Espacios Euclídeos

Natalia Boal
María Luisa Sein-Echaluce
Universidad de Zaragoza

A lo largo de todo el capítulo consideraremos que V un espacio vectorial real de dimensión finita.

1 Producto escalar

Definición. Se dice que $F : V \times V \longrightarrow \mathbb{R}$ es un *producto escalar sobre* V si es una forma bilineal simétrica y definida positiva.

Notación. $F(u, v) = (u, v)$.

Es inmediato observar que

1. $(u, v) = (v, u), \quad \forall u, v \in V,$
2. $(u_1 + u_2, v) = (u_1, v) + (u_2, v), \quad \forall u_1, u_2, v \in V,$
3. $(\lambda u, v) = \lambda(u, v), \quad \forall u, v \in V,\ \forall \lambda \in \mathbb{R},$
4. $(u, u) > 0, \quad \forall u \in V,\ u \neq 0_V.$

Definición. Un *espacio euclídeo* es un espacio vectorial con un producto escalar, escribiremos, (V, F) o bien $(V, (\cdot\,,\cdot)\,)$.

Definición. Dado el espacio euclídeo $(V, (\cdot\,,\cdot)\,)$. Se define la *norma asociada o inducida* por el producto escalar $(\cdot\,,\cdot)$ a la aplicación $\|\cdot\| : V \to \mathbb{R}^+$ tal que

$$\|v\| = \sqrt{(v, v)}.$$

Propiedades.

1. $\|\lambda\, v\| = |\lambda|\, \|v\|, \quad \forall \lambda \in \mathbb{R},\ \forall v \in V.$
2. $\|v\| \geq 0, \quad \forall v \in V.$
3. Desigualdad triangular: $\|u + v\| \leq \|u\| + \|v\|,\ \forall u, v \in V.$
4. Desigualdad Cauchy-Schwarz: $|\,(u, v)\,| \leq \|u\|\, \|v\|,\ \forall u, v \in V.$

Definición. Una aplicación $\|\cdot\| : V \to \mathbb{R}^+$ verificando las propiedades (1), (2) y (3) anteriores se dice *norma en* V y $(V, \|\cdot\|)$ es un *espacio normado.*

Observación. A partir de todo espacio euclídeo podemos construir un espacio normado. Sin embargo el recíproco no es cierto. Existen espacio normados que no vienen inducidos por ningún espacio euclídeo.

2 Bases ortonormales

Definición. Los vectores $u, v \in V$ se dicen *ortogonales* si $(u, v) = 0$. Lo denotaremos $u \perp v$.

Definición. Una familia de vectores de V, $\{v_1, \ldots, v_m\}$, se dice *ortogonal* si los vectores son ortogonales dos a dos, esto es, $(v_i, v_j) = 0$, $\forall i \neq j$.

Definición. Un vector $v \in V$ se dice *unitario* si $\|v\| = 1$.

Definición. Una familia de vectores de V, $\{v_1, \ldots, v_m\}$, se dice *ortonormal* si es una familia de vectores ortogonales y unitarios. Así, $(v_i, v_j) = 0$, $\forall i \neq j$ y $\|v_i\| = 1$, $\forall i$.

Proposición. Toda familia de vectores de V no nulos y ortogonales es una familia libre.

2.1 Método de Gram-Schmidt

Proposición. Sean $(V, (\cdot\,, \cdot)\,)$ espacio euclídeo y S subespacio de V de dimensión finita. Entonces, S posee una base ortonormal.

Demostración. Método de Gram-Schmidt, este método permite a partir de una base de S dada, construir una base ortonormal de S. $\square$

MÉTODO DE GRAM-SCHMIDT. Sea $\{v_1, v_2, \ldots, v_m\}$ una base de S subespacio de V. El método de Gram-Schmidt consiste en construir una familia ortonormal equivalente, es decir, una familia $\{w_1, w_2, \ldots, w_m\}$ tal que

- $w_i \perp w_j$, $\forall i \neq j$,
- $\mathbb{R}\langle\{v_1, \ldots, v_k\}\rangle = \mathbb{R}\langle\{w_1, \ldots, w_k\}\rangle$, $k = 1, \ldots, m$.

Proceso.

1. Definimos $w_1 = v_1/||v_1||$ (normalizamos, $v_1 = \widetilde{v}_1$ en general no es unitario).
2. Buscamos $\widetilde{v}_2$ tal que
   - $\mathbb{R}\langle\{v_1, v_2\}\rangle = \mathbb{R}\langle\{w_1, \widetilde{v}_2\}\rangle$, por ello, elegimos $\widetilde{v}_2 = v_2 + \alpha\, w_1$. Entonces,
   $$(\widetilde{v}_2, w_1) = (v_2, w_1) + \alpha\,(w_1, w_1).$$
   Además se ha de satisfacer
   - $\widetilde{v}_2 \perp w_1$, esto es, $(\widetilde{v}_2, w_1) = 0$. Como w_1 es unitario se tiene que $\alpha = -(v_2, w_1)$. Luego
   $$\widetilde{v}_2 = v_2 - (v_2, w_1)\, w_1.$$
   - Definimos $w_2 = \widetilde{v}_2/||\widetilde{v}_2||$ (normalizamos, $\widetilde{v}_2$ en general no es unitario).
3. Buscamos $\widetilde{v}_3$ tal que

- $\mathbb{R}\langle\{v_1, v_2, v_3\}\rangle = \mathbb{R}\langle\{w_1, w_2, \widetilde{v}_3\}\rangle$, por ello, elegimos $\widetilde{v}_3 = v_3 + \beta_1 w_1 + \beta_2 w_2$. Entonces,
$$(\widetilde{v}_3, w_1) = (v_3, w_1) + \beta_1 (w_1, w_1) + \beta_2 (w_2, w_1),$$
$$(\widetilde{v}_3, w_2) = (v_3, w_2) + \beta_1 (w_1, w_2) + \beta_2 (w_2, w_2).$$
Además se ha de satisfacer
- $\widetilde{v}_3 \perp w_1$, $\widetilde{v}_3 \perp w_2$, esto es, $(\widetilde{v}_3, w_1) = 0$ y $(\widetilde{v}_3, w_2) = 0$. Como w_1, w_2 son unitarios se tiene que $\beta_1 = -(v_3, w_1)$ y $\beta_2 = -(v_3, w_2)$. Luego
$$\widetilde{v}_3 = v_3 - (v_3, w_1)\, w_1 - (v_3, w_2)\, w_2.$$
- Definimos $w_3 = \widetilde{v}_3/||\widetilde{v}_3||$ (normalizamos, $\widetilde{v}_3$ en general no es unitario).

4. Reiterando este proceso, para $k = 4, \ldots, m$
   - Definimos: $\widetilde{v}_k = v_k - (v_k, w1)\, w_1 - \cdots - (v_k, w_{k-1})\, w_{k-1}$,
   - Normalizamos: $w_k = \widetilde{v}_k/||\widetilde{v}_k||$.

Observaciones.

- $\{\widetilde{v}_1, \ldots, \widetilde{v}_k\}$ es una familia ortogonal $k = 2, \ldots, m$.
- $\{w_1, \ldots, w_k\}$ es una familia ortonormal $k = 2, \ldots, m$.
- $\mathbb{R}\langle\{v_1, \ldots, v_k\}\rangle = \mathbb{R}\langle\{\widetilde{v}_1, \ldots, \widetilde{v}_k\}\rangle = \mathbb{R}\langle\{w_1, \ldots, w_k\}\rangle$, $k = 1, \ldots, m$.
- El proceso de ortonormalización de Gram-Schmidt permite construir bases ortonormales de V con el producto escalar $(\cdot\,,\cdot)$.
- Si $\{w_i\}_{i=1}^n$ es una base de V ortonormal respecto de $(\cdot\,,\cdot)$, $u = (w_i)^t X$ y $v = (w_i)^t Y$. La expresión coordenada respecto de esta base será
$$(u, v) = X^t I_n Y$$
donde I_n es la identidad de orden n.

2.2 Cambio de coordenadas entre bases ortonormales

Sean (V, F) espacio euclideo y $\{a_i\}_{i=1}^n$ y $\{b_j\}_{j=1}^n$ dos bases ortonormales de V. Entonces, si

- $u = (a_i)^t X = (b_i)^t \widetilde{X}$,
- $v = (a_i)^t Y = (b_i)^t \widetilde{Y}$,

se tiene

- expresión coordenada respecto de $\{a_i\}_{i=1}^n : (u, v) = X^t I_n Y$,
- expresión coordenada respecto de $\{b_i\}_{i=1}^n : (u, v) = \widetilde{X}^t I_n \widetilde{Y}$.

Sabemos que existe una matriz de cambio de base P (regular) tal que $(b_j)^t = (a_i)^t P$, luego $X = P\widetilde{X}$ e $Y = P\widetilde{Y}$. De donde se deduce que

$$P^t P = I_n.$$

Definición. Una matriz $Q \in \mathcal{M}_{m\times n}(\mathbb{R})$ de rango n se dice *ortogonal* si

$$Q^t Q = I_n.$$

Observaciones.

- Si $Q \in \mathcal{M}_n(\mathbb{R})$ ortogonal, se tiene que Q es regular y además $Q^{-1} = Q^t$.
- La matriz de cambio entre bases ortonormales es ortogonal.

3 Clasificación de formas cuadráticas

Definición. Dos matrices simétricas reales $A, B \in \mathcal{M}_n(\mathbb{R})$ se dicen *congruentes ortogonales* si existe $P \in \mathcal{M}_n(\mathbb{R})$ ortogonal tal que $B = P^t A P$.

Sea $A \in \mathcal{M}_n(\mathbb{R})$ simétrica entonces se verifica:

- Todos los valores propios de A son reales.
- A es diagonalizable, es decir, existe una base de $\mathbb{R}^n$ de vectores propios $\{v_i\}_{i=1}^n$.
- $\mathbb{R}^n = V(\lambda_1) \oplus \cdots \oplus V(\lambda_r)$ con $\lambda_1, \ldots, \lambda_r$ los distintos valores propios de A.
- Sean α y β dos valores propios de A (distintos), v_α, v_β vectores propios asociados a α y β, respectivamente. Entonces, $\{v_\alpha, v_\beta\}$ libre y además

  $$v_\alpha \perp v_\beta \quad \text{(respecto del producto escalar ordinario de } \mathbb{R}^n\text{)}.$$

  Luego los vectores propios asociados a valores propios distintos son ortogonales (respecto del producto escalar ordinario de $\mathbb{R}^n$).

Propiedades.

1. Toda matriz simétrica y real es congruente ortogonal con una matriz diagonal.

   - Existe una base de vectores propios $\{v_i\}_{i=1}^n$.
   - Partiendo de la base $\{v_i\}_{i=1}^n$ aplicamos el proceso de ortonormalización de Gram-Schmidt para obtener una base de vectores propios ortonomal (respecto del producto escalar ordinario de $\mathbb{R}^n$), $\{w_i\}_{i=1}^n$.

$$\begin{array}{ccc} \text{Base vect. propios} & & \text{Base ortonormal vect. propios} \\ \{v_i\}_{i=1}^n & -\text{ Gram-Schmidt } \rightarrow & \{w_i\}_{i=1}^n \end{array}$$

- Entonces, como $\{e_i\}_{i=1}^n$ es una base ortonormal (respecto del producto escalar ordinario de $\mathbb{R}^n$) y la matriz de cambio entre bases ortonormales es ortogonal se tiene que existe $P \in \mathcal{M}_n(\mathbb{R})$ ortogonal con

$$(w_i)' = (e_i)'P.$$

- La matriz coordenada respecto de $\{w_i\}_{i=1}^n$ es diagonal, D con $D = P^tAP$ con $P^t = P^{-1}$.

2. Toda forma cuadrática puede clasificarse (rango y signatura) a partir de los valores propios de su matriz coordenada A.

   - Rango Q = número de valores propios no nulos de A.
   - Signatura Q = número de valores propios positivos de A.

3. Una matriz simétrica A es:

   - Definida positiva $\Longleftrightarrow$ todos sus valores propios son positivos.
   - Definida negativa $\Longleftrightarrow$ todos sus valores propios son negativos.
   - Semidefinida positiva $\Longleftrightarrow$ todos sus valores propios son no negativos.
   - Semidefinida negativa $\Longleftrightarrow$ todos sus valores propios son no positivos.
   - Indefinida $\Longleftrightarrow$ existen algún valor propio positivo y algún valor propio negativo.

4 Factorización QR

Toda matriz $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ de rango n admite *factorización QR*, es decir, descomposición en producto de una matriz $Q \in \mathcal{M}_{m\times n}(\mathbb{R})$ ortogonal por una matriz $R \in \mathcal{M}_n(\mathbb{R})$ triangular superior con $r_{ii} > 0$.

Factorización QR

Escribimos $A = (A_1|A_2|\dots|A_n)$ con las columnas $A_i \in \mathbb{R}^m$.

1. <u>Construcción de Q:</u>

   Tomamos $\{A_1, \dots, A_n\}$, sistema libre de vectores de $\mathbb{R}^m$. Consideramos en $\mathbb{R}^m$ el producto escalar estándar y aplicamos el proceso de ortonormalización de Gram-Schmidt para obtener una familia ortonormal $\{Q_1, \dots, Q_n\}$. Entonces definimos

$$Q = (Q_1|Q_2|\dots|Q_n).$$

   Observación. Cuando multiplicamos Q^tQ tenemos que el elemento (i,j) será $Q_i^tQ_j$. Como $\{Q_1, \dots, Q_n\}$ es sistema ortonormal en $\mathbb{R}^n$ respecto del producto escalar ordinario, se tiene

$$Q_i^tQ_j = \begin{cases} 0, & \text{si } i \neq j \\ 1, & \text{si } i = j. \end{cases}$$

2. Construcción de R:

   Al aplicar el proceso de ortonormalización de Gram-Schmidt hemos obtenido

$$\begin{array}{rcl} A_1 & = & \|\widetilde{Q}_1\|\, Q_1 \\ A_2 & = & (A_2, Q_1)\, Q_1 + \|\widetilde{Q}_2\|\, Q_2 \\ & \vdots & \\ A_m & = & (A_m, Q_1)\, Q_1 + (A_m, Q_2)\, Q_2 + \cdots + \|\widetilde{Q}_m\|\, Q_m \end{array}$$

   siendo $\{\widetilde{Q}_i\}$ la familia ortogonal que se va generando en el proceso de ortonormalización. Luego $A = QR$ donde

$$R = \begin{pmatrix} \|\widetilde{Q}_1\| & (A_2, Q_1) & \dots & (A_m, Q_1) \\ 0 & \|\widetilde{Q}_2\| & \dots & (A_m, Q_2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \|\widetilde{Q}_m\| \end{pmatrix}$$

   Observación. Como Q es ortogonal y $A = QR$ se tiene que $R = Q^t A$.

5 La mejor aproximación

Definición. Sea S subespacio vectorial del espacio euclídeo $(V, (\cdot\,,\cdot)\,)$. Se llama *complemento ortogonal* de S al conjunto

$$S^{\perp} = \{v \in V \,/\, v \perp s,\ \forall s \in S\}.$$

Proposición. Sea $\{s_1, \dots, s_m\}$ una base de S subespacio vectorial del espacio euclídeo $(V, (\cdot\,,\cdot)\,)$. Entonces, dado $v \in V$

$$v \perp s, \quad \forall s \in S \ \Leftrightarrow\ v \perp s_i, \quad i = 1, \dots, m.$$

Proposición. Dados S subespacio vectorial del espacio euclídeo $(V, (\cdot\,,\cdot)\,)$ y $v \in V$ entonces existe un único $\widetilde{s} \in S$ tal que $(v - \widetilde{s}) \in S^{\perp}$.

Definición. Este vector $\widetilde{s}$ se dice *proyección ortogonal de v en S*.

Observaciones.

- Problemas equivalentes
  - Hallar $\widetilde{s} \in S$ tal que $(v - \widetilde{s}) \in S^{\perp}$.
  - Hallar $\widetilde{s} \in S$ tal que $\|v - \widetilde{s}\| \le \|v - s\|$, $\forall s \in S$.

  Por ello, la proyección ortogonal de v en S, $\widetilde{s}$, es la *mejor aproximación* o *aproximación cuadrática* de v por vectores de S.

- Obviamente, si $v \in S$ su proyección ortogonal en S es él mismo.

6 Mínimos cuadrados

Definición. Sean $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $\mathbf{x} \in \mathbb{R}^n$ y $\mathbf{b} \in \mathbb{R}^m$. El sistema lineal $A\mathbf{x} = \mathbf{b}$ se dice *sobredeterminado* si posee más ecuaciones que incógnitas, es decir, $m > n$.

Observación. En general los sistemas sobredeterminados suelen ser incompatibles, luego para todo $\mathbf{x} \in \mathbb{R}^n$, $A\mathbf{x} \neq \mathbf{b}$.

Consideramos $A_1, \ldots, A_n \in \mathbb{R}^m$ las n columnas de A. El sistema $A\mathbf{x} = \mathbf{b}$ podemos verlo como

$$x_1 A_1 + x_2 A_2 + \ldots x_n A_n = \mathbf{b}.$$

De modo que, la existencia de solución del sistema anterior equivale a que el vector $\mathbf{b}$ se pueda expresar como combinación lineal de los vectores $A_1, \ldots, A_n$. Por tanto,

- si $\mathbf{b} \in S = \mathbb{R}\langle A_1, \ldots, A_n\rangle$ entonces el sistema $A\mathbf{x} = \mathbf{b}$ es compatible.

En caso contrario, $A\mathbf{x} = \mathbf{b}$ no tiene solución. Sin embargo, buscaremos $\mathbf{x}_0 \in \mathbb{R}^n$ tal que $A\mathbf{x}_0 \in \mathbb{R}^m$ sea la mejor aproximación de $\mathbf{b}$ en $S = \mathbb{R}\langle A_1, \ldots, A_n\rangle$. Luego, $(A\mathbf{x}_0 - \mathbf{b}) \in S^\perp$, es decir, para todo $\mathbf{z} \in \mathbb{R}^n$ se ha de satisfacer

$$A\mathbf{z} \perp (A\mathbf{x}_0 - \mathbf{b}) \Leftrightarrow (A\mathbf{z})^t (A\mathbf{x}_0 - \mathbf{b}) = 0$$

y por tanto

$$\mathbf{z}^t (A^t A\mathbf{x}_0 - A^t\mathbf{b}) = 0, \quad \forall \mathbf{z} \in \mathbb{R}^n.$$

De donde se deduce *el sistema de ecuaciones normales* asociado a $A\mathbf{x} = \mathbf{b}$:

$$A^t A\mathbf{x}_0 = A^t\mathbf{b}.$$

Observación. Este sistema de ecuaciones normales siempre es compatible. El vector $\mathbf{x}_0 \in \mathbb{R}^n$ se llama *solución por mínimos cuadrados.*

Usamos la factorización QR:

Supongamos que $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ tiene rango n y que admite factorización $A = QR$. Entonces

$$A^t A\mathbf{x}_0 = A^t\mathbf{b} \Rightarrow R^t Q^t Q R\mathbf{x}_0 = R^t Q^t \mathbf{b}$$

como $Q^t Q = I_m$ (Q es ortogonal) se tiene que la solución por mínimos cuadrados $\mathbf{x}_0$ será la solución del sistema triangular

$$R\mathbf{x}_0 = Q^t\mathbf{b}.$$

Observación. Este sistema es equivalente a $A^t A\mathbf{x}_0 = A^t\mathbf{b}$.