Capítulo 3
EL ESPACIO EUCLíDEO $\mathbb{R}^N$. ORTOGONALIDAD Y MíNIMOS CUADRADOS

Este capítulo se dedica al estudio de la estructura euclídea de $\mathbb{R}^n$. Definiremos los conceptos de producto escalar ordinario y norma, que nos permitirán medir distancias y ángulos entre vectores, y definir los conceptos de complemento ortogonal de un subespacio vectorial y proyección ortogonal de un vector sobre un subespacio, que nos conducirán al método de resolución de ecuaciones en el sentido de los mínimos cuadrados.

3.1. Producto escalar y norma. Ortogonalidad

3.1.1. Definiciones y propiedades

Definición 111 *Si u y v son dos vectores cualesquiera de $\mathbb{R}^n$ su producto escalar ordinario es:*

$$u \cdot v = (u_1, ..., u_n) \cdot (v_1, ..., v_n) = u_1 v_1 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

$\mathbb{R}^n$ con este producto escalar se llama espacio euclídeo n–dimensional. Si u y v se representan por columnas, se verifica que: $u \cdot v = v^T u = u^T v$

Proposición 112 *Sean $u, v \in \mathbb{R}^n$ y $c \in \mathbb{R}$, cualesquiera. Se verifica:*

a) $u \cdot v = v \cdot u$.

b) $(u + v) \cdot w = u \cdot w + v \cdot w$.

c) $(cu) \cdot v = c(u \cdot v)$

d) $u \cdot u \geq 0$, *y* $u \cdot u = 0 \Leftrightarrow u = 0$.

Demostración:

a) $u \cdot v = (u_1, ..., u_n) \cdot (v_1, ..., v_n) = u_1 v_1 + \cdots + u_n v_n = v_1 u_1 + \cdots + v_n u_n = v \cdot u$.

b) $(u + v) \cdot w = ((u_1, ..., u_n) + (v_1, ..., v_n)) \cdot (w_1, ..., w_n)$

$$= (u_1 + v_1, ..., u_n + v_n) \cdot (w_1, ..., w_n)$$

$$= (u_1 + v_1)w_1 + \cdots + (u_n + v_n)w_n$$

$$= (u_1w_1 + \cdots + u_nw_n) + (v_1w_1 + \cdots + v_nw_n)$$

$$= u \cdot w + v \cdot w.$$

c) $(cu) \cdot v = c(u_1, ..., u_n) \cdot (v_1, ..., v_n) = (cu_1, ..., cu_n) \cdot (v_1, ..., v_n)$

$$= cu_1v_1 + \cdots + cu_nv_n = c(u_1v_1 + \cdots + u_nv_n)$$

$$= c(u \cdot v).$$

d) $u \cdot u = (u_1, ..., u_n) \cdot (u_1, ..., u_n) = u_1^2 + \cdots + u_n^2 \geq 0.$

$u \cdot u = 0 \Leftrightarrow u_1^2 + \cdots + u_n^2 = 0 \Leftrightarrow u_1 = \cdots = u_n = 0 \Leftrightarrow u = 0.$ □

Definición 113 *La norma de un vector* $u \in \mathbb{R}^n$ *es el número real*

$$\|u\| = \sqrt{u \cdot u} = \sqrt{u_1^2 + \cdots + u_n^2} = \sqrt{\sum_{i=1}^{n} u_i^2} = \left(\sum_{i=1}^{n} u_i^2\right)^{\frac{1}{2}}.$$

Proposición 114 *Sean* $u, v \in \mathbb{R}^n$, $k \in \mathbb{R}$, *cualesquiera. Se verifica:*

a) $\|u\| = 0 \Leftrightarrow u = \mathbf{0}$.

b) $\|ku\| = |k| \|u\|$.

c) $|u \cdot v| \leq \|u\| \|v\|$ *(desigualdad de Cauchy-Schwarz)*

d) $\|u + v\| \leq \|u\| + \|v\|$ *(desigualdad triangular)*

Demostración:

a) $\|u\| = 0 \Leftrightarrow \sqrt{u \cdot u} = \mathbf{0} \Leftrightarrow u \cdot u = 0 \Leftrightarrow u = 0.$

b) $\|ku\| = \sqrt{(ku) \cdot (ku)} = \sqrt{k^2}\sqrt{u \cdot u} = |k| \|u\|$.

c) Si $u = \mathbf{0}$, entonces $\|u\| = 0$, $u \cdot v = 0$ y el resultado es cierto.

Si $u \neq \mathbf{0}$, consideramos el vector $xu + v$, siendo x un escalar no nulo.

$0 \leq \|xu + v\|^2 = (xu + v) \cdot (xu + v) = x^2 u \cdot u + 2xu \cdot v + v \cdot v = ax^2 + 2bx + c = p(x)$

siendo, $a = u \cdot u > 0$, $b = u \cdot v$, $c = v \cdot v$.

Como $p(x) \geq 0$ la ecuación $p(x) = 0$ o no tiene raices reales o tiene una raíz doble:

$$x = \frac{-2b \pm \sqrt{4b^2 - 4ac}}{2}$$

Esto implica que $4b^2 - 4ac \leq 0$. Por tanto,

$$b^2 \leq ac \Leftrightarrow |b| \leq \sqrt{a}\sqrt{c} \Leftrightarrow |u \cdot v| \leq \sqrt{u \cdot u}\sqrt{v \cdot v} \Leftrightarrow |u \cdot v| \leq \|u\| \, \|v\|$$

d)
$$\begin{aligned} \|u+v\|^2 &= (u+v)(u+v) = u \cdot u + 2u \cdot v + v \cdot v \\ &= \|u\|^2 + 2u \cdot v + \|v\|^2 \leq \|u\|^2 + 2\,|u \cdot v| + \|v\|^2 \\ &\leq \|u\|^2 + 2\,\|u\|\,\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2 . \end{aligned}$$

Por tanto, $\|u+v\|^2 \leq (\|u\| + \|v\|)^2 \Rightarrow \|u+v\| \leq \|u\| + \|v\|$. □

Definición 115 *La distancia entre dos vectores $u, v \in \mathbb{R}^n$ viene dada por:*

$$d(u,v) = \|u - v\| = \sqrt{(u_1 - v_1)^2 + \cdots + (u_n - v_n)^2} = \sqrt{\sum_{i=1}^{n} (u_i - v_i)^2}.$$

Proposición 116 *Sean $u, v, w \in \mathbb{R}^n$, cualesquiera. Se verifica:*

a) $d(u,v) = 0 \Leftrightarrow u = v$.

b) $d(u,v) = d(v,u)$.

c) $d(u,v) \leq d(u,w) + d(w,u)$.

Demostración:

a) $d(u,v) = 0 \Leftrightarrow \sqrt{(u_1 - v_1)^2 + \cdots + (u_n - v_n)^2} = 0 \Leftrightarrow (u_1 - v_1)^2 = \cdots = (u_n - v_n)^2 = 0$

$\Leftrightarrow (u_1 - v_1) = \cdots = (u_n - v_n) = 0 \Leftrightarrow u_i = v_i$ para todo $i : 1, ...n \Leftrightarrow u = v$.

b) $d(u,v) = \sqrt{(u_1 - v_1)^2 + \cdots + (u_n - v_n)^2} = \sqrt{(v_1 - u_1)^2 + \cdots + (v_n - u_n)^2} = d(v,u)$.

c) $d(u,v) = \|u - v\| = \|u - w + w - v\| = \|(u - w) + (w - v)\| \leq \|u - w\| + \|w - v\|$
$= d(u,w) + d(w,v)$. □

Definición 117 *Dos vectores $u, v \in \mathbb{R}^n$ son ortogonales si $u \cdot v = 0$, y se denota $u \perp v$.*

Definición 118 *Dados dos vectores no nulos $u, v \in \mathbb{R}^n$, se dice que forman un ángulo $\alpha \in [0, \pi]$, si*

$$\cos\alpha = \frac{u \cdot v}{\|u\| \, \|v\|}.$$

Teorema 119 (de Pitágoras) *Sean $u, v \in \mathbb{R}^n$ dos vectores ortogonales, entonces:*

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2 .$$

Demostración:

$\|u+v\|^2 = (u+v)\cdot(u+v) = u\cdot u + u\cdot v + v\cdot u + v\cdot v = \|u\|^2 + \|v\|^2$
pues $u \cdot v = 0$ al ser ortogonales. □

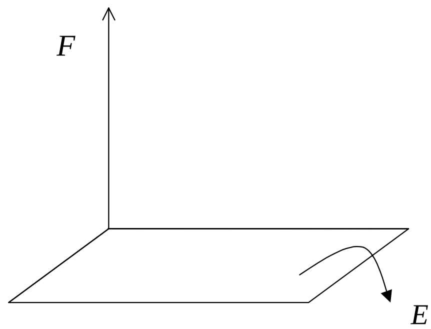
Corolario 120 *Si $u_1, ..., u_p \in \mathbb{R}^n$ son vectores ortogonales dos a dos, se tiene:*

$$\|u_1 + \cdots + u_p\|^2 = \|u_1\|^2 + \cdots + \|u_p\|^2 .$$

3.1.2. Complemento ortogonal

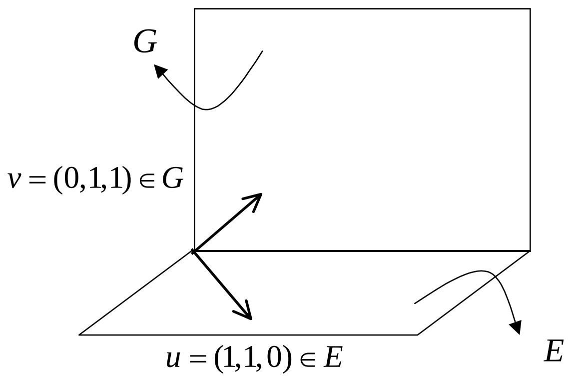
Definición 121 *Sea E un subespacio vectorial de $\mathbb{R}^n$.*
a) Se dice que un vector $x \in \mathbb{R}^n$ es ortogonal a E si es ortogonal a todos los vectores de E.
b) El conjunto de todos los vectores de $\mathbb{R}^n$ que son ortogonales a E se llama complemento ortogonal y se denota por $E^{\perp}$.

Ejemplo 122 *Sean $E = \{(u_1, u_2, u_3) \in \mathbb{R}^3 : u_3 = 0\}$ y $F = \{u \in \mathbb{R}^3 : u_1 = 0, u_2 = 0\}$, se verifica que $E^{\perp} = F$.*



Ejemplo 123 *Los subespacios $E = \{(u_1, u_2, u_3) \in \mathbb{R}^3 : u_3 = 0\}$ y $G = \{(u_1, u_2, u_3) \in \mathbb{R}^3 : u_1 = 0\}$ no son subespacios ortogonales, pues hay vectores de E y G que no son ortogonales entre si.*

$$u = (1,1,0) \in E, v = (0,1,1) \in G \Rightarrow u \cdot v = 1 \neq 0$$



Esto nos indica que la "pared" y el "suelo" no son ortogonales en el sentido que aquí se ha definido.

Proposición 124 *Sea E un subespacio vectorial de $\mathbb{R}^n$, dotado del producto escalar usual. Se verifica:*

a) *$E^\perp$ es un subespacio vectorial.*

b) $E \cap E^\perp = \{\mathbf{0}\}$.

c) $\dim E + \dim E^\perp = n$

d) $(E^\perp)^\perp = E$.

Demostración:

a) Sean $u_1, u_2 \in E^\perp$ y $\lambda, \mu \in \mathbb{R}$ cualesquiera, y sea v un vector cualquiera de E. Se verifica que:

$$(\lambda u_1 + \mu u_2) \cdot v = (\lambda u_1) \cdot v + (\mu u_2) \cdot v = \lambda(u_1 \cdot v) + \mu(u_2 \cdot v) = \lambda 0 + \mu 0 = \mathbf{0}.$$

Luego $(\lambda u_1 + \mu u_2) \in E^\perp$ y $E^\perp$ es un subespacio vectorial de $\mathbb{R}^n$.

b) Sea $u \in E \cap E^\perp$, enontces $u \in E$ y $u \in E^\perp$, luego $u \cdot u = 0 \Rightarrow \|u\|^2 = 0 \Rightarrow u = \mathbf{0}$. Luego $E \cap E^\perp = \{\mathbf{0}\}$.

c) Sea $E = L\{v_1, ..., v_p\}$, con $\dim E = p$, entonces $E^\perp = \{x \in \mathbb{R}^n | x \cdot v_i = 0, i = 1, ..., p\}$, y

como los vectores v_i son l.i. se tiene que $\dim E^\perp = n - p$, y por tanto: $\dim E + \dim E^\perp = n$.

d) Si $F = E^\perp$ se verifica que $\dim F + \dim F^\perp = n$, luego $\dim (E^\perp)^\perp = \dim F^\perp = n - \dim F = n - (n - p) = p$.

Además, para todo $x \in E$, se cumple que $x \cdot y = 0$, siendo y cualquier vector de $E^\perp$, luego $x \in (E^\perp)^\perp$, por tanto $E \subseteq (E^\perp)^\perp$, como $\dim E = p$ resulta que $(E^\perp)^\perp = E$. □

Ejemplo 125 *Sea $E = L\{v_1, v_2\}, v_1 = (1,1,0,1), v_2 = (0,-1,1,1)$ un subespacio vectorial de $\mathbb{R}^4$. Determinar una base de $E^\perp$.*

$$x \in E^\perp \text{ si y sólo si } \left\{ \begin{array}{l} x \cdot v_1 = 0 \\ x \cdot v_2 = 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{rrrr} x_1 + x_2 & & + x_4 = 0 \\ & -x_2 + x_3 & + x_4 = 0 \end{array} \right. \Rightarrow \boxed{\begin{array}{l} x_1 = -\alpha - 2\beta \\ x_2 = \alpha + \beta \\ x_3 = \alpha \\ x_4 = \beta \end{array}}$$

con $\alpha, \beta \in \mathbb{R}$.

Luego $E^{\perp} = L\{u_1, u_2\}$, *siendo* $u_1 = (-1, 1, 1, 0), u_2 = (-2, 1, 0, 1)$.

La base pedida es $B_{E^{\perp}} = \{u_1, u_2\}$, *por ser un sistema generador l.i.*

Nota: Obsérvese que $u_1 \cdot v_1 = u_1 \cdot v_2 = u_2 \cdot v_1 = u_2 \cdot v_2 = 0$, *y que* $\dim E + \dim E^{\perp} = 4$.

Ejemplo 126 *Sea* $W = L\{w_1, w_2, w_3, w_4\}$, *donde* $w_1 = (2, -1, 1, 3, 0)$, $w_2 = (1, 2, 0, 1, -2)$, $w_3 = (4, 3, 1, 5, -4)$, $w_4 = (3, 1, 2, -1, 1)$. *Determinar una base de* $W^{\perp}$.

En primer lugar buscamos una base de W, *para ello disponemos los vectores* w_i *como columnas de una matriz y estudiamos su dependencia lineal mediante un proceso de escalonamiento.*

$$\begin{bmatrix} 2 & 1 & 4 & 3 \\ -1 & 2 & 3 & 1 \\ 1 & 0 & 1 & 2 \\ 3 & 1 & 5 & -1 \\ 0 & -2 & -4 & 1 \end{bmatrix} \underset{\substack{F_{21}(\frac{1}{2}) \\ F_{31}(-\frac{1}{2}) \\ F_{41}(-\frac{3}{2})}}{\sim} \begin{bmatrix} 2 & 1 & 4 & 3 \\ 0 & \frac{5}{2} & 5 & \frac{5}{2} \\ 0 & -\frac{1}{2} & -1 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & -1 & -\frac{11}{2} \\ 0 & -2 & -4 & 1 \end{bmatrix} \underset{\substack{F_{32}(\frac{1}{5}) \\ F_{42}(\frac{1}{5}) \\ F_{52}(\frac{4}{5})}}{\sim} \begin{bmatrix} 2 & 1 & 4 & 3 \\ 0 & \frac{5}{2} & 5 & \frac{5}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

$$\underset{F_{53}}{\sim} \begin{bmatrix} 2 & 1 & 4 & 3 \\ 0 & \frac{5}{2} & 5 & \frac{5}{2} \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \underset{F_{43}(\frac{5}{3})}{\sim} \begin{bmatrix} 2 & 1 & 4 & 3 \\ 0 & \frac{5}{2} & 5 & \frac{5}{2} \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Por tanto, $\dim W = 3$, *y los vectores* $\{w_1, w_2, w_4\}$ *constituyen una base de* W.

Es decir: $W = L\{w_1, w_2, w_4\}$

$w_1 = (2, -1, 1, 3, 0)$, $w_2 = (1, 2, 0, 1, -2)$, $w_3 = (4, 3, 1, 5, -4)$, $w_4 = (3, 1, 2, -1, 1)$.

$$x \in W^{\perp} \text{ si y sólo si } x \cdot w_i = 0 \ (i = 1, 2, 4) \Rightarrow \begin{cases} 2x_1 - x_2 + x_3 + 3x_4 = 0 \\ x_1 + 2x_2 + x_4 - 2x_5 = 0 \\ 3x_1 + x_2 + 2x_3 - x_4 + x_5 = 0 \end{cases}$$

y la solución es $\boxed{\begin{array}{l} x_1 = -\frac{17}{5}\alpha + \frac{8}{5}\beta \\ x_2 = \frac{6}{5}\alpha + \frac{1}{5}\beta \\ x_3 = 5\alpha - 3\beta \\ x_4 = \alpha \\ x_5 = \beta \end{array}}$

$$W^{\perp} = L\left\{(-\tfrac{17}{5}, \tfrac{6}{5}, 5, 1, 0), (\tfrac{8}{5}, \tfrac{1}{5}, -3, 0, 1)\right\} = L\{(-17, 6, 25, 5, 0), (8, 1, -15, 0, 5)\}$$

$B_{W^\perp} = \{(-17, 6, 25, 5, 0), (8, 1, -15, 0, 5)\}$ *es la base pedida.*

Nota: Obsérvese que las componentes de los vectores $\{w_1, w_2, w_4\}$ *, base de* W*, coinciden con los coeficientes del sistema de ecuaciones implícitas que determina* $W^\perp$*.*

El siguiente resultado relaciona el ortogonal del espacio columna de una matriz con el espacio nulo de la matriz traspuesta, y lo emplearemos más adelante con ocasión del estudio del método de los mínimos cuadrados.

Teorema 127 *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$*, se verifica que* $N(A^T) = R(A)^\perp$*.*

Demostración:
Si $rg(A) = r \Rightarrow \dim R(A) = r$ y $\dim R(A)^\perp = m - r$, siendo $R(A) \subseteq \mathbb{R}^m$. Por otra parte, $N(A^T) \subseteq \mathbb{R}^m$ y $\dim N(A^T) = m - r$. Para probar que $R(A)$ y $N(A^T)$ coinciden bastará demostrar que: $N(A^T) \subseteq R(A)^\perp$.

Si $y \in N(A^T)$, entonces $A^T y = \mathbf{0}$, y representando la matriz A^T por filas:

$$\begin{bmatrix} a_1^T \\ \hline \vdots \\ \hline a_n^T \end{bmatrix} y = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} a_1^T y \\ \hline \vdots \\ \hline a_n^T y \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow a_i^T y = 0 \Rightarrow a_i \cdot y = 0$$

Por tanto, $a_i \perp y$ para todo $i = 1, ..., n \Rightarrow y \in R(A)^\perp \Rightarrow N(A^T) \subseteq R(A)^\perp$. Como $\dim N(A^T) = \dim R(A)^\perp$ se tiene finalmente: $N(A^T) = R(A)^\perp$. □

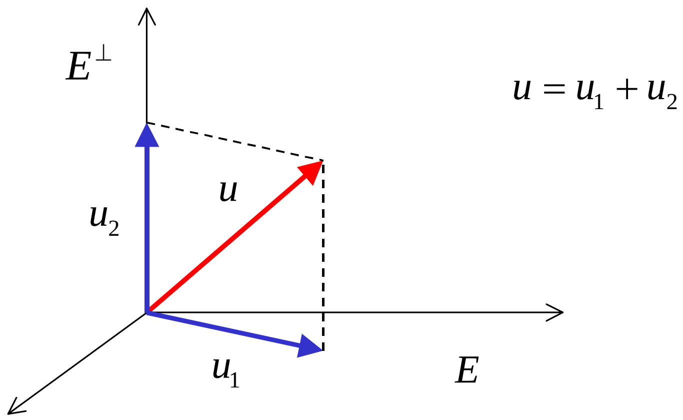
3.2. Proyección ortogonal y Método de mínimos cuadrados

El objetivo de esta sección es introducir el método de mínimos cuadrados para la resolución de sistemas de ecuaciones lineales. Este método permite dar una *"solucion aproximada"* a los sistemas de ecuaciones lineales que son incompatibles. El ajuste de una curva a una nube de puntos lo estudiaremos como una aplicación de este método. La idea clave en la que se basa la resolución de sistemas de ecuaciones por el método de mínimos cuadrados es la proyección ortogonal.

3.2.1. Proyección ortogonal

En $\mathbb{R}^3$ sabemos que si E es una recta o un plano que pasa por el origen, entonces todo vector $u \in \mathbb{R}^3$ se puede expresar como suma de dos vectores $u = u_1 + u_2$, donde $u_1 \in E$ y $u_2 \in E^\perp$.

En la figura se ha representado la descomposición de una vector u como suma de dos vectores, u_1 que está contenido en el plano $E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0\}$, y u_2 que está contenido en la recta $E^\perp = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = x_2 = 0\}$.



Esta situación se puede generalizar al espacio vectorial $\mathbb{R}^n$. Si E es un subespacio vectorial de $\mathbb{R}^n$, entonces todo vector $u \in \mathbb{R}^n$ se puede expresar de manera única como $u = u_1 + u_2$, donde $u_1 \in E$ y $u_2 \in E^\perp$. Se suele escribir $\mathbb{R}^n = E \oplus E^\perp$, indicando con ello que

$$\mathbb{R}^n = E + E^\perp \text{ y } E \cap E^\perp = \{\mathbf{0}\}.$$

El vector u_1 se denomina proyección ortogonal de u sobre E, y se escribe: $u_1 = \operatorname{proy}_E u$. Análogamente $u_2 = \operatorname{proy}_{E^\perp} u$. Es decir: $u = \operatorname{proy}_E u + \operatorname{proy}_{E^\perp} u$.

Además, si B_E es una base de E y $B_{E^\perp}$ es una base de $E^\perp$ se cumple que $E + E^\perp = L\{B_E \cup B_{E^\perp}\}$.

Ejemplo 128 *Expresar el vector* $u = (1,1,1)$ *como suma de dos vectores* $u_1 \in E$ *y* $u_2 \in E^\perp$, *donde* $E = L\{v_1, v_2\}$, *siendo* $v_1 = (2,1,1)$, $v_2 = (-3,0,1)\}$.

Se verifica que $\dim E = 2$, *luego* $\dim E^\perp = 1$.

Si $(x,y,z) \in E^\perp \Rightarrow \left\{ \begin{array}{c} (x,y,z)\cdot(2,1,1) = 0 \\ (x,y,z)\cdot(-3,0,1) = 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{l} 2x + y + z = 0 \\ -3x \qquad + z = 0 \end{array} \right. \Rightarrow \boxed{\begin{array}{l} x = \frac{1}{3}\alpha \\ y = -\frac{5}{3}\alpha \\ z = \alpha \end{array}}$

luego $E^\perp = L\{(1,-5,3)\} = L\{v_3\}$

Reuniendo las bases de E *y de* $E^\perp$ *se tiene una base de* $\mathbb{R}^3$

$$B = \{v_1, v_2, v_3\} = \{(2,1,1), (-3,0,1), (1,-5,3)\}.$$

A continuación calculamos las coordenadas del vector $u = (1,1,1)$ *en esta base:*

$$u = Pu_B, \text{ donde } P = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 0 & -5 \\ 1 & 1 & 3 \end{bmatrix} \text{ y } u_B = (\alpha_1, \alpha_2, \alpha_3)_B.$$

Resolviendo el sistema $Pu_B = u$, $\begin{bmatrix} 2 & -3 & 1 \\ 1 & 0 & -5 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \boxed{\begin{array}{l} \alpha_1 = \frac{6}{7} \\ \alpha_2 = \frac{8}{35} \\ \alpha_3 = -\frac{1}{35} \end{array}}$

se tiene: $u_B = (\frac{6}{7}, \frac{8}{35}, -\frac{1}{35})_B$

sustituyendo: $u_B = \underbrace{\frac{6}{7}(2,1,1) + \frac{8}{35}(-3,0,1)}_{u_1 \in E} \underbrace{- \frac{1}{35}(1,-5,3)}_{u_2 \in E^\perp}$

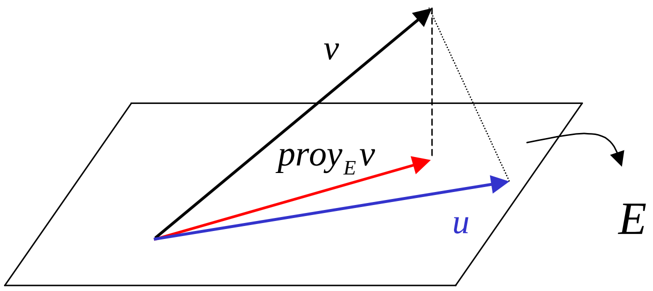
$$u_1 = \text{proy}_E\, u = \tfrac{6}{7}(2,1,1) + \tfrac{8}{35}(-3,0,1) = (\tfrac{36}{35}, \tfrac{6}{7}, \tfrac{38}{35}) \in E$$

$$u_2 = \text{proy}_{E^\perp}\, u = -\tfrac{1}{35}(1,-5,3) = (-\tfrac{1}{35}, \tfrac{1}{7}, -\tfrac{3}{35}) \in E^\perp$$

Nota 1: Si se conoce P^{-1}, $u_B = P^{-1}u = \begin{bmatrix} \frac{1}{7} & \frac{2}{7} & \frac{3}{7} \\ -\frac{8}{35} & \frac{1}{7} & \frac{11}{35} \\ \frac{1}{35} & -\frac{1}{7} & \frac{3}{35} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{6}{7} \\ \frac{8}{35} \\ -\frac{1}{35} \end{bmatrix}$

Nota 2: Comprobar que $u_1 + u_2 = u$, $u_1 \cdot u_2 = 0$.

Dado un vector $v \in \mathbb{R}^n$ y un subespacio E de $\mathbb{R}^n$, la proyección ortogonal de v sobre E es precisamente el vector de E más próximo a v. Esto se pone de manifiesto en el siguiente resultado.



Proposición 129 *Sea E un subespacio vectorial de $\mathbb{R}^n$, y $v \in \mathbb{R}^n$. Entonces, para todo vector $u \in E, u \neq \text{proy}_E\, v$ se verifica:* $\|v - \text{proy}_E\, v\| < \|v - u\|$.

Demostración: Sea $u \in E, u \neq \text{proy}_E\, v$, entonces:

$$v - u = \underbrace{(v - \text{proy}_E\, v)}_{\in E^\perp} + \underbrace{(\text{proy}_E\, v - u)}_{\in E}$$

Si aplicamos el teorema de Pitágoras:

$$\|v-u\|^2 = \|v - \text{proy}_E\, v\|^2 + \|\text{proy}_E\, v - u\|^2$$

Como $u \neq \text{proy}_E\, v$ se tiene que $\|\text{proy}_E\, v - u\|^2 > 0$, y en consecuencia

$$\|v-u\|^2 > \|v - \text{proy}_E\, v\|^2 \Rightarrow \|v-u\| > \|v - \text{proy}_E\, v\|$$

$\square$

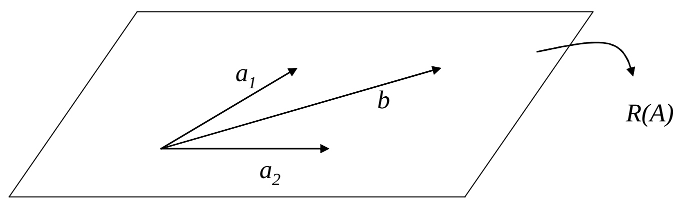
La proyección ortogonal de un vector sobre un subespacio puede calcularse de diversas maneras. A continuación describimos *el método de los mínimos cuadrados* para la resolución de sistemas de ecuaciones lineales, que nos proporciona un algoritmo para el cálculo de la proyección ortogonal, además de darnos una *"solución aproximada"* para los sistemas de ecuaciones lineales que sean incompatibles.

3.2.2. Método de los mínimos cuadrados

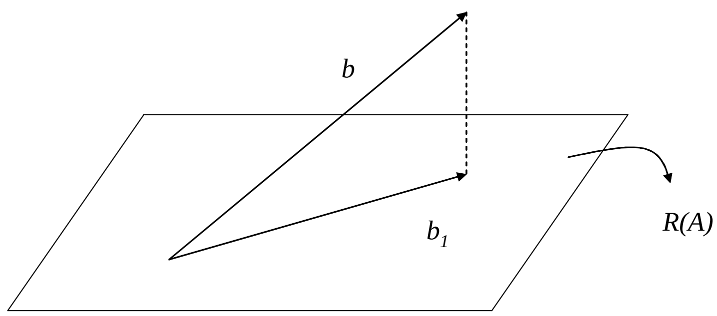
Para introducir este método consideraremos un sistema de ecuaciones lineales $Ax = b$. En relación con la compatibilidad del sistema pueden darse dos situaciones:

- Cuando el sistema $Ax = b$ es compatible y $b = \beta_1 a_1 + \cdots + \beta_n a_n$, siendo $\beta = (\beta_1, \ldots, \beta_n)$ una solución del sistema. Si denotamos por $R(A)$ el espacio columna de la matriz A, podemos escribir $b \in R(A)$.

  Geométricamente podemos considerar la situación de la siguiente figura.



- Si $Ax = b$ no es compatible, el sistema no tiene solución y entonces $b \notin R(A)$.



Ahora bien la proyección ortogonal de b sobre $R(A)$ sí pertenece a este subespacio:

$$b_1 = \mathrm{proy}_{R(A)}\, b \in R(A)$$

y podemos escribir:

$$b_1 = \alpha_1 a_1 + \cdots + \alpha_n a_n \Leftrightarrow b_1 = A\alpha.$$

Obviamente, el vector $\alpha = (\alpha_1, \ldots, \alpha_n)$ no es una solución del sistema $Ax = b$, pero sí podemos tomarla como una *solución aproximada*. Surge la pregunta, ¿cuál es el error que se comete al hacer esta aproximación? Como estimación del error vamos a considerar la norma del vector $\|b - b_1\|$.

Para determinar los coeficientes de la proyección ortogonal, es decir el vector $\alpha = (\alpha_1, \ldots, \alpha_n)$, procedemos del siguiente modo:

$$b_1 \in R(A) \Rightarrow b - b_1 \in R(A)^\perp \Rightarrow b - b_1 \in N(A^T) \Rightarrow A^T(b - b_1) = 0 \Rightarrow A^T b_1 = A^T b$$

Sustituyendo $b_1 = A\alpha$, resulta el sistema:

$$A^T A\alpha = A^T b$$

que se denomina *sistema normal asociado*. La solución de este sistema nos determina los coeficientes de la proyección ortogonal de b sobre el espacio columna $R(A)$. En general, se tiene el siguiente resultado.

Teorema 130 *Sea $(S) : Ax = b$ un sistema de ecuaciones lineales, la solución del sistema $(SN) : A^T Ax = A^T b$ determina los coeficientes de la proyección ortogonal de b sobre el espacio columna $R(A)$.*

Observaciones:

1) $b_1 = \text{proy}_{R(A)}\, b = Ax$, donde x es solución de (SN).

2) La solución del sistema (SN) se denomina *pseudosolución* o *solución en el sentido de los mínimos cuadrados* de (S).

3) El sistema (SN) es un sistema cuadrado con n incógnitas.

4) El sistema (SN) siempre es compatible, pues la proyección ortogonal siempre está determinada.

5) Si las columnas de A son linealmente independientes $(rg(A^T A) = n)$ la solución de (SN) es única.

6) Si las columnas de A son linealmente dependientes $(rg(A^T A) < n)$ la solución de (SN) no es única.

7) Si (S) es compatible, sus soluciones coinciden con las del sistema normal asociado (SN).

8) Si (S) es incompatible, podemos usar la solución del sistema (SN) como *solución aproximada* del sistema (S), siendo el error cometido $E = \|b - b_1\| = \|b - Ax\|$.

Ejemplo 131 *Resolver el sistema $(S): Ax = b$ en el sentido de los mínimos cuadrados. A continuación calcular la proyección ortogonal del vector b sobre el espacio columna de A.*

$$(S): Ax = b \qquad \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$$

$$(SN): A^T A x = A^T b \qquad \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix} \Rightarrow \boxed{\begin{array}{l} x_1 = \frac{5}{7} \\ x_2 = \frac{6}{7} \end{array}}$$

El sistema (S) es incompatible pues $rg(A|b) = 3$ y $rg(A) = 2$.

Por tanto, la solución de (SN) es una solución aproximada de (S).

La proyección ortogonal de b sobre $R(A)$ viene dada por:

$$b_1 = \operatorname{proy}_{R(A)} b = Ax = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} \frac{5}{7} \\ \frac{6}{7} \end{bmatrix} = \begin{bmatrix} \frac{11}{7} \\ -\frac{1}{7} \\ \frac{16}{7} \end{bmatrix}$$

finalmente, $b_1 = (\frac{11}{7}, -\frac{1}{7}, \frac{16}{7})$

Nota: El error cometido al tomar $x = (\frac{5}{7}, \frac{6}{7})$ como solución aproximada es:

$E = \|b - b_1\| = \frac{1}{7}\sqrt{14}$

Ejemplo 132 *Sea $E = L\{(1,1,2), (1,-1,1)\}$ y sea $u = (2,0,2)$, hallar $u_1 = \operatorname{proy}_E u$.*

Para calcular u_1 basta resolver el sistema $(SN): A^T A x = A^T u$

siendo $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 1 \end{bmatrix}$ $u = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$,

La solución del sistema (SN) es $x = \begin{bmatrix} \frac{5}{7} \\ \frac{6}{7} \end{bmatrix}$, pues se trata del mismo sistema del *ejemplo anterior. Por tanto,*

$$u_1 = \mathrm{proy}_E\, u = \mathrm{proy}_{R(A)}\, u = Ax = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} \frac{5}{7} \\ \frac{6}{7} \end{bmatrix} = \begin{bmatrix} \frac{11}{7} \\ -\frac{1}{7} \\ \frac{16}{7} \end{bmatrix}$$

es decir, $u_1 = (\frac{11}{7} - \frac{1}{7}, \frac{16}{7})$.

Ejemplo 133 *Resolver el sistema* (S) *con el método de mínimos cuadrados el sistema*

$$(S) : \begin{cases} x_1 + x_2 + x_3 + x_4 = 2 \\ x_1 + x_2 + x_3 + x_4 = 3 \\ x_1 + x_2 + x_3 + x_4 = 4 \end{cases}$$

De manera trivial se observa que el sistema (S) *es incompatible. Consideramos* (SN)*, sistema normal asociado.*

$$(S) : Ax = b \qquad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

$(SN) : A^T A x = A^T b$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 9 \\ 9 \\ 9 \\ 9 \end{bmatrix}$$

El sistema (SN) *es compatible indeterminado, y basta resolver la ecuación*

$$3x_1 + 3x_2 + 3x_3 + 3x_4 = 9.$$

cuya solución es: $\boxed{\begin{array}{l} x_1 = 3 - \alpha - \beta - \gamma \\ x_2 = \alpha \\ x_3 = \beta \\ x_4 = \gamma \end{array}}$ *con* α, β *y* $\gamma \in \mathbb{R}$.

Ejemplo 134 *Para el sistema $Ax = b$ del ejemplo anterior, calcula la proyección ortogonal del vector b sobre el espacio columna de A.*

Puesto que conocemos la pseudosolución del sistema $Ax = b$,

$$b_1 = \text{proy}_{R(A)}\, b = Ax = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3-\alpha-\beta-\gamma \\ \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

es decir, $b_1 = (3, 3, 3)$.

3.2.3. Ajuste de datos por el método de mínimos cuadrados

El método de los mínimos cuadrados es útil para resolver problemas de ajuste de curvas (rectas o parábolas, por ejemplo) a nubes de puntos. Estos problemas dan lugar a sistemas sobredeterminados que en general son incompatibles. A continuación consideramos dos ejemplos.

Ejemplo 135 *Dados los puntos $(-1,-1), (0,1), (1,2), (2,2)$. Ajustar a una recta por el método de los mínimos cuadrados. Repetir la operación ajustando a una parábola.*

a) <u>Ajuste a una recta</u>: $y = mx + n$.

Imponiendo que los puntos de la nube pertenezcan a dicha recta, obtenemos el sistema de ecuaciones

$$(S): \left\{ \begin{aligned} -m + n &= -1 \\ n &= 1 \\ m + n &= 2 \\ 2m + n &= 2 \end{aligned} \right. \qquad A\alpha = b, \text{ siendo } A = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}, \; b = \begin{bmatrix} -1 \\ 1 \\ 2 \\ 2 \end{bmatrix}$$

El sistema normal asociado es $(SN): A^T A\alpha = A^T b$

$$\begin{bmatrix} -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 2 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \end{bmatrix} \leftrightarrow \left\{ \begin{aligned} 6m + 2n &= 7 \\ 2m + 4n &= 4 \end{aligned} \right.$$

Resolviendo el sistema anterior, la pseudosolución es $m = 1, n = \frac{1}{2}$ y la recta de *mejor aproximación en el sentido de los mínimos cuadrados es* $y = x + \frac{1}{2}$.

El error cometido es: $E = \|b - b_1\|$

siendo $b_1 = \operatorname{proy}_{R(A)} b = A\alpha = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{3}{2} \\ \frac{5}{2} \end{bmatrix}$

sustituyendo $E = \|b - b_1\| = \left\| \begin{bmatrix} -1 \\ 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{3}{2} \\ \frac{5}{2} \end{bmatrix} \right\| = \left\| \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} \right\| = 1$

Nota: La recta obtenida es, precisamente, la que minimiza la suma de los cuadrados de los errores que se cometen al estimar la ordenada de un punto.

$$E^2 = \sum_{i=1}^{4} \left| y_i - (x_i + \tfrac{1}{2}) \right|^2 = \varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 + \varepsilon_4^2 = \tfrac{1}{4} + \tfrac{1}{4} + \tfrac{1}{4} + \tfrac{1}{4} = 1 \Rightarrow E = 1$$

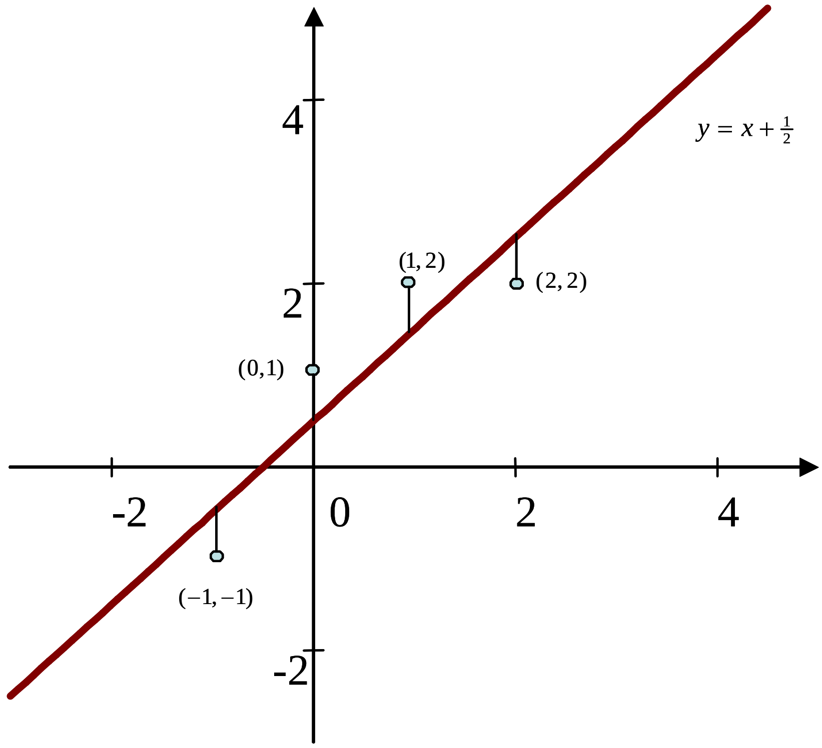
$$\varepsilon_1 = y_1 - (x_1 + \tfrac{1}{2}) = (-1) - (-1 + \tfrac{1}{2}) = -\tfrac{1}{2}$$

$$\varepsilon_2 = y_2 - (x_2 + \tfrac{1}{2}) = 1 - (0 + \tfrac{1}{2}) = \tfrac{1}{2}$$

$$\varepsilon_3 = y_3 - (x_3 + \tfrac{1}{2}) = 2 - (1 + \tfrac{1}{2}) = \tfrac{1}{2}$$

$$\varepsilon_4 = y_4 - (x_4 + \tfrac{1}{2}) = 2 - (2 + \tfrac{1}{2}) = -\tfrac{1}{2}$$

En general, para una nube de p puntos: $E^2 = \sum_{i=1}^{p} |y_i - (mx_i + n)|^2 = \|b - b_1\|^2 \Rightarrow E = \|b - b_1\|$.



b) Ajuste a una parábola: $y = mx^2 + nx + p$

Imponiendo que la parábola pase por todos los puntos de la nube, se obtiene el sistema de ecuaciones

$$(S): \left\{ \begin{aligned} m - n + p &= -1 \\ p &= 1 \\ m + n + p &= 2 \\ 4m + 2n + p &= 2 \end{aligned} \right.$$

$$(S): \ A\alpha = b, \ \textit{siendo}\ A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \ y \ \ b = \begin{bmatrix} -1 \\ 1 \\ 2 \\ 2 \end{bmatrix}$$

El sistema normal asociado es $(SN): A^T A\alpha = A^T b$

$$\begin{bmatrix} 1 & 0 & 1 & 4 \\ -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} m \\ n \\ p \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 4 \\ -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 2 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{bmatrix} \begin{bmatrix} m \\ n \\ p \end{bmatrix} = \begin{bmatrix} 9 \\ 7 \\ 4 \end{bmatrix}, \text{ cuya solución es: } \boxed{\begin{array}{l} m = -\frac{1}{2} \\ n = \frac{3}{2} \\ p = 1 \end{array}}$$

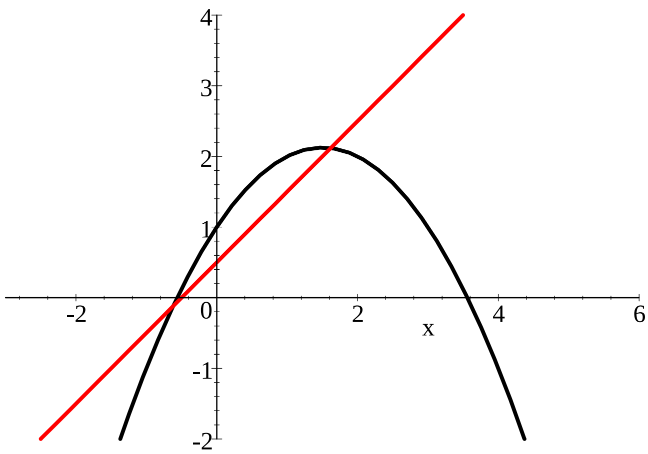
Luego la parábola de mejor ajuste es $y = -\frac{1}{2}x^2 + \frac{3}{2}x + 1$.

El error cometido es: $E = \|b - b_1\|$

siendo $b_1 = \operatorname{proy}_{R(A)} b = A\alpha = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ \frac{3}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \\ 2 \end{bmatrix} = b$

sustituyendo $E = \|b - b_1\| = 0$.

La recta y parábola obtenidas se representan en la siguiente figura. Obsérvese que en este caso, la parábola pasa por todos los puntos.



Ejemplo 136 *Cinco niños tienen las edades (años) y pesos (kg.) siguientes:*

Edad	2	3	5	7	8
Peso	10	26	30	40	44

Ajustar una recta a estos datos mediante el método de mínimos cuadrados, y estimar, haciendo uso de ella, el peso para niños de 4 y 10 años.

Sea: $y = mx + n$

$$(S) : A\alpha = b \qquad \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 5 & 1 \\ 7 & 1 \\ 8 & 1 \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} 10 \\ 26 \\ 30 \\ 40 \\ 44 \end{bmatrix}$$

$(SN) : A^T A\alpha = A^T b$

$$\begin{bmatrix} 2 & 3 & 5 & 7 & 8 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 5 & 1 \\ 7 & 1 \\ 8 & 1 \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} 2 & 3 & 5 & 7 & 8 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 10 \\ 26 \\ 30 \\ 40 \\ 44 \end{bmatrix}$$

$$\begin{bmatrix} 151 & 25 \\ 25 & 5 \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} 880 \\ 150 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 151 & 25 & 880 \\ 25 & 5 & 150 \end{array}\right] \underset{F_{21}(-\frac{25}{151})}{\sim} \left[\begin{array}{cc|c} 151 & 25 & 880 \\ 0 & \frac{130}{151} & \frac{650}{151} \end{array}\right] \underset{F_{21}(151)}{\sim} \left[\begin{array}{cc|c} 151 & 25 & 880 \\ 0 & 130 & 650 \end{array}\right]$$

La solución que se obtiene es $\boxed{\begin{matrix} m = 5 \\ n = 5 \end{matrix}}$. *Luego la recta pedida es:* $y = 5x + 5$.

A una edad de 4 años le corresponde un peso de 25 *Kg., y a una edad de 10 años le corresponde un peso de 55 Kg.*

3.3. Bases ortonormales

En $\mathbb{R}^n$ determinados cálculos se simplifican cuando usamos la base canónica, que es una base caracterizada porque su vectores son unitarios y ortogonales entre sí. Dado un subespacio E de $\mathbb{R}^n$, vamos a ver que para dicho subespacio se puede encontrar bases con propiedades similares a las de la base canónica; es decir, podemos mostrar que E contiene al menos una base B cuyos vectores son de norma uno y ortogonales dos a dos. El procedimiento para obtener dicha base es el *Método de ortonormalización de Gram-Schmid*t que describiremos más adelante.

3.3.1. Definición y propiedades

Definición 137 *a) Un conjunto de vectores de* $\mathbb{R}^n$ *se llama conjunto ortogonal si sus vectores son ortogonales dos a dos.*
b) Un conjunto ortogonal en el que cada vector tiene norma 1 *se denomina conjunto ortonormal.*

Ejemplo 138 *a) La base canónica de* $\mathbb{R}^n$ *es un conjunto ortonormal, pues* $e_i \cdot e_j = 0$ *y* $\|e_i\| = 1$ *para cada* i, j.

b) $H = \{u_1, u_2, u_3\}$, *con* $u_1 = (0,1,0), u_2 = (1,0,1)$ *y* $u_3 = (1,0,-1)$ *es ortogonal pues* $u_1 \cdot u_2 = u_2 \cdot u_3 = u_3 \cdot u_1 = 0$.

A partir de H *se puede obtener un conjunto o sistema ortonormal dividiendo cada vector por su norma:*

$$v_1 = \frac{u_1}{\|u_1\|} = (0,1,0) \quad v_2 = \frac{u_2}{\|u_2\|} = (\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}) \quad v_3 = \frac{u_3}{\|u_3\|} = (\tfrac{1}{\sqrt{2}}, 0, -\tfrac{1}{\sqrt{2}})$$

El conjunto $H' = \{v_1, v_2, v_3\}$ *es ortonormal.*

Teorema 139 *Todo conjunto* $H = \{u_1, ..., u_p\}$ *ortogonal de vectores no nulos es linealmente independiente.*

Demostración: Hemos de probar que la ecuación $h_1u_1 + \cdots + h_pu_p = \mathbf{0}$ se verifica si y sólo si $h_i = 0$ para todo $i = 1, ...p$. Puesto que H es ortogonal se verifica que: $u_i \cdot u_j = 0$ para cada i, j. Entonces,

$$h_1u_1 + \cdots + h_pu_p = \mathbf{0} \Rightarrow u_i \cdot (h_1u_1 + \cdots + h_pu_p) = 0 \Rightarrow h_i(u_i \cdot u_i) = 0 \Rightarrow$$

$$h_i \|u_i\|^2 = 0 \underset{\|u_i\| \neq 0}{\Rightarrow} h_i = 0 \text{ para todo } i = 1, ...p. \Rightarrow S \text{ es un conjunto l.i.} \qquad \square$$

Corolario 140 *Todo conjunto* $H = \{u_1, ..., u_p\}$ *ortonormal de vectores es linealmente independiente.*

Definición 141 *Una base* B *de un subespacio vectorial de* $\mathbb{R}^n$ *es ortogonal (ortonormal) si* B *es un conjunto ortogonal (ortonormal).*

Ejemplo 142 *a)* $B = \{(2,0), (0,2)\}$ *es una base ortogonal de* $\mathbb{R}^2$.
b) $B' = \{(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}), (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})\}$ *es un base ortonormal de* $\mathbb{R}^2$.

Teorema 143 *Si* $B = \{v_1, ..., v_p\}$ *es una base ortonormal de un subespacio vectorial* S *de* $\mathbb{R}^n$ *entonces para todo vector* $u \in S$, *se verifica que* $u = (\alpha_1, ..., \alpha_p)_B$, *siendo* $\alpha_i = u \cdot v_i$ *para* $i = 1, ..., p$.

Demostración: Si $u = (\alpha_1, ..., \alpha_p)_B$, sabemos que $u = \alpha_1v_1 + \cdots + \alpha_pv_p$. Veamos que $\alpha_i = u \cdot v_i$.

$$u \cdot v_i = (\alpha_1v_1 + \cdots + \alpha_pv_p) \cdot v_i = \alpha_1(v_1 \cdot v_i) + \cdots + \alpha_p(v_p \cdot v_i) = \alpha_i(v_i \cdot v_i) = \alpha_i$$

pues $v_i \cdot v_j = 0$ si $i \neq j$ y $v_i \cdot v_i = 1$. $\square$

3.3.2. Bases ortonormales y proyección ortogonal

Cuando se conoce una base ortonormal de un subespacio vectorial S se puede calcular fácilmente la proyección ortogonal de un vector u sobre dicho subespacio.

Sea $u \in \mathbb{R}^n$ y $S = L(B)$, siendo $B = \{w_1, ..., w_p\}$ una base ortonormal de S. Se tiene que $u = u_1 + u_2$, siendo $u_1 = \operatorname{proy}_S u$ y $u_2 = \operatorname{proy}_{S^\perp} u$, entonces:

$$u_1 = \operatorname{proy}_S u = \alpha_1 w_1 + \cdots + \alpha_p w_p, \text{ donde } \alpha_i = u \cdot w_i \text{ para cada } i = 1, ..., p.$$

La proyección sobre el ortogonal se calcularía por diferencia: $u_2 = \operatorname{proy}_{S^\perp} u = u - u_1$.

Nota: si $\{v_1, ..., v_p\}$ es una base ortogonal de E se tiene que $\{\frac{v_1}{\|v_1\|}, ..., \frac{v_p}{\|v_p\|}\}$ es una base ortonormal, y entonces:

$$u_1 = \operatorname{proy}_S u = \alpha_1 \tfrac{v_1}{\|v_1\|} + \cdots + \alpha_p \tfrac{v_1}{\|v_1\|}, \text{ donde } \alpha_i = u \cdot \tfrac{v_i}{\|v_i\|} \text{ para cada } i = 1, ..., p.$$

y sustituyendo se tiene:

$$u_1 = \operatorname{proy}_S u = (u \cdot \tfrac{v_1}{\|v_1\|})\tfrac{v_1}{\|v_1\|} + \cdots + (u \cdot \tfrac{v_1}{\|v_1\|})\tfrac{v_1}{\|v_1\|} = \frac{u \cdot v_1}{v_1 \cdot v_1} v_1 + \cdots + \frac{u \cdot v_p}{v_p \cdot v_p} v_p$$

Ejemplo 144 *Dado el subespacio* $S = L\left\{(0,1,0), (-\frac{4}{5}, 0, \frac{3}{5})\right\}$, *calcular la proyección del vector* $u = (1,1,1)$ *sobre* $S^\perp$.

Se tiene que $u = u_1 + u_2$, *siendo* $u_1 = \operatorname{proy}_S u, \quad u_2 = \operatorname{proy}_{S^\perp} u$.

Por otra parte, S *viene determinado por una base ortonormal:*

$v_1 = (0,1,0), v_2 = (-\frac{4}{5}, 0, \frac{3}{5}) \qquad v_1 \cdot v_2 = 0, \quad \|v_1\| = \|v_2\| = 1$

Se verifica que:

$u_1 = \operatorname{proy}_S u = \alpha_1 v_1 + \alpha_2 v_2 = v_1 - \frac{1}{5} v_2 = (0,1,0) - \frac{1}{5}(-\frac{4}{5}, 0, \frac{3}{5}) = (\frac{4}{25}, 1, -\frac{3}{25})$

$\alpha_1 = u \cdot v_1 = (1,1,1) \cdot (0,1,0) = 1$

$\alpha_2 = u \cdot v_2 = (1,1,1) \cdot (-\frac{4}{5}, 0, \frac{3}{5}) = -\frac{1}{5}$

Para calcular $u_2 = \operatorname{proy}_{S^\perp} u$, *consideramos que*

$u_2 = u - u_1 = (1,1,1) - (\frac{4}{25}, 1, -\frac{3}{25}) = (\frac{21}{25}, 0, \frac{28}{25})$

Ejemplo 145 *Calcular la proyección ortogonal del vector* $b = (2, 3, 4)$ *sobre el subespacio* $S = L\{(1, 1, 1)\}$.

Para determinar el subespacio S *mediante una base ortonormal basta normalizar el vector* $(1, 1, 1)$*:* $S = L\{(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3})\} = L\{z\}$.

Aplicando la proposición anterior, se tiene que:

$b_1 = \text{proy}_S\, b = \alpha_1 z$, *siendo* $\alpha_1 = b \cdot z = (2, 3, 4) \cdot (\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}) = 3\sqrt{3}$

es decir, $b_1 = \text{proy}_S\, b = \alpha_1 z = 3\sqrt{3}(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}) = (3, 3, 3)$.

Otro resultado que simplifica el cálculo de la proyección ortogonal cuando se trabaja con una base ortonormal es el siguiente.

Proposición 146 *Sea* $S = L(B)$*, siendo* $B = \{w_1, ..., w_p\}$*, una base ortonormal, y sea* $u \in \mathbb{R}^n$ *entonces* $u_1 = \text{proy}_S\, u = AA^T u$*, siendo* $A = [w_1|\cdots|w_p]$.

Demostración: $u_1 = \text{proy}_S\, u = \text{proy}_{R(A)}\, u = Ax$. Si B es una base ortonormal, entonces $A^T A = I$, y $(SN) : A^T Ax = A^T u \sim x = A^T u$, con lo que $u_1 = \text{proy}_S\, u = Ax = AA^T u$. □

Ejemplo 147 *Dado el subespacio* $S = L\{(0, 1, 0), (-\frac{4}{5}, 0, \frac{3}{5})\}$*, calcular la proyección del vector* $u = (1, 1, 1)$ *sobre* $S^\perp$.

Se tiene que $u = u_1 + u_2$*, siendo* $u_1 = \text{proy}_S\ u$*,* $\quad u_2 = \text{proy}_{S^\perp}\ u$*. Por otra parte,* S *viene determinado por una base ortonormal:*

$v_1 = (0, 1, 0), v_2 = (-\frac{4}{5}, 0, \frac{3}{5})$*, y se cumple:* $v_1 \cdot v_2 = 0, \quad \|v_1\| = \|v_2\| = 1$

Se verifica que: $u_1 = \text{proy}_S\ u = AA^T u$*, siendo* $A = \begin{bmatrix} 0 & -\frac{4}{5} \\ 1 & 0 \\ 0 & \frac{3}{5} \end{bmatrix}$

sustituyendo:

$$u_1 = \text{proy}_S\, u = AA^T u = \begin{bmatrix} 0 & -\frac{4}{5} \\ 1 & 0 \\ 0 & \frac{3}{5} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -\frac{4}{5} & 0 & \frac{3}{5} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{4}{25} \\ 1 \\ -\frac{3}{25} \end{bmatrix}$$

$u_1 = \text{proy}_S\, u = (\frac{4}{25}, 1, -\frac{3}{25})$

finalmente,

$u_2 = \text{proy}_{S^\perp}\ u = u - u_1 = (1, 1, 1) - (\frac{4}{25}, 1, -\frac{3}{25}) = (\frac{21}{25}, 0, \frac{28}{25})$.

A continuación exponemos el Método de Gram-Schmidt para obtener una base ortonormal de un subespacio a partir de una base cualquiera del mismo.

3.3.3. Método de ortonormalización de Gram-Schmidt

Es un método que permite obtener una base ortonormal de un subespacio S de $\mathbb{R}^n$ a partir de una base cualquiera $B = \{u_1, ..., u_p\}$ de dicho subespacio. El proceso de cálculo se divide en dos etapas: en la primera se halla una base ortogonal de $S : B' = \{v_1, ..., v_p\}$, y a continuación, diviendo cada vector por v_i por su norma $w_i = \frac{v_i}{\|v_i\|}$, se obtiene la base ortonormal buscada $B'' = \{w_1, ..., w_p\}$.

Ejemplo 148 *Sea $B = \{(1,0),(2,2)\}$. A partir de B hallar una base ortonormal de $\mathbb{R}^2$ con el método de Gram-Schmidt.*

1) Cálculo de la base ortogonal: $B' = \{v_1, v_2\}$.

$$\begin{aligned} &v_1 = u_1 = (1,0) \\ &v_2 = u_2 + \alpha v_1 = (2,2) + \alpha(1,0) \end{aligned}$$

imponemos la condición de que v_1 y v_2 sean ortogonales: $v_2 \cdot v_1 = 0$

$$\begin{aligned} &v_2 \cdot v_1 = (u_2 + \alpha v_1) \cdot v_1 = u_2 \cdot v_1 + \alpha(v_1 \cdot v_1) = 0 \Rightarrow \alpha = -\frac{u_2 \cdot v_1}{v_1 \cdot v_1} \\ &u_2 \cdot v_1 = (2,2) \cdot (1,0) = 2 \\ &v_1 \cdot v_1 = (1,0) \cdot (1,0) = 1 \end{aligned}$$

sustituyendo: $\alpha = -\dfrac{2}{1} = -2$, luego $v_2 = u_2 - 2v_1 = (2,2) - 2(1,0) = (0,2)$

así $B' = \{v_1, v_2\}$, con $v_1 = (1,0)$, $v_2 = (0,2)$, es una base ortogonal.

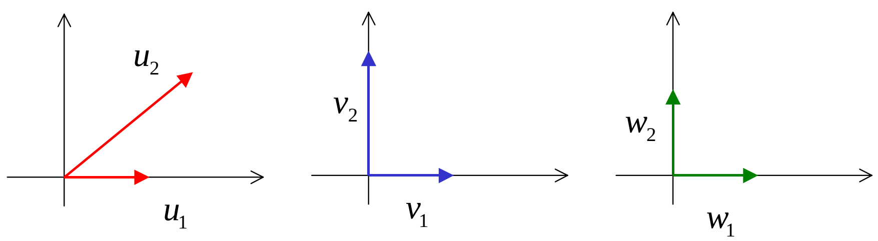
Obsérvese que $v_2 = u_2 - \mathrm{proy}_{L\{v_1\}} u_2$. Luego, $v_2 \in (L\{v_1\})^\perp$.

2) Cálculo de la base ortonormal: $B'' = \{w_1, w_2\}$.

Basta dividir cada vector v_i por su norma:

$$w_1 = \tfrac{v_1}{\|v_1\|} = (1,0), \quad w_1 = \tfrac{v_2}{\|v_2\|} = (0,1)$$

A continuación se resume el proceso seguido, obsérvese que para el cálculo del vector v_2 lo que se ha hecho es restarle al vector u_2 su proyección sobre el vector v_1.



$$B = \{u_1, u_2\} \quad \longrightarrow \quad B' = \{v_1, v_2\} \quad \longrightarrow \quad B'' = \{w_1, w_2\}$$

Ejemplo 149 *Sea $W = \{(x, y, z)|x + y + z = 0\}$, un subespacio vectorial de $\mathbb{R}^3$. Determinar una base ortonormal de dicho subespacio.*

En primer lugar resolvemos el sistema: $x + y + z = 0 \Rightarrow \begin{cases} x = -\alpha - \beta \\ y = \alpha \\ z = \beta \end{cases}$

lo que permite determinar W con un sistema generador.

$$W = L\{(-1, 1, 0), (-1, 0, 1)\} = L\{u_1, u_2\}$$

Los vectores u_1 y u_2 son l.i., por tanto $B = \{u_1, u_2\}$ es una base de W.

Buscamos primero una base ortogonal: $B'_W = \{v_1, v_2\}$

$$v_1 = u_1 \Rightarrow v_1 = (-1, 1, 0)$$

$$v_2 = u_2 + \lambda v_1$$

imponemos que $v_1 \perp v_2 \Rightarrow v_1 \cdot v_2 = 0 \Rightarrow v_1 \cdot (u_2 + \lambda v_1) = 0 \Rightarrow \lambda = -\dfrac{u_2 \cdot v_1}{v_1 \cdot v_1}$

$$\left.\begin{array}{l} v_1 \cdot v_1 = (-1, 1, 0) \cdot (-1, 1, 0) = 2 \\ u_2 \cdot v_1 = (-1, 0, 1) \cdot (-1, 1, 0) = 1 \end{array}\right. \Rightarrow \lambda = -\tfrac{1}{2}$$

$$v_2 = u_2 - \frac{1}{2}v_1 = (-1, 0, 1) - \frac{1}{2}(-1, 1, 0) = (-\tfrac{1}{2}, -\tfrac{1}{2}, 1)$$

luego $B_W = \{(-1, 1, 0), (-\tfrac{1}{2}, -\tfrac{1}{2}, 1)\}$ *es la base ortogonal.*

Nota: Obsérvese que $v_2 = u_2 - \operatorname{proy}_{L\{v_1\}} u_2$. *Es decir,* $v_2 \in (L\{v_1\})^{\perp}$.

Para hallar una base ortonormal $B''_W = \{z_1, z_2\}$, *basta dividir los vectores v_1 y v_2 por su norma.*

$$z_1 = \frac{v_1}{\|v_1\|} = \frac{(-1, 1, 0)}{\sqrt{2}} = (-\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0) = (-\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}, 0)$$

$$\|v_1\| = \sqrt{v_1 \cdot v_1} = \sqrt{2}$$

$$z_2 = \frac{v_2}{\|v_2\|} = \frac{(-\frac{1}{2}, -\frac{1}{2}, 1)}{\sqrt{\frac{3}{2}}} = \sqrt{\tfrac{2}{3}}(-\tfrac{1}{2}, -\tfrac{1}{2}, 1) = \tfrac{\sqrt{6}}{3}(-\tfrac{1}{2}, -\tfrac{1}{2}, 1) = (-\tfrac{\sqrt{6}}{6}, -\tfrac{\sqrt{6}}{6}, \tfrac{\sqrt{6}}{3})$$

$$v_2 \cdot v_2 = (-\tfrac{1}{2}, -\tfrac{1}{2}, 1) \cdot (-\tfrac{1}{2}, -\tfrac{1}{2}, 1) = \tfrac{1}{4} + \tfrac{1}{4} + 1 = \tfrac{3}{2}$$

$$\|v_2\| = \sqrt{v_2 \cdot v_2} = \sqrt{\tfrac{3}{2}}$$

$B''_W = \{(-\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}, 0), (-\tfrac{\sqrt{6}}{6}, -\tfrac{\sqrt{6}}{6}, \tfrac{\sqrt{6}}{3})\}$ *es una base ortonormal de W.*

A continuación se resume el método de ortonormalización de Gram-Schmidt para obtención de una base ortonormal en un subespacio S de $\mathbb{R}^n$.

Método de Gram-Schmidt Sea $S = L(B)$ un subespacio vectorial de $\mathbb{R}^n$, determinado por la base $B = \{u_1, ..., u_p\}$.

1) Se calcula una base ortogonal de S: $B' = \{v_1, ..., v_p\}$

$$v_1 = u_1$$

$$v_i = u_i + \alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1}, \quad i = 2, ..., p$$

$$\text{donde } \alpha_i = -\frac{u_i \cdot v_j}{v_j \cdot v_j}, \quad j = 1, ..., i-1$$

2) Se halla la base ortonormal de S: $B'' = \{z_1, ..., z_p\}$, siendo $z_i = \dfrac{v_i}{\|v_i\|}, i = 1, ..., p$.

Nota:

a) Obsérvese que $v_i = u_i - \text{proy}_{L\{v_1,...,v_{i-1}\}} u_i$. Por tanto, $v_i \in (L\{v_1, ..., v_{i-1}\})^{\perp}$.

b) Puesto que en cada etapa de cálculo de los v_i o de los z_i se han realizado transformaciones elementales. con los vectores, resulta que:

$$L\{u_1\} = L\{v_1\} = L\{z_1\}$$

$$L\{u_1, ..., u_k\} = L\{v_1, ..., v_k\} = L\{z_1, ..., z_k\} \text{ para } k = 2, ..., p$$

Ejemplo 150 *Sea $W = L\{u_1, u_2, u_3\}$ un subespacio vectorial de $\mathbb{R}^4$, donde $u_1 = (1, -2, 0, 1)$, $u_2 = (-1, 0, 0, -1)$ y $u_3 = (1, 1, 0, 0)$. Hallar una base ortonormal de W.*
Buscamos primero una base ortogonal: $B'_W = \{v_1, v_2, v_3\}$

$$v_1 = u_1 \Rightarrow v_1 = (1, -2, 0, 1)$$

$$v_2 = u_2 + \alpha_1 v_1 = u_2 - \frac{u_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$v_1 \cdot v_1 = (1, -2, 0, 1) \cdot (1, -2, 0, 1) = 6$$

$$u_2 \cdot v_1 = (-1, 0, 0, -1) \cdot (1, -2, 0, 1) = -2$$

luego $v_2 = u_2 - (\frac{-2}{6})v_1 = (-1, 0, 0, -1) + \dfrac{1}{3}(1, -2, 0, 1) = (-\frac{2}{3}, -\frac{2}{3}, 0, -\frac{2}{3})$

$$v_3 = u_3 + \alpha_1 v_1 + \alpha_2 v_2 = u_3 - \frac{u_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{u_3 \cdot v_2}{v_2 \cdot v_2} v_2$$

$$u_3 \cdot v_1 = (1,1,0,0) \cdot (1,-2,0,1) = -1$$

$$v_2 \cdot v_2 = (-\tfrac{2}{3}, -\tfrac{2}{3}, 0, -\tfrac{2}{3}) \cdot (-\tfrac{2}{3}, -\tfrac{2}{3}, 0, -\tfrac{2}{3}) = \tfrac{4}{3}$$

$$u_3 \cdot v_2 = (1,1,0,0) \cdot (-\tfrac{2}{3}, -\tfrac{2}{3}, 0, -\tfrac{2}{3}) = -\tfrac{4}{3}$$

luego $v_3 = u_3 + \frac{1}{6}v_1 + v_2$

$v_3 = (1,1,0,0) + \frac{1}{6}(1,-2,0,1) + (-\frac{2}{3}, -\frac{2}{3}, 0, -\frac{2}{3}) = (\frac{1}{2}, 0, 0, -\frac{1}{2})$

$B'_W = \{(1,-2,0,1), (-\frac{2}{3}, -\frac{2}{3}, 0, -\frac{2}{3}), (\frac{1}{2}, 0, 0, -\frac{1}{2})\}$

Para hallar una base ortonormal $B''_W = \{z_1, z_2, z_3\}$, *basta dividir los vectores* v_1,v_2 *y* v_3 *por su norma.*

$$z_1 = \frac{v_1}{\|v_1\|} = \frac{(1,-2,0,1)}{\sqrt{6}} = (\tfrac{\sqrt{6}}{6}, -\tfrac{\sqrt{6}}{3}, 0, \tfrac{\sqrt{6}}{6})$$

$$\|v_1\| = \sqrt{v_1 \cdot v_1} = \sqrt{6}$$

$$z_2 = \frac{v_2}{\|v_2\|} = \frac{(-\frac{2}{3}, -\frac{2}{3}, 0, -\frac{2}{3})}{\sqrt{\frac{4}{3}}} = (-\tfrac{\sqrt{3}}{3}, -\tfrac{\sqrt{3}}{3}, 0, -\tfrac{\sqrt{3}}{3})$$

$$\|v_2\| = \sqrt{v_2 \cdot v_2} = \sqrt{\tfrac{4}{3}}$$

$$z_3 = \frac{v_3}{\|v_3\|} = \frac{(\frac{1}{2}, 0, 0, -\frac{1}{2})}{\sqrt{\frac{1}{2}}} = (\tfrac{\sqrt{2}}{2}, 0, 0, -\tfrac{\sqrt{2}}{2})$$

$$v_3 \cdot v_3 = (\tfrac{1}{2}, 0, 0, -\tfrac{1}{2}) \cdot (\tfrac{1}{2}, 0, 0, -\tfrac{1}{2}) = \tfrac{1}{2}$$
$$\|v_3\| = \sqrt{v_3 \cdot v_3} = \sqrt{\tfrac{1}{2}}$$

$B''_W = \{(\frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{3}, 0, \frac{\sqrt{6}}{6}), (-\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}, 0, -\frac{\sqrt{3}}{3}), (\frac{\sqrt{2}}{2}, 0, 0, -\frac{\sqrt{2}}{2})\}$

es la base ortonormal buscada.

3.3.4. Matrices ortogonales

Sea $B = \{v_1, v_2, v_3\}$ una base ortonormal de $\mathbb{R}^3$, con $v_1 = (0,1,0)$, $v_2 = (-\frac{4}{5}, 0, \frac{3}{5})$, $v_3 = (\frac{3}{5}, 0, \frac{4}{5})$. ¿Cómo es la matriz de cambio de base B a la base canónica?

La expresión del cambio de base sería: $x = Px_B \Rightarrow P = [v_1|v_2|v_3] = \begin{bmatrix} 0 & -\frac{4}{5} & \frac{3}{5} \\ 1 & 0 & 0 \\ 0 & \frac{3}{5} & \frac{4}{5} \end{bmatrix}$

Observamos que, en este caso, la matriz del cambio de base verifica que $P^{-1} = P^T$. Esto sucede siempre que planteamos un cambio de base entre bases ortonormales. A las matrices que tienen esta propiedad se denominan *matrices ortogonales.*

Definición 151 *Una matriz $Q \in \mathcal{M}_n(\mathbb{R})$ es ortogonal si $QQ^T = Q^TQ = I$.*

Observaciones:

a) Si $Q \in \mathcal{M}_n(\mathbb{R})$ es ortogonal entonces: $Q^{-1} = Q^T$

b) Las columnas de una matriz ortogonal Q de orden n constituyen una base ortonormal de $\mathbb{R}^n$.