TEMA 6

FORMAS BILINEALES Y PRODUCTO ESCALAR

Índice

6.1. Formas bilineales

Hasta ahora hemos visto aplicaciones lineales entre espacios vectoriales V y E ambos sobre un mismo cuerpo $\mathbb{R}$. Ahora bien, $\mathbb{R}$ es también un espacio vectorial (sobre sí mismo) de dimensión 1.Las aplicaciones lineales $f\colon V \to \mathbb{R}$ reciben el nombre de *formas lineales* (o 1-formas).

Ejercicio 6.1. *¿De qué orden es la matriz de una forma lineal definida sobre un espacio vectorial de dimensión n?*

Podemos extender este concepto a funciones con más de una entrada. Si V y E son espacios vectoriales, y una función $f\colon V \times V \to E$ es lineal para cada una de las entradas, se dice que f es una *aplicación bilineal*. Si además $E = \mathbb{R}$, entonces se dice que es una *forma bilineal*. Damos la definición explícita.

Una forma es bilineal si tiene dos entradas de vectores y, fijada una de ellas, la forma que queda es lineal, es decir, para cada vector $\vec{x}_0 \in V$ obtenemos dos formas lineales

$$\begin{array}{rccc} f_{\vec{x}_0}\colon & V & \to & \mathbb{R} \\ & \vec{x} & \to & f(\vec{x}_0, \vec{x}) \end{array}$$

$$\begin{array}{rccc} f^{\vec{x}_0}\colon & V & \to & \mathbb{R} \\ & \vec{x} & \to & f(\vec{x}, \vec{x}_0) \end{array}$$

Definición 6.2. Sea V un espacio vectorial de dimensión n sobre $\mathbb{R}$. Una *forma bilineal* es una función $f\colon V \times V \to \mathbb{R}$ tal que para cualquier combinación lineal $\sum_{i=1}^{k} \lambda_i \vec{v}_i$ de vectores de V se tiene:

1. $f(\sum_{i=1}^{k} \lambda_i \vec{v}_i, \vec{y}) = \sum_{i=1}^{k} \lambda_i f(\vec{v}_i, \vec{y})$ para cada $\vec{y} \in V$.
2. $f(\vec{x}, \sum_{i=1}^{k} \lambda_i \vec{v}_i) = \sum_{i=1}^{k} \lambda_i f(\vec{x}, \vec{v}_i)$ para cada $\vec{x} \in V$.

Ejemplo 6.3. La función $f\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ definida analíticamente

$$f\left((x_1, x_2, x_3), (y_1, y_2, y_3)\right) = x_1 y_2 - x_2 y_3$$

es una forma bilineal, puesto que, si fijamos el vector $\vec{x}_0 = (a, b, c)$, las funciones

$$f_{\vec{x}_0}(y_1, y_2, y_3) = a y_2 - b y_3 \quad \text{y} \quad f^{\vec{x}_0}(x_1, x_2, x_3) = b x_1 - c x_2$$

son formas lineales.

Definición 6.4. Una forma bilineal $f\colon V \times V \to \mathbb{R}$ se dice que es:

- *Simétrica*, si $f(\vec{x}, \vec{y}) = f(\vec{y}, \vec{x})$ para cada para de vectores $\vec{x}$ e $\vec{y}$.
- *Antisimétrica*, si $f(\vec{x}, \vec{y}) = -f(\vec{y}, \vec{x})$ para cada para de vectores $\vec{x}$ e $\vec{y}$.

A las formas bilineales (especialmente a las antisimétricas) también se les llama 2-formas.

Veamos que las formas bilineales antisimétricas quedan caracterizadas por las imágenes de los pares $(\vec{x}, \vec{x})$:

Proposición 6.5. *Una forma bilineal f definida en V es antisimétrica si y solo si $f(\vec{x}, \vec{x}) = 0$ para cada $\vec{x} \in V$.*

Demostración. Si f es antisimétrica $f(\vec{x},\vec{x}) = -f(\vec{x},\vec{x})$ luego $f(\vec{x},\vec{x}) = 0$. Inversamente, supongamos que $f(\vec{x},\vec{x}) = 0$, para todo vector $\vec{x}$, entonces

$$0 = f(\vec{v}+\vec{w},\vec{v}+\vec{w}) = f(\vec{v},\vec{w}) + f(\vec{w},\vec{v})$$

luego f es antisimétrica. □

Ejercicio 6.6. *Usa la proposición anterior para encontrar lo valores del parámetro "a" para la siguiente forma bilineal en $\mathbb{R}^2$ es antisimétrica.*

$$f((x_1,x_2),(y_1,y_2)) = ax_1y_1 + 3x_1y_2 - 2x_2y_1$$

6.1.1. Representación matricial de una forma bilineal

Sea f una forma bilineal definida en un espacio V de dimensión finita n y sea $\mathcal{B} = \{\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n\}$ una base de V. Tenemos

$$f(\vec{x},\vec{y}) = f(\sum_{i=1}^{n} x_i\vec{e}_i, \sum_{j=1}^{n} y_j\vec{e}_j) = \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} f(\vec{e}_i,\vec{e}_j)y_j\right) = \sum_{i,j}^{n} x_i f(\vec{e}_i,\vec{e}_j)y_j$$

que matricialmente se puede expresar del siguiente modo

$$f(\vec{x},\vec{y}) = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} f(\vec{e}_1,\vec{e}_1) & f(\vec{e}_1,\vec{e}_2) & \dots & f(\vec{e}_1,\vec{e}_n) \\ f(\vec{e}_2,\vec{e}_1) & f(\vec{e}_2,\vec{e}_2) & \dots & f(\vec{e}_2,\vec{e}_n) \\ \vdots & \vdots & & \vdots \\ f(\vec{e}_n,\vec{e}_1) & f(\vec{e}_n,\vec{e}_2) & \dots & f(\vec{e}_n,\vec{e}_n) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

por tanto, continuando con la notación habitual, si X e Y son las matrices columnas formadas por las coordenadas de los vectores $\vec{x}$ e $\vec{y}$, respectivamente, una forma bilineal se puede escribir de una manera compacta como

$$\boxed{f(\vec{x},\vec{y}) = X^T \cdot A \cdot Y.}$$

Y a la matriz A se le denomina *matriz de la forma bilineal f respecto de la base $\mathcal{B}$.*

Ejercicio 6.7. *Sea $f\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ la forma bilineal*

$$f((x_1,x_2),(y_1,y_2)) = 2x_1y_1 - 3x_1y_2 + x_2y_2$$

Determina la matriz de f respecto a las bases:

1. *Canónica.*
2. $\mathcal{B} = \{(1,-1),(3,-2)\}$

Formas bilineales degeneradas

Una forma bilineal se anula cuando una de las entradas es el vector $\vec{0}$, veamos

$$f(\vec{x},\vec{0}) = f(\vec{x},\vec{0}+\vec{0}) = f(\vec{x},\vec{0}) + f(\vec{x},\vec{0}) \Rightarrow f(\vec{x},\vec{0}) = 0$$

análogamente, $f(\vec{0},\vec{y}) = 0$.

Pero también puede ocurrir que $f(\vec{x},\vec{y})$ se anule siendo $\vec{x}$ e $\vec{y}$ vectores distintos de cero.

El vector $\vec{x}_0$ es *degenerado* para f cuando alguna de las formas lineales $f_{\vec{x}_0}$ o $f^{\vec{x}_0}$ es nula (véase definición 6.2 y ejemplo 6.3).

Definición 6.8. Decimos que un vector $\vec{x}_0 \neq \vec{0}$ es *degenerado* para la forma bilineal f si $f(\vec{x}_0,\vec{y}) = 0$ para cada $\vec{y} \in V$ o bien $f(\vec{x},\vec{x}_0) = 0$ para cada $\vec{x} \in V$.

Ejercicio 6.9. *Encuentra una forma bilineal en $\mathbb{R}^2$ que tenga un vector degenerado.*

Definición 6.10. Se dice que una forma bilineal f es *degenerada* si posee algún vector degenerado. Por tanto, una forma bilineal es *no degenerada* si cumple las dos condiciones siguientes (ambas)

- Para todo $\vec{x}$, $f(\vec{x},\vec{y}) = 0 \Longrightarrow \vec{y} = \vec{0}$
- Para todo $\vec{y}$, $f(\vec{x},\vec{y}) = 0 \Longrightarrow \vec{x} = \vec{0}$

Ejercicio 6.11. *Determina si la forma bilineal f del ejemplo 6.3 es o no es degenerada.*

Podemos clasificar las formas bilineales por su matriz, como vemos en el siguiente teorema.

Teorema 6.12. *Sea f una forma bilineal definida sobre un espacio vectorial finito y A su matriz respecto a alguna base, entonces*

1. *f es* simétrica *si y solo si A es simétrica, es decir, $A = A^T$.*
2. *f es* antisimétrica *si y solo si A es antisimétrica, es decir, $A = -A^T$.*
3. *f es* degenerada *si y solo si A es singular, es decir* $\det A = 0$

Demostración. Las dos primeras equivalencias son triviales y se dejan como ejercicio. Veamos la tercera equivalencia.

Supongamos que f es degenerada, o sea, existe $\vec{x}_0 \neq \vec{0}$ de forma que $f(\vec{x},\vec{x}_0) = 0$ (o bien $f(\vec{x}_0,\vec{x}) = 0$) para cualquier vector $\vec{x}$. Esto se traduce en que

$$X^T A X_0 = 0 \text{ para alguna matriz columna } X_0 \neq 0$$

y esto ocurre para cualquier X, por tanto la matriz columna AX_0 tiene que ser 0. Luego el sistema homogéneo $AX = 0$ es (compatible) indeterminado, por tanto, como sabemos, $\det A = 0$.

Inversamente, si $\det A = 0$, el sistema de $AX = 0$ tiene solución X_0 distinta de 0, por tanto $X^T(AX_0) = 0$, que significa $f(\vec{x}, \vec{x}_0) = 0$ para cualquier vector $\vec{x}$, es decir, f es degenerada. □

Clasificación por el signo

Definición 6.13 (Formas definidas)**.** Una forma bilineal f se dice

- *Definida positiva* si $f(\vec{x}, \vec{x}) > 0$ para todo $\vec{x} \neq \vec{0}$.
- *Definida negativa* si $f(\vec{x}, \vec{x}) < 0$ para todo $\vec{x} \neq \vec{0}$.

Proposición 6.14. *Las formas bilineales definidas (positiva o negativa) son no degeneradas.*

Demostración. Si f es degenerada existe $\vec{x}_0 \neq \vec{0}$ y $f(\vec{x}_0, \vec{x}) = 0$ para cada vector $\vec{x} \in V$, entonces $f(\vec{x}_0, \vec{x}_0) = 0$, que contradice que f sea definida. □

Definición 6.15 (Formas semidefinidas)**.** Una forma bilineal f es

- *Semidefinida positiva* si $f(\vec{x}, \vec{x}) \geq 0$ para todo $\vec{x}$ y no es definida positiva.
- *Semidefinida negativa* si $f(\vec{x}, \vec{x}) \leq 0$ para todo $\vec{x}$ y no es definida negativa.

Proposición 6.16. *Las formas bilineales semidefinidas (positiva o negativa) son degeneradas.*

Demostración. Evidente. □

El resto de las formas bilineales reales no tienen signo, así:

Definición 6.17 (Formas indefinidas)**.** Una forma bilineal f es *indefinida* si existen $\vec{x}, \vec{y} \in V$ con $f(\vec{x}, \vec{x}) > 0$ y $f(\vec{y}, \vec{y}) < 0$.

Nota. Las formas bilineales indefinidas pueden ser degeneradas o no degeneradas.

Ejercicio 6.18. *Da ejemplos de formas bilineales reales en $\mathbb{R}^2$ que sean definidas, semidefinidas e indefinidas.*

Matrices definidas y semidefinidas

Diremos que una matriz cuadrada A es definida positiva si la forma bilineal que define es definida positiva. Es decir,

$$X^T AX > 0 \text{ para cada } X \neq 0$$

Análogamente para matriz definida negativa, semidefinida positiva y semidefinida negativa.

Cambio de base

Proposición 6.19. *Sea f una forma bilineal y sean $\mathcal{B}_1$ y $\mathcal{B}_2$ dos bases distintas de V. Si A y B son las matrices de f asociadas a las bases $\mathcal{B}_1$ y $\mathcal{B}_2$, respectivamente, entonces existe una matriz invertible P tal que*

$$B = P^T AP \tag{6.1}$$

La congruencia de matrices es una relación de equivalencia en el conjunto de las matrices cuadradas. Es un caso particular de la equivalencia de matrices, puesto que existe una matriz Q de forma que $P^T = Q^{-1}$, pero es distinta de la semejanza.

Las matrices A y B que se relacionan mediante (6.1) se dice que son *congruentes*.

Demostración. Sea P la matriz del cambio de base de $\mathcal{B}_2$ a $\mathcal{B}_1$. Dados dos vectores $\vec{x}$ e $\vec{y}$ con coordenadas X e Y respecto a $\mathcal{B}_1$ y $\widehat{X}$ e $\widehat{Y}$ respecto a $\mathcal{B}_2$, tenemos $X = P\widehat{X}$ e $Y = P\widehat{Y}$, entonces

$$f(\vec{x},\vec{y}) = X^T AY = (P\widehat{X})^T A(P\widehat{Y}) = \widehat{X}^T \overbrace{(P^T AP)}^{B} \widehat{Y} = \widehat{X}^T B\widehat{Y}$$

de donde $B = P^T AP$. □

Ejercicio 6.20. *Comprueba que las matrices de la aplicación bilineal obtenidas en el ejercicio 6.7 son congruentes.*

6.1.2. Formas multilineales reales

El concepto de forma bilineal se extiende a más de dos entradas, así

Definición 6.21. Llamamos forma multilineal (de orden m) a una aplicación $f\colon V\times \overset{m \text{ veces}}{\cdots} \times V \to \mathbb{R}$ que es lineal en cada una de las m entradas, es decir en cada entrada $i = 1, \dots, m$

$$f(\vec{x}_1, \dots, \sum_k \alpha_k \vec{x}_{i_k}, \dots, \vec{x}_m) = \sum_k \alpha_k f(\vec{x}_1, \dots, \vec{x}_{i_k}, \dots, \vec{x}_m).$$

El equivalente a las formas bilineales antisimétricas son las llamadas formas multilineales alternadas.

Definición 6.22. Una forma multilineal f es alternada, si para cualesquiera i, j se tiene

$$f(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_m) = -f(\vec{x}_1, \dots, \vec{x}_j, \dots, \vec{x}_i, \dots, \vec{x}_m)$$

Es decir, permutando dos entradas, la forma cambia de signo.

Teorema 6.23. *Dada una base ordenada $\{\vec{e}_1, \vec{e}_2, \dots . \vec{e}_n\}$ de un espacio vectorial V, existe una* única *forma multilineal alternada D de orden n que verifica*

$$D(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n) = 1$$

A dicha n-forma se le llama determinante.

Así, si $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$ son vectores de V, expresadas sus coordenadas respecto a la base $\mathcal{B}$ del teorema anterior en la siguiente matriz cuadrada $M = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}$ formada por las coordenadas de los vectores puestos en columna, se tiene

$$\boxed{D(\vec{x}_1, \dots, \vec{x}_n) = \det \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}}$$

De la definición y del teorema anterior se derivan todas las propiedades que conoces de los determinantes.

6.2. Formas cuadráticas

Definición 6.24. Sea V un espacio vectorial *real* de dimensión n. Una *forma cuadrática* es una función $q\colon V \to \mathbb{R}$ que verifica

1. $q(\lambda \vec{x}) = \lambda^2 q(\vec{x}) \quad \forall \lambda \in \mathbb{R}, \vec{x} \in V$
2. La función $f_q\colon V \times V \to \mathbb{R}$ definida

$$\boxed{f_q(\vec{x}, \vec{y}) = \frac{1}{2}\left(q(\vec{x} + \vec{y}) - q(\vec{x}) - q(\vec{y})\right)}$$

   es una forma bilineal que recibe el nobre de *forma polar asociada a la forma cuadrática*.

Es fácil comprobar que la forma polar f_q es *simétrica*. Efectivamente,

Demostración.

$$\begin{aligned} f_q(\vec{x}, \vec{y}) &= \frac{1}{2}\left(q(\vec{x} + \vec{y}) - q(\vec{x}) - q(\vec{y})\right) = \frac{1}{2}\left(q(\vec{y} + \vec{x}) - q(\vec{y}) - q(\vec{x})\right) = \\ &= f_q(\vec{y}, \vec{x}) \end{aligned}$$ □

Equivalentemente, si f es una forma bilineal simétrica, entonces $q(\vec{x}) = f(\vec{x},\vec{x})$ es una forma cuadrática, tal que $f_q = f$.

Demostración. Efectivamente,

$$\begin{aligned} f_q(\vec{x},\vec{y}) &= \frac{1}{2}\left(f(\vec{x}+\vec{y},\vec{x}+\vec{y}) - f(\vec{x},\vec{x}) - f(\vec{y},\vec{y})\right) = \\ &= \frac{1}{2}\left(f(\vec{x},\vec{x}) + f(\vec{x},\vec{y}) + f(\vec{y},\vec{x}) + f(\vec{y},\vec{y}) - f(\vec{x},\vec{x}) - f(\vec{y},\vec{y})\right) = \\ &= \frac{1}{2}\left(f(\vec{x},\vec{y}) + f(\vec{y},\vec{x})\right) = \\ &= f(\vec{x},\vec{y}). \end{aligned}$$

□

Ejemplo 6.25. La función $q\colon \mathbb{R}^3 \to \mathbb{R}$ definida

$$q(x_1,x_2,x_3) = 2x_1^2 - x_1x_3 - 3x_2x_3 - x_2^2$$

es una forma cuadrática. Vamos a comprobarlo:

1. $q(\lambda\vec{x}) = q(\lambda x_1,\lambda x_2,\lambda x_3) = 2(\lambda x_1)^2 - \lambda x_1\lambda x_3 - 3\lambda x_2\lambda x_3 - (\lambda x_2)^2 = \lambda^2 q(\vec{x})$

2. Veamos su forma polar:

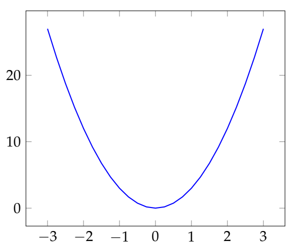
$$\begin{aligned} &f_q((x_1,x_2,x_3),(y_1,y_2,y_3)) = \\ &= \frac{2\,(x_1+y_1)^2 - (x_1+y_1)\,(x_3+y_3) - 3\,(x_2+y_2)\,(x_3+y_3) - (x_2+y_2)^2}{2} - \\ &\quad - \frac{(2x_1^2 - x_1x_3 - 3x_2x_3 - x_2^2)}{2} - \frac{(2y_1^2 - y_1y_3 - 3y_2y_3 - y_2^2)}{2} \end{aligned}$$

que, simplificando, queda una expresión analítica de una forma bilineal

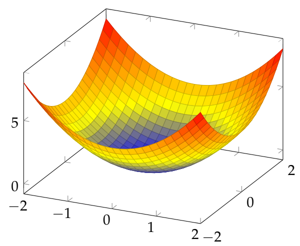
$$\begin{aligned} f_q((x_1,x_2,x_3),&(y_1,y_2,y_3)) = \\ &= 2x_1y_1 - \frac{1}{2}x_1y_3 - x_2y_2 - \frac{3}{2}x_2y_3 - \frac{1}{2}x_3y_1 - \frac{3}{2}x_3y_2 \end{aligned}$$

claramente simétrica.

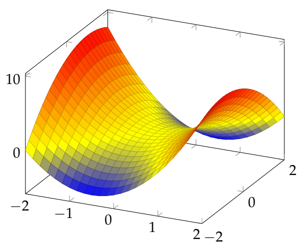
Ejemplo 6.26. Las formas cuadráticas en $\mathbb{R}$ son todas de la forma $q(x) = ax^2$ y en $\mathbb{R}^2$ son de la forma $q(x,y) = ax^2 + bxy + cy^2$. En ambos casos se pueden "dibujar". Aquí tenemos algún ejemplo.



$q(x) = 3x^2$



$q(x,y) = x^2 + y^2$



$q(x,y) = 2x^2 - xy - y^2$

6.2.1. Representación matricial de una forma cuadrática

Si fijamos una base $\mathcal{B}$ obtenemos una matriz cuadrada de orden n que representa la forma polar A, es decir, $f_q(\vec{x}, \vec{y}) = X^T AY$. Esta misma matriz A (simétrica) representa a la forma cuadrática (respecto de la base $\mathcal{B}$).

Definición 6.27. Dada una forma cuadrática q definida en V llamamos *matriz de q respecto de una base $\mathcal{B}$* a la matriz de la forma polar de q respecto de dicha base.

Por tanto, fijada una base del espacio V una forma cuadrática toma la expresión matricial

$$\boxed{q(\vec{x}) = X^T AX}$$

siendo A una *matriz simétrica* que corresponde a la forma cuadrática q.

Ejemplo 6.28. La matriz asociada a la forma cuadrática definida en el ejemplo 6.25 respecto de la base canónica es la siguiente:

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

Ejercicio 6.29. *Calcula las formas analítica y matricial de una forma cuadrática q en $\mathbb{R}^3$ sabiendo que $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ es una base y conociendo algunas imágenes de q:*

$$q(\vec{e}_1) = 2,\ q(\vec{e}_2) = 0,\ q(\vec{e}_3) = -1,$$
$$q(\vec{e}_1 + \vec{e}_2) = 5,\ q(\vec{e}_1 + \vec{e}_3) = 3,\ q(\vec{e}_2 + \vec{e}_3) = 0$$

6.2.2. Clasificación

Las formas cuadráticas se clasifican según su forma polar. Así, decimos que q es definida, semidefinida o indefinida si lo es f_q según las definiciones 6.13 y 6.15. Así queda reflejado en el cuadro 6.1:

Ejemplo 6.30. La forma cuadrática del ejemplo 6.25 anterior, $q(x_1, x_2, x_3) = 2x_1^2 - x_1x_3 - 3x_2x_3 - x_2^2$ es indefinida, puesto que $q(1, 0, 0) = 2 > 0$ y $q(0, 1, 0) = -1 < 0$.

La forma cuadrática $q(x_1, x_2, x_3) = 2x_1^2 + x_3^2$ es semidefinida positiva, pues, claramente $q \geq 0$ y el cero se alcanza: $q(0, 1, 0) = 0$.

No siempre es fácil clasificar una forma cuadrática. En la section 6.6 de este tema pretendemos dar herramientas para hacerlo, pero antes necesitamos el concepto de producto escalar.

Sea$q : V \to \mathbb{R}$ una forma cuadrática. Decimos que q es:

- *definida positiva* si $q(\vec{x}) > 0$ para todo $\vec{x} \neq 0$.
- *definida negativa* si $q(\vec{x}) < 0$ para todo $\vec{x} \neq 0$.
- *semidefinida positiva* si $q(\vec{x}) \geq 0$ para todo $\vec{x} \neq 0$ y no es definida positiva.
- *semidefinida negativa* si $q(\vec{x}) \leq 0$ para todo $\vec{x} \neq 0$ y no es definida negativa.
- *indefinida* cuando existen $\vec{x}, \vec{y} \in V$ con $q(\vec{x}) > 0$ y $q(\vec{y}) < 0$.

Nota. Una forma cuadrática semidefinida positiva y semidefinida negativa necesariamente es la forma cuadrática nula $q = 0$.

Cuadro 6.1: Clasificación de las formas cuadráticas.

6.3. Producto escalar

Definición 6.31. Llamamos producto escalar (o producto interno) de un espacio vectorial real V a cualquier forma bilineal en V simétrica y definida positiva.

Notación: Si $g \colon V \times V \to \mathbb{R}$ es un producto escalar, suele denotarse de las siguientes formas:

$$g(\vec{x}, \vec{y}) = \langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot \vec{y} = (\vec{x} | \vec{y})$$

Durante este curso usaremos casi siempre la notación $\langle \, , \, \rangle$ para representar el producto interno, aunque se puede usar cualquiera de ellas.

Propiedades: Como consecuencia directa de la definición se tiene el producto interno tiene las siguientes propiedades:

- $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$ (simétrico).
- $\langle \vec{x}, \vec{x} \rangle = 0 \iff \vec{x} = \vec{0}$ (no degenerado).
- $\langle \vec{x}, \vec{x} \rangle > 0 \iff \vec{x} \neq \vec{0}$ (definido positivo).

Definición 6.32. Llamaremos *espacio vectorial euclídeo* a cualquier espacio vectorial real V de dimensión finita sobre el que se ha definido un producto interno.

Ejemplos 6.33. Los siguientes espacios vectoriales son euclídeos.

1. Sea $V = \mathbb{R}^n$ y sean $\vec{x}$ e $\vec{y}$ vectores de V cuyas coordenadas respecto a una base fijada se representan por las matrices columnas

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ e } Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

entonces el *producto escalar usual o estándar* está definido como

$$\vec{x} \cdot \vec{y} = X^T Y = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

2. Sea $V = \{f\colon [a,b] \to \mathbb{R} \mid f \text{ es continua en } [a,b]\}$. Definimos un producto interno en V de la siguiente forma:

$$\langle f, g \rangle = \int_a^b f(x) g(x)\, dx$$

3. Sea $V = \mathcal{M}_{m\times n}(\mathbb{R})$, definimos un producto interno en este espacio del siguiente modo:

$$\langle A, B \rangle = \operatorname{tr}(B^T A)$$

donde $\operatorname{tr} M$ es la traza de la matriz cuadrada M, es decir, la suma de los elementos de su diagonal principal.

6.3.1. Norma de un vector

Definición 6.34. Dado un espacio vectorial euclídeo V, llamaremos *norma o módulo* de un vector $\vec{x}$ al número real (positivo o cero) siguiente

$$\|\vec{x}\| = +\sqrt{\langle \vec{x}, \vec{x} \rangle}$$

Nota. En un contexto de vectores libres, la norma de un vector se suele interpretar como su *longitud*. También $\|\vec{x}\|$ se interpreta como "*la distancia*" de un punto al origen de coordenadas, cuando se consideran los vectores como puntos de $\mathbb{R}^n$ (véase definición 6.41).

Ejemplos 6.35.

1. En caso del producto escalar usual $\vec{x} \cdot \vec{y} = x_1 y_1 + \cdots + x_n y_n$ en $\mathbb{R}^n$ se tiene la norma usual o estándar
$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

2. Cualquier producto interno en $\mathbb{R}^n$ define una norma en dicho espacio. Por ejemplo, en $\mathbb{R}^2$, el producto interno
$$(\vec{x} \mid \vec{y}) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$
da la norma siguiente $\|\vec{x}\| = \sqrt{x_1{}^2 - 2x_1x_2 + 2x_2{}^2}$ que es distinta de la norma usual.

Ejercicio 6.36. *Del producto escalar definido en el ejemplo 6.33-item 2, calcula las normas de $f(x) = e^x$ y $g(x) = x - x^2$ definidas en el intervalo $[0,1]$.*

Ejercicio 6.37. *Da una expresión para la norma de $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{R})$ según el producto escalar definido en el ejemplo 6.33-item 3.*

Solución: $\|A\|^2 = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2$

Teorema 6.38 (Propiedades de las normas)**.** *Si V es un espacio vectorial euclídeo, entonces*

1. $\|\vec{x}\| \geq 0$ *y además,* $\|\vec{x}\| = 0 \iff \vec{x} = \vec{0}$.
2. $\|\lambda\vec{x}\| = |\lambda|\,\|\vec{x}\|\ \forall \vec{x} \in V,\ \forall \lambda \in \mathbb{R}$.
3. *(Desigualdad de Schwarz)* $|\langle \vec{x}, \vec{y}\rangle| \leq \|\vec{x}\|\,\|\vec{y}\| \quad \forall\, \vec{x}, \vec{y} \in V$
4. *(Desigualdad triangular)* $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| \quad \forall \vec{x}, \vec{y} \in V$
5. $\|\vec{x} - \vec{y}\| \geq |\|\vec{x}\| - \|\vec{y}\|| \quad \forall \vec{x}, \vec{y} \in V$

Demostración. Vemos los más complicados, dejando el resto de ejercicios:

3. (Desigualdad de Schwarz) Si $\vec{x} = \vec{0}$ o $\vec{y} = \vec{0}$, el resultado es evidente. Supongamos entonces que ambos vectores $\vec{x}$ e $\vec{y}$ no son nulos, definimos entonces el nuevo vector
$$\vec{z} = \frac{\vec{x}}{\|\vec{x}\|} - \frac{\vec{y}}{\|\vec{y}\|}$$
que sabemos que $\langle \vec{z}, \vec{z}\rangle \geq 0$ tenemos:
$$0 \leq \left\langle \frac{\vec{x}}{\|\vec{x}\|} - \frac{\vec{y}}{\|\vec{y}\|}, \frac{\vec{x}}{\|\vec{x}\|} - \frac{\vec{y}}{\|\vec{y}\|} \right\rangle = \frac{\langle \vec{x},\vec{x}\rangle}{\|\vec{x}\|^2} + \frac{\langle \vec{y},\vec{y}\rangle}{\|\vec{y}\|^2} - 2\frac{\langle \vec{x},\vec{y}\rangle}{\|\vec{x}\|\|\vec{y}\|} = 2 - 2\frac{\langle \vec{x},\vec{y}\rangle}{\|\vec{x}\|\|\vec{y}\|}$$
luego $\langle \vec{x}, \vec{y}\rangle \leq \|\vec{x}\|\|\vec{y}\|$.
Si hacemos $\vec{z} = \frac{\vec{x}}{\|\vec{x}\|} + \frac{\vec{y}}{\|\vec{y}\|}$, por el mismo procedimiento obtenemos $\langle \vec{x}, \vec{y}\rangle \geq -\|\vec{x}\|\|\vec{y}\|$. Lo que prueba el resultado.

4. (Desigualdad triangular) Se prueba usando la desigualdad de Schwarz.

$$\begin{aligned}\|\vec{x}+\vec{y}\|^2 = \langle \vec{x}+\vec{y}, \vec{x}+\vec{y}\rangle = \|\vec{x}\|^2 + \|\vec{y}\|^2 + 2\,\langle \vec{x}, \vec{y}\rangle \leq \\ \leq \|\vec{x}\|^2 + \|\vec{y}\|^2 + 2\|\vec{x}\|\,\|\vec{y}\| = (\|\vec{x}\| + \|\vec{y}\|)^2\end{aligned}$$

y, como la norma es positiva (o cero), podemos quitar los cuadrados. □

Definición 6.39. Un vector $\vec{u}$ es *unitario* si $\|\vec{u}\| = 1$.

A partir de cualquier vector $\vec{x}$ distinto de $\vec{0}$ se puede obtener un vector unitario sin más que multiplicarlo por el inverso de su norma

$$\vec{u} = \frac{1}{\|\vec{x}\|}\,\vec{x} \quad \text{(vector normalizado).}$$

A esta operación la llamaremos *normalizar*.

Ejercicio 6.40. *Normaliza la siguiente base de* $\mathbb{R}^3$*:*

$$\mathcal{B} = \{(1,0,1), (0,2,1), (\sqrt{2},1,1)\}$$

Todo espacio euclídeo es también un espacio métrico, es decir, se puede definir una distancia entre cada par de elementos.

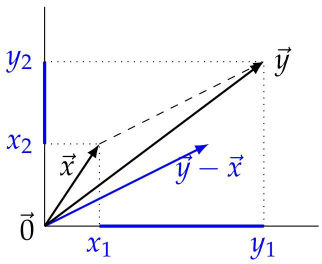
Distancia entre vectores

Definición 6.41 (Distancia)**.** Se define la *distancia entre vectores* $\vec{x}$ e $\vec{y}$ como

$$d(\vec{x},\vec{y}) = \|\vec{y} - \vec{x}\|$$

Nota. Obsérvese que hubiese dado igual definir $d(\vec{x},\vec{y}) = \|\vec{x} - \vec{y}\|$. Se hace así por conveniencia de notación con respecto al espacio afín euclídeo (sección 7.2).

En dicha sección 7.2 veremos como, de un modo general, se puede trasladar el concepto de distancia entre vectores a *distancia entre puntos* de un espacio sin más que considerar los vectores que definen dichos puntos.



Ejemplo 6.42. Si consideramos los vectores $\vec{x}$ e $\vec{y}$ del plano $\mathbb{R}^2$, el producto escalar usual definido en el ejemplo 6.33 nos da la llamada *distancia euclídea* entre dichos vectores del plano

$$d\left((x_1,x_2),(y_1,y_2)\right) = \|(x_1,x_2) - (y_1,y_2)\| = \sqrt{(y_1-x_1)^2 + (y_2-x_2)^2}$$

Ángulos

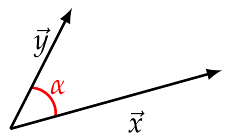
Por otro lado, si dos vectores $\vec{x}$ y $\vec{y}$ no son nulos, la desigualdad de Schwarz se puede traducir como

$$\frac{|\langle \vec{x}, \vec{y} \rangle|}{\|\vec{x}\| \, \|\vec{y}\|} \leq 1 \Longleftrightarrow -1 \leq \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \, \|\vec{y}\|} \leq 1$$

que nos permite identificar este valor con el coseno de un cierto ángulo.

Definición 6.43. Dados dos vectores $\vec{x} \neq \vec{0}$ e $\vec{y} \neq \vec{0}$, definimos el *coseno del ángulo* α formado entre ellos de la forma:

$$\cos \alpha = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \, \|\vec{y}\|} \tag{6.2}$$



El ángulo α que forman dos vectores no nulos, será, por tanto, uno que tome como coseno el valor de la expresión (6.2). De entre todos estos, nos quedamos con el único que verifica $0 \leq \alpha \leq \pi$, es decir, el menor positivo posible.

Según lo anterior, si dos vectores no nulos tienen producto interno 0, forman un ángulo de $\pi/2$ radianes. Estos vectores se dicen que son "*perpendiculares*". Este concepto se generaliza con el de ortogonalidad

6.3.2. Ortogonalidad

Definición 6.44. Dos vectores $\vec{x}$ e $\vec{y}$ de un espacio euclídeo son ortogonales si y solo si $\langle \vec{x}, \vec{y} \rangle = 0$. A veces representamos por $\vec{x} \perp \vec{y}$ para representar que dos vecores son ortogonales.

El símbolo $\perp$ se lee "es ortogonal a" y establece una relación simétrica en V.

Proposición 6.45. *En un espacio euclídeo*

1. *El vector $\vec{0}$ es el único ortogonal a sí mismo.*
2. *El vector $\vec{0}$ es el único ortogonal a todos los vectores del espacio.*

Demostración. Es trivial, puesto que el producto interno es una forma bilineal no degenerada. □

Los conceptos de producto interno y ortogonalidad nos permite demostrar de una manera fácil el teorema más famoso de las matemáticas.

Teorema 6.46 (de Pitágoras generalizado)**.** *Los vectores $\vec{x}$ e $\vec{y}$ son ortogonales si y solo si*

$$\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$$

Demostración. La dejamos como ejercicio. Sigue los mismos pasos que para probar la desigualdad triangular en el teorema 6.38. □

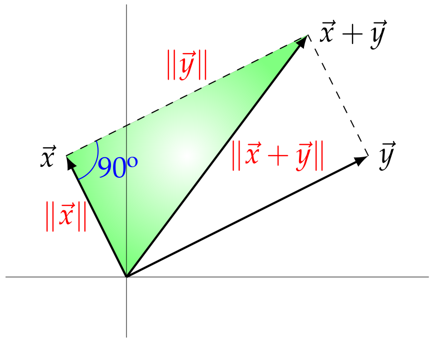


Figura 6.1: El teorema de Pitágoras.

Bases ortonormales

Definición 6.47. Decimos que un sistema de vectores $X = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ es un *sistema ortogonal* si los vectores son ortogonales dos a dos, es decir, $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ siempre que $i \neq j$.

Teorema 6.48. *Si $X = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ es un sistema ortogonal de vectores no nulos, entonces es linealmente independiente.*

Demostración. Supongamos $\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_n\vec{v}_n = \vec{0}$. Entonces para cada i multiplicamos por $\vec{v}_i$:

$$\langle \vec{v}_i, \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_n\vec{v}_n \rangle = \lambda_i \langle \vec{v}_i, \vec{v}_i \rangle = 0$$

y por tanto (como $\vec{v}_i \neq \vec{0}$) se tiene que cada $\lambda_i = 0$. □

Definición 6.49.

1. Llamaremos *base ortogonal* a una base que además es un sistema ortogonal.
2. Llamaremos *base ortonormal* a una base ortogonal de vectores unitarios.

Ejercicio 6.50. *Comprueba que la base canónica en $\mathbb{R}^n$ es ortonormal para el producto escalar usual en cambio no es ortonormal para otros productos internos.*

Encontrar una base ortonormal para un espacio euclídeo es un problema importante para muchas aplicaciones. En caso de dimensión finita, el siguiente teorema nos garantiza su existencia, además de darnos un método para encontrarla. Este método se conoce con el nombre de *método de Gram-Schmidt*.

Teorema 6.51 (Gram-Schmidt). *Todo espacio vectorial euclídeo de dimensión finita admite una base ortogonal.*

Demostración. Partiendo de una base cualquiera $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$, construimos una base ortogonal $\mathcal{B}^* = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ definida en el cuadro 6.2 que prueba el teorema. $\square$

$$
\begin{aligned}
\vec{e}_1 &= \vec{v}_1 \\
\vec{e}_2 &= \vec{v}_2 - \frac{\langle \vec{e}_1, \vec{v}_2 \rangle}{\langle \vec{e}_1, \vec{e}_1 \rangle}\vec{e}_1 \\
\vec{e}_3 &= \vec{v}_3 - \frac{\langle \vec{e}_1, \vec{v}_3 \rangle}{\langle \vec{e}_1, \vec{e}_1 \rangle}\vec{e}_1 - \frac{\langle \vec{e}_2, \vec{v}_3 \rangle}{\langle \vec{e}_2, \vec{e}_2 \rangle}\vec{e}_2 \\
&\vdots \\
\vec{e}_n &= \vec{v}_n - \frac{\langle \vec{e}_1, \vec{v}_n \rangle}{\langle \vec{e}_1, \vec{e}_1 \rangle}\vec{e}_1 - \frac{\langle \vec{e}_2, \vec{v}_n \rangle}{\langle \vec{e}_2, \vec{e}_2 \rangle}\vec{e}_2 - \cdots - \frac{\langle \vec{e}_{n-1}, \vec{v}_n \rangle}{\langle \vec{e}_{n-1}, \vec{e}_{n-1} \rangle}\vec{e}_{n-1}
\end{aligned}
$$

Cuadro 6.2: Método de Gram-Schmidt.

Corolario 6.52. *Todo espacio vectorial euclídeo de dimensión finita admite una base ortonormal.*

Demostración. Cada vector de la base ortogonal $\mathcal{B}^*$ se puede normalizar de la forma $\vec{e}_i^{\,\mathrm{Nor}} = \dfrac{\vec{e}_i}{\|\vec{e}_i\|}$, así obtenemos una nueva base

$$\mathcal{B}^{\mathrm{Nor}} = \{\vec{e}_1^{\,\mathrm{Nor}}, \vec{e}_2^{\,\mathrm{Nor}}, \ldots, \vec{e}_n^{\,\mathrm{Nor}}\}$$

que es ortonormal. $\square$

Ejemplo 6.53. Consideremos el espacio vectorial real $V = P_3[\lambda]$ de los polinomios de coeficientes reales (con variable λ) de grado menor o igual que 3 con el siguiente producto interno

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\, dx$$

Vamos a construir una base ortonormal de este espacio euclídeo.

SOLUCIÓN: Partimos de la base $\mathcal{B} = \{1, \lambda, \lambda^2, \lambda^3\}$ para construir una base ortogonal siguiendo el método de Gram-Schmidt que hemos visto en el teorema anterior.

Entonces,

$$p_1 = 1\,;$$

$$p_2 = \lambda - \frac{\langle 1,\lambda\rangle}{\langle 1,1\rangle}1 = \lambda - \frac{\int_{-1}^{1} x dx}{\int_{-1}^{1} 1 dx} = \lambda - \frac{0}{2} = \lambda\,;$$

$$p_3 = \lambda^2 - \frac{\langle 1,\lambda^2\rangle}{\langle 1,1\rangle}1 - \frac{\langle \lambda,\lambda^2\rangle}{\langle \lambda,\lambda\rangle}\lambda = \lambda^2 - \frac{\int_{-1}^{1} x^2 dx}{\int_{-1}^{1} 1 dx}1 - \frac{\int_{-1}^{1} x^3 dx}{\int_{-1}^{1} x^2 dx}\lambda = \lambda^2 - \frac{1}{3}\,;$$

$$p_4 = \lambda^3 - \frac{\langle 1,\lambda^3\rangle}{\langle 1,1\rangle}1 - \frac{\langle \lambda,\lambda^3\rangle}{\langle \lambda,\lambda\rangle}\lambda - \frac{\langle \lambda^2 - \frac{1}{3},\lambda^3\rangle}{\langle \lambda^2 - \frac{1}{3},\lambda^2 - \frac{1}{3}\rangle}(\lambda^2 - \frac{1}{3}) = \lambda^3 - \frac{3}{5}\lambda$$

Es decir, la base $\mathcal{B}^* = \{1, \lambda, (\lambda^2 - \frac{1}{3}), (\lambda^3 - \frac{3}{5}\lambda)\}$ es ortogonal. Para dar una base ortonormal, calculamos la norma de cada uno de los vectores (polinomios)

$$\|p_1\| = \|1\| = \sqrt{\int_{-1}^{1} 1^2\, dx} = \sqrt{2}\,;$$

$$\|p_2\| = \|\lambda\| = \sqrt{\int_{-1}^{1} x^2\, dx} = \frac{\sqrt{2}}{\sqrt{3}}\,;$$

$$\|p_3\| = \left\|\lambda^2 - \frac{1}{3}\right\| = \sqrt{\int_{-1}^{1}\left(x^2 - \frac{1}{3}\right)^2 dx} = \sqrt{\frac{8}{45}} = \frac{2\sqrt{2}}{3\sqrt{5}}\,;$$

$$\|p_4\| = \left\|\lambda^3 - \frac{3}{5}\lambda\right\| = \sqrt{\int_{-1}^{1}\left(x^3 - \frac{3}{5}x\right)^2 dx} = \sqrt{\frac{8}{175}} = \frac{2\sqrt{2}}{5\sqrt{7}}$$

y normalizamos cada uno de los vectores de la base $\mathcal{B}^*$ para dar la siguiente base ortonormal

$$\boxed{\mathcal{B}^{\text{Nor}} = \left\{\frac{1}{\sqrt{2}}, \frac{\sqrt{3}\lambda}{\sqrt{2}}, \frac{\sqrt{5}\,(3\lambda^2 - 1)}{2\sqrt{2}}, \frac{\sqrt{7}\,(5\lambda^3 - 3\lambda)}{2\sqrt{2}}\right\}}$$

Las bases ortonormales simplifican el cálculo del producto interno. Esto se espresa en la siguiente proposición.

Proposición 6.54. *Sea $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ una base ortonormal. Dados dos vectores $\vec{x}, \vec{y}$ de coordenas $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ e $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ respectivamente, respecto de la base $\mathcal{B}$, entonces el producto escalar toma la expresión del producto escalar usual*

$$\langle \vec{x}, \vec{y}\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n = X^TY = \vec{x}\cdot\vec{y}$$

Demostración. Tenemos que $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n$ y $\vec{y} = y_1\vec{e}_1 + y_2\vec{e}_2 + \cdots + y_n\vec{e}_n$, luego

$$\langle \vec{x}, \vec{y} \rangle = \langle x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n, y_1\vec{e}_1 + y_2\vec{e}_2 + \cdots + y_n\vec{e}_n \rangle$$

y haciendo uso de la bilinealidad de producto escalar junto con

$$\langle \vec{e}_i, \vec{e}_j \rangle = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

obtenemos el resultado requerido. □

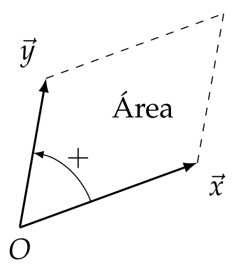
6.3.3. Orientación de un espacio vectorial euclídeo

En teorema 6.23 hemos definido, para un espacio vectorial real V, la forma multilineal determinante $D\colon V \times V \times \cdots \times V \to \mathbb{R}$.

Fijada una base ortonormal $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ y definido el determinante, $D(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) = 1$, decimos *que $\mathcal{B}$ define una orientación*.

Cualquier base tendrá determinante positivo o negativo (nunca nulo), de aquí, se establece una relación de equivalencia en el conjunto de todas las bases:

Nota. Dos bases están relacionadas cuando sus determinantes tienen el mismo signo. Elegida, por tanto, una base tenemos una orientación de V: *positiva* si está en la clase de la base $\mathcal{B}$ (determinante positivo) y *negativa* en caso contrario.



En el espacio vectorial $V = \mathbb{R}^2$ la base $\{(1,0), (0,1)\}$ es positiva. Dada otra base $\{\vec{x}, \vec{y}\} = \{(x_1, x_2), (y_1, y_2)\}$ su determinante es

$$\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1$$

que será positivo cuando el giro del primer vector hacia el segundo se hace en sentido contrario a las agujas del reloj.

Además, el determinante 2×2, cuando la orientación de la base es positiva, nos da el área del paralelogramo definido por dicha base.

Demostración. Como podemos ver en la figura **??** el área del paralelogramo es

$$\text{Area}(\vec{x}, \vec{y}) = \|\vec{x}\| \, \|\vec{y}\| \operatorname{sen} \alpha$$

□

En el espacio vectorial $V = \mathbb{R}^3$, una base $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ tendrá orientación positiva si sigue la regla del sacacorchos: *"al girar $\vec{x}_1$ hacia $\vec{x}_2$ (por el ángulo menor) el sacacorchos seguiría la dirección de $\vec{x}_3$"*.

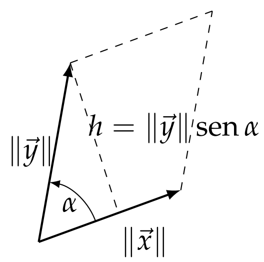
También se suele utilizar la llamada *regla de la mano derecha*. Consiste en señalar la dirección y el sentido del vector $\vec{x}_3$ con el pulgar de dicha mano y el sentido que siguen los restantes dedos indican el sentido positivo de la base si señalan el giro desde el vector $\vec{x}_1$ hacia el vector $\vec{x}_2$.

http://youtu.be/faevbjfs28I

Así ocurre, por ejemplo con la base canónica $\{(1,0,0),(0,1,0),(0,0,1)\}$.

6.3.4. Producto vectorial

En un espacio vectorial euclídeo *tridimensional* orientado tenemos un caso especial de operación de vectores que vamos a describir.



En primer lugar observamos que el *seno del ángulo* (en sentido positivo) de dos vectores independientes, junto con la norma de cada uno de ellos, nos determina el área del paralelogramo, $\operatorname{Area}(\vec{x}, \vec{y})$, que definen (véase figura al margen). Así, si α es el ángulo entre ellos, tenemos

$$\operatorname{Area}(\vec{x}, \vec{y}) = \|\vec{x}\| \, \|\vec{y}\| \operatorname{sen} \alpha$$

Es de notar, también, que (en $\mathbb{R}^3$) el determinante de una base orientada positivamente es el volumen del paralelepípedo que determinan.

$$\operatorname{Volumen}(\vec{x}, \vec{y}, \vec{z}) = D(\vec{x}, \vec{y}, \vec{z})$$

Definición 6.55. Definimos el *producto vectorial* de dos vectores $\{\vec{x}, \vec{y}\}$ del espacio tridimensional como el único vector $\vec{x} \times \vec{y}$ que verifica:

1. $\vec{x} \times \vec{y}$ es ortogonal a ambos vectores,
2. Si $\vec{x}, \vec{y}$ son independientes, $\{\vec{x}, \vec{y}, \vec{x} \times \vec{y}\}$ es una base orientada positivamente.
3. $\|\vec{x} \times \vec{y}\| = \|\vec{x}\| \, \|\vec{y}\| \operatorname{sen} \alpha$.

Si $\vec{x}, \vec{y}$ son independientes, $\|\vec{x} \times \vec{y}\| = \operatorname{Area}(\vec{x}, \vec{y})$.

Del punto 3 se deduce que si los vectores $\{\vec{x}, \vec{y}\}$ son linealmente dependientes, entonces $\vec{x} \times \vec{y} = \vec{0}$. Además son inmediatas las siguientes propiedades del producto vectorial

Proposición 6.56.

1. *El producto vectorial es antisimétrico: $\vec{x} \times \vec{y} = -\vec{y} \times \vec{x}$.*
2. *Si $\vec{x} \neq \vec{0}$ y $\vec{y} \neq 0$, entonces*

$$\vec{x} \times \vec{y} = \vec{0} \Rightarrow \vec{x}, \vec{y} \textit{ son linealmente dependientes.}$$

Expresión analítica del producto vectorial

Dada una base ortonormal de $V = \mathbb{R}^3$ orientada positivamente, los vectores $\vec{x}, \vec{y}$ tendrán unas coordenadas, $\vec{x} = (x_1, x_2, x_3)$ e $\vec{y} = (y_1, y_2, y_3)$. Entonces, el producto escalar tiene las siguientes coordenadas, respecto a dicha base ortonormal

$$\boxed{\vec{x} \times \vec{y} = \left(\begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix}, - \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix}, \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \right)}$$

Demostración. Tenemos que probar que se verican los puntos de la definición 6.55.

1. Como las coordenadas están expresadas en una base ortonormal, se tiene que

$$\langle \vec{x}, \vec{x} \times \vec{y} \rangle = x_1 \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} - x_2 \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} + x_3 \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 & x_2 & y_2 \\ x_2 & x_2 & y_2 \\ x_3 & x_3 & y_3 \end{vmatrix} = 0$$

   luego $\vec{x} \perp \vec{x} \times \vec{y}$. Igualmente $\vec{y} \perp \vec{x} \times \vec{y}$.

2. Basta comprobar que

$$D(\vec{x}, \vec{y}, \vec{x} \times \vec{y}) = \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix}^2 + \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix}^2 + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}^2 > 0 \qquad (6.3)$$

3. Por (6.3) tenemos $\|\vec{x} \times \vec{y}\|^2 = D(\vec{x}, \vec{y}, \vec{x} \times \vec{y}) = \text{Volumen}(\vec{x}, \vec{y}, \vec{x} \times \vec{y})$. Pero se puede comprobar (véase figura 6.2) que

$$\text{Volumen}(\vec{x}, \vec{y}, \vec{x} \times \vec{y}) = \text{Area}^2(\vec{x}, \vec{y})$$

que prueba el teorema. □

Como consecuencia directa de este resultado se tiene las siguientes propiedades del producto vectorial:

1. El producto vectorial es una aplicación bilineal $V \times V \to V$, ya que

$$\vec{x} \times (\vec{y} + \vec{z}) = \vec{x} \times \vec{y} + \vec{x} \times \vec{z}$$
$$\vec{x} \times \lambda \vec{y} = \lambda (\vec{x} \times \vec{y})$$

2. Regla de exclusión: $\vec{x} \times (\vec{y} \times \vec{z}) = \langle \vec{x}, \vec{z} \rangle \vec{y} - \langle \vec{x}, \vec{y} \rangle \vec{z}$.

3. Identidad de Jacobi $\vec{x} \times (\vec{y} \times \vec{z}) + \vec{z} \times (\vec{x} \times \vec{y}) + \vec{y} \times (\vec{z} \times \vec{x}) = 0$.

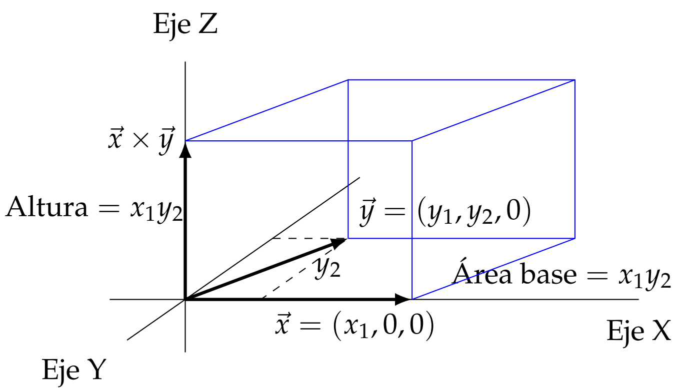


Figura 6.2: El volumen $= (x_1y_2)^2$ es el cuadrado del área de la base.

6.3.5. Aplicaciones ortogonales

Definición 6.57. Diremos que una aplicación $f\colon V \to E$ entre espacios vectoriales euclídeos es *ortogonal* si conserva el producto escalar, es decir

$$\langle f(\vec{x}), f(\vec{y})\rangle = \langle \vec{x}, \vec{y}\rangle \text{ para cada } \vec{x}, \vec{y} \in V \tag{6.4}$$

Evidentemente, las aplicaciones ortogonales mantienen la norma, la distancia y ángulo entre vectores.

En realidad, la condición 6.4 sobre cualquier función f entre espacios vectoriales euclídeos implica automáticamente la linealidad de f.

Demostración. Basta probar: 1. $\|f(\vec{x}+\vec{y}) - f(\vec{x}) - f(\vec{y})\|^2 = 0$, y 2. $\|f(\lambda\vec{x}) - \lambda f(\vec{x})\|^2 = 0$ y, ambas pruebas las dejamos como ejercicio. □

Proposición 6.58. *Las aplicaciones ortogonales son monomorfismos de espacios vectoriales.*

Demostración. Sea $\vec{x} \in \ker f$, entonces $\langle \vec{x}, \vec{x}\rangle = \langle f(\vec{x}), f(\vec{x})\rangle = 0$. Por tanto, la no degeneración del producto escalar implica $\vec{x} = \vec{0}$. □

Especial interés tienen las aplicaciones ortogonales entre espacios vectoriales de la misma dimensión, así

Teorema 6.59. *Una aplicación lineal $f\colon E \to V$ entre espacios vectoriales euclídeos de la misma dimensión es ortogonal si y solo si es lineal y transforma bases ortonormales en bases ortonormales.*

Demostración. Trivial □

Este resultado tiene como consecuencia inmediata que las aplicaciones ortogonales entre espacios de la misma dimensión son invertibles.

Ejercicio 6.60. *Sea V un espacio vectorial euclídeo de dimensión 2 con base ortonormal $\{\vec{e}_1, \vec{e}_2\}$ y f un endomorfismo en V que verifica*

$$f(\vec{e}_1) = \vec{e}_1 + \alpha\vec{e}_2 \text{ y } f(\vec{e}_2) = \beta\vec{e}_1 + \vec{e}_2$$

determina para que valores de α y β es f una aplicación ortogonal.

6.4. Subespacios ortogonales. Complemento ortogonal

Extendemos el concepto de ortogonalidad a un subespacio. Así, en un espacio euclídeo V, diremos que un vector $\vec{x}$ es *ortogonal a un subespacio vectorial* S, si es ortogonal a todos los vectores de S, y lo representamos $\vec{x} \perp S$.

Si S_1 y S_2 son subespacios vectoriales, diremos que son ortogonales, representado $S_1 \perp S_2$, si todos los vectores de S_1 son ortogonales a todos los vectores de S_2.

Proposición 6.61. *Si V es un espacio euclídeo de dimensión n, entonces el conjunto de todos los vectores ortogonales a un vector $\vec{x}_0$ forman un subespacio vectorial de V, que representaremos por $\vec{x}_0^{\perp}$.*

Además, si $\vec{x}_0$ es un vector no nulo, entonces $\dim \vec{x}_0^{\perp} = n - 1$.

Claramente, si $\vec{0}^{\perp} = V$, en otro caso, $\vec{x}_0^{\perp}$ está una dimensión por debajo de V, que es lo que se llama hiperplano.

Demostración. La función $f\colon V \to \mathbb{R}$ definida $f(\vec{x}) = \langle \vec{x}, \vec{x}_0 \rangle$ es lineal y $\vec{x}_0^{\perp}$ es su núcleo, por tanto es subespacio vectorial. Por otro lado, por teorema 5.21, si $\vec{x}_0 \neq \vec{0}$,

$$n = \dim \operatorname{Img} f + \dim \ker f = 1 + \dim x_0^{\perp}$$

prueba la segunda parte. □

Ejemplo 6.62. Si $\vec{x}_0 = (-1, 2, -3)$ es un vector de $\mathbb{R}^3$, vamos a calcular una base de $\vec{x}_0^{\perp}$.

SOLUCIÓN: De $(x_1, x_2, x_3) \cdot (-1, 2, -3) = 0$, tenemos la siguiente ecuación (cartesiana) de $\vec{x}_0^{\perp}$

$$-x_1 + 2x_2 - 3x_3 = 0$$

que se expresa

$$\boxed{\vec{x}_0^{\perp} = L(\{(2,1,0), (-3,0,1)\})}$$

Definición 6.63. Si S es un subespacio vectorial de V, representamos por $S^{\perp}$ el conjunto de los vectores ortogonales a S y que recibe el nombre de *complemento ortogonal* de S.

Teorema 6.64. *El complemento ortogonal $S^{\perp}$ es un subespacio de V y además*

$$V = S \oplus S^{\perp}.$$

Demostración. Consideremos $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_r\}$ una base ortonormal de S. Es fácil ver que

$$S^{\perp} = \bigcap_{i=1}^{r} \vec{e}_i^{\perp}$$

y, por 6.61, $S^{\perp}$ es un subespacio vectorial.

Por otro lado, si $\vec{x} \in S \cap S^{\perp}$, entonces $\langle \vec{x}, \vec{x} \rangle = 0 \Rightarrow \vec{x} = \vec{0}$. Por tanto,

$$S \cap S^{\perp} = \{\vec{0}\}$$

Y, por último, si $\vec{x}$ es un vector de V, definimos el vector

$$\vec{x}_S = \langle \vec{e}_1, \vec{x} \rangle \vec{e}_1 + \langle \vec{e}_2, \vec{x} \rangle \vec{e}_2 + \cdots + \langle \vec{e}_r, \vec{x} \rangle \vec{e}_r \tag{6.5}$$

que claramente es de S. Entonces $\vec{x}_N = \vec{x} - \vec{x}_S$ es ortogonal a S porque

$$\langle \vec{x}_N, \vec{e}_i \rangle = \langle \vec{x}, \vec{e}_i \rangle - \langle \vec{x}_S, \vec{e}_i \rangle = \langle \vec{x}, \vec{e}_i \rangle - \langle \vec{x}, \vec{e}_i \rangle = 0$$

para cada $i = 1, 2, \ldots r$, que nos dice que $\vec{v} = \vec{x}_S + \vec{x}_N \in S + S^{\perp}$. Por tanto, $V = S + S^{\perp}$. □

Ejercicio 6.65. *Si $V = \mathbb{R}^4$, calcula es complemento ortogonal del subespacio vectorial S generado por los vectores $(1, 0, 1, 0)$ y $(1, 1, 0, -1)$.*

Corolario 6.66. *Si S es subespacio vectorial de V, entonces*

$$\left(S^{\perp}\right)^{\perp} = S.$$

Demostración. Si $\vec{x} \in S$ entonces $\vec{x} \perp S^{\perp}$, luego $S \subseteq \left(S^{\perp}\right)^{\perp}$. Recíprocamente, si $\vec{x} \in \left(S^{\perp}\right)^{\perp}$, por el teorema anterior lo expresamos de la forma $\vec{x} = \vec{x}_S + \vec{x}_N \in S \oplus S^{\perp}$. Entonces,

$$0 = \langle \vec{x}, \vec{x}_N \rangle = \langle \vec{x}_S, \vec{x}_N \rangle + \langle \vec{x}_N, \vec{x}_N \rangle = \langle \vec{x}_N, \vec{x}_N \rangle$$

luego $\vec{x}_N = \vec{0}$ y de aquí que $\vec{x} = \vec{x}_S \in S$. □

6.4.1. Proyecciones ortogonales

Una endomorfismo $f\colon V \to V$ se dice que es una proyección si y solo si

$$f^2 = f \circ f = f$$

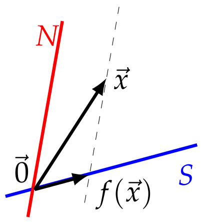
Si $S = \operatorname{Img} f$, entonces la aplicación $f|_S$ es la identidad, pues $f(\vec{x}) = \vec{x}$ para cada $\vec{x} \in S$. Además si $N = \ker f$ tenemos que $S \cap N = \{\vec{0}\}$ y, como $\dim V = \dim N + \dim S$ (véase 5.21) se deduce

$$V = S \oplus N$$

En los términos anteriores, si S es la imagen y N el núcleo de una proyección, se dice f proyecta V en el subespacio S a lo largo del subespacio N.

¡Ojo! una proyección ortogonal no es una aplicación ortogonal.

Definición 6.67. Diremos que f es una *proyección ortogonal* si $N \perp S$, en caso contrario se dice que es una *proyección oblicua*.



En la figura al margen representamos una proyección en el plano en una recta S a lo largo de otra N. Esta proyección sería oblicua, puesto que S y N no son ortogonales.

Definición 6.68. Un endomorfismo $f\colon V \to V$ en un espacio euclídeo se dice que es *autoadjunto* si

$$\langle f(\vec{x}), \vec{y}\rangle = \langle \vec{x}, f(\vec{y})\rangle$$

para cada para de vectores $\vec{x}, \vec{y} \in V$.

Teorema 6.69. *Sea $f\colon V \to V$ una proyección. Entonces, f es* ortogonal *si y solo si es* autoadjunta*.*

Demostración. Veamos que toda proyección ortogonal es autoadjunta, para ello observamos que $\vec{y} - f(\vec{y}) \in \ker f = N$, luego

$$\langle f(\vec{x}), \vec{y} - f(\vec{y})\rangle = 0 \Rightarrow \langle f(\vec{x}), \vec{y}\rangle = \langle f(\vec{x}), f(\vec{y})\rangle$$

Análogamente se prueba que $\langle \vec{x}, f(\vec{y})\rangle = \langle f(\vec{x}), f(\vec{y})\rangle$.

Inversamente, supongamos que f es una proyección autoadjunta. Sean $\vec{x} \in N$ e $\vec{y} \in S = \operatorname{Img} f$, entonces existe un vector $\vec{v}$ tal que $f(\vec{v}) = \vec{y}$, de aquí

$$\langle \vec{x}, \vec{y}\rangle = \langle \vec{x}, f(\vec{v})\rangle = \langle f(\vec{x}), \vec{v}\rangle = \langle \vec{0}, \vec{v}\rangle = 0$$

por tanto, tenemos $N \perp S$, que prueba que f es una proyección ortogonal. $\square$

Obsérvese que si S es un subespacio vectorial de V, podemos definir el endomorfismo $Proy_S\colon V \to V$ como $Proy_S(\vec{x}) = \vec{x}_S$, conforme a (6.5). Claramente $Proy_S$ es una proyección ortogonal, cuyo núcleo $N = \ker(Proy_S)$ es precisamente el complemento ortogonal de $S^\perp$ de S.

Nota: $Proy_N = I - Proy_S$, donde I es la identidad en V.

Ejercicio 6.70. *Prueba que si f es una proyección ortogonal en S a lo largo de N, entonces $f = Proy_S$ (y, por tanto, $N = S^\perp$).*

La proyección ortogonal $Proy_S$ aplicada a un vector de V nos devuelve otro vector que tiene la siguiente propiedad:

Teorema 6.71 (de aproximación óptima)**.** *El vector $Proy_S(\vec{x})$ es el vector de S que más se aproxima a $\vec{x}$.*

Demostración. Para cualquier $\vec{v} \in S$ sabemos que $\vec{v} - \vec{x}_S \in S$ y que $\vec{x}_S - \vec{x} \in S^\perp$, aplicando el teorema 6.46 de Pitágoras generalizado

$$d^2(\vec{x},\vec{v}) = \|\vec{v}-\vec{x}\|^2 = \|\vec{v}-\vec{x}_S+\vec{x}_S-\vec{x}\|^2 = \|\vec{v}-\vec{x}_S\|^2 + \|\vec{x}_S-\vec{x}\|^2 \geq \\ \geq d^2(\vec{x},\vec{x}_S)$$

que prueba que $\vec{x}_S = Proy_S(\vec{x})$ es el vector que más se aproxima a $\vec{x}$ en S. □

Matriz de una proyección ortogonal

Consideremos una base ortonormal $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ de un espacio vectorial euclídeo V. Si f es una proyección en V, entonces, como hemos visto, f es un endomormismo autoadjunto. Esto significa que su matriz (respecto a la mencionada base ortonormal) es simétrica. Veamos

Teorema 6.72. *Si A es la matriz asociada a un endomorfismo f* respecto de una base ortonormal, *entonces f* es autoadjunto si y solo si A es simétrica.

Demostración. Dados dos vectores $\vec{x}, \vec{y}$ con coordenadas X e Y respecto a la base ortonormal dada, la proposición 6.54 nos dice que el producto escalar se reduce al producto escalar usual, es decir $\langle \vec{x}, \vec{y}\rangle = X^TY$. Por tanto,

- f autoadjunto $\Rightarrow \langle \vec{x}, f(\vec{y})\rangle = \langle f(\vec{x}), \vec{y}\rangle \Rightarrow X^T(AY) = (AX)^TY \Rightarrow X^TAY = X^TA^TY \Rightarrow A = A^T$.
- Recíprocamente, A simétrica $\Rightarrow A = A^T$, luego

$$\langle \vec{x}, f(\vec{y})\rangle = X^T(AY) = X^TA^TY = (AX)^TY = \langle f(\vec{x}), \vec{y}\rangle$$

es decir, f es autoadjunto. □

Sea que S es un subespacio vectorial de V de base $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_r\}$ (no necesariamente ortonormal). Consideremos la matriz $A \in \mathcal{M}_{n\times r}$ formada por las coordenadas de dichos vectores respecto de una base ortornormal. Usaremos dicha matriz A para dar una expresión de la matriz P del endomorfismo $Proy_S$. En primer lugar tenemos el siguiente lema.

Lema 6.73. *La matriz $A^T A \in \mathcal{M}_{r\times r}$ es invertible, por tanto existe $(A^T A)^{-1}$.*

Demostración. Si no fuese invertible, el sitema $(A^T A)X = 0$ tendría solución $X_0 \neq 0$, luego el producto escalar

$$\langle AX_0, AX_0\rangle = (AX_0)^T(AX_0) = X_0^T(A^T A)X_0 = X_0^T 0 = 0$$

de donde $AX_0 = 0$. Esto implica que el rango de A es menor que r, pero esto es imposible porque sus r columnas son linealmente independientes (A está construida a partir de una base). $\square$

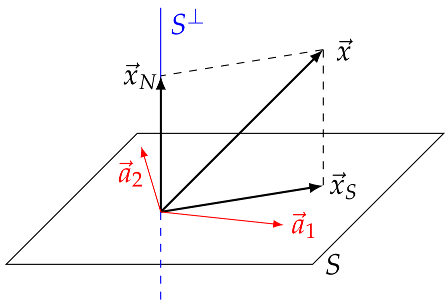


Figura 6.3: Proyección ortogonal del espacio en un plano S.

Dado un vector $\vec{x} = \vec{x}_S + \vec{x}_N \in S \oplus S^\perp$, como cada vector de la base $\vec{a}_i$ de la base de S es ortogonal a $\vec{x}_N$, en coordenadas respecto de la base de V (ortogonal) podemos exscribir $A^T x_N = 0$, luego

$$A^T x_N = A^T(x - x_S) = 0 \Rightarrow A^T x = A^T x_S$$

Ahora bien, como $\vec{x}_S \in S$, existen r números reales α_i tales que

$$\vec{x}_S = \alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \cdots + \alpha_r\vec{a}_r$$

que, en coordenadas, nos da la expresión $x_S = A\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_r\end{pmatrix} = A\alpha$, luego

$$A^T x = A^T A\alpha \Rightarrow (A^T A)^{-1}A^T x = \alpha$$

entonces $A(A^T A)^{-1}A^T x = A\alpha = x_S$ y, como $Proy_X(\vec{x}) = \vec{x}_S$, tenemos la matriz P de la proyección ortogonal $Proy_S$,

Teorema 6.74. *La matriz de la proyección ortogonal en un subespacio vectorial S es*

$$P = A(A^T A)^{-1} A^T$$

siendo A la matriz cuyas columnas son las coordenadas de una base de S respecto de una base ortonormal del espacio.

Si elegimos como base de S una base ortonormal, entonces la matriz $P = AA^T$.

Ejemplo 6.75. La proyección de $\mathbb{R}^3$ en el plano $S = L(\{1,1,0),(-1,2,1)\}$ está definida por la matriz

$$P = \begin{pmatrix} 1 & -1 \\ 1 & 2 \\ 0 & 1 \end{pmatrix} \left[\begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 2 \\ 0 & 1 \end{pmatrix} \right]^{-1} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{10}{11} & \frac{1}{11} & -\frac{3}{11} \\ \frac{1}{11} & \frac{10}{11} & \frac{3}{11} \\ -\frac{3}{11} & \frac{3}{11} & \frac{2}{11} \end{pmatrix}$$

6.5. Diagonalización ortogonal

6.5.1. Endomorfismos ortogonales

En la subsección 6.3.5 ya hemos estudiado las aplicaciones ortogonales. En el caso que una aplicación ortogonal sea un endomorfismo en un espacio vectorial euclídeo de dimensión finita, respecto a una base, que podemos suponer ortonormal, le corresponde una matriz cuadrada invertible, que diremos que es ortogonal, más concretamente:

Definición 6.76. Decimos que una matriz cuadrada P de orden n es *ortogonal* si y solo si

$$P^T P = I_n.$$

P es ortogonal sii $P^{-1} = P^T$.

Del teorema 6.59 se dedude inmediatamente el siguiente resultado que nos permite reconocer fácilmente las matrices ortogonales.

Teorema 6.77. *Las matrices de los endomorfismos ortogonales respecto a bases ortonormales son la matrices ortogonales. Por tanto una matriz P es ortogonal si y solo si las columnas (resp. filas) de P forman una base ortonormal respecto al producto escalar usual.*

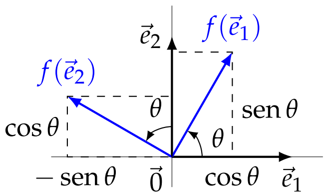
Demostración. Es trivial, puesto que las imágenes de los vectores de una base ortonormal también lo es, luego al multiplicar $P^T P$ estamos realizando los productos escalares de los elementos de dicha base imagen. $\square$

Corolario 6.78. *Si P es ortogonal, entonces* $\det P = \pm 1$.

Demostración. $\det(P^T P) = \det I_n \Rightarrow \det P^T \det P = (\det P)^2 = 1$. $\square$

Endomorfismos ortogonales en $\mathbb{R}^2$

Giros (o rotaciones) de vectores. Un endomorfismo f en $\mathbb{R}^2$ se dice que es un *giro vectorial* si cada vector $\vec{x}$ forma un ángulo constante α con $f(\vec{x})$.



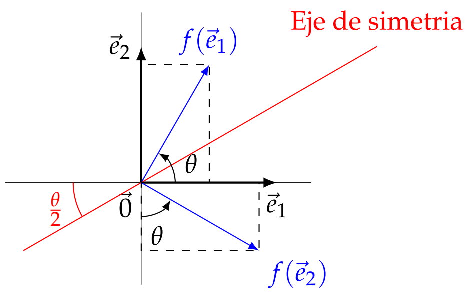
Es muy fácil comprobar que la matriz de un giro en $\mathbb{R}^2$, respecto a una base ortonormal, es de la forma

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

donde θ es el ángulo de giro.

Ejemplo 6.79. La identidad $I\colon \mathbb{R}^2 \to \mathbb{R}^2$ es un giro de 0 radianes.

Simetrías axiales o reflexiones. Llamaremos simetría axial en $\mathbb{R}^2$ a una aplicación lineal que deja invariantes los vectores de una recta (vectorial) y "refleja" al otro lado de dicha recta al resto de vectores.



Así, como se puede comprobar en el dibujo, para una base ortonormal, las coordenadas de las imágenes de los vectores serán $(\cos\theta, \operatorname{sen}\theta)$ y $(\operatorname{sen}\theta, -\cos\theta)$, siendo $\theta/2$ el ángulo del eje de simetría con el primer vector de la base.

Por tanto, la matriz de una simetría axial será de la forma

$$\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix}$$

Los giros y las simetrías axiales determinan todos los endomorfismos ortogonales en $\mathbb{R}^2$ como reflejamos en el siguiente teorema que damos sin demostración:

Teorema 6.80. *Los únicos endomorfismos ortogonales en* $\mathbb{R}^2$ *son los giros y las reflexiones. Además, si* $P \in M_{2\times 2}$ *es una matriz ortogonal entonces:*

1. *P es un giro si y solo si* $\det P = 1$
2. *P es una reflexión si y solo si* $\det P = -1$.

Endomorfismos ortogonales en $\mathbb{R}^3$

Giros axiales de vectores Un endomorfismo f en $\mathbb{R}^3$ es un giro axial si existe una recta $r = L(\vec{u}_1)$ (con $\|\vec{u}_1\| = 1$) invariante[1] para f, y si $\vec{x} \notin r$, entonces las proyecciones ortogonales en $S = \vec{u}_1^{\perp}$ de los vectores $\vec{x}$ y $f(\vec{x})$ forman un ángulo (constante) θ en S. A la recta r se le llama *eje de giro*.

Si el vector $\vec{u}_1$ es el primero de una base ortonormal $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, en dicha base, la matriz de f toma la expresión

$\vec{u}_2$ y $\vec{u}_3$ son base ortononormal de $S = \vec{u}_1^{\perp}$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

Simetría especular o reflexión Un endomorfismo f en $\mathbb{R}^3$ es una simetría especular si existe un plano $\pi = L(\{\vec{u}_2, \vec{u}_3\})$ (con $\{\vec{u}_2, \vec{u}_3\}$ base ortonormal de π) invariante[2] para f de forma que si $\vec{x} \in \pi^{\perp}$ se tiene $f(\vec{x}) = -\vec{x}$. La matriz de f respecto a la base ortonormal $\{\vec{u}_2 \wedge \vec{u}_3, \vec{u}_2, \vec{u}_3\}$ es

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Composición de giro y simetría Si componemos un giro axial y una simetría especular obtenemos un endomorfismo ortogonal, que respecto a la base adecuada, tiene por matriz

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

Estos tres tipos de endomorfismos caracterizan los endomorfismos ortogonales en $\mathbb{R}^3$, es decir

Teorema 6.81. *Los únicos endomorfismos ortogonales en $\mathbb{R}^3$ son los giros axiales, las simetrías especulares o las composiciones de ambas.*

Nota. Un caso distinguido es la identidad que es un giro axial de 0 radianes. Otro caso distinguido es la composición de una simetría especular con un giro de π radianes que, respecto a la base adecuada, tiene por matriz

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

que recibe el nombre de *simetría central o respecto del un punto*.

[1] $\vec{u}_1$ es autovector para de f para el autovalor $\lambda = 1$.
[2] π es subespacio propio de f para $\lambda = 1$.

El grupo ortogonal

En el tema de estructuras algebraicas ya mencionamos el *grupo general lineal* de grado n, representado $GL(n)$ definido como el conjunto de las matrices cuadradas invertibles junto con la operación producto. Este grupo es isomorfo al grupo de los automorfimos $V \to V$ en un espacio de dimensión n.

El conjunto de las matrices ortogonales de orden n forman un subgrupo de grupo $GL(n)$ que recibe el nombre de *grupo ortogonal* y se representa $O(n)$. Las matrices ortogonales con determinante 1 forman a su vez un subgrupo del grupo $O(n)$ que recibe el nombre de *grupo ortogonal especial* que se representa $SO(n)$.

6.5.2. Diagonalización de matrices simétricas

Definición 6.82. Sea f un endomorfismo de un espacio vectorial euclídeo V. Se dice que f es *simétrico (o autoadjunto)* si y solo si

$$\boxed{\langle \vec{x}, f(\vec{y}) \rangle = \langle f(\vec{x}), \vec{y} \rangle \qquad \forall \vec{x}, \vec{y} \in V.}$$

Los endomorfismos autoadjuntos se relacionan con las matrices simétricas mediante el siguiente teorema.

Teorema 6.83. *Si A es la matriz asociada a un endomorfismo f respecto de una base ortonormal, entonces f es autoadjunto si y solo si A es simétrica.*

Demostración. Dados dos vectores $\vec{x}, \vec{y}$ con coordenadas X e Y respecto a la base ortonormal dada, la proposición 6.54 nos dice que el producto escalar se reduce al producto escalar usual, es decir $\langle \vec{x}, \vec{y} \rangle = X^T Y$. Por tanto,

- f autoadjunto $\Rightarrow$ $\langle \vec{x}, f(\vec{y}) \rangle = \langle f(\vec{x}), \vec{y} \rangle \Rightarrow X^T(AY) = (AX)^T Y \Rightarrow X^T AY = X^T A^T Y \Rightarrow A = A^T$.

- Recíprocamente, A simétrica $\Rightarrow A = A^T$, luego

$$\langle \vec{x}, f(\vec{y}) \rangle = X^T(AY) = X^T A^T Y = (AX)^T Y = \langle f(\vec{x}), \vec{y} \rangle$$

es decir, f es autoadjunto. □

Es decir, en el caso que la matriz A que se quiere diagonalizar sea simétrica, la podemos considerar como la matriz asociada a un endomorfismo autoadjunto f y el siguiente teorema nos prueba que los autovectores asociados a autovalores distintos de una matriz simétrica, no sólo sean linealmente independientes, sino que además son ortogonales.

Teorema 6.84. *Los autovectores correspondientes a autovalores distintos de un endomorfismo autoadjunto son ortogonales.*

Demostración. Sean $\lambda_1 \neq \lambda_2$ autovalores y $\vec{x}, \vec{y}$ autovectores asociados a ellos respectivamente. Entonces $f(\vec{x}) = \lambda_1 \vec{x}$ y $f(\vec{y}) = \lambda_2 \vec{y}$. Así tenemos:

$$\left.\begin{array}{r}\langle f(\vec{x}), \vec{y}\rangle = \lambda_1 \langle \vec{x}, \vec{y}\rangle \\ \langle \vec{x}, f(\vec{y})\rangle = \lambda_2 \langle \vec{x}, \vec{y}\rangle\end{array}\right\}$$

restando, tenemos $(\lambda_1 - \lambda_2)\langle \vec{x}, \vec{y}\rangle = 0$, y como $\lambda_1 \neq \lambda_2$, queda probado que $\vec{x}$ e $\vec{y}$ son ortogonales. □

Ejercicio 6.85. *Comprueba que el anterior teorema calculando los autovalores y autovectores de la matriz simétrica* $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

6.6. Diagonalización de formas cuadráticas

Dada la expresión matricial de una forma cuadrática $q(\vec{x}) = X^T A X$, como siempre, nos hacemos la siguiente pregunta:

> *¿Cómo cambia la matriz de una forma cuadrática q cuando se cambia la base?*

La respuesta es la misma que hemos visto cuando analizábamos el cambio de base en formas bilineales (tema anterior).

Si P es la matriz del cambio de base de $\mathcal{B}_2$ a $\mathcal{B}_1$, las coordenadas de un vector $\vec{x}$ en cada base se relacionan de la forma $X = P\overline{X}$. Por tanto

$$q(\vec{x}) = X^T A X = (P\overline{X})^T A (P\overline{X}) = \overline{X}^T (P^T A P)\overline{X}$$

de donde vemos que dos matrices (simétricas) de la misma forma cuadrática son *congruentes*, es decir, $B = P^T A P$, con P invertible.

Ejercicio 6.86. *Prueba que, efectivamente, si una matriz A es simétrica, cualquier matriz congruente con A también es simétrica.*

Proposición 6.87. *Dada una forma cuadrática* $q\colon V \to \mathbb{R}$, *existe una base de V tal la que la matriz de q es diagonal.*

Demostración. Sea A la matriz simétrica de q respecto de una base $\mathcal{B}$. Consideramos el endomorfismo simétrico $f\colon V \to V$ cuya matriz respecto a dicha base es $A = \mathcal{M}(f; B)$. Diagonalizando ortogonalmente, existe una base ortonormal $\mathcal{B}^{\text{Nor}}$ de autovectores de f de forma que $D = \mathcal{M}(f; \mathcal{B}^{\text{Nor}})$ es diagonal. Puesto que la matriz de paso P es ortogonal, tenemos $D = P^T A P$.

Por tanto, respecto de la base ortonormal $\mathcal{B}^{\text{Nor}}$ la matriz de q es D, y esto prueba la proposición. □

Definición 6.88. Una forma cuadrática q se dice que está *diagonalizada* cuando está expresada matricialmente con una matriz D diagonal, es decir

$$q(\vec{x}) = X^T D X$$

Esto quiere decir que existe una base de V para la cual $\vec{x}$ se expresa analíticamente como

$$q(\vec{x}) = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_n x_n^2$$

donde $d_i \in \mathbb{R}$ y $(x_1, x_2, \ldots, x_n)$ son las coordenadas de $\vec{x}$ en dicha base.

6.6.1. Método de los autovectores

La prueba de la proposición 6.87 nos está dando un método de diagonalización. Si partimos de una matriz simétrica A de q,

$$q(\vec{x}) = X^T A X$$

diagonalizando ortogonalmente A podemos encontrar una matriz ortogonal P tal que $D = P^T A P$. Haciendo el cambio de base $X = P\overline{X} \iff X^T = \overline{X}^T P^T$, tenemos

$$q(\vec{x}) = X^T A X = \overline{X}^T (P^T A P) \overline{X} = \overline{X}^T D \overline{X}$$

que nos da la diagonalización.

Este método, en teoría, funciona siempre, pero el problema estriba en que, para formas cuadráticas de tres o más variables, la diagonalización ortogonal de su matriz implica encontrar raíces de polinomios de grado mayor o igual que tres, y esto no siempre es realizable (sin usar métodos numéricos).

Damos un método alternativo que no involucra resolución de ecuaciones polinómicas.

6.6.2. Método de Lagrange

Este método consiste en ir completando cuadrados perfectos con las variables. Partimos de una forma cuadrática en forma analítica

$$q(x_1, x_2, \ldots, x_n) = a_{11} x_1^2 + a_{12} x_1 x_2 + \cdots + a_{ij} x_i x_j + \cdots + a_{nn} x_n^2$$

y pueden darse dos situaciones:

1. Si algún coeficiente $a_{ii} \neq 0$. Agrupando todos los sumandos donde interviene x_i podemos expresar q de la siguiente manera:

$$q(x_1, x_2, \ldots, x_n) = a_{ii} \left(\sum_{j=1}^{n} b_j x_j \right)^2 + q_1(x, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

hacemos, entonces, el cambio de variable $\overline{x_i} = \sum_{j=1}^{n} b_j x_j$ y volvemos aplicar el método a q_1 que tiene una variable menos.

2. Si todos los coeficientes a_{ii} son ceros, elegimos $a_{ij} \neq 0$ (debe existir, claro, en caso contrario $q = 0$). Hacemos entonces los cambios de variable

$$x_i = \overline{x_i} + \overline{x_j}, \quad x_j = \overline{x_i} - \overline{x_j}, \quad x_k = \overline{x_k} \text{ si } k \neq i, j$$

entonces $a_{ij}x_ix_j = a_{ij}\overline{x_i}^2 - a_{ij}\overline{x_j}^2$ y pasamos al caso 1 anterior.

Ejemplos de diagonalización de formas cuadráticas

Explicamos lo anterior con algunos ejemplos.

Ejemplo 6.89. Encuentra una base para la que la forma cuadrática en $\mathbb{R}^2$ siguiente

$$q_1(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 - 4x_1x_2 \tag{6.6}$$

sea diagonalizable.

SOLUCIÓN:

Método de los autovectores Llamando $A = \left(\begin{smallmatrix} 1 & -2 \\ -2 & 0 \end{smallmatrix}\right)$, y calculando sus autovalores

$$\lambda_1 = \frac{1}{2} - \frac{1}{2}\sqrt{17} \approx -1{,}56 \text{ y } \lambda_2 = \frac{1}{2} + \frac{1}{2}\sqrt{17} \approx 2{,}56$$

obtenemos la forma cuadrática diagonalizada que toma la expresión

$$\boxed{q_1(\overline{x_1}, \overline{x_2}) = \left(\tfrac{1}{2} - \tfrac{1}{2}\sqrt{17}\right)\overline{x_1}^2 + \left(\tfrac{1}{2} + \tfrac{1}{2}\sqrt{17}\right)\overline{x_2}^2}$$

Los sistemas de ecuaciones $(A - \lambda_i I) = 0$ nos dan los subespacios propios

$$N_{\lambda_1} = L\left(1, \frac{1}{4}\sqrt{17} + \frac{1}{4}\right) \text{ y } N_{\lambda_2} = L\left(1, -\frac{1}{4}\sqrt{17} + \frac{1}{4}\right)$$

y tomando un vector de cada subespacio propio obtenemos una base de autovectores que sabemos que es ortogonal. Dividiendo cada vector por su módulo obtenemos la base ortonormal buscada:

$$\mathcal{B} = \left\{ \left(\frac{1}{\sqrt{\frac{1}{16}\left(\sqrt{17}+1\right)^2+1}}, \frac{1}{4}\,\frac{\left(\sqrt{17}+1\right)}{\sqrt{\frac{1}{16}\left(\sqrt{17}+1\right)^2+1}} \right), \right.$$
$$\left. \left(\frac{1}{\sqrt{\frac{1}{16}\left(\sqrt{17}-1\right)^2+1}}, \ -\frac{1}{4}\,\frac{\left(\sqrt{17}-1\right)}{\sqrt{\frac{1}{16}\left(\sqrt{17}-1\right)^2+1}} \right) \right\}$$

En este caso se ha podido obtener la solución al problema explícitamente de una forma algebraica (el polinomio característico es de segundo grado) y lo hemos hecho, claramente, usando un programa de cálculo simbólico. En la mayoría de ocasiones, incluso esto, es imposible y hay que recurrir a métodos numéricos para dar una solución aproximada.

Método de Lagrange Obsérvese que $(x_1 - 2x_2)^2 = x_1^2 - 4x_1x_2 + 4x_2^2$, luego para obtener este cuadrado perfecto sumamos y restamos $4x_2^2$ en la expresión (6.6), así

$$q_1(x_1, x_2) = x_1^2 - 4x_1x_2 + 4x_2^2 - 4x_2^2 = (x_1 - 2x_2)^2 - 4x_2^2$$

Así haciendo

$$\begin{aligned} \overline{x_1} &= x_1 - 2x_2 \\ \overline{x_2} &= x_2 \end{aligned}$$

tenemos la forma cuadrática diagonalizada

$$\boxed{q_1(\overline{x_1}, \overline{x_2}) = \overline{x_1}^2 - 4\overline{x_2}^2 = \begin{pmatrix} \overline{x_1} & \overline{x_2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} \begin{pmatrix} \overline{x_1} \\ \overline{x_2} \end{pmatrix}}$$

Para calcular la base respecto de la cual hemos diagonalizado hacemos lo siguiente:

$$\left.\begin{aligned} \overline{x_1} &= x_1 - 2x_2 \\ \overline{x_2} &= x_2 \end{aligned}\right\} \iff \left.\begin{aligned} x_1 &= \overline{x_1} + 2\overline{x_2} \\ x_2 &= \overline{x_2} \end{aligned}\right\}$$

que matricialmente se expresa

$$X = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \overline{X}$$

luego la base buscada es $\boxed{\mathcal{B} = \{(1,0), (2,1)\}}$.

La ventaja del primer método es que obtenemos directamente una base ortonormal para la cual diagonaliza la forma cuadrática. La ventaja del segundo está en la facilidad del cálculo.

Observa que los valores de la diagonal por los dos métodos coinciden en los signos (es decir, igual cantidad de positivos que de negativos que de ceros). Esto es siempre así, son *invariantes*, que veremos más adelante.

Ejemplo 6.90. Igual para

$$q(x_1, x_2, x_3) = x_1x_3 - 2x_2x_3$$

SOLUCIÓN:

Método de Lagrange En este caso no existe ningún término cuadrado. Recurrimos a hacer el cambio introduciendo sumas por diferencias: $x_1 = \overline{x_1} + \overline{x_3}$, $x_3 = \overline{x_1} - \overline{x_3}$, $x_2 = \overline{x_2}$.

$$q(\overline{x_1}, \overline{x_2}, \overline{x_3}) = \overline{x_1}^2 - 2\overline{x_1}\,\overline{x_2} + 2\overline{x_2}\,\overline{x_3} - \overline{x_3}^2$$

Buscamos, entonces, un cuadrado perfecto para $\overline{x_1}$, observamos que $(\overline{x_1} - \overline{x_2})^2 = \overline{x_1}^2 - 2\overline{x_1}\,\overline{x_2} + \overline{x_2}^2$, quedando entonces la forma cuadrática

$$q(\overline{x_1}, \overline{x_2}, \overline{x_3}) = (\overline{x_1} - \overline{x_2})^2 - \overline{x_2}^2 + 2\overline{x_2}\,\overline{x_3} - \overline{x_3}^2$$

Igualmente

$$q(\overline{x_1}, \overline{x_2}, \overline{x_3}) = (\overline{x_1} - \overline{x_2})^2 - (\overline{x_2} - \overline{x_3})^2$$

haciendo los cambios de variable siguientes:

$$\begin{aligned} \overline{\overline{x_1}} &= \overline{x_1} - \overline{x_2} \\ \overline{\overline{x_2}} &= \overline{x_2} - \overline{x_3} \\ \overline{\overline{x_3}} &= \overline{x_3} \end{aligned}$$

q aparece diagonalizada

$$\boxed{q(\overline{\overline{x_1}}, \overline{\overline{x_2}}, \overline{\overline{x_3}}) = \overline{\overline{x_1}}^2 - \overline{\overline{x_2}}^2 + 0\overline{\overline{x_3}}^2}$$

Por último, para calcular la base para la cual la forma cuadrática q está diagonalizada, trabajamos con los cambios de base matricialmente, así

$$\begin{pmatrix} \overline{\overline{x_1}} \\ \overline{\overline{x_2}} \\ \overline{\overline{x_3}} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \overline{x_1} \\ \overline{x_2} \\ \overline{x_3} \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} \overline{x_1} \\ \overline{x_2} \\ \overline{x_3} \end{pmatrix}$$

por tanto,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \overline{\overline{x_1}} \\ \overline{\overline{x_2}} \\ \overline{\overline{x_3}} \end{pmatrix}$$

que, operando las matrices, podemos expresar

$$X = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \overline{\overline{X}}$$

de donde la base es $\boxed{\mathcal{B} = \{(1,0,1), (1,1,1), (2,1,0)\}}$.

Método de los autovalores La forma cuadrática se expresa matricialmente de la forma

$$q(x_1,x_2,x_3) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & -1 \\ \frac{1}{2} & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

En este ejemplo usamos software de cálculo de matemáticas para encontrar los autovalores y autovectores, así, los autovalores son

$$\lambda_1 = 0, \lambda_2 = -1{,}118033988749895, \lambda_3 = 1{,}118033988749895$$

que nos da la expresión diagonalizada aproximada de q

$$\boxed{q(\overline{x_1},\overline{x_2},\overline{x_3}) \approx -1{,}118\,\overline{x_2}^2 + 1{,}118\,\overline{x_3}^2}$$

y los autovectores, en el orden adecuado nos producen la base

$$\{(1,\tfrac{1}{2},0),(1,-2,-2{,}23606797749979),(1,-2,2{,}236067977499790)\}$$

pero nos interesa la base ortonormal. Sin más que normalizar dicha base obtenemos la base (aproximada) que buscamos:

$$\begin{aligned} \mathcal{B} = \{&(0{,}8944, 0{,}4472, 0), (0{,}3162, -0{,}6324, \ -0{,}7071), \\ &\qquad\qquad (0{,}3162, -0{,}6324, 0{,}7071)\} \end{aligned}$$

Ejemplo 6.91. Vamos a diagonalizar la forma cuadrática siguiente

$$q(x_1,x_2,x_3) = 3x_1^2 + 4x_1x_2 + 2x_1x_3 + x_3^2 + 4x_2x_3$$

SOLUCIÓN:

Método de Lagrange Comenzamos encontrando un cuadrado perfecto para la primera variable. Observamos que

$$3(x_1 + \frac{2}{3}x_2 + \frac{1}{3}x_3)^2 = 3{x_1}^2 + 4x_1x_2 + 2x_1x_3 + \frac{{x_3}^2}{3} + \frac{4x_2x_3}{3} + \frac{4{x_2}^2}{3}$$

por tanto, q se expresa

$$\begin{aligned} q(x_1,x_2,x_3) &= 3(x_1 + \frac{2}{3}x_2 + \frac{1}{3}x_3)^2 - \frac{4}{3}x_2^2 + \frac{8}{3}x_2x_3 + \frac{2}{3}x_3^2 \\ &= 3(x_1 + \frac{2}{3}x_2 + \frac{1}{3}x_3)^2 - \frac{4}{3}(x_2 - x_3)^2 + 2x_3^2 \end{aligned}$$

Haciendo los cambios de variables

$$\begin{aligned} \overline{x_1} &= x_1 + \tfrac{2}{3}x_2 + \tfrac{1}{3}x_3 \\ \overline{x_2} &= x_2 - x_3 \\ \overline{x_3} &= x_3 \end{aligned}$$

obtenemos la forma cuadrática diagonalizada

$$\boxed{q(\vec{x}) = 3\overline{x_1}^2 - \frac{4}{3}\overline{x_2}^2 + 2\overline{x_3}^2}$$

Para calcular la base observamos que

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

y obtenemos una base que diagonaliza a q.

$$\boxed{\mathcal{B} = \{(1,0,0), (\tfrac{-2}{3}, 1, 0), (-1,1,1)\}}$$

6.6.3. Clasificación por invariantes de una forma cuadrática

Sabemos que el rango de una matriz es un invariante para matrices equivalentes, por tanto, para también para matrices congruentes. Obsérvese que, si una forma cuadrática está diagonalizada, *el* rango *es el número de elementos de la diagonal principal no nulos*, independientemente de la diagonalización elegida. Es, lo que se llama, un *invariante* de la forma cuadrática.

Definición 6.92. Si q es una forma cuadrática diagonalizada, llamaremos *signatura* al número de elementos positivos de su diagonal principal.

Para el siguiente teorema no damos demostración y nos remitimos a los manuales de Álgebra Lineal.

Teorema 6.93 (Ley de inercia de Sylvester)**.** *Sea q una forma cuadrática real definida en un espacio vectorial V.* El rango y la signatura son invariantes *de q.*

Ejercicio 6.94. *Determina el rango y la signatura de las formas cuadráticas de todos los ejemplos de la sección 6.6 anterior.*

Clasificación de las formas cuadráticas

Sea q una forma cuadrática diagonalizada

$$q(\vec{x}) = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_n x_n^2$$

entonces:

1. q es *definida positiva* $\iff d_i > 0 \quad \forall i = 1, 2, \ldots, n$.
2. q es *definida negativa* $\iff d_i < 0 \quad \forall i = 1, 2, \ldots, n$.
3. q es *semidefinida positiva* $\iff d_i \geq 0 \quad \forall i = 1, 2, \ldots, n$ con algún $d_j = 0$.

4. q es *semidefinida negativa* $\iff$ $d_i \leq 0 \quad \forall i = 1, 2, \ldots, n$ con algún $d_j = 0$.

5. En cualquier otro caso q es *indefinida*. Dicho de otro modo, existen $d_i < 0$ y $d_j > 0$.

Ejercicio 6.95. *Clasifica la siguiente forma cuadrática, usando los autovalores y por el método de Lagrange.*

$$q(x_1, x_2, x_3) = x_3\,(7x_3 + x_1) + x_1\,(x_3 + 7x_1) + 4x_2^2$$

Criterio de Sylvester

Existe un criterio que caracteriza las formas cuadráticas definidas, aunque no clasifica las semidefinidas ni las indefinidas.

Teorema 6.96 (Criterio de Sylvester)**.** *Sea $q(\vec{x}) = X^T AX$ una forma cuadrática real y sean $\Delta_1, \Delta_2, \cdots, \Delta_n$ los menores principales de A. Entonces:*

1. *q es definida positiva* $\iff$ $\Delta_i > 0 \quad \forall i = 1, 2, \ldots, n.$

2. *q es definida negativa* $\iff$ $(-1)^i \Delta_i > 0 \quad \forall i = 1, 2, \ldots, n.$

En cualquier otro caso este criterio no clasifica nada, es decir podría ser semidefinida o indefinida.

Demostración. Supongamos que la forma cuadrática está diagonalizada. Si q es definida positiva, entonces los elementos de la diagonal principal son todos positivos, esto es equivalente a que los Δ_i son todos positivos. Si q es definida negativa, entonces los elementos de la diagonal principal son todos negativos, de aquí que los Δ_i alternan su signo en la forma indicada en el teorema. □

Clasificación de formas cuadráticas en espacios bidimensionales

Para el caso de formas cuadráticas en espacios vectoriales reales de dimensión 2 (como sabemos, isomorfos a $\mathbb{R}^2$) se puede hacer una clasificación completa independiente de los teoremas 6.93 y 6.96. Sea

$$q(x_1, x_2) = (x_1 x_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = ax_1^2 + 2bx_1x_2 + cx_2^2$$

tenemos $\Delta_1 = a$ y $\Delta_2 = \begin{vmatrix} a & b \\ b & c \end{vmatrix} = ac - b^2$ y, aplicando el método de Lagrange se da una de las tres siguientes situaciones:

1. Si $\Delta_1 = a \neq 0$, entonces $\Delta_2 = a\left(c - \frac{b^2}{a}\right)$ y

$$q(x_1, x_2) = a\left(x_1 + \frac{b}{a}x_2\right)^2 + \left(c - \frac{b^2}{a}\right)x_2^2 = \Delta_1 \overline{x_1}^2 + \frac{\Delta_2}{\Delta_1}\overline{x_2}^2$$

2. Si $\Delta_1 = a = 0$ y $c \neq 0$, necesariamente $\Delta_2 = -b^2 \leq 0$ y

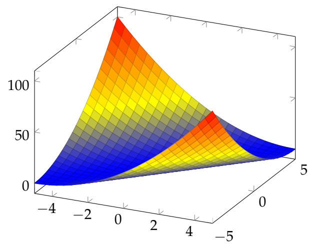
$$q(x_1, x_2) = 2bx_1x_2 + cx_2^2 = c\left(x_2 + \frac{b}{c}x_1\right)^2 + \left(-\frac{b^2}{c}\right)x_1^2 = \frac{\Delta_2}{c}\overline{x_1}^2 + c\overline{x_2}^2$$

3. Si $\Delta_1 = a = 0$ y $c = 0$, $\Delta_2 \leq 0$ y haciendo los cambios de variables $x_1 = |b|^{1/2}(\overline{x_1} + \overline{x_2})$ y $x_2 = |b|^{1/2}(\overline{x_1} - \overline{x_2})$ obtenemos

$$q(x_1, x_2) = 2bx_1x_2 = 2b|b|\left(\overline{x_1}^2 - \overline{x_2}^2\right) = (\mp 2)\Delta_2\overline{x_1}^2 + (\pm 2\Delta_2)\overline{x_2}^2$$

Teorema 6.97 (Criterio de Sylvester para dim=2)**.** *Sea $q(\vec{x}) = X^T AX$ una forma cuadrática real de dmensión 2 y sean Δ_1, Δ_2 los menores principales de A. Entonces:*

1. *q es definida positiva* $\iff$ $\Delta_1 > 0$ *y* $\Delta_2 > 0$.
2. *q es definida negativa* $\iff$ $\Delta_1 < 0$ *y* $\Delta_2 > 0$.
3. *q es indefinida* $\iff$ $\Delta_2 < 0$.
4. *q es semidefinida*[3] $\iff$ $\Delta_2 = 0$.



Veamos un ejemplo. La siguiente forma cuadrática

$$f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1x_2$$

tiene por matriz $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, por tanto, $\Delta_1 = 1$ y $\Delta_2 = 0$. El teorema 6.97 anterior nos dice que esta forma cuadrática es semidefinida, pero no nos dice si es (*semidefinida*) positiva o negativa. Buscando un vector $\vec{v}$ para el que $q(\vec{v})$ no sea cero, nos dará el signo. En este ejemplo, $q(1, 0) = 1 > 0$, luego q es semidefinida positiva. En la siguiente gráfica podemos ver que, efectivamente, lo es.

Ejercicio 6.98. *Da un ejemplo de forma cuadrática definida en $\mathbb{R}^2$ de cada uno de las clases: definida positiva, definida negativa, semidefinida ....*

[3] En este caso es nula cuando también $\Delta_1 = 0$ y $c = 0$.

Ejercicios Propuestos

Ej. 6.1 — Comprueba que la forma bilineal $f((x_1, x_2), (y_1, y_2)) = x_2(y_1 - y_2) - 2x_1(y_1 - y_2)$ es degenerada encontrando al menos un vector degenerado para f. Comprueba, además, que es indefinida.

Ej. 6.2 — Sea $f\colon \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ la forma bilineal definida como: $f((x_1, x_2), (y_1, y_2)) = 2x_1y_1 - 3x_1y_2 + x_2y_2$

1. Halla la matriz A de f en la base $\{u_1 = (1, 0), u_2 = (1, 1)\}$.
2. Halla la matriz B de f en la base $\{v_1 = (2, 1), v_2 = (1, -1)\}$.
3. Encuentra la matriz de cambio de base P de la base $\{v_1, v_2\}$ a la base $\{u_1, u_2\}$ y comprueba que $B = P^T AP$.

Ej. 6.3 — Si $A = \begin{pmatrix} 1 & -1 & 3 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$ es la matriz de una forma bilineal respecto a la base canónica, se pide:

1. Calcula su matriz respecto a la base $\mathcal{B} = \{(1, 1, 0), (0, 1, 0), (1, 1, 1)\}$.
2. Determina si es o no es degenerada.
3. Da las expresiones analíticas de la forma cuadrática asociada a f respecto de la base canónica y respecto de la base $\mathcal{B}$ anterior.

Ej. 6.4 — Dada la siguiente matriz simétrica $A = \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}$, encuentra las expresiones analíticas de:

1. El endomorfismo asociado a A.
2. La forma bilineal asociada a A.
3. La forma cuadrática asociada a A.

Ej. 6.5 — Sea V el espacio vectorial real de los polinomios de grado menor o igual que 2. Para $f, g \in V$, se define el siguiente producto escalar: $\langle f, g \rangle = \int_0^1 f(t)g(t)\,dt$. Se pide:

1. Da una base ortonormal para este producto escalar
2. Dados los polinomios $p(t) = t + 2$ y $q(t) = 2t - 3$; halla sus módulos y el coseno del ángulo que forman.

Ej. 6.6 — ¿Qué ángulo deben formar los vectores unitarios $\vec{u}$ y $\vec{v}$ para que la norma del vector suma tome el valor $\|\vec{u} + \vec{v}\| = \sqrt{3}$.

Ej. 6.7 — Sea V el espacio vectorial sobre $\mathbb{R}$ de los polinomios de grado ≤ 2. Sean a, b, c tres números reales distintos. Se considera la aplicación

$f : V \times V \to \mathbb{R}$, definida de la forma:

$$f(p(x), q(x)) = p(a)q(a) + p(b)q(b) + p(c)q(c)$$

1. Probar que es una forma bilineal simétrica.
2. Hallar la matriz asociada a dicha forma bilineal en la base $B = \{1, x, x^2\}$.
3. Para $a = -1$, $b = 1$ y $c = 0$, f define un producto escalar. Encontrar una base ortonormal, para dicho producto escalar, del subespacio de los polinomios sin término independiente.

Ej. 6.8 — Sabemos que dos vectores $\vec{u}, \vec{v} \in \mathbb{R}^3$ verifican $\|\vec{v}\| = 2\|\vec{u}\|$ y que $\|\vec{u} \times \vec{v}\| = 2$. Expresa el ángulo entre $\vec{u}$ y $\vec{v}$ en función del módulo de $\vec{u}$.

1. Encuentra para que valor de $\|\vec{u}\|$ se tiene que el ángulo entre ambos vectores sea $\frac{\pi}{4}$ radianes.
2. Encuentra, si existen, vectores $\vec{u}$ y $\vec{v}$ en las anteriores condiciones con la condición de que $\|\vec{v}\|^2 = 2$.

Ej. 6.9 — En el espacio vectorial euclídeo $\mathbb{R}^4$ con el producto usual se pide:

1. Obtener mediante el método de Gram-Schmidt una base de vectores ortonormales para el subespacio:

$$S = L\{(1, 2, -1, 0),\ (0, 1, 1, 0),\ (1, 0, -2, 1)\}$$

2. Determina una base ortonormal del espacio ortogonal $S^\perp$ del subespacio anterior.

Ej. 6.10 — De un endomorfismo F de $\mathbb{R}^4$ se conoce :

a) $\Im F = \ker F$.

b) $F(1, 2, -2, 1) = (0, 1, 0, 1)$ y $F(3, 3, -1, 2) = (1, 0, 1, 0)$

Se pide:

1. Calcula la matriz de F.
2. Ecuaciones paramétricas e implícitas de $\Im F$ (o del núcleo) y una base ortonormal de la misma.

Ej. 6.11 — Consideramos el producto escalar definido por la matriz:

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 5 \end{pmatrix}.$$

Se pide:

1. Comprueba que, efectivamente, es la matriz de un producto escalar.

2. A partir de la base canónica de $\mathbb{R}^3$, obtén una base ortonormal respecto de dicho producto escalar y halla la matriz del cambio de base.

Ej. 6.12 — Consideramos el espacio euclídeo tridimensional y el subespacio vectorial S generado por la base $\mathcal{B} = \{(1,0,2),(0,1,1)\}$. Se pide:

1. Matriz A del producto escalar restringido a S.
2. Sustituye el segundo vector de la base $\mathcal{B}$ por otro vector $\vec{x}$ de forma que $\{(1,0,2),\vec{x}\}$ sea una base ortogonal de S.
3. Encuentra la matriz B asociada a este producto escalar en esta nueva base y la matriz P que hace $B = P^T A P$.

Ej. 6.13 — Calcula una base, así como las ecuaciones, del subespacio ortogonal en $\mathbb{R}^4$ del subespacio

$$S \equiv \begin{cases} x - y + 2t = 0 \\ z + t = 0 \end{cases}$$

Ej. 6.14 — Diagonalizar la matriz A y encontrar la matriz de paso ortogonal.

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

Ej. 6.15 — Clasifica las siguientes formas cuadráticas:

1. $3x^2 - 6xy + 11y^2$
2. $3x^2 + 2xy + 3y^2 + 2xz + 2yz + 3z^2$
3. $xy + yz + zx$

Ej. 6.16 — Clasifica la siguiente forma cuadrática para los distintos valores de a.

$$q(x,y,z) = \alpha x^2 + \alpha y^2 + (\alpha - 1)z^2 + 2xy \text{ con } \alpha \in \mathbb{R} \text{ fijo.}$$

Ej. 6.17 — Diagonaliza las formas cuadráticas del ejercicio 6.15 anterior.

Ej. 6.18 — Dada la forma cuadrática cuya matriz asociada es

$$\begin{pmatrix} a & 0 & a \\ 0 & a+c & 0 \\ a & 0 & a \end{pmatrix}$$

1. Exprésala como suma de cuadrados usando dos métodos diferentes.
2. Clasifícala en función de los parámetros a y c.
3. Pon un ejemplo, si existe, de valores de a y c para los que la forma cuadrática es definida positiva.
4. Pon un ejemplo, si existe, de valores de a y c para los que la forma cuadrática es semidefinida positiva.
5. Pon un ejemplo, si existe, de valores de a y c para los que la forma cuadrática es indefinida.

Ej. 6.19 — Diagonaliza de dos maneras distintas las formas cuadráticas siguientes:

1. $q(x,y) = -6xy + 8y^2$
2. $q(x,y,z,t) = x^2 + 2y^2 - 6xz + z^2 - 4t^2$

Ej. 6.20 — Se consideran en $\mathbb{R}^3$ y con respecto a la base canónica, las formas cuadráticas

$$q_a(x,y,z) = x^2 + ay^2 + 3z^2 + 2xy + 2xz + 2ayz$$

donde a es un parámetro real. Diagonaliza y estudia el carácter de q_a para los distintos valores de a.