Consideramos el espacio vectorial $\mathbb{R}^n$ con el producto escalar canónico.

1 Bases ortogonales y ortonormales. Matriz ortogonal

Un conjunto de vectores $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ de $\mathbb{R}^n$ se dice que es un **conjunto ortogonal** si cada par de vectores distintos del conjunto es ortogonal, es decir, si $\vec{v}_i \cdot \vec{v}_j = 0\ \forall\ i \neq j$.

Un conjunto de vectores $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ de $\mathbb{R}^n$ es un **conjunto ortonormal** si es un conjunto ortogonal de vectores unitarios.

$$\vec{v}_i \cdot \vec{v}_j = \delta_{ij} = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

TEOREMA. Si $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ es un conjunto ortogonal de vectores no nulos de $\mathbb{R}^n$, entonces S es un conjunto linealmente independiente y por tanto es una base del subespacio generado por S.

Una **base ortogonal** de un subespacio W de $\mathbb{R}^n$ es una base de W que es además conjunto ortogonal.

Una **base ortonormal** de un subespacio W de $\mathbb{R}^n$, es una base de W que es además conjunto ortonormal.

TEOREMA. De cualquier subespacio W de $\mathbb{R}^n$ se puede obtener una base ortogonal, y mediante normalización de sus vectores, una base ortonormal. La excepción es obviamente el subespacio cero.

TEOREMA. La base canónica de $\mathbb{R}^n$ es base ortonormal.

Ejemplo 1. *Muestra que* $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ *es un conjunto ortogonal de* $\mathbb{R}^3$*, siendo*

$$\vec{v}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

Consideremos los tres posibles pares de vectores distintos, a saber $\{\vec{v}_1, \vec{v}_2\}$,$\{\vec{v}_1, \vec{v}_3\}$ *y* $\{\vec{v}_2, \vec{v}_3\}$.

$\vec{v}_1 \cdot \vec{v}_2 = 3(-1) + 1(2) + 1(1) = 0$
$\vec{v}_1 \cdot \vec{v}_3 = 3(-1/2) + 1(-2) + 1(7/2) = 0$
$\vec{v}_2 \cdot \vec{v}_3 = -1(-1/2) + 2(-2) + 1(7/2) = 0$

Cada par de vectores distintos es ortogonal, por tanto $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ *es un conjunto ortogonal.*

Ejemplo 2. *Muestra que* $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ *es una base ortonormal de* $\mathbb{R}^3$*, siendo:*

$$\vec{u}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \vec{u}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

$$\vec{u}_1 \cdot \vec{u}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 & 1 & 1 \end{bmatrix} \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = \frac{(9+1+1)}{11} = \frac{11}{11}$$

$$\vec{u}_2 \cdot \vec{u}_2 = \frac{(1+4+1)}{6} = \frac{6}{6} = 1$$

$$\vec{u}_3 \cdot \vec{u}_3 = \frac{(1+16+49)}{66} = \frac{66}{6} = 1$$

$$\vec{u}_1 \cdot \vec{u}_2 = \frac{(-3+2+1)}{\sqrt{11}\sqrt{6}} = 0$$

$$\vec{u}_1 \cdot \vec{u}_3 = \frac{(-3-4+7)}{\sqrt{11}\sqrt{66}} = 0$$

$$\vec{u}_2 \cdot \vec{u}_3 = \frac{(1-8+7)}{\sqrt{11}\sqrt{66}} = 0$$

El conjunto $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ *es un conjunto ortonormal, luego es linealmente independiente, y añadido que tiene rango 3, resulta ser base de* $\mathbb{R}^3$. *Por ser un conjunto ortonormal es una base ortonormal.*

Cuando los vectores de un conjunto ortogonal se "normalizan" para tener longitud unidad, el nuevo conjunto sigue siendo ortogonal, y por tanto, al tener longitud unidad, forma un conjunto ortonormal.

Estos son de hecho los mismos vectores del ejemplo anterior, que como habíamos comprobado son ortogonales, pero ahora normalizados.

TEOREMA. Una matriz cuadrada A_n tiene columnas ortonormales si y solo si $A^tA = I$.

Demostración: Podemos escribir $A = [\ \vec{a}_1\ \vec{a}_2\ \dots\ \vec{a}_n\ ]$ quedando por tanto $A^t = \begin{bmatrix} \vec{a}_1^t \\ \vec{a}_2^t \\ \vdots \\ \vec{a}_n^t \end{bmatrix}$

Desarrollando A^tA obtenemos $A^tA = \begin{bmatrix} \vec{a}_1^t \\ \vec{a}_2^t \\ \vdots \\ \vec{a}_n^t \end{bmatrix} [\ \vec{a}_1\ \vec{a}_2\ \dots\ \vec{a}_n\ ] = \begin{bmatrix} \vec{a}_1^t\vec{a}_1 & \dots & \vec{a}_1^t\vec{a}_n \\ \vdots & \ddots & \vdots \\ \vec{a}_n^t\vec{a}_1 & \dots & \vec{a}_n^t\vec{a}_n \end{bmatrix}$

La última expresión es igual a I si y solo si $\vec{a}_i^t\vec{a}_j = \delta_{ij} = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$

$\vec{a}_i^t\ \vec{a}_j$ es el producto escalar de $\vec{a}_i$ y $\vec{a}_j$, por tanto $A^tA = I$ si y sólo si las columnas de A son vectores unitarios ortogonales entre sí. $\square$

TEOREMA. Dada A_n, $A^tA = I$ si y sólo si $AA^t = I$.

Demostración: "$\Rightarrow$" $A^tA = I$ implica que el determinante de A es 1 o -1, en efecto:

$A^tA = I \Rightarrow |A^t||A| = 1 \Rightarrow |A|^2 = 1 \Rightarrow |A|$ es 1 o -1

$|A|$ distinto de cero, por tanto A es invertible.

Premultiplicando $A^tA = I$ por A por la izquierda y por A^{-1} por la derecha queda: $AA^t = I$

"$\Leftarrow$" $AA^t = I$ implica que el determinante de A es 1 o -1, con una demostración similar a la anterior. Por tanto A invertible.

Premultiplicando $AA^t = I$ por A^{-1} por la izquierda y por A por la derecha queda: $A^tA = I$ $\square$

Se dice que una matriz A_n dada es **matriz ortogonal** si cumple que $AA^t = A^tA = I$, o lo que es lo mismo, <u>si su inversa coincide con su traspuesta</u>, o lo que es lo mismo, <u>si sus columnas son base ortonormal de $\mathbb{R}^n$</u>.

Es importante recordar que la base debe ser ortonormal (el nombre "matriz ortogonal" puede inducirnos a pensar, erróneamente, que bastaría con que los vectores fueran ortogonales entre sí).

Resultados sobre matrices ortogonales:

- Si A es ortogonal también lo es A^t, por tanto las filas de A también son base ortonormal de $\mathbb{R}^n$.
- Las matrices ortogonales tienen determinante 1 o -1.

  $AA^t = I \Rightarrow |A||A^t| = 1 \Rightarrow |A|^2 = 1$
- El producto de matrices ortogonales es ortogonal.

  Lo comprobamos para A y B ortogonales.

  $(A\,B)^t\,A\,B = B^t\,A^t\,A\,B = B^t\,I\,B = B^t\,B = I$

A las aplicaciones o transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^n$ cuya matriz estándar asociada es ortogonal las denominamos **aplicaciones o transformaciones ortogonales**.

2 Subespacios ortogonales

Se dice que $\vec{z}$ es **ortogonal a un subespacio** W de $\mathbb{R}^n$ si $\vec{z}$ es ortogonal a todo vector $\vec{w} \in W$.

TEOREMA. $\vec{z}$ es ortogonal al subespacio W de $\mathbb{R}^n$ si y sólo si $\vec{z}$ es ortogonal a una base de W.

Se dice que el subespacio H es ortogonal a W si $\forall \vec{z} \in H$, $\vec{z}$ es ortogonal a W.

TEOREMA. H es ortogonal al subespacio W de $\mathbb{R}^n$ si y sólo si los vectores de una base de H son ortogonales a los de una base de W.

Si H es ortogonal a W, W es a su vez ortogonal a H (por la simetría del producto escalar), y se dice de W y H que son **subespacios** de $\mathbb{R}^n$ **ortogonales entre sí**.

Ejemplo 3. *Considera en* $\mathbb{R}^3$ *las rectas* $\mathbf{r} =< (1, a, 2) >$ *y* $\mathbf{s} =< (1, -2, 0) >$. *Determina el o los valores de* a *tales que* $\mathbf{r}$ *y* $\mathbf{s}$ *sean subespacios ortogonales.*

<u>*Sol:*</u> *Las bases de* r *y* s *son respectivamente* $\{(1, a, 2)\}$ *y* $\{(1, -2, 0)\}$.

El producto escalar de los dos vectores es $1 - 2a$, *por tanto las rectas son ortogonales si y sólo si* $a = 1/2$.

La recta $\mathbf{r}$ *es la siguiente:* $< (1, 1/2, 2) >$.

TEOREMA. La suma de dos subespacios H y W ortogonales entre sí es suma directa.

En efecto todo par de vectores $\vec{z}$, $\vec{w}$ no nulos, cada uno perteneneciente a un subespacio, es un conjunto linealmente independiente, por ser los vectores ortogonales entre sí, y ello garantiza que la suma es directa.

En relación con el **Tema 5** “Autovalores, autovectores y diagonalización” se tiene el siguiente resultado:

TEOREMA. A real es simétrica $\Leftrightarrow$ es diagonalizable y todos sus subespacios propios son ortogonales entre sí.

Las matrices estándar A de las simetrías ortogonales en $\mathbb{R}^2$ y en $\mathbb{R}^3$ respecto de rectas o planos respectivamente, son simétricas porque estos endomorfismos son diagonalizables y los subespacios propios son ortogonales entre sí. Lo mismo sucede con las matrices estándar A de las proyecciones ortogonales en $\mathbb{R}^2$ o $\mathbb{R}^3$ sobre rectas o planos respectivamente.

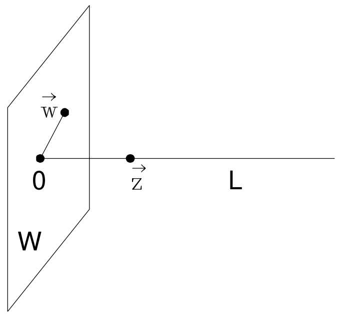
---

3 Complemento ortogonal

El conjunto de todos los vectores $\vec{z}$ que son ortogonales a W se denomina **complemento ortogonal** de W y se denota como $W^{\perp}$.

$W^{\perp} = \{\vec{z} \in \mathbb{R}^n \,/\, \vec{z} \cdot \vec{w} = 0 \quad \forall\, \vec{w} \in W\}$

Ejemplo 4. *Consideremos en $\mathbb{R}^3$ el subespacio W identificado con un plano que contiene el origen y el subespacio L identificado con la recta que pasa por el origen y perpendicular al plano anterior. Se tiene entonces que $\forall \vec{z} \in L$ y $\forall \vec{w} \in W$, $\vec{z} \cdot \vec{w} = 0$. Ver la figura. En efecto, L está formado por* **todos** *los vectores que son ortogonales a los $\vec{w}$ de W y recíprocamente W está formado por* **todos** *los vectores ortogonales a los $\vec{z}$ de L. Es decir, $L = W^{\perp}$ y $W = L^{\perp}$*



Tomemos por ejemplo el caso de $W =< (1,0,0),(0,1,0) >$.

Los vectores ortogonales a W serán los $(x,y,z) \in \mathbb{R}^3$ ortogonales a la base de W, es decir, tales que:

$$\begin{cases} (x,y,z)\cdot(1,0,0) = 0 \\ (x,y,z)\cdot(0,1,0) = 0 \end{cases} \Rightarrow \begin{cases} x = 0 \\ y = 0 \end{cases} \qquad \textit{Por tanto } W^{\perp} =< (0,0,1) >$$

TEOREMA. Se cumplen los siguientes resultados sobre $W^{\perp}$, siendo W un subespacio del espacio euclídeo $\mathbb{R}^n$.

1. $W^{\perp}$ es un subespacio de $\mathbb{R}^n$
2. $(W^{\perp})^{\perp} = W$
3. $W \bigoplus W^{\perp} = \mathbb{R}^n$

OBSERVACIÓN:

Todo subespacio $W \subset \mathbb{R}^n$ (salvo el $\{\vec{0}\}$ y el propio $\mathbb{R}^n$) admite infinitos subespacios complementarios, pero solo uno de ellos es complemento ortogonal.

Por ejemplo, en $\mathbb{R}^3$ cualquier recta pasando por el origen y no incluida en el plano XY es subespacio complementario del subespacio formado por el plano XY, pero el complemento ortogonal del plano XY es un subespacio único, que es la recta que define el eje Z.

Nótese en el ejemplo anterior que para obtener los vectores ortogonales a (a, b, c) en $\mathbb{R}^3$ hay que resolver la ec. $(a, b, c) \cdot (x, y, z) = 0$, es decir, la ec. lineal homogénea $ax + by + cz = 0$.

$ax+by+cz = 0$ es la forma implícita de un subespacio bidimensional F de $\mathbb{R}^3$. Dicha forma expresa que los vectores (x, y, z) de F y los vectores $< (a, b, c) >$[1] son ortogonales entre sí, y que por tanto F y $< (a, b, c) >$ son complementos ortogonales. $F^\perp =< (a, b, c) >$.

Complemento ortogonal de un subespacio W dado en implícitas: $A\vec{x} = \vec{0}$

De forma más general, para un subespacio W de dimensión d de $\mathbb{R}^n$, su forma implícita $A_{\ n-d\ ,\ n}\ \vec{x} = \vec{0}$ expresa que los $\vec{x} \in W$ — que son las soluciones y por tanto Nul$A$ — son ortogonales a las filas de A.

La base de W la forman las d soluciones independientes del SLH, o lo que es lo mismo la base de Nul$A$.

La base de $W^\perp$ la forman las filas de A.

Por ejemplo, para el subespacio $W = \{(x, y) \in \mathbb{R}^2\ /\ 2x + 3y = 0\}$

$B = \{(-3, 2)\}$ es base de W, porque $(-3, 2)$ es solución de la ecuación implícita.

$C = \{(2, 3)\}$ es base de $W^\perp$.

Es importante darse cuenta de que el "vector que aparece" en la ecuación, en este caso $(2, 3)$, es precisamente el ortogonal, y por tanto no perteneciente al subespacio que define la ecuación.

Complemento ortogonal de un subespacio W con base B conocida

Si partimos de un subespacio W de dimensión d del cual conocemos una base $B = \{\vec{b}_1, \dots, \vec{b}_d\}$, tomando $B = [\ \vec{b}_1\ \vec{b}_2\ \dots\ \vec{b}_d\ ]$ se tiene que la base de $W^\perp$ son las soluciones independientes del SLH $B^t\vec{x} = \vec{0}$, o lo que es lo mismo la base de Nul(B^t).

Por ejemplo partiendo de $W =< (1, 0, 0, 0), (2, 1, 0, 0) >$, $W^\perp$ son las soluciones del SLH $\begin{cases} x_1 = 0 \\ 2x_1 + x_2 = 0 \end{cases}$.

$$B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \Rightarrow \quad W^\perp = \text{Nul}(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{bmatrix}) \quad .$$

La matriz tiene rango 2, por tanto el SL tiene dos parámetros libres y la base de $W^\perp$ tendrá dos vectores. Un ejemplo de base de $W^\perp$ es $C = \{(0, 0, 1, 0), (0, 0, 0, 1)\}$.

Nótese que la base obtenida, puesta por filas, produce la matriz de coeficientes de la forma implícita de W.

$$W = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4\ /\ x_3 = 0\ ,\ x_4 = 0\}$$

[1] Si (a, b, c) cumple la ecuación también la cumple todo múltiplo de (a, b, c)

Ejemplo 5. *Sea* $B = \{(3,2,2,4),(1,0,0,2),(1,-1,-1,1)\}$ *una base de* F*, subespacio de* $\mathbb{R}^4$*. Halla una base del complemento ortogonal de* F*.*

Sol:

Los vectores ortogonales a los dados serán los (x,y,z,t) *que cumplan las siguientes ecuaciones:*

$(3,2,2,4)\cdot(x,y,z,t)=0$
$(1,0,0,2)\cdot(x,y,z,t)=0$
$(1,-1,-1,1)\cdot(x,y,z,t)=0$

forma impl. de $F^\perp$*:* $\begin{cases} 3x+2y+2z+4t=0 \\ x+2t=0 \\ x-y-z+t=0 \end{cases}$

Las 3 ecuaciones anteriores forman un SLH por lo que ya estamos viendo que $F^\perp$ *es un subespacio vectorial. Para obtener la base de* $F^\perp$ *hay que resolver el sistema.*

$$\begin{bmatrix} 3 & 2 & 2 & 4 & | \ 0 \\ 1 & 0 & 0 & 2 & | \ 0 \\ 1 & -1 & -1 & 1 & | \ 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -1 & 1 & | \ 0 \\ 1 & 0 & 0 & 2 & | \ 0 \\ 3 & 2 & 2 & 4 & | \ 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -1 & 1 & | \ 0 \\ 0 & 1 & 1 & 1 & | \ 0 \\ 0 & 5 & 5 & 1 & | \ 0 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & -1 & -1 & 1 & | \ 0 \\ 0 & 1 & 1 & 1 & | \ 0 \\ 0 & 0 & 0 & -4 & | \ 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 2 & | \ 0 \\ 0 & 1 & 1 & 1 & | \ 0 \\ 0 & 0 & 0 & -4 & | \ 0 \end{bmatrix}$$

Tenemos 3 ecuaciones, rango 3, y 4 incógnitas. Por tanto tenemos $4-3=1$ *parámetro libre. Tomando* z *como parámetro libre deducimos:*

$-4t = 0 \ \Rightarrow \ \boxed{t=0}$
$y+z+t=0 \ \Rightarrow \ y+z=0 \ \Rightarrow \ \boxed{y=-z}$
$x=-2t \ \Rightarrow \ \boxed{x=0}$

El vector solución es de la forma $(x,y,z,t) = (0,-z,z,0) \quad \forall z \in \mathbb{R}$
El conjunto de vectores ortogonales a los tres dados es un subespacio vectorial de dimensión 1. Una posible base es: $\{(0,1,-1,0)\}$

Nótese que $0x_1 + x_2 - x_3 + 0x_4 = 0$*, o lo que es lo mismo,* $x_2 - x_3 = 0$ *es la ecuación implícita de* F*.*

Ejemplo 6. *Se consideran en* $\mathbb{R}^3$ *los subespacios* $W_1 =< (1,1,0),(0,3,6) >$*,* $W_2 =< (1,2,1) >$ *y* $W_3 =< (7,8,5),(6,3,1),(1,3,6) >$*.*
Halla una base de cada uno de los subespacios ortogonales correspondientes, $W_1^\perp$*,* $W_2^\perp$*,* $W_3^\perp$*.*

Sol:

a) $W_1^\perp = (x,y,z) \in \mathbb{R}^3$ */*(x,y,z) *es ortogonal a (1,1,0) y a (0,3,6)*

$(x,y,z)\cdot(1,1,0) = x+y = 0$
$(x,y,z)\cdot(0,3,6) = 3y+6z = 0$

Tenemos dos ecuaciones y tres incógnitas. $\begin{bmatrix} 1 & 1 & 0 & | \ 0 \\ 0 & 3 & 6 & | \ 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 & | \ 0 \\ 0 & 1 & 2 & | \ 0 \end{bmatrix} \sim$

$\begin{bmatrix} 1 & 0 & -2 & | \ 0 \\ 0 & 1 & 2 & | \ 0 \end{bmatrix}$ *El sistema se encuentra ya en la forma escalonada reducida.*

Tomamos z *como parámetro libre.* $\quad x=2z \quad y=-2z \quad z=z$

El vector solución es de la forma $(x,y,z) = (2z,-2z,z) \quad \forall z \in \mathbb{R}$
El conjunto de vectores ortogonales a los dos dados es un subespacio vectorial de dimensión 1. Una posible base es $\{(2,-2,1)\}$

$W_1^\perp$ se expresaría como $W_1^\perp = < (2,-2,1) >$

*NOTA: Se puede obtener $W_1^\perp = < \vec{n} >$, siendo $\vec{n} = (1,1,0) \times (0,3,6)$ (***producto vectorial***).*

b) $W_2^\perp = (x,y,z) \in \mathbb{R}^3$ */(x,y,z) es ortogonal a (1,2,1)*

$(x,y,z) \cdot (1,2,1) = x + 2y + z = 0$

La última ecuación es la ecuación implícita de $W_2^\perp$

Tenemos una sola ecuación y tres incógnitas, por tanto dos parámetros libres. Dejando como parámetros libres y y z tendremos

$x + 2y + z = 0 \quad \Rightarrow x = -2y - z \quad y = y, \quad z = z$

El vector solución en forma paramétrica es $(x,y,z) = (-2y - z, y, z) \quad \forall y \in \mathbb{R} \quad \forall z \in \mathbb{R}$

El conjunto de vectores ortogonales al dado es un subespacio vectorial de dimensión 2.

Una posible base es: $\{(-2,1,0),(-1,0,1)\}$.

$W_2^\perp = < (-2,1,0),(-1,0,1) >$

c) $W_3^\perp = (x,y,z) \in \mathbb{R}^3$ */(x,y,z) es ortogonal a (7,8,5), (6,3,1) y (1,3,6)*

$$A = \begin{bmatrix} 1 & 3 & 6 & | \, 0 \\ 6 & 3 & 1 & | \, 0 \\ 7 & 8 & 5 & | \, 0 \end{bmatrix} \quad , \quad \begin{vmatrix} 1 & 3 & 6 \\ 6 & 3 & 1 \\ 7 & 8 & 5 \end{vmatrix} = 100$$

$detA \neq 0$, por tanto tenemos un sistema de Cramer, con solución única, y como el sistema es homogéneo, la solución única es la trivial.

Por tanto $W_3^\perp = \{(0,0,0)\}$.

W_3 representa todo el espacio $\mathbb{R}^3$, por lo que era de esperar que $W_3^\perp = \{\vec{0}\}$.

Habíamos visto anteriormente como $\vec{0}$ es ortogonal a todos los vectores.

Ejemplo 7. *Se considera el subespacio W de $\mathbb{R}^3$ dado por $2x + y - z = 0$. Determina una base de $W^\perp$.*

Sol:

Podemos interpretar la ecuación $2x + y - z = 0$ como la expresión del producto escalar de dos vectores, igualado a cero, siendo los vectores $(2,1,-1)$ y (x,y,z). Los vectores (x,y,z) contenidos en el plano que representa el subespacio vectorial W de $\mathbb{R}^3$ son los vectores ortogonales al vector $(2,1,-1)$, y obviamente a los múltiplos de éste.

En efecto si $(2,1,-1) \cdot (x,y,z) = 0$, entonces $\quad \lambda(2,1,-1) \cdot (x,y,z) = 0$

Por tanto $W^\perp = < (2,1,-1) >$, y una posible base $B = \{(2,1,-1)\}$

Ejemplo 8. *Se considera el espacio euclídeo canónico $\mathbb{R}^3$ y el subespacio W dado por la forma implícita siguiente:* $\begin{cases} x + 4y + 8z = 0 \\ x - y + z = 0 \end{cases}$ *. Determina una base de $W^\perp$.*

Sol:

$$W = \{(x,y,z) \,/\, \begin{cases} x + 4y + 8z = 0 \\ x - y + z = 0 \end{cases} \quad \}$$

Las ecuaciones implícitas expresan que los vectores (x,y,z) de W son a la vez ortogonales al vector $(1,4,8)$ y al vector $(1,-1,1)$, por tanto una base de $W^\perp$ es $\{(1,4,8),(1,-1,1)\}$.

4 Descomposición ortogonal y proyección ortogonal

El resultado $W \bigoplus W^{\perp} = \mathbb{R}^n$, significa que cada $\vec{y} \in \mathbb{R}^n$ se puede escribir de forma única como suma de un vector $\hat{\vec{y}} \in W$ y un vector $\vec{z} \in W^{\perp}$

$$\vec{y} = \hat{\vec{y}} + \vec{z} \qquad \text{con } \hat{\vec{y}} \in W \text{ y } \vec{z} \in W^{\perp}$$

$\hat{\vec{y}}$ es la **proyección ortogonal de $\vec{y}$ sobre** W, que denotamos también como proy $_W$ $\vec{y}$

También se puede dar esta definición equivalente: La proyección ortogonal de $\vec{y}$ sobre W es el vector $\hat{\vec{y}} \in W$ tal que $\vec{y} - \hat{\vec{y}} \ \in W^{\perp}$

¿ Cómo obtener $\text{proy}_W \vec{y}$?

Ya que $W \bigoplus W^{\perp} = \mathbb{R}^n$ podemos obtener una base de $\mathbb{R}^n$ $B = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_d, \vec{b}_{d+1}, \ldots, \vec{b}_n\}$, con $\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_d\}$ base de W y $\{\vec{b}_{d+1}, \ldots, \vec{b}_n\}$ base de $W^{\perp}$.

Entonces $\vec{y} = \underbrace{c_1\vec{b}_1 + c_2\vec{b}_2 + \ldots + c_d\vec{b}_d}_{\text{proy}_W \vec{y} \in W} + \underbrace{c_{d+1}\vec{b}_{d+1} + \ldots + c_n\vec{b}_n}_{\vec{z} \in W^{\perp}}$

con $\text{proy}_W \vec{y} = c_1\vec{b}_1 + c_2\vec{b}_2 + \ldots + c_d\vec{b}_d \in W$ y

$\vec{z} = c_{d+1}\vec{b}_{d+1} + \ldots + c_n\vec{b}_n \in W^{\perp}$

Una vez obtenidas las coordenadas c_1, c_2, ... c_n, que son únicas (las coordenadas respecto de una base dada son únicas), podremos determinar el vector único $\text{proy}_W \vec{y} \in W$ y el vector único $\vec{z} \in W^{\perp}$ tales que $\vec{y} = \text{proy}_W \vec{y} + \vec{z}$.

Esquemas en el espacio euclídeo canónico $\mathbb{R}^3$ (Linear Algebra and its Applications, Lay, Quinta edición. p. 350). Proyección ortogonal del vector $\vec{y}$ sobre el subespacio $< W >$ de dimensión 2:

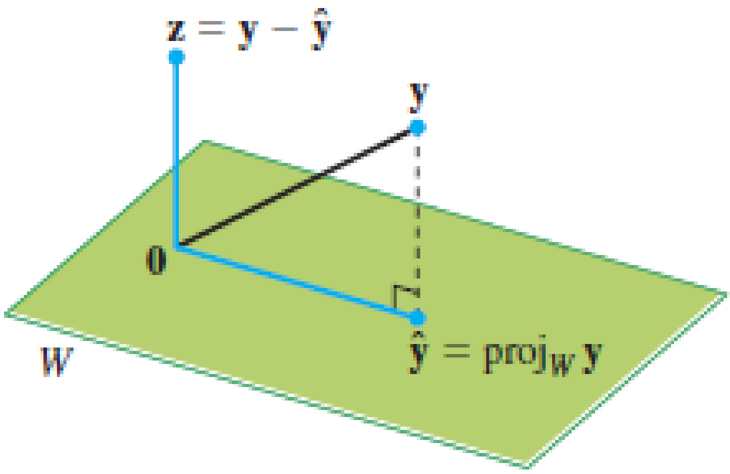


FIGURE 2 The orthogonal projection of $\mathbf{y}$ onto W.

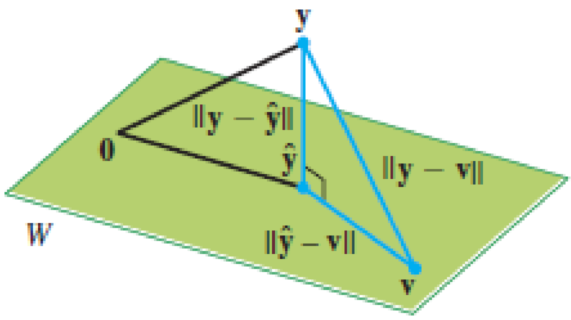


FIGURE 4 The orthogonal projection of **y** onto W is the closest point in W to **y**.

Por otra parte, $\hat{\vec{y}}$ tiene la propiedad de ser el vector de W más cercano a $\vec{y}$

$$\| \vec{y} - \text{proy}_W \ \vec{y} \| \quad < \quad \| \vec{y} - \vec{v} \| \quad \forall \vec{v} \in W \qquad \text{con} \quad \vec{v} \neq \hat{\vec{y}}$$

Decimos entonces que dado $\vec{y} \in \mathbb{R}^n$, la **mejor aproximación** de $\vec{y}$ que puedo obtener mediante un vector de W, subespacio de $\mathbb{R}^n$, es $\text{proy}_W \vec{y}$

¿En qué sentido es mejor aproximación?. En el sentido de menor distancia.

La **distancia** de $\vec{y} \in \mathbb{R}^n$ **al subespacio** W se define como la distancia desde $\vec{y}$ al punto más cercano de W. Dicho de otra forma, la distancia de $\vec{y}$ a W es igual a $\| \vec{y} - \text{proy}_W \ \vec{y} \| = \| \vec{z} \|$

OBSERVACIONES

- $\vec{z} = \text{proy}_{W^\perp} \vec{y}$, es decir, $\vec{z}$ es la proyección ortogonal de $\vec{y}$ sobre $W^\perp$.
- $\| \text{proy}_W \vec{y} \|$ es la distancia de $\vec{y}$ a $W^\perp$.
- Si $\vec{y} \in W$, entonces $\text{proy}_W \vec{y} = \vec{y}$
- Si $\vec{y} \in W^\perp$, entonces $\text{proy}_W \vec{y} = \vec{0}$

En temas anteriores hemos estudiado proyecciones ortogonales en $\mathbb{R}^2$ y en $\mathbb{R}^3$, determinando a partir de las características de la transformación cuál era la matriz asociada referida a la "base natural" de la transformación, o base de autovectores de la transformación, y cual era la matriz estándar asociada, obtenida mediante los cambios de base. Con procedimientos similares pudimos obtener la matriz asociada a la simetría ortogonal. A partir de las matrices resultaba sencillo obtener la proyección ortogonal o el simétrico de cualquier vector.

En este tema estudiamos la proyección ortogonal a partir de la descomposición ortogonal y extendemos el concepto de proyección ortogonal al espacio vectorial $\mathbb{R}^n$.

5 Proyección ortogonal conocida una base ortogonal

TEOREMA. Sea W un subespacio de $\mathbb{R}^n$, $\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_d\}$ una base ortogonal de W e $\vec{y} \in \mathbb{R}^n$. Entonces

$$proy_W \vec{y} = \frac{\vec{y} \cdot \vec{b}_1}{\vec{b}_1 \cdot \vec{b}_1}\,\vec{b}_1 + \frac{\vec{y} \cdot \vec{b}_2}{\vec{b}_2 \cdot \vec{b}_2}\,\vec{b}_2 + \ldots + \frac{\vec{y} \cdot \vec{b}_d}{\vec{b}_d \cdot \vec{b}_d}\,\vec{b}_d \qquad [\mathbf{1a}]$$

Cada sumando corresponde a la proyección ortogonal sobre un subespacio unidimensional $< \vec{b}_i >$. Por tanto la proyección ortogonal de $\vec{y}$ sobre $< \vec{b}_1, \vec{b}_2, \ldots, \vec{b}_d >$, siendo $\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_d\}$ una base ortogonal, es igual a la suma de las d proyecciones ortogonales sobre subespacios unidimensionales, mutuamente ortogonales, en las direcciones de $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_d$.

• Si la base $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_d\}$ es ortonormal la expresión [**1a**] queda cómo:

$$proy_W \vec{y} = (\vec{y} \cdot \vec{u}_1)\,\vec{u}_1 + (\vec{y} \cdot \vec{u}_2)\,\vec{u}_2 + \ldots + (\vec{y} \cdot \vec{u}_d)\,\vec{u}_d \qquad [\mathbf{1b}]$$

Para cálculos a mano se recomienda utilizar la fórmula [**1a**] ya que en la [**1b**] aparecerán en general raíces cuadradas.

Se muestra un esquema en $\mathbb{R}^3$ de la proyección sobre un subespacio W de dimensión 2 genérico del que conocemos una base ortogonal B genérica. (Lay, Linear Algebra and its Applications. Quinta edición, p. 351).

De acuerdo con la notación que usamos en esta sección, ha de tenerse en cuenta:

- La base $\{\vec{u}_1, \vec{u}_2\}$ de la figura ha de entenderse como una base ortogonal, no necesariamente ortonormal. En efecto en la fórmula de la figura aparecen explícitamente los productos escalares $\vec{u}_i \cdot \vec{u}_i$, que no serían necesarios si la base fuera ortonormal. Para interpretar la figura de acuerdo con la notación usada en esta Sección, esta base ortogonal sería la base $B = \{\vec{b}_1, \vec{b}_2\}$.

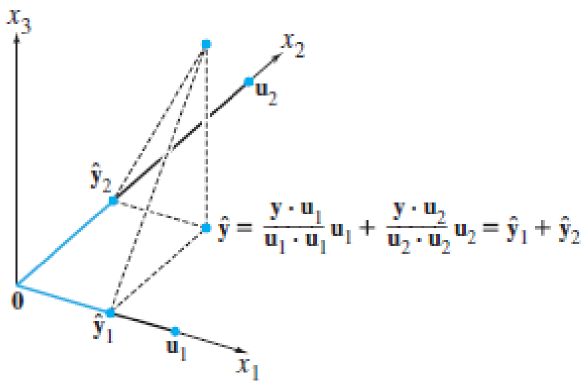


FIGURE 3 The orthogonal projection of **y** is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

Ejemplo 9. *Considera en* $\mathbb{R}^2$ *los vectores* $\vec{y} = (7,6)$ *y* $\vec{b} = (2,1)$*. Encuentra la proyección ortogonal de* $\vec{y}$ *sobre* $< \ \vec{b} \ >$ *y la distancia de* $\vec{y}$ *a la recta* $< \vec{b} >$*.*

Sol:

• *Puesto que se proyecta sobre un subespacio de dimensión 1, su base puede considerarse base ortogonal.*

$$\text{proy}_{<\vec{b}>} \ \vec{y} = \frac{\vec{y}.\vec{b}}{\vec{b}.\vec{b}} \ \vec{b}$$

$$\text{proy}_{<\vec{b}>} \ \vec{y} = \frac{(7,6).(2,1)}{(2,1).(2,1)} \ (2,1) = \frac{20}{5} \ (2,1) = 4(2,1) = (8,4)$$

• *La distancia de* $\vec{y}$ *a* $< \vec{b} >$ *es igual a la norma de* $\vec{z} = \vec{y} - proy_{<\vec{b}>} \ \vec{y} = (7,6) - (8,4) = (-1,2)$

$$d = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$$

OBSERVACIÓN: Este procedimiento es el más sencillo, pero $\text{proy}_{<\vec{b}>} \ \vec{y}$ *se podría haber obtenido con otros métodos.*

• *Mediante el procedimiento visto en la Sección 4:*

Calculamos en primer lugar un vector $\vec{z}_1$ *ortogonal a* $\vec{b} = (2,1)$*. Un ejemplo es* $\vec{z}_1 = (1,-2)$*.*

Considerada la base $B = \{\vec{b}, \vec{z}_1\}$*, con un vector en la dirección de* $\vec{b}$ *y otro en la dirección ortogonal* $\vec{z}_1$*, se calculan las coordenadas de* $(7,6)$ *respecto a esta base, resultando ser* $(4,-1)$*, es decir:*

$$(7,6) = 4(2,1) + -1(1,-2) = \underbrace{(8,4)}_{\in< \vec{b} >} + \underbrace{(-1,2)}_{\in< \vec{z}_1 >}$$

Por tanto la proyección de $(7,6)$ *sobre* $< (2,1) >$ *es el vector* $\text{proy}_{<\vec{b}>} \ \vec{y} = (8,4)$

• *A partir de la matriz asociada a la aplicación lineal vista en temas anteriores (matriz de la proyección ortogonal en* $\mathbb{R}^2$*), usando la base anterior:*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \left(\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} \right)^{-1} = ...$$

$$\text{proy}_{<\vec{b}>} \ \vec{y} = A \begin{bmatrix} 7 \\ 6 \end{bmatrix}$$

6 Base ortogonal mediante Gram-Schmidt

El algoritmo de Gram-Schmidt permite obtener una base ortogonal de cualquier subespacio de $\mathbb{R}^n$ a partir de una base dada.

TEOREMA. Teorema de la Ortogonalización de Gram-Schmidt

Dada una base $\{\vec{a}_1, \ldots, \vec{a}_d\}$ de un subespacio W de $\mathbb{R}^n$, y definidos:

$$\vec{b}_1 = \vec{a}_1$$

$$\vec{b}_2 = \vec{a}_2 - \frac{\vec{a}_2 \cdot \vec{b}_1}{\vec{b}_1 \cdot \vec{b}_1}\vec{b}_1$$

$$\vec{b}_3 = \vec{a}_3 - \frac{\vec{a}_3 \cdot \vec{b}_1}{\vec{b}_1 \cdot \vec{b}_1}\vec{b}_1 - \frac{\vec{a}_3 \cdot \vec{b}_2}{\vec{b}_2 \cdot \vec{b}_2}\vec{b}_2$$

$$\vdots$$

$$\vec{b}_d = \vec{a}_d - \frac{\vec{a}_d \cdot \vec{b}_1}{\vec{b}_1 \cdot \vec{b}_1}\vec{b}_1 - \frac{\vec{a}_d \cdot \vec{b}_2}{\vec{b}_2 \cdot \vec{b}_2}\vec{b}_2 - \ldots - \frac{\vec{a}_d \cdot \vec{b}_{d-1}}{\vec{b}_{d-1} \cdot \vec{b}_{d-1}}\vec{b}_{d-1} \quad ,$$

$\{\vec{b}_1, \ldots, \vec{b}_d\}$ resulta ser una base ortogonal de W. Además,

$< \vec{b}_1, \ldots, \vec{b}_k > = < \vec{a}_1, \ldots, \vec{a}_k >$, para $1 \leq k \leq d$

Ejemplo 10. *Sea $W = < (0,1,1), (1,1,2) >$ un subespacio de $\mathbb{R}^3$. Obtén una base ortogonal de este subespacio.*

Sol.:

A partir de los vectores originales $\vec{a}_1 = (0,1,1)$ y $\vec{a}_2 = (1,1,2)$, que forman la base B, construiremos la base $B' = \{\vec{b}_1, \vec{b}_2\}$, siendo $\vec{b}_1$ y $\vec{b}_2$ vectores de W ortogonales entre sí .

Tomamos $\vec{b}_1 = \vec{a}_1$ y a continuación construimos $\vec{b}_2$.

Consideremos el vector $\operatorname{proy}_{<\vec{b}_1>} \vec{a}_2$,

$\vec{b}_2 = \vec{a}_2 - \operatorname{proy}_{<\vec{b}_1>} \vec{a}_2$ *es ortogonal a $\vec{b}_1$ y pertenece a W, pués es combinación lineal de $\vec{b}_1$ y $\vec{a}_2$.*

Por tanto, una base ortogonal de W es: $\begin{cases} \vec{b}_1 = \vec{a}_1 \\ \vec{b}_2 = \vec{a}_2 - \operatorname{proy}_{<\vec{b}_1>} \vec{a}_2 = \vec{a}_2 - \frac{\vec{a}_2 . \vec{b}_1}{\vec{b}_1 . \vec{b}_1}\vec{b}_1 \end{cases}$

$\vec{b}_1 = (0,1,1)$

$$\vec{b}_2 = (1,1,2) - \frac{(1,1,2)\cdot(0,1,1)}{(0,1,1)\cdot(0,1,1)}(0,1,1) = (1,1,2) - 3/2(0,1,1) = (1, 1-3/2, 2-3/2) = (1, -1/2, 1/2)$$

Podemos tomar $\vec{b}_2 = (2,-1,1)$ para evitar fracciones.

Comprobamos que $\vec{b}_1$ y $\vec{b}_2$ forman una base ortogonal: $(0,1,1)\cdot(2,-1,1) = 0 - 1 + 1 = 0$

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

7 Ejercicios resueltos

Ejemplo 11. *Supongamos dos clases de alimento, A y B, con las cantidades de vitamina C, calcio y magnesio dadas en la tabla. Las cantidades corresponden a miligramos por unidad de alimento.*

	A	*B*
Vitamina C	1	1
Calcio	5	2
Magnesio	3	1

a) Demuestra que combinando las dos clases de alimentos no podemos obtener el siguiente aporte exacto:

$\vec{v} = (17 \text{ } mg \text{ } de \text{ } vitamina \text{ } C, 54 \text{ } mg \text{ } de \text{ } calcio, 31 \text{ } mg \text{ } de \text{ } magnesio)$

b) Determina el aporte más cercano al aporte exacto que se podría conseguir combinando los dos alimentos.

Sol.:

a) Para determinar si el aporte $(17, 54, 31)$ *se puede obtener combinando* x *unidades de alimento A con aporte* $(1, 5, 3)$ *e* y *unidades de B con aporte* $(1, 2, 1)$ *hay que resolver la siguiente ecuación:*

$x(1, 5, 3) + y(1, 2, 1) = (17, 54, 31)$

Por tanto el SL: $\begin{cases} x + y = 17 \\ 5x + 2y = 54 \\ 3x + y = 31 \end{cases}$ *, con matriz ampliada* $A^* = \begin{bmatrix} 1 & 1 & \mid 17 \\ 5 & 2 & \mid 54 \\ 3 & 1 & \mid 31 \end{bmatrix}$

$$A^* \sim \begin{bmatrix} 1 & 1 & \mid & 17 \\ 0 & -3 & \mid & -31 \\ 0 & -2 & \mid & -20 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & \mid & 17 \\ 0 & -3 & \mid & -31 \\ 0 & 1 & \mid & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & \mid & 17 \\ 0 & 1 & \mid & 10 \\ 0 & -3 & \mid & -31 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & \mid & 17 \\ 0 & 1 & \mid & 10 \\ 0 & 0 & \mid & -1 \end{bmatrix}$$

El SL es incompatible pues rgA=2 y rgA=3*

Esto significa que $\vec{v}$ *no es combinación lineal de* $\vec{a} = (1, 5, 3)$ *y* $\vec{b} = (1, 2, 1)$*, o dicho de otra forma, que* $\vec{v}$ *no pertenece al plano* $< \vec{a}, \vec{b} >$.

b) El vector más cercano a $\vec{v} = (17, 54, 31)$ *que se puede obtener combinando los alimentos A y B será la proyección ortogonal de* $\vec{v}$ *sobre el subespacio generado por* $\vec{a}$ *y* $\vec{b}$*. Por tanto tenemos que obtener:*

$proy_W\vec{v}$*, siendo* $W = < \vec{a}, \vec{b} >$

- *Procedimiento según la Sección 5:*

  Obtendremos en primer lugar la proyección ortogonal sobre $W^\perp$*, por ser más sencillo aplicar la fórmula sobre un subespacio de dimensión 1. Además la base de W que conocemos no es ortogonal.*

  Tomamos $W = < (1, 5, 3), (1, 2, 1) >$*, y nos falta el vector* $\vec{n} = (n_1, n_2, n_3) \in W^\perp$.

  $\begin{matrix} (1, 5, 3).(n_1, n_2, n_3) = 0 \\ (1, 2, 1).(n_1, n_2, n_3) = 0 \end{matrix} \begin{cases} n_1 + 5n_2 + 3n_3 = 0 \\ n_1 + 2n_2 + n_3 = 0 \end{cases} \quad A^* = \begin{bmatrix} 1 & 5 & 3 & \mid 0 \\ 1 & 2 & 1 & \mid 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 3 & \mid 0 \\ 0 & -3 & -2 & \mid 0 \end{bmatrix}$

  $$\sim \begin{bmatrix} 1 & 5 & 3 & \mid 0 \\ 0 & 1 & 2/3 & \mid 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -10/3 + 3 & \mid 0 \\ 0 & 1 & 2/3 & \mid 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1/3 & \mid 0 \\ 0 & 1 & 2/3 & \mid 0 \end{bmatrix} \Rightarrow (1/3n_3, -2/3n_3, n_3)$$

$W^\perp = <(1/3,-2/3,1)>$ *o más simple* $W^\perp = <(1,-2,3)>$*,* $\quad \vec{n}=(1,-2,3)$.

Este vector también se podría haber obtenido mediante el producto vectorial.

La proyección de $\vec{v}$ *sobre la recta* $<\vec{n}> = W^\perp$ *es:*

$$\text{proy}_{W^\perp}\vec{v} = \frac{\vec{v}.\vec{n}}{\vec{n}.\vec{n}}\,\vec{n} = \frac{(17,54,31)\cdot(1,-2,3)}{14}\begin{bmatrix}1\\-2\\3\end{bmatrix} = \frac{(17-108+93)}{14}\begin{bmatrix}1\\-2\\3\end{bmatrix} = \frac{1}{7}\begin{bmatrix}1\\-2\\3\end{bmatrix}$$

$\text{proy}_W\vec{v} = \vec{v} - \text{proy}_{W^\perp}\vec{v} = (17,54,31) - \frac{1}{7}(1,-2,3) = (118/7, 380/7, 214/7) = 2/7(59,190,107)$.

La expresión decimal es $proy_W\vec{v} = (16.8571, 54.2857, 30.5714)$

- *Procedimiento según la Sección 4 (suma de subespacios ortogonales):*

  En el procedimiento anterior ya se presentó la obtención de la base de $W^\perp$*,* $\{\vec{n} = (1,-2,3)\}$.

  Expresando (17,54,31) como combinación lineal de los vectores de la base completa de $\mathbb{R}^3$ $B=\{(1,5,3),(1,2,1),(1,-2,3)\}$ *obtenemos:*

$$(17,54,31) = \underbrace{\frac{48}{7}(1,5,3)+10(1,2,1)}_{proy_W\vec{v}} \; + \; \underbrace{\frac{1}{7}(1,-2,3)}_{proy_{W^\perp}\vec{v}},$$

  sin más que resolver el SL $[\ \vec{a}\ \vec{b}\ \vec{n}\ |\ \vec{v}\ ]$*, y tras encontrar que dicha solución es* $(48/7, 10, 1/7)$.

  $proy_W\vec{v} = 48/7(1,5,3)+10(1,2,1) = (118/7, 380/7, 214/7)$ *es el aporte más cercano posible.*

  Este procedimiento nos permite ver que el aporte más cercano posible se obtiene tomando $48/7$ *unidades del alimento tipo A y 10 unidades del alimento tipo* B.

Ejemplo 12. *a) Demuestra que no es posible expresar el vector $\vec{v} = (15, 6, 4, -5)$ como combinación lineal de los vectores del conjunto $S = \{(-1, 0, 0, 1), (4, 1, -1, -2)\}$.*
b) Obtén los <u>*coeficientes de la combinación lineal*</u> *de los vectores de S que dan el vector más cercano a $\vec{v}$.*

<u>*Sol.:*</u>

a)
$$\begin{bmatrix} -1 & 4 & | & 15 \\ 0 & 1 & | & 6 \\ 0 & -1 & | & 4 \\ 1 & -2 & | & -5 \end{bmatrix}$$

Una forma de ver que el sistema es incompatible es sumar las ecuaciones 2 y 3, ya que se obtiene una ecuación de la forma $[\,0\ 0\,|\,\neq 0\,]$

b) El vector de $< S >$ más cercano a $\vec{v}$ es la proyección de $\vec{v}$ sobre $< S >$. Obtendremos la proyección $\hat{\vec{v}}$ y las coordenadas (los coeficientes del enunciado) de esta respecto de la base S, utilizando dos métodos distintos.

- <u>*Aplicando la fórmula de proyección, cuando se conoce base ortogonal del subespacio (Sección 5)*</u>

  Los dos vectores de S forman base de $< S >$, pero esa base no es ortogonal. Aplicamos Gram-Schmidt:

  $\vec{b}_1 = (-1, 0, 0, 1)$,

$$\vec{b}_2 = (4, 1, -1, -2) - \frac{(4, 1, -1, -2) \cdot (-1, 0, 0, 1)}{2}(-1, 0, 0, 1) =$$

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

  Ya tenemos una base ortogonal en $< S >$, $B' = \{(-1, 0, 0, 1), (1, 1, -1, 1)\}$, y podemos aplicar la fórmula de la Sección 5.

$$\text{proy}_{<S>}\vec{v} = \frac{(15, 6, 4, -5) \cdot (-1, 0, 0, 1)}{2} \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + \frac{(15, 6, 4, -5) \cdot (1, 1, -1, 1)}{4} \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} =$$

$$-10 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 13 \\ 3 \\ -3 \\ -7 \end{bmatrix}$$

  Obtenemos ahora los coeficientes de la combinación lineal de los vectores de $< S >$, que llamaremos $\vec{a}_1$ y $\vec{a}_2$, que dan el vector más cercano a $\vec{v}$. Para ello hay que resolver el siguiente sistema:

$$\begin{bmatrix} -1 & 4 & | & 13 \\ 0 & 1 & | & 3 \\ 0 & -1 & | & -3 \\ 1 & -2 & | & -7 \end{bmatrix} \sim \begin{bmatrix} -1 & 4 & | & 13 \\ 0 & 1 & | & 3 \\ 0 & 0 & | & 0 \\ 0 & 2 & | & 6 \end{bmatrix} \sim \begin{bmatrix} -1 & 4 & | & 13 \\ 0 & 1 & | & 3 \\ 0 & 0 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix} \sim \begin{bmatrix} -1 & 0 & | & 1 \\ 0 & 1 & | & 3 \\ 0 & 0 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}$$

$$\boxed{\begin{cases} c_1 = -1 \\ c_2 = 3 \end{cases}}$$

- <u>*Método de descomposición ortogonal de $\vec{v}$ usando suma de subespacios ortogonales (Sección 4)*</u>

  Para obtener una base del complemento ortogonal de $< S >=< (-1,0,0,1),(4,1,-1,2) >$ *tenemos que resolver el SL homogéneo con matriz de coeficientes* $\begin{bmatrix} -1 & 0 & 0 & 1 \\ 4 & 1 & -1 & -2 \end{bmatrix}$

$$\begin{bmatrix} -1 & 0 & 0 & 1 & \mid 0 \\ 4 & 1 & -1 & -2 & \mid 0 \end{bmatrix} \sim \begin{bmatrix} -1 & 0 & 0 & 1 & \mid 0 \\ 0 & 1 & -1 & 2 & \mid 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -1 & \mid 0 \\ 0 & 1 & -1 & 2 & \mid 0 \end{bmatrix}$$

La solución general en forma paramétrica, tomando como parámetros z y t es:

$(t, z-2t, z, t)$

Y la base más sencilla de $< S >^{\perp}$ es $B = \{(1,-2,0,1),(0,1,1,0)\}$.

Puede resultar más claro modificar la forma paramétrica a $(x, z-2x, z, x)$, teniendo en cuenta que $t = x$.

Resolviendo el sistema: $\begin{bmatrix} -1 & 4 & 0 & 1 & \mid & 15 \\ 0 & 1 & 1 & -2 & \mid & 6 \\ 0 & -1 & 1 & 0 & \mid & 4 \\ 1 & -2 & 0 & 1 & \mid & -5 \end{bmatrix}$ *, obtenemos la solución* $(-1,3,7,2)$.

Base original de $< S >$ Base de $< S >^{\perp}$

$\boxed{\textit{Los coeficientes buscados son } -1 \textit{ y } 3}$, *ya que*

$$\text{proy}_{<S>}\vec{v} = -1 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 4 \\ 1 \\ -1 \\ -2 \end{bmatrix}$$ *es el vector de $< S >$ más cercano a $\vec{v}$*

$$\text{proy}_{<S>}\vec{v} = \begin{bmatrix} 13 \\ 3 \\ -3 \\ -7 \end{bmatrix}$$

8 Matriz de proyección ortogonal

<u>Método 1</u>

La proyección ortogonal de $\vec{v} \in \mathbb{R}^n$ sobre W es un endomorfismo en $\mathbb{R}^n$, con matriz estándar asociada A_{proy} tal que

$$A_{\text{proy}}\vec{v} = \hat{\vec{v}} \qquad\qquad A_{\text{proy}} \text{ es cuadrada de orden } n.$$

Conocida una base $B = \{\vec{b}_1, ..., \vec{b}_d\}$ de W, puede obtenerse A_{proy} mediante la siguiente expresión[2]:

$$\boxed{A_{\text{proy}} = B(B^tB)^{-1}B^t \qquad [\mathbf{2a}]}$$

Es obvio que la matriz B^tB es simétrica[3]. Esta matriz es invertible debido a la condición de que las columnas de B sean linealmente independientes[2].

Desarrollamos seguidamente el producto B^tB:

$$B^tB = \begin{bmatrix} \vec{b}_1^t \\ \vec{b}_2^t \\ \vdots \\ \vec{b}_d^t \end{bmatrix}_{d\times n} \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_d \end{bmatrix}_{n\times d} = \begin{bmatrix} \vec{b}_1\cdot\vec{b}_1 & \vec{b}_1\cdot\vec{b}_2 & \dots & \vec{b}_1\cdot\vec{b}_d \\ \vec{b}_2\cdot\vec{b}_1 & \vec{b}_2\cdot\vec{b}_2 & \dots & \vec{b}_2\cdot\vec{b}_d \\ \vdots & \dots & \ddots & \vdots \\ \vec{b}_d\cdot\vec{b}_1 & \vec{b}_d\cdot\vec{b}_2 & \dots & \vec{b}_d\cdot\vec{b}_d \end{bmatrix}$$

(también podríamos deducir de aquí que como el producto escalar es simétrico la matriz B^tB es simétrica.)

En el desarrollo de B^tB vemos de forma inmediata que si la base B es ortogonal B^tB es una matriz diagonal, y que <u>si la base es ortonormal</u> $B^tB = I_d$.

En este último caso la matriz de proyección queda: $\boxed{A_{\text{proy}} = BB^t \qquad [\mathbf{2b}]}$.

Esta última ecuación se puede presentar como $A_{\text{proy}} = UU^t$, entendiéndose entonces que U representa una base ortonormal de W.

Se puede demostrar fácilmente, utilizando la expresión [**2a**] o la [**2b**], que la matriz es simétrica e idempotente. La última propiedad significa que $A_{\text{proy}}^2 = A_{\text{proy}}$.

La ventaja de las expresiones obtenidas es que sólo requieren calcular una base B de W. En la expresión [**2a**] una base cualquiera y en la expresión [**2b**] una base ortonormal.

Señalamos también el siguiente resultado:

$$A_{\text{proyW}} + A_{\text{proyW}^\perp} = I$$

[2] No presentamos su deducción

[3] $(M^tM)^t = M^t(M^t)^t = M^tM$

Método 2: Obtención de A_{proy} a partir de su semejante diagonal

Este procedimiento se usó en el Tema 5 al estudiar la proyección ortogonal sobre una recta en $\mathbb{R}^2$ y sobre un plano en $\mathbb{R}^3$. Aquí lo extendemos a subespacios de $\mathbb{R}^n$ de cualquier dimensión.

En primer lugar obtenemos, a partir de la base B del subespacio considerado W, una base C de $W^{\perp}$, resolviendo el SL $B^t\vec{x} = \vec{0}$, o lo que es lo mismo, obteniendo una base de $\text{Nul}(B^t)$.

Así tenemos ya una base de cada uno de los dos subespacios propios ortogonales entre sí :

$\text{Col}(B)$ o $< B >$ es el subespacio propio de autovalor 1, V_1.
V_1 es W: los vectores de W se tranforman como $f(\vec{v}) = \vec{v}$.

$\text{Col}(C)$ o $< C >$ es el subespacio propio de autovalor 0, V_0.
V_0 es $W^{\perp}$. Los vectores de $W^{\perp}$ se transforman como $f(\vec{v}) = \vec{0}$

La matriz cuadrada $P = [B\ C] = [\ \vec{b}_1\ \vec{b}_2\ \dots\ \vec{b}_d\ \ \vec{b}_{d+1}\ \vec{b}_{d+2}\ \dots\ \vec{b}_n\ ]$, tiene como columnas una base de $\mathbb{R}^n$ formada por los autovectores de la proyección ortogonal.

$$\text{Por tanto } A_{\text{proy}} = PDP^{-1} = P \begin{bmatrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & 0 & 0 & \dots & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & 0 & 0 & \ddots & \vdots \\ \underbrace{0 \quad \dots \quad 0} & & & \underbrace{0 \quad \dots \quad 0} & & \end{bmatrix} P^{-1}$$

Al ser los subespacios propios ortogonales, podemos obtener una matriz P, más "avanzada", cuyas columnas sean base ortonormal de autovectores, sin más que unir una base ortonormal B de W y una base ortonormal C de $W^{\perp}$. Podemos denotar esa matriz como Q, que es la letra que se usa en general en Álgebra para denotar matrices ortogonales. $Q^{-1} = Q^t$, por tanto la matriz A_{proy} se obtendría, usando Q, como:
$A_{\text{proy}} = QDQ^t$.

Este procedimiento requiere determinar la base del complemento ortogonal de W si queremos obtener P. Si queremos obtener Q debemos obtener bases ortogonales, dentro de los dos subespacios propios, y normalizarlas.

Relación entre las ecuaciones [**1b**] y [**2b**]

$$UU^t\vec{v} = [\ \vec{u}_1\ \vec{u}_2\ \dots\ \vec{u}_d\ ]_{n\times d} \begin{bmatrix} \vec{u}_1^t \\ \vec{u}_2^t \\ \vdots \\ \vec{u}_d^t \end{bmatrix}_{d\times n} \vec{v}_{n\times 1} = [\ \vec{u}_1\ \vec{u}_2\ \dots\ \vec{u}_d\ ]_{n\times d} \begin{bmatrix} \vec{u}_1\cdot\vec{v} \\ \vec{u}_2\cdot\vec{v} \\ \vdots \\ \vec{u}_d\cdot\vec{v} \end{bmatrix}_{d\times 1} =$$

$$[\ \vec{u}_1\ \vec{u}_2\ \dots\ \vec{u}_d\ ]_{n\times d} \begin{bmatrix} \vec{v}\cdot\vec{u}_1 \\ \vec{v}\cdot\vec{u}_2 \\ \vdots \\ \vec{v}\cdot\vec{u}_d \end{bmatrix}_{d\times 1} = \vec{v}\cdot\vec{u}_1\,\vec{u}_1 + \vec{v}\cdot\vec{u}_2\,\vec{u}_2 + \dots + \vec{v}\cdot\vec{u}_d\,\vec{u}_d$$

9 Diagonalización ortogonal de las matrices simétricas

En la Sección 2 enunciamos el siguiente **TEOREMA:** "A real es simétrica si y sólo si es diagonalizable y todos sus subespacios propios son ortogonales entre sí".

Como a su vez, cada subespacio propio admite base ortonormal, tenemos el siguiente resultado:

- Si A es una matriz real, cuadrada y simétrica, entonces existen Q ortogonal y D diagonal tales que $A = QDQ^{-1}$. D es la matriz de autovalores, y Q tiene por columnas una base de autovectores que es además base ortonormal, en el orden de sus correspondientes autovalores en D.

  Teniendo en cuenta que $Q^{-1} = Q^t$, podemos expresar la factorización anterior en la forma $A = QDQ^t$.

  Por existir la matriz ortogonal Q tal que $A = QDQ^{-1}$ se dice de A que es 'ortogonalmente diagonalizable'.

  El teorema anterior y el resultado que acabamos de ver, para toda matriz A real y simétrica de orden n, los podemos desglosar en las siguientes afirmaciones:

  1. La ecuación característica de A tiene n raíces reales contando multiplicidades, por tanto n autovalores reales contando multiplicidades.
  2. La dimensión de los subespacios propios es igual a la multiplicidad algebraica de los correspondientes autovalores.
  3. Los subespacios propios correspondientes a autovalores distintos son ortogonales entre sí respecto del producto escalar canónico en $\mathbb{R}^n$.
  4. Obteniendo para cada subespacio propio una base ortogonal[4], y uniendo estas bases, lograremos una base ortogonal de $\mathbb{R}^n$ formada por autovectores de A. Dividiendo cada vector por su norma obtendremos una base ortonormal.

     $A = QDQ^{-1} = QDQ^t$ con $Q = [\ \vec{u}_1 \ \dots \ \vec{u}_n\ ]$ siendo $\{\ \vec{u}_1 \ \dots \ \vec{u}_n\ \}$ una base ortonormal de $\mathbb{R}^n$ formada por autovectores de A.

- Se demuestra fácilmente que se cumple el resultado recíproco, o lo que es lo mismo, la implicación inversa del teorema anterior: A ortogonalmente diagonalizable implica A simétrica. En efecto, considerando $A = QDQ^t$ y tomando traspuestas obtenemos $A^t = (QDQ^t)^t = QDQ^t = A$.

TEOREMA. A es simétrica si y sólo si A es ortogonalmente diagonalizable.

[4] Está demostrado que los vectores de subespacios propios distintos son ortogonales entre sí, pero dentro de un mismo subespacio propio de dimensión mayor o igual que 2 debemos determinar una base ortogonal.

Ejemplo 13. *Obtén la matriz de la proyección ortogonal en* $\mathbb{R}^3$ *sobre el plano* W *de ecuación implícita* $x+y+z=0$.

Sol.:

Presentamos la solución mediante los dos procedimientos distintos que acabamos de presentar.

Método 1. Utilizando la sencilla expresión [2b], $A_{\text{proy}} = UU^t$, *con* U *base ortonormal de* W.
(U *es una designación habitual en Álgebra para matrices cuyas columnas son ortonormales).*

$B = \{(1,-1,0),(1,0,-1)\}$ *es base de* W, *pero no es base ortogonal. Vamos a buscar una base ortonormal de* W

Aplicamos Gram-Schmidt tomando $\vec{b}_1 = (1,-1,0)$

$$(1,0,-1) - \frac{(1,0,-1)\cdot(1,-1,0)}{2}(1,-1,0) = (1,0,-1) - \frac{1}{2}(1,-1,0) = (1/2,1/2,-1)$$

Tomamos $\vec{b}_2 = (1,1,-2)$

La base $U = \{(1,-1,0)/\sqrt{2},(1,1,-2)/\sqrt{6}\}$ *es base ortonormal de* W.

$$A_{\text{proy}} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix}$$

Método 2. Solución mediante diagonalización ortogonal.

$$A_{\text{proy}} = QDQ^t$$

Las dos primeras columnas de Q *son los vectores de* U *obtenidos arriba. La tercera columna es el vector normal al plano, normalizado:* $\vec{u}_3 = (1,1,1)/\sqrt{(3)}$.

$$A_{\text{proy}} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}^t = \begin{bmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix}$$

Observa que la matriz es simétrica, como era de esperar. También puedes comprobar que es idempotente. Además cumple que su determinante es 0 *(tendrías que realizar los cálculos) y su traza 2 (respectivamente producto y suma de sus autovalores).*

Esta es la misma matriz que la obtenida en el Ejercicio 4.6.

10 Ejercicios

Ejercicio 1. *Se considera el subespacio* $H = \{(x, y, z, t) \in \mathbb{R}^4 / x - z - t = 0\}$

a) *Calcula una base ortonormal de* H.

b) *Calcula una base ortonormal de* $H^{\perp}$.

c) *Calcula la proyección ortogonal del vector* $\vec{v} = (1, 0, 1, 1)$ *sobre* H.

d) *Calcula la distancia de* $\vec{v}$ *a* H.

Ejercicio 2. *Considerado en* $\mathbb{R}^4$ *el subespacio vectorial* $W =< (1, 0, -1, 0), (1, 1, 0, -1) >$ *y el vector* $\vec{v} = (1, 0, 0, 13)$, *se pide:*

a) *Una base de* $W^{\perp}$.

b) *El vector* $\vec{v_1} \in W$ *más cercano a* $\vec{v}$.

c) *El vector* $\vec{v_2} \in W^{\perp}$ *más cercano a* $\vec{v}$.

d) *La distancia de* $\vec{v}$ *a* W *(la expresión más simplificada posible).*

Ejercicio 3. *Obtén la* **solución de mínimos cuadrados** *del SL* $A\vec{x} = \vec{b}$, *con* $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}$ *y* $\vec{b} = (8, 7, 3, -4)$.

La solución de mínimos cuadrados es la del sistema compatible determinado $A\vec{x} = \hat{\vec{b}}$, *siendo* $\hat{\vec{b}}$ *la proyección ortogonal de* $\vec{b}$ *sobre el subespacio generado por las columnas de* A, *ColA.*

$\hat{\vec{b}} = \mathrm{proy}_{\mathrm{ColA}}\vec{b} = A(A^tA)^{-1}A^t\,\vec{b}$

Por tanto $A\vec{x} = A(A^tA)^{-1}A^t\,\vec{b}$

$A(\vec{x} - (A^tA)^{-1}A^t\vec{b}) = \vec{0}$, *sacando factor común A.*

$\vec{x} - (A^tA)^{-1}A^t\vec{b} = \vec{0}$ *pues las columnas de A son l.i., y por tanto no existen soluciones distintas de la trivial. por tanto:*

$\vec{x} = (A^tA)^{-1}A^t\vec{b}$

Podemos reescribir la ec. como $(A^tA)\vec{x} = A^t\vec{b}$

Por tanto resolviendo el SL de matriz ampliada $[\,A^tA \,|\, A^t\,\vec{b}\,]$ *obtendríamos, de la forma más directa posible, la solución de mínimos cuadrados.*

Ejercicio 4. *Diagonaliza ortogonalmente la matriz* $A = \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, *es decir, obtén Q ortogonal y D diagonal tales que* $A = QDQ^t$. *Realiza la comprobación.*

Ejercicio 5. *Considera* $A = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$, $\vec{b_1} = (-2, 2, 1)$ *y* $\vec{b_2} = (1, 1, 0)$. *Comprueba que* $\vec{b_1}$ *y* $\vec{b_2}$ *son autovectores de A y diagonaliza ortogonalmente A.*

Ejercicio 6. *a) Obtén la matriz de proyección sobre el subespacio W generado por las columnas de la matriz* $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}$*, utilizando la expresión* $A_{\text{proy}} = UU^t$ *siendo U una base ortonormal de W. b) Obtén la proyección ortogonal sobre W del vector* $\vec{b} = (8, 7, 3, -4)$*, usando la matriz anterior. (Este ejercicio tiene relación con el Ejercicio 3).*

11 Polinomio interpolador

Dado un conjunto de datos experimentales (x_i, y_i) con i desde 1 hasta n y tales que $x_i \neq x_j$ para $i \neq j$, llamamos **polinomio interpolador** de esos datos al polinomio de menor grado cuyo gráfico pasa por todos los puntos. En trabajos científicos un polinomio de este tipo puede usarse, por ejemplo, para estimar valores entre puntos conocidos, y de ahí su nombre de polinomio interpolador.

Veremos a continuación que dicho polinomio es único y que su grado es menor o igual que $n-1$.

El modelo de un polinomio de grado $n-1$ es:

$$y = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1}$$

La condición de que todos los puntos cumplan el modelo significa que se ha de cumplir el siguiente sistema de ecuaciones:

$$\begin{cases} a_0 + a_1 x_1 + a_2 x_1^2 + \ldots + a_{n-1} x_1^{n-1} = y_1 \\ a_0 + a_1 x_2 + a_2 x_2^2 + \ldots + a_{n-1} x_2^{n-1} = y_2 \\ \ldots \\ a_0 + a_1 x_n + a_2 x_n^2 + \ldots + a_{n-1} x_n^{n-1} = y_n \end{cases},$$

$$\text{con matriz de coeficientes} \quad A = \begin{bmatrix} 1 & x_1 & {x_1}^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & {x_2}^2 & \ldots & x_2^{n-1} \\ \ldots & \ldots & \ldots & & \\ 1 & x_3 & {x_n}^2 & \ldots & x_n^{n-1} \end{bmatrix}$$

Esta matriz es la llamada matriz de Vandermonde de orden n, que tiene rango n si y solo si $x_i \neq x_j$ para $i \neq j$ (visto en el Ejercicio 13 del Tema 2). Al ser cuadrada de orden n y rango completo, el SL es compatible determinado. Por tanto obtendremos una solución única $(a_o, a_1, a_2, \ldots, a_{n-1})$. El grado del polinomio lo determinará el mayor coeficiente no nulo. En efecto, aunque el modelo que estamos usando es un polinomio de grado $n-1$, los polinomios de menor grado están incluidos en el modelo, y simplemente tienen los coeficientes más altos nulos.

Por ejemplo si tenemos $n = 3$ y los tres puntos están alineados y con pendiente no nula el polinomio interpolador que deduciremos será de orden 1 (recta), y si además de estar alineados tienen pendiente nula, entonces el polinomio es de orden 0 ($y = constante$).

Si consideráramos un polinomio de grado $t > n-1$, tendríamos una columna más por cada grado que aumentamos, mientras que la matriz A sigue teniendo rango n, igual al número de ecuaciones (=filas), por tanto el sistema pasa a ser indeterminado, con infinitos polinomios de grado menor o igual que t que pasan por los puntos.

Ejemplo 14. Encuentra el polinomio interpolador para el conjunto de datos (x, y) siguiente:

(0,1), (3,4), (6,5)

Método de solución: Resolver el SL $A\vec{a} = \vec{y}$, siendo

A la matriz de Vandermonde para la primera variable $\vec{x} = (0, 3, 6)$

$\vec{y} = (1, 4, 5)$ los valores de la segunda variable.

$\vec{a} = (a_0, a_1, a_2)$ los coeficientes del polinomio de grado 2, que son las incógnitas.

12 Ajuste de datos (x, y) mediante polinomios

En una serie de datos científicos (x, y), esperamos errores experimentales, por lo que la relación verdadera entre las variables en general va a quedar mejor descrita por un polinomio de grado más bajo que el del polinomio interpolador.

Se describe a continuación como se obtienen las modelizaciones o ajustes, mediante mínimos cuadrados, con un polinomio de grado $p < n$, siendo n el número de datos.

Modelo polinomio de grado p: $y = a_0 + a_1 x + a_2 x^2 + \ldots + a_p x^p$

Datos: $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ con $x_i \neq x_j \ \forall i \neq j$

Si el conjunto de valores y_i cumpliera el modelo, el siguiente SL sería compatible.

$$\begin{cases} a_0 \ + \ a_1 \ x_1 \ + \ a_2 \ (x_1)^2 \ + \ ... \ + \ a_p \ (x_1)^p = y_1 \\ a_0 \ + \ a_1 \ x_2 \ + \ a_2 \ (x_2)^2 \ + \ ... \ + \ a_p \ (x_2)^p = y_2 \\ ... \\ a_0 \ + \ a_1 \ x_n \ + \ a_2 \ (x_n)^2 \ + \ ... \ + \ a_p \ (x_n)^p = y_n \end{cases}$$

Sistema en notación matricial:

$$\underbrace{\begin{bmatrix} 1 & x_1 & {x_1}^2 & ... & (x_1)^p \\ 1 & x_2 & {x_2}^2 & ... & (x_2)^p \\ \vdots & \ldots & \ldots & \ddots & \vdots \\ 1 & x_n & {x_n}^2 & ... & (x_n)^p \end{bmatrix}}_{A} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Si el conjunto de puntos no cumple el modelo, la solución de mínimos cuadrados, es decir, el vector de coeficientes del polinomio de grado p que mejor ajusta los datos, se obtiene resolviendo el sistema $A^t \ A \ \vec{a} = A^t \ \vec{y}$ (Ver la justificación en el enunciado del ejercicio 3. La condición de que las columnas de A sean independientes se cumple por haber tomado todos los x_i distintos).

Ejemplo 15. Encuentra el mejor ajuste por mínimos cuadrados mediante una función lineal para el conjunto de datos (x, y) siguiente:

(0,1), (3,4), (6,5)