{Q_2(x)=1} Q_1(x).$$

Lo mismo podremos pensar para los puntos más alejados del origen y el máximo que nos interesa.

Entonces, para graficar el conjunto de nivel $Q_2(x) = 1$, diagonalizemos ortogonalmente a la matriz B. Ya vimos que $\lambda_1 = 3$ y $\lambda_2 = -7$ son sus autovalores y los autoespacios asociados son:

$$\text{nul}(B - 3I) = \text{nul}(\begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}) = gen\{\begin{bmatrix} 3 \\ 1 \end{bmatrix}\},$$

$$\text{nul}(B+7I) = \text{nul}(\begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}) = gen\{\begin{bmatrix} -1 \\ 3 \end{bmatrix}\}.$$

Entonces, $\{\frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{10}}\begin{bmatrix} -1 \\ 3 \end{bmatrix}\}$ es una bon de $\mathbb{R}^2$ y si tomamos $P = \begin{bmatrix} \frac{3}{\sqrt{10}} & -\frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}$, entonces $\det(P) = 1$ y una diagonalización ortogonal de B es

$$B = P\begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}P^T.$$

Por lo tanto, con el cambio de variables $x = Py$, nos queda $Q(x) = \tilde{Q}(y) = 3y_1^2 - 7y_2^2$. Entonces $Q_2(x) = 1$ si y sólo $3y_1^2 - 7y_2^2 = 1$, ó equivalentemente,

$$\frac{y_1^2}{[\frac{1}{\sqrt{3}}]^2} - \frac{y_2^2}{[\frac{1}{\sqrt{7}}]^2} = 1.$$

Por lo tanto el conjunto de nivel $Q_2(x) = 1$ es una hipérbola en $\mathbb{R}^2$ en los nuevos ejes y_1, y_2 determinados por los versores $v_1 := \frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ y $v_2 := \frac{1}{\sqrt{10}}\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ que son las columnas de P y que son una rotación de un ángulo (positivo) de los ejes originales x_1, x_2. Ver Figura 6.5.

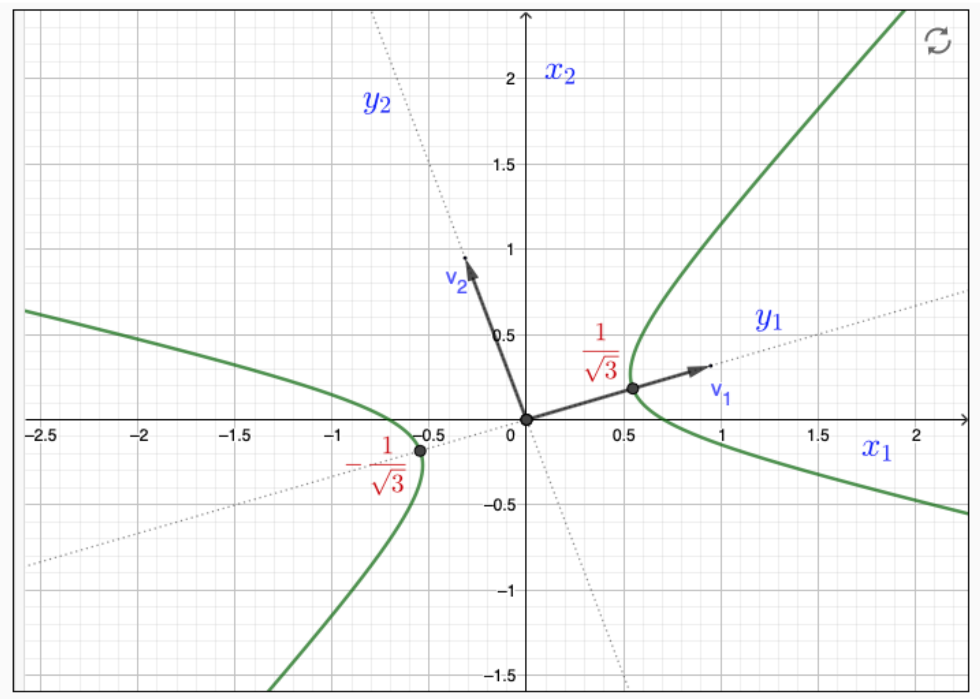

Figura 6.5: Ejercicio 6.C

Como se puede observar, los puntos del conjunto de nivel $Q_2(x) = 1$ (que son los puntos de la hipérbola de color verde) más cercanos al origen, son los puntos $y_{m1} = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ 0 \end{bmatrix}$ e $y_{m2} = -\begin{bmatrix} \frac{1}{\sqrt{3}} \\ 0 \end{bmatrix}$ (en los ejes y_1, y_2). Por lo tanto, usando el cambio de variables $x = Py$, tenemos que, los puntos más cercanos al origen son

$$x_{m1} = Py_{m1} = \begin{bmatrix} \frac{3}{\sqrt{10}} & -\frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{3}} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{30}} \\ \frac{1}{\sqrt{30}} \end{bmatrix} \text{ y } x_{m2} = Py_{m2} = \begin{bmatrix} \frac{-3}{\sqrt{30}} \\ \frac{-1}{\sqrt{30}} \end{bmatrix}.$$

Entonces

$$\min_{Q_2(x)=1} Q_1(x) = \min_{Q_2(x)=1} \|x\|^2 = (\min_{Q_2(x)=1} \|x\|)^2 = \|x_m\|^2 = \|y_m\|^2 = (\frac{1}{\sqrt{3}})^2 = \frac{1}{3}$$

y dicho mínimo, como acabamos de ver, se alcanza en los puntos $x_{m1} = \begin{bmatrix} \frac{3}{\sqrt{30}} \\ \frac{1}{\sqrt{30}} \end{bmatrix}$ y $x_{m2} = \begin{bmatrix} \frac{-3}{\sqrt{30}} \\ \frac{-1}{\sqrt{30}} \end{bmatrix}$.

Por último, es claro que no existe el máximo de $\|x\|^2$ restringido a $Q_2(x) = 1$. ■

> Tal como hicimos en el **Ejercicio 6.I**, podemos determinar gráficamente (si existen) $\min_{Q_2(x)=1} Q_1(x)$ y $\max_{Q_2(x)=1} Q_1(x)$ y los puntos donde se alcanzan esos extremos, para cualquier forma cuadrática Q_2 (sin importar si es definida positiva) pero cuando Q_1 es una forma cuadrática muy particular, por ejemplo $Q_1(x) = \|x\|^2$ o formas cuadráticas similares.

Veamos otro ejemplo:

Ejercicio 6.J. Hallar si existen, el máximo y el mínimo y (los puntos donde se alcanzan) de la forma cuadrática $Q_1 : \mathbb{R}^2 \to \mathbb{R}$, $Q_1(x) = 5\|x\|^2$, sujeto a la restricción

$$Q_2(x) = 2x_1^2 - 6x_2^2 + 6x_1x_2 = 1.$$

Dem. Notar que

$$5\ (\min_{Q_2(x)=1} \|x\|)^2 = \min_{Q_2(x)=1} 5\|x\|^2 = \min_{Q_2(x)=1} Q_1(x).$$

Lo mismo podremos pensar para los puntos más alejados del origen y el máximo que nos interesa. Por lo tanto, operando de manera similar al **Ejercicio 6.I**, tenemos que

$$\min_{Q_2(x)=1} Q_1(x) = \min_{Q_2(x)=1} 5\|x\|^2 = 5(\min_{Q_2(x)=1} \|x\|)^2 = 5\|x_m\|^2 = 5\|y_m\|^2 = 5(\frac{1}{\sqrt{3}})^2 = \frac{5}{3}$$

y dicho mínimo, se alcanza en los puntos $x_{m1} = \begin{bmatrix} \frac{3}{\sqrt{30}} \\ \frac{1}{\sqrt{30}} \end{bmatrix}$ y $x_{m2} = \begin{bmatrix} \frac{-3}{\sqrt{30}} \\ \frac{-1}{\sqrt{30}} \end{bmatrix}$.

Por último, es claro que no existe el máximo de $5\|x\|^2$ restringido a $Q_2(x) = 1$. ■

Ejercicio 6.13. Sea $Q : \mathbb{R}^2 \to \mathbb{R}$ la forma cuadrática definida por $Q(x) := ax_1^2 + ax_2^2 + 2bx_1x_2$, donde $a, b \in \mathbb{R}$.

a) Hallar y graficar el conjunto de todos los pares (a, b) para los cuales Q es definida positiva.

b) Hallar y graficar el conjunto de todos los pares (a, b) para los cuales

$$\min_{\|x\|^2=1} Q(x) = 0 \text{ y } \max_{\|x\|^2=1} Q(x) = 4.$$

c) Determinar los valores de a y b para los cuales existe un cambio de variables ortogonal $x = Py$ tal que $\tilde{Q}(y) := Q(Py) = 4y_1^2 + 9y_2^2$.

d) Para $a = \frac{5}{2}$ y $b = \frac{3}{2}$ hallar los puntos de la curva de nivel $Q(x) = 1$ más cercanos al origen. ¿Qué puede decirse de los más lejanos?

e) Para $a = \frac{3}{2}$ y $b = \frac{5}{2}$ hallar los puntos de la curva de nivel $Q(x) = 1$ más cercanos al origen. ¿Qué puede decirse de los más lejanos?

Dem. $a)$: Notar que $Q(x) = x^T Ax$, con $A := \begin{bmatrix} a & b \\ b & a \end{bmatrix}$. Por lo tanto, tal como vimos en el **Teorema 5.4.2**, A es definida positiva si y sólo si el determinate de todas sus submatrices principales es positivo. En este caso, tenemos dos submatrices principales: $A^{(1)} = [a]$ cuyo determinante es $\det(A^{(1)}) = a$ y $A^{(2)} = A$ cuyo determinante es $\det(A^{(2)}) = \det(A) = a^2 - b^2$. Entonces A es definida positiva, si y sólo si $a > 0$ y $a^2 > b^2$. Entonces, tomando raíz en la última desigualdad y usando que $a > 0$, tenemos que $|a| = a > |b|$. Por lo tanto, A es definida positiva, si y sólo $a > 0$ y $-a < b < a$.

$b)$: El polinomio característico de A es $\chi_A(t) = (a-t)^2 - b^2$, cuyas raíces son $\lambda_1 = a + b$ y $\lambda_2 = a - b$. Veamos distintos casos:

- **Caso 1: b>0**. En este caso, $a + b > a - b$, el mayor autovalor de A es $a + b$ y el menor es $a - b$. Por lo tanto, por el Teorema de Rayleigh, $\min\limits_{\|x\|^2=1} Q(x) = a - b$. Si buscamos que $\min\limits_{\|x\|^2=1} Q(x) = a - b = 0$, tenemos que $a = b$. De la misma manera, $\max\limits_{\|x\|^2=1} Q(x) = a + b$. Si buscamos que $\min\limits_{\|x\|^2=1} Q(x) = a + b = 4$. Despejando, nos queda que $2b = 4$, por lo tanto $b = 2$ y $a = 2$.

- **Caso 2: b<0**. En este caso, $a - b > a + b$, el mayor autovalor de A es $a - b$ y el menor es $a + b$. Por lo tanto, por el Teorema de Rayleigh, $\min\limits_{\|x\|^2=1} Q(x) = a + b$. Si buscamos que $\min\limits_{\|x\|^2=1} Q(x) = a + b = 0$, tenemos que $a = -b$. De la misma manera, $\max\limits_{\|x\|^2=1} Q(x) = a - b$. Si buscamos que $\min\limits_{\|x\|^2=1} Q(x) = a - b = 4$. Despejando, nos queda que $-2b = 4$, por lo tanto $b = -2$ y $a = 2$.

- **Caso 3: b=0**. En este caso, $a - b = a + b = a$ y por lo tanto $\min\limits_{\|x\|^2=1} Q(x) = \max\limits_{\|x\|^2=1} Q(x) = a$. Si buscamos que $\min\limits_{\|x\|^2=1} Q(x) = 0$, tenemos que $a = 0$. De la misma manera, si buscamos que $\min\limits_{\|x\|^2=1} Q(x) = a = 4$. Entonces $a = 4$. Lo que es absurdo y este caso no puede suceder.

Conclusión $\min\limits_{\|x\|^2=1} Q(x) = 0$ y $\max\limits_{\|x\|^2=1} Q(x) = 4$ si y sólo si $(a, b) \in \{(2, 2), (2, -2)\}$.

$c)$: Notar que existe un cambio de variables ortogonal $x = Py$ tal que $\tilde{Q}(y) := Q(Py) = 4y_1^2 + 9y_2^2$ si y sólo si

$$Q(Py) = y^T \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} y = Q(x) = y^T P^T APy.$$

Entonces

$$y^T (\begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} - P^T AP)y = \left\langle y, (\begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} - P^T AP)y \right\rangle = 0$$

para todo $y \in \mathbb{R}^2$. Por lo tanto $P^T AP = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$ o lo que es equivalente, $A = P \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} P^T$. Es decir, los autovalores de A son $\lambda_1 = 4$ y $\lambda_2 = 9$. Entonces, tal como hicimos en $b)$ tenemos que analizar 3 casos recordando que los autovalores de A son $a + b$ y $a - b$. Entonces:

- **Caso 1: b>0**. En este caso, $a + b > a - b$, el mayor autovalor de A es $a + b$ y el menor es $a - b$. Por lo tanto, por el Teorema de Rayleigh, $\min\limits_{\|x\|^2=1} Q(x) = a - b = 4$. De la misma manera, $\max\limits_{\|x\|^2=1} Q(x) = a + b = 9$. Despejando, nos queda que $a = \frac{13}{2}$ y $b = \frac{5}{2}$.

- **Caso 2: b<0**. En este caso, $a - b > a + b$, el mayor autovalor de A es $a - b$ y el menor es $a + b$. Por lo tanto, por el Teorema de Rayleigh, $\min\limits_{\|x\|^2=1} Q(x) = a + b = 4$. De la misma manera, $\max\limits_{\|x\|^2=1} Q(x) = a - b = 9$. Despejando, nos queda que $a = \frac{13}{2}$ y $b = -\frac{5}{2}$.

- **Caso 3: b=0**. En este caso A tiene sólo 1 autovalor doble $\lambda = a - b = a + b = a$ y por lo tanto no existen valores de a y b tales que los autovalores de A sean 4 y 9.

Conclusión, existe un cambio de variables ortogonal $x = Py$ tal que $\tilde{Q}(y) := Q(Py) = 4y_1^2 + 9y_2^2$ si y sólo si $(a, b) \in \{(\frac{13}{2}, \frac{5}{2}), (\frac{13}{2}, \frac{-5}{2})\}$.

$d)$: Si $a = \frac{5}{2}$ y $b = \frac{3}{2}$ entonces los autovalores de A son $\lambda_M = a + b = 4$ y $\lambda_m = a - b = 1$ y los autoespacios asociados son $\mathcal{S}_{\lambda_M} = \text{nul}(A-4I) = gen\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$ y $\mathcal{S}_{\lambda_m} = \text{nul}(A-I) = gen\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$. Entonces $\{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$ es una bon de $\mathbb{R}^2$. Si tomamos $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ tenemos que $\det(P) = 1$ y $A = P\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}P^T$. En este caso, con el cambio de variables $x = Py$, vale que

$$1 = Q(x) = \tilde{Q}(y) = 4y_1^2 + 1y_2^2 = \frac{y_1^2}{(\frac{1}{2})^2} + y_2^2.$$

En los nuevos ejes rotados y_1, y_2 determinados por los versores de $\mathbb{R}^2$: $v_1 := \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $v_2 := \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ el conjunto de nivel $\mathcal{N}_1(Q) = \{x \in \mathbb{R}^2 : Q(x) = 1\}$ es una elipse con centro en el origen de coordenadas y ejes coincidentes con los ejes y_1, y_2, ver Figura 6.6. En los nuevos ejes, los puntos $y_m := \pm\begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix}$ son los puntos de $\mathcal{N}_1(Q)$ más cercanos al origen. En la variable original, estos puntos son $x_m = Py_m = \pm\frac{1}{2\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y el valor de la distancia mínima es $\|y_m\| = \|x_m\| = \|\pm\frac{1}{2\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\| = \frac{1}{2}$. De manera similar, los puntos $y_M := \pm\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ son los puntos de $\mathcal{N}_1(Q)$ más alejados del origen. En la variable original, estos puntos son $x_M = Py_M = \pm\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ y el valor de la distancia máxima es $\|y_m\| = \|x_m\| = \|\pm\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}\| = 1$.

$e)$: Si $a = \frac{3}{2}$ y $b = \frac{5}{2}$ entonces los autovalores de A son $\lambda_M = a + b = 4$ y $\lambda_m = a - b = -1$ y los autoespacios asociados son $\mathcal{S}_{\lambda_M} = \text{nul}(A-4I) = gen\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$ y $\mathcal{S}_{\lambda_m} = \text{nul}(A+I) = gen\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$.

Entonces $\{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$ es una bon de $\mathbb{R}^2$. Si tomamos $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ tenemos que $\det(P) = 1$ y $A = P\begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}P^T$. En este caso, con el cambio de variables $x = Py$, vale que

$$1 = Q(x) = \tilde{Q}(y) = 4y_1^2 - 1y_2^2 = \frac{y_1^2}{(\frac{1}{2})^2} - y_2^2.$$

En los nuevos ejes rotados y_1, y_2 determinados por los versores de $\mathbb{R}^2$: $v_1 := \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $v_2 := \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ el conjunto de nivel $\mathcal{N}_1(Q) = \{x \in \mathbb{R}^2 : Q(x) = 1\}$ es una hipérbola, ver Figura 6.7. Por lo tanto, en los nuevos ejes, los puntos $y_m := \pm\begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix}$ son los puntos de $\mathcal{N}_1(Q)$ más cercanos al origen. En la variable original, estos puntos son $x_m = Py_m = \pm\frac{1}{2\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y el valor de la distancia mínima es $\|y_m\| = \|x_m\| = \|\pm\frac{1}{2\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\| = \frac{1}{2}$. Cómo se puede observar en la Figura 6.7, el conjunto no posee máximo (no está acotado superiormente). ■

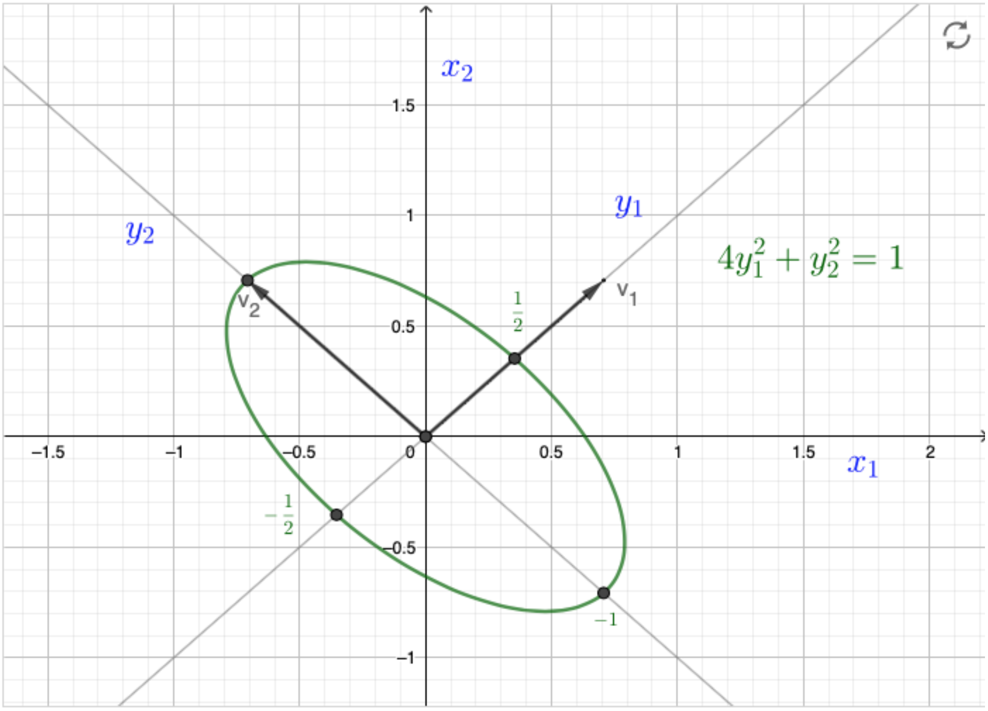

Figura 6.6: Ejercicio 6.13 c)

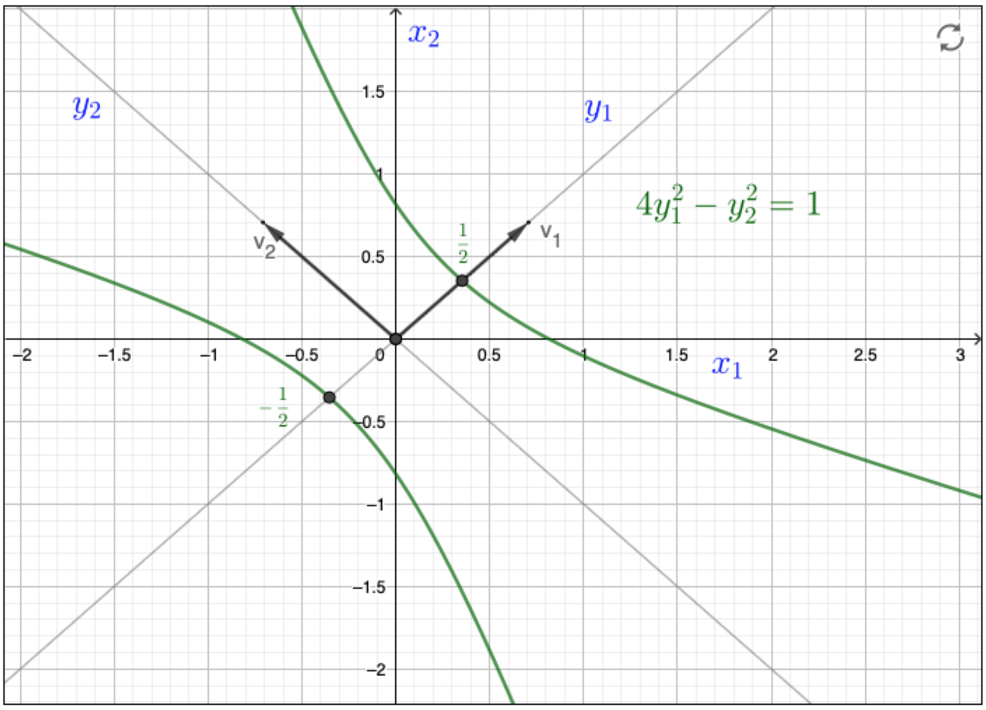

Figura 6.7: Ejercicio 6.13 d)

Bibliografía

[1] Grossman S., *Álgebra lineal con aplicaciones*, Mac Graw-Hill, 3era. Edición, 1990.

[2] Hoffman K., Kunze R., *Álgebra Lineal*, Prentice Hall, 1984.

[3] Jeronimo G., Sabia J., Tesauri S., *Álgebra Lineal*, Departamento de Matemática - FCEyN - Universidad de Buenos Aires, 2008.

[4] Maltsev A.I., *Fundamentos de Álgebra Lineal*, MIR, 1978.

[5] Mancilla Aguilar J.L., *Apuntes sobre Álgebra Lineal*, Departamento de Matemática - FIUBA - Universidad de Buenos Aires.