Espacios Euclídeos (Resumen)

Índice

En lo que sigue, se **asume** que $\mathscr{V}$ es un espacio vectorial real; existe una teoría paralela con espacio vectoriales complejos que da lugar a los espacios unitarios. No obstante, en esta asignatura no se desarrolla ese tema.

1. Conceptos básicos.

Definición. Sea $\mathscr{V}$ un espacio vectorial real y F una forma bilineal sobre $\mathscr{V}$. Se dice que F es un producto escalar sobre $\mathscr{V}$ si F es simétrica y su forma cuadrática asociada es definida positiva.

Notación. En lo que sigue, si F es un producto escalar sobre $\mathscr{V}$ se utiliza la notación:

$$\begin{array}{rccl} < >_F: & \mathscr{V}\times\mathscr{V} & \longrightarrow & \mathbb{R} \\ & (u,v) & \longmapsto & < u,v >_F = F(u,v). \end{array}$$

Si no hay problemas de ambigüedad, se utilizará $< >$ en lugar de $< >_F$.

Definición. Un espacio euclídeo es un par $(\mathscr{V}, < >)$ donde $\mathscr{V}$ es un espacio vectorial real y $< >$ un producto escalar sobre $\mathscr{V}$

Observación. [El espacio euclídeo usual]
Es fácil demostrar que $(\mathbb{R}^n, < >_u)$, donde

$$\begin{array}{rccl} < >_u: & \mathscr{V}\times\mathscr{V} & \longrightarrow & \mathbb{R} \\ & ((x_1,\dots,x_n),(y_1,\dots,y_n)) & \longmapsto & < (x_1,\dots,x_n),(y_1,\dots,y_n) >_u = \displaystyle\sum_{i=1}^{n} x_i y_i, \end{array}$$

es un espacio euclídeo. A $(\mathbb{R}^n, < >_u)$ se le llama espacio euclídeo usual y a $< >_u$ producto escalar usual. Las nociones de espacio euclídeo usual y producto escalar usual se extienden de forma directa a los espacio vectoriales reales $\mathcal{M}_{n\times m}(\mathbb{R})$ y $\mathbb{R}_n[t]$.

Proposición. Sea $(\mathscr{V}, < >)$ un espacio euclídeo, entonces se verifica:

(1.) $\forall\ \lambda,\mu\in\mathbb{R}$ y $\forall\ u,v,w\in\mathscr{V},\ < \lambda u+\mu v, w > = \lambda < u,w > + \mu < v,w >$.

(2.) $\forall\ \lambda,\mu\in\mathbb{R}$ y $\forall\ u,v,w\in\mathscr{V},\ < u, \lambda v+\mu w > = \lambda < u,v > + \mu < u,w >$.

(3.) $\forall\ u,v\in\mathscr{V},\ < u,v > = < v,u >$.

(4.) $\forall\ u\in\mathscr{V}\setminus\{\bar{0}\},\ < u,u > > 0$.

(5.) $\forall\ u\in\mathscr{V},\ < \bar{0},u > = 0$.

Definición. Sea $(\mathscr{V}, < >)$ un espacio euclídeo. Se define la norma en $(\mathscr{V}, < >)$ como la aplicación:

$$\begin{array}{rccl} \|\ \|: & \mathscr{V} & \longrightarrow & \mathbb{R} \\ & u & \longmapsto & \|u\| = +\sqrt{< u,u >}. \end{array}$$

Asimismo, se define la distancia en $(\mathscr{V}, < >)$ como

$$\begin{array}{rccl} d: & \mathscr{V}\times\mathscr{V} & \longrightarrow & \mathbb{R} \\ & (u,v) & \longmapsto & d(u,v) = \|u-v\|. \end{array}$$

Observación. Obsérvese que

(1.) la nociones de norma y distancia están bien definidas,

(2.) $\|u\| = 0 \Longleftrightarrow u = \bar{0}$,

(3.) $\|\lambda u\| = |\lambda|\cdot\|u\|$,

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

(5.) $d(u,v) = 0 \Longleftrightarrow u = v$.

2. Ortogonalidad

Teorema. (Desigualdad de Cauchy–Schwarz).
Sea $(\mathscr{V}, < \ >)$ un espacio euclídeo. Entonces,

$$\forall\ u, v \in \mathscr{V}, \text{ se verifica que } | < u, v > | \leq \|u\|\,\|v\|.$$

En virtud del teorema anterior se introduce la noción de ángulo como sigue:

Definición. Sea $(\mathscr{V}, < \ >)$ un espacio euclídeo y $u, v \in \mathscr{V} \setminus \{\bar{0}\}$. Se define el ángulo que forman los vectores no nulos u y v como

$$\angle(u, v) = \arccos_{[0,\pi]}\left(\frac{< u, v >}{\|u\|\,\|v\|}\right).$$

Definición. Sea $(\mathscr{V}, < \ >)$ un espacio euclídeo y $u, v \in \mathscr{V}$. Se dice que los vectores u y v son ortogonales si $< u, v >= 0$.

Observación.

(1.) $\bar{0}$ es ortogonal a cuaquier vector.

(2.) La ortogonalidad en $(\mathscr{V}, < \ >)$ se representa como $u \perp v$ o como $u \perp_{< \ >} v$.

Proposición. Toda colección finita de vectores no nulos de un espacio euclídeo, ortogonales dos a dos, es linealmente independiente.

Definición. Sea $(\mathscr{V}, < \ >)$ un espacio euclídeo y $\mathscr{W}$ un subespacio vectorial de $\mathscr{V}$. El complemento ortogonal de $\mathscr{W}$ se define como:

$$\mathscr{W}^{\perp} = \{u \in \mathscr{V} \mid\ < u, v >= 0\ \forall\, v \in \mathscr{W}\}.$$

Proposición. Sea $(\mathscr{V}, < \ >)$ un espacio euclídeo y $\mathscr{W}$ un subespacio vectorial de $\mathscr{V}$, entonces $\mathscr{W}^{\perp}$ es un subespacio vectorial de $\mathscr{V}$.

Proposición. Sea $(\mathscr{V}, < \ >)$ un espacio euclídeo de tipo finito y $\mathscr{W}$ un subespacio vectorial $\mathscr{V}$, entonces

$$\mathscr{V} = \mathscr{W} \oplus \mathscr{W}^{\perp}.$$

Proposición. Sea $(\mathscr{V}, < \ >)$ un espacio euclídeo y $\mathscr{W} = L(\{v_1, \ldots, v_r\}) \subset \mathscr{V}$, entonces

$$\mathscr{W}^{\perp} = \{v \in \mathscr{V} \mid\ < v, v_i >= 0\ \forall\, i \in \{1, \ldots, r\}\}.$$

3. Proceso de ortonormalización de Gram-Schimdt

3.1. Bases ortonormales y bases ortogonales

Definición. Sea $(\mathscr{V}, < \ >)$ un espacio euclídeo de tipo finito.

(1.) Se dice que una base $\mathscr{B} = \{u_1, \ldots, u_n\}$ es ortogonal si $< u_i, u_j >= 0$ para todo $i, j \in \{1, \ldots, n\}$ con $i \neq j$.

(2.) Se dice que una base $\mathscr{B} = \{u_1, \ldots, u_n\}$ es ortonormal si es ortogonal y $\forall\, i \in \{1, \ldots, n\}$ se cumple que $\|u_i\| = 1$.

3.2. Planteamiento

Se trata de resolver el siguiente problema.

Dado un subespacio vectorial $\mathscr{W}$ de un espacio euclídeo $(\mathscr{V}, < \; >)$ de tipo finito, **determinar** una base ortonormal de $\mathscr{W}$.

3.3. Proceso

A continuación se expone el método de ortonormalización de Gram-Schmidt, que resuelve el problema. El proceso consta esencialmente de dos partes. En la primera se obtiene una base ortogonal de $\mathscr{W}$ y en la segunda se determina la base ortonormal. La descripción que se hace del método da las ideas básicas y a partir de ellas (si se desea) se pueden encontrar fórmulas cerradas.

1. (Paso inicial) Determinar una base $\{u_1, \ldots, u_r\}$ de $\mathscr{W}$.

2. (Paso de ortogonalización)

   2.1. $u_1^* := u_1$

   2.2. $u_2^* := u_2 + \lambda_1 u_1^*$ donde λ_1 es un parámetro indeterminado. A partir de la ecuación

   $$< u_2^*, u_1^* >= 0$$

   obtener λ_1. Substituir el valor de λ_1 en la expresión anterior para conseguir el vector u_2^*.

   2.3. $u_3^* := u_3 + \lambda_1 u_1^* + \lambda_2 u_2^*$ donde λ_1, λ_2 son parámetros indeterminados. A partir del sistema de ecuaciones lineales

   $$\left.\begin{array}{r} < u_3^*, u_1^* >= 0 \\ < u_3^*, u_2^* >= 0 \end{array}\right\}$$

   obtener λ_1 y λ_2. Substituir el valor de λ_1 y λ_2 en la expresión anterior para conseguir el vector u_3^*.

   2.4. $u_4^* := u_4 + \lambda_1 u_1^* + \lambda_2 u_2^* + \lambda_4 u_3^*$ donde $\lambda_1, \lambda_2, \lambda_3$ son parámetros indeterminados. A partir del sistema de ecuaciones lineales

   $$\left.\begin{array}{r} < u_4^*, u_1^* >= 0 \\ < u_4^*, u_2^* >= 0 \\ < u_4^*, u_3^* >= 0 \end{array}\right\}$$

   obtener λ_1, λ_2 y λ_3. Substituir el valor de λ_1, λ_2 y λ_3 en la expresión anterior para conseguir el vector u_4^*.

   2.5. Repetir el proceso hasta obtener u_r^*.

   $\{u_1^*, \ldots, u_r^*\}$ es una base ortogonal de $\mathscr{W}$.

3. (Paso de normalización)

   $$\tilde{u}_1 := \frac{u_1^*}{\|u_1^*\|}, \ldots, \tilde{u}_r := \frac{u_r^*}{\|u_r^*\|}$$

   $\{\tilde{u}_1, \ldots, \tilde{u}_r\}$ es una base ortonormal de $\mathscr{W}$

4. Factorización A=QR

4.2. Planteamiento

En lo que sigue se considera una matriz

$$A \in \mathscr{M}_{m\times n}(\mathbb{R}) \quad \text{y} \quad \text{rg}(A) = n \leq m.$$

En esta sección, se presenta el método de factorización QR de Gram-Schmidt; existe un método alternativa, denominado método de Householder, que no se desarrolla en este resumen. Para ello, previamente recordamos que una matriz $Q \in \mathscr{M}_{r\times r}(\mathbb{R})$ se dice que es ortogonal si $Q\,Q^T = \mathrm{I}$ (véase [Resumen sobre Análisis Matricial]). En particular se cumple que si Q es ortogonal, entonces $\det(Q) = \pm 1$ y que $Q^T = Q^{-1}$. Un propiedad importante sobre matrices ortogonales es la siguiente caracterización.

Teorema. Sea $Q \in \mathscr{M}_{r\times r}(\mathbb{R})$. Las siguiente afirmaciones son equivalentes

1. Q es ortogonal.
2. Las columnas de Q forman una base ortonormal de $(\mathbb{R}^r, < >_u)$.
3. Las filas de Q forman una base ortonormal de $(\mathbb{R}^r, < >_u)$.

La factorización QR de Gram-Schmidt proporciona una factorización de A de la forma

$$A = \begin{array}{c} \boxed{\text{Col. ortonormales.}} \\ \downarrow \\ Q \\ \uparrow \\ \boxed{m \times n} \end{array} \times \begin{array}{c} \boxed{\text{Triang. Sup. Regular}} \\ \downarrow \\ R \\ \uparrow \\ \boxed{n \times n} \end{array}$$

mientras que la factorización QR de Householder proporciona una factorización de A de la forma

$$A = \begin{array}{c} \boxed{\text{Ortogonal}} \\ \downarrow \\ Q \\ \uparrow \\ \boxed{m \times m} \end{array} \times \begin{array}{c} \boxed{\text{Triang. Sup.}} \\ \downarrow \\ R \\ \uparrow \\ \boxed{m \times n} \end{array}$$

Observación.

1. En el caso cuadrado, $m = n$, las dos factorizaciones coinciden.
2. Si $A \in \mathscr{M}_{r\times r}(\mathbb{R})$, no-singular, y se conoce una factorización QR, para resolver el sistema $A \cdot \bar{x} = \bar{b}$ basta resolver el sistema triangular $R \cdot \bar{x} = Q^T \bar{b}$.

4.2. Método QR de Gram-Schmidt.

Sea $A \in \mathscr{M}_{m\times n}(\mathbb{R})$, $\text{rg}(A) = n \leq m$, (es decir, las columnas de A son linealmente independientes) entonces existen

- $Q \in \mathscr{M}_{m\times n}(\mathbb{R})$ con columnas ortonormales
- $R \in \mathscr{M}_{n\times n}(\mathbb{R})$ triangular superior invertible

tal que
$$A = Q \cdot R$$
Las columnas de Q son el resultado de aplicar el proceso de ortonormalización de Gram-Schmidt a las columnas de A y las columnas de R son las coordenadas de cada columna de A respecto a las columnas de Q.

5. Diagonalización mediante transformaciones ortogonales

5.1. Teorema Espectral

Teorema Espectral. (Caso simétrico)
Toda matriz real simétrica es diagonalizable. Es más, siempre se puede diagonalizar de forma que la matriz de paso sea una matriz ortogonal.

5.2. Proceso

Sea $A \in \mathscr{M}_{n\times n}(\mathbb{R})$ simétrica. Se obtiene una matriz diagonal D y una matriz ortogonal P tal que
$$D = P^{-1}AP = P^TAP.$$

1. Obtener $P_A(t)$; supóngase que $P_A(t) = (t-\lambda_1)^{n_1}\cdots(t-\lambda_s)^{n_s}$ (el Teorema Espectral asegura que $\lambda_1,\dots,\lambda_s \in \mathbb{R}$)
2. Calcular una base $\mathscr{B}_{\lambda_i}$ de cada autoespacio $\mathscr{W}_{\lambda_i}$.
3. Aplicar el Proceso de Ortonormalización de Gram-Schmidt a cada una de las bases $\mathscr{B}_{\lambda_i}$ respecto al producto escalar usual. Sea $\mathscr{B}^{\perp}_{\lambda_i}$ el output correspondiente a $\mathscr{B}_{\lambda_i}$.
$$\begin{array}{ccc} \mathscr{B}_{\lambda_1} & \to & \mathscr{B}^{\perp}_{\lambda_1} \\ & \vdots & \\ \mathscr{B}_{\lambda_s} & \to & \mathscr{B}^{\perp}_{\lambda_s} \end{array}$$
4. Sea $\mathscr{B}^{\perp}_a := \mathscr{B}^{\perp}_{\lambda_1} \cup \cdots \cup \mathscr{B}^{\perp}_{\lambda_s}$
5. $P := \mathcal{M}(\mathscr{B}^{\perp}_a, \mathscr{B}_c)$ y
$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_s \end{pmatrix}$$

7. Proyecciones ortogonales.

En toda la sección, por simplicidad, asumimos que $(\mathscr{V}, < >)$ es un espacio euclídeo de tipo finito, con $\mathscr{V}$ uno de los espacios vectoriales prototipo y $< >$ el correspondiente producto escalar usual.

6.1. Conceptos y resultados básicos.

Sea $(\mathscr{V}, < >)$ un espacio euclídeo de tipo finito y sea $\mathscr{W}$ un subespacio vectorial de $\mathscr{V}$. Entonces, según se ha visto en la Sección 2, $\mathscr{V}$ se puede expresar como
$$\mathscr{V} = \mathscr{W} \oplus \mathscr{W}^{\perp}.$$

Esto implica que si $u \in \mathscr{V}$, como $\mathscr{V} = \mathscr{W} + \mathscr{W}^{\perp}$, existen $u_1 \in \mathscr{W}$ y $u_2 \in \mathscr{W}^{\perp}$ tal que

$$u = \underbrace{\overbrace{u_1}^{\text{vector de } \mathscr{W}}}_{\mathrm{Proy}_{\mathscr{W}}(u)} + \underbrace{\overbrace{u_2}^{\text{vector de } \mathscr{W}^{\perp}}}_{\mathrm{Proy}_{\mathscr{W}^{\perp}}(u)}.$$

Es más, teniendo en cuenta que $\mathscr{W} \cap \mathscr{W}^{\perp} = \{\overline{\mathbf{0}}\}$, se demuestra que esta descomposición es única[1]. En esta situación a u_1 se le llama proyección de u sobre $\mathscr{W}$ y se representa como $\mathrm{Proy}_{\mathscr{W}}(u)$.

Teorema de aproximación óptima. Sea $(V, < \ >)$ un espacio euclídeo de tipo finito, sea $u \in \mathscr{V}$ y sea $\mathscr{W}$ un subespacio vectorial de $\mathscr{V}$. Entonces, la mejor aproximación de u en $\mathscr{W}$ es $\mathrm{Proy}_{\mathscr{W}}(u)$. Es decir,

$$\|u - \mathrm{Proy}_{\mathscr{W}}(u)\| = \mathrm{mín}\{\|u - w\| \mid w \in \mathscr{W}\}$$

6.2. Matriz de Gram.

Sea $(\mathscr{V}, < \ >)$ un espacio euclídeo de tipo finito y $\{v_1, \ldots, v_r\}$ vectores linealmente independientes. Se define la matriz de Gram asociada a $\{v_1, \ldots, v_r\}$ como la matriz $r \times r$

$$\begin{pmatrix} < v_1, v_1 > & \cdots & < v_1, v_r > \\ \vdots & & \vdots \\ < v_r, v_1 > & \cdots & < v_r, v_r > \end{pmatrix}$$

La matriz de Gram se puede visualizar como la representación matricial de producto escalar $< \ >$ restringido a $\mathscr{W} = L(\{v_1, \ldots, v_r\})$; véase [Resumen sobre Formas Bilineales y Cuadráticas]. Se deduce por tanto que el determinante de la matriz de Gram, asociada a una colección de vectores linealmente independiente, es no nulo. Este hecho también se puede razonar como sique. Consideramos la ecuación vectorial

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = \overline{\mathbf{0}}.$$

Entonces

$$< \lambda_1 v_1 + \cdots + \lambda_r v_r, v_i > = < \overline{\mathbf{0}}, v_i > = 0, \ i = 1, \ldots, r.$$

Por tanto,

$$\lambda_1 < v_1, v_i > + \cdots + \lambda_r < v_r, v_i > = 0, \ i = 1, \ldots, r.$$

Matricialmente, se obtiene el sistema homegéneo

$$\begin{pmatrix} < v_1, v_1 > & \cdots & < v_1, v_r > \\ \vdots & & \vdots \\ < v_r, v_1 > & \cdots & < v_r, v_r > \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Como $\{v_1, \ldots, v_r\}$ son l.i., la única solución es la nula. En consecuencia el determinante de la matriz de Gram es distinto de 0.

Teorema. El determinante de la matriz de Gram, asociada a una colección finita de vectores linealmente independiente, es distinto de 0.

[1] En efecto, si existiese otra descomposición $u = v_1 + v_2$ con $v_1 \in \mathscr{W}$ y $v_2 \in \mathscr{W}^{\perp}$ entonces $u_1 + u_2 = u = v_1 + v_2$. De donde $u_1 - v_1 = -u_2 + v_2 \in \mathscr{W} \cap \mathscr{W}^{\perp} = \{\overline{\mathbf{0}}\}$. Se deduce entonces que $u_1 = v_1$ y $u_2 = v_2$.

6.3. Obtención de $\text{Proy}_{\mathscr{W}}(u)$

Sea $\mathscr{B}_{\mathscr{W}} = \{v_1, \ldots, v_r\}$ una base de $\mathscr{W}$. Entonces, como $\text{Proy}_{\mathscr{W}}(u) \in \mathscr{W}$, existen $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ tales que

$$\text{Proy}_{\mathscr{W}}(u) = \lambda_1 v_1 + \cdots + \lambda_r v_r \quad \text{¿Cómo obtener } \lambda_i\text{?}$$

Teniendo en cuenta que $u = \text{Proy}_{\mathscr{W}}(u) + \text{Proy}_{\mathscr{W}^\perp}(u)$, se tiene que

$$\begin{aligned}
< u, v_i > &= < \text{Proy}_{\mathscr{W}}(u), v_i > + < \text{Proy}_{\mathscr{W}^\perp}(u), v_i > \\
&= < \text{Proy}_{\mathscr{W}}(u), v_i > \qquad\qquad [\text{Proy}_{\mathscr{W}^\perp}(u) \in \mathscr{W}^\perp \text{ y } v_i \in \mathscr{W}] \\
&= < \lambda_1 v_1 + \cdots + \lambda_r v_r, v_i > \\
&= \lambda_1 < v_1, v_i > + \cdots + \lambda_r < v_r, v_i >
\end{aligned}$$

Matricialmente

$$\begin{pmatrix} < u, v_1 > \\ \vdots \\ < u, v_r > \end{pmatrix} = \overbrace{\begin{pmatrix} < v_1, v_1 > & \cdots & < v_1, v_r > \\ \vdots & & \vdots \\ < v_r, v_1 > & \cdots & < v_r, v_r > \end{pmatrix}}^{\text{Matrix de Gram}} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix}$$

Por tanto, teniendo en cuenta que $\mathscr{B}_{\mathscr{W}}$ es una base, la matriz de Gram es no singular y se deduce que

$$\boxed{\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} = \begin{pmatrix} < v_1, v_1 > & \cdots & < v_1, v_r > \\ \vdots & & \vdots \\ < v_r, v_1 > & \cdots & < v_r, v_r > \end{pmatrix}^{-1} \begin{pmatrix} < u, v_1 > \\ \vdots \\ < u, v_r > \end{pmatrix}}$$

Si tomamos una base $\mathscr{B}_{\mathscr{W}}^\perp = \{w_1, \ldots, w_r\}$ ortonormal de $\mathscr{W}$, entonces la matriz de Gram es la identidad y la igualdad anterior se traduce en

$$\boxed{\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} = \begin{pmatrix} < u, w_1 > \\ \vdots \\ < u, w_r > \end{pmatrix}}$$

Se obtienen así dos métodos para conseguir $\text{Proy}_{\mathscr{W}}(u)$.

Opción-1: Sea $u \in \mathscr{V}$ y $\mathscr{W}$ un subespacio vectorial de un espacio euclídeo $(\mathscr{V}, < \ >)$.

1. Determinar una base $\mathscr{B}_{\mathscr{W}} = \{v_1, \ldots, v_r\}$ de $\mathscr{W}$.
2. Construir la matriz de Gram asociada a $\mathscr{W}$; la llamamos G.
3. Sea $(\lambda_1, \ldots, \lambda_r)^T = G^{-1}(< u, v_1 >, \ldots, < u, v_r >)^T$
4. $\text{Proy}_{\mathscr{W}}(u) = \lambda_1 v_1 + \cdots + \lambda_r v_r$.

Opción-2: Sea $u \in \mathscr{V}$ y $\mathscr{W}$ un subespacio vectorial de un espacio euclídeo $(\mathscr{V}, < \ >)$.

1. Determinar una base $\mathscr{B}_{\mathscr{W}} = \{v_1, \ldots, v_r\}$ de $\mathscr{W}$.
2. Aplicar el proceso de ortonormalización de Gram-Schimdt, respecto a $< \ >$, a $\mathscr{B}_{\mathscr{W}}$. Sea $\mathscr{B}_{\mathscr{W}}^\perp = \{w_1, \ldots, w_r\}$ la base ortonormal obtenida
3. $\text{Proy}_{\mathscr{W}}(u) =< u, w_1 > w_1 + \cdots + < u, w_r > w_r$.

6.4. Matriz de proyección

Consideramos ahora el espacio euclídeo usual $(\mathscr{V}, < >) = (\mathbb{R}^n, < >_u)$ y $\mathscr{W}$ un subespacio vectorial de $\mathbb{R}^n$, con $\dim(\mathscr{W}) = r$; las ideas que aquí aparecen se pueden extender al caso más general de espacios euclídeos de tipo finito. Sea $\mathrm{Proy}_{\mathscr{W}}$ la aplicación

$$\begin{array}{rccc} \mathrm{Proy}_{\mathscr{W}}: & \mathbb{R}^n & \longrightarrow & \mathscr{W} \\ & u & \longmapsto & \mathrm{Proy}_{\mathscr{W}}(u) \end{array}$$

La aplicación $\mathrm{Proy}_{\mathscr{W}}$ es de hecho un homomorfismo. Sea $\mathscr{B}_{\mathscr{W}}$ una base de $\mathscr{W}$. Sea $A \in \mathscr{M}_{n\times r}(\mathbb{R})$ la matriz cuyas columnas son los vectores de $\mathscr{B}_{\mathscr{W}}$. Entonces (aquí, por simplicidad en la notación, X representa al vector de $\mathbb{R}^n$ escrito como columna en lugar de como túpla (fila), como se ha hecho en los capítulos anteriores; véase en particular el [Resumen sobre Espacios Vectoriales]):

$$\begin{array}{rccl} \mathrm{Proy}_{\mathscr{W}}: & \mathbb{R}^n & \longrightarrow & \mathscr{W} \\ & X & \longmapsto & \mathrm{Proy}_{\mathscr{W}}(X) = \underbrace{A(A^TA)^{-1}A^T}_{\text{Matriz de Proyección}}\ X \end{array}$$

Observación.

1. Veamos que la afirmación anterior es correcta. Sean $\lambda_1, \ldots, \lambda_r$ las coordenadas de $\mathrm{Proy}_{\mathscr{W}}(X)$ respecto a $\mathscr{B}_{\mathscr{W}} = \{v_1, \ldots, v_r\}$. Entonces

$$\begin{array}{rcll} \mathrm{Proy}_{\mathscr{W}}(X) & = & \lambda_1 v_1 + \cdots + \lambda_r v_r & \\ & = & A\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} & \\ & = & A\,G^{-1}\begin{pmatrix} < X, v_1 > \\ \vdots \\ < X, v_r > \end{pmatrix} & G \text{ es la matriz de Gram de } \mathscr{B}_{\mathscr{W}} \\ & = & A\,G^{-1}\,A^T X & \\ & = & A\,(A^TA)^{-1}A^T X & \end{array}$$

2. Obsérvese que, como se trabaja con $< >_u$, se cumple que A^TA es la matriz de Gram asociada a $\mathscr{B}_{\mathscr{W}}$ y por tanto es invertible.

3. La matriz de proyección no varía al cambiar la base de $\mathscr{W}$ utilizada.

   En efecto, sean $\mathscr{B}^1_{\mathscr{W}}$ y $\mathscr{B}^2_{\mathscr{W}}$ dos bases distintas de $\mathscr{W}$. Sea $M = \mathscr{M}(\mathscr{B}^1_{\mathscr{W}}, \mathscr{B}^2_{\mathscr{W}}) \in \mathscr{M}_{r\times r}(\mathbb{R})$ la matriz del cambio de base. Obsérvese que $\det(M) \neq 0$. Sea A_i la matriz cuyas columnas son las coordenadas de los vectores de $\mathscr{B}^i_{\mathscr{W}}$, $i = 1, 2$, respecto a $\mathscr{B}_c$ de $\mathbb{R}^n$. Entonces, $A_2 = A_1M$. Ahora

$$\begin{array}{rcl} \overbrace{A_2(A_2^TA_2)^{-1}A_2^T}^{\text{Matriz de Proy. de } \mathscr{B}^2_{\mathscr{W}}} & = & A_1M((A_1M)^TA_1M)^{-1}(A_1M)^T \\ & = & A_1M(M^TA_1^TA_1M)^{-1}M^TA_1^T \\ & = & A_1MM^{-1}(A_1^TA_1)^{-1}(M^T)^{-1}M^TA_1^T \\ & = & \underbrace{A_1(A_1^TA_1)^{-1}A_1^T}_{\text{Matriz de Proy. de } \mathscr{B}^1_{\mathscr{W}}} \end{array}$$

4. Si la base $\mathscr{B}_{\mathscr{W}}^{\perp}$ es ortonormal entonces la matriz de Gram $A^T A$ es la identidad y la correspondiente matriz de proyección se expresa como $AI^{-1}A^T = AA^T$.

7. El problema de mínimos cuadrados

En toda la sección, se trabaja con el producto escalar usual.

7.1. Planteamiento.

En esta sección, por simplicidad en la notación, los vectores de $\mathbb{R}^n$ y $\mathbb{R}^m$ los representamos como columnas en lugar de túplas (filas) como se ha hecho en los capítulos anteriores; véase en particular el [Resumen sobre Espacios Vectoriales].

Se considera el problema de *resolver* el sistema de ecuaciones

$$A\,\mathbf{x} = \mathbf{b} \qquad \text{donde } A \in \mathscr{M}_{m\times n}(\mathbb{R}) \text{ y } \mathbf{b} \in \mathbb{R}^m. \tag{1}$$

Si $\mathrm{rg}(A) = \mathrm{rg}(A|\mathbf{b})$, el sistema tiene solución(es) y ya se ha estudiado como determinar dichas soluciones; véase [Resumen de Análisis Matricial]. Nos centramos, por tanto, en la situación en la que el sistema (1) no tiene solución, es decir $\mathrm{rg}(A) \neq \mathrm{rg}(A|\mathbf{b})$. En este caso, buscamos $\mathbf{s} \in \mathbb{R}^n$ tal que el error cometido sea mínimo. Es decir, se busca $\mathbf{s} \in \mathbb{R}^n$ tal que

$$\|A\mathbf{s} - \mathbf{b}\|_u \leq \|A\mathbf{x} - \mathbf{b}\|_u \quad \forall\, \mathbf{x} \in \mathbb{R}^n.$$

A la solución, o soluciones, del problema anterior se les llama solución(es) por mímimos cuadrados de (1).

7.2. Resolución.

Sea $\mathscr{W}_A$ el subespacio de $\mathbb{R}^m$ generado por las columnas de A

$$\mathscr{W}_A = \{A\,\mathbf{x} \,|\, \mathbf{x} \in \mathbb{R}^n\} \subset \mathbb{R}^m.$$

Buscamos $\mathbf{s} \in \mathbb{R}^n$ tal que $\|A\mathbf{s}-\mathbf{b}\|_u \leq \|A\mathbf{x}-\mathbf{b}\|_u$ para todo $\mathbf{x} \in \mathbb{R}^n$. Aplicando el Teorema de aproximación óptima (véase Sección 6.1.), sabemos que

$$\mathbf{s} = \mathrm{Proy}_{\mathscr{W}_A}(\mathbf{b}).$$

Por tanto, la(s) solución(es), por mínimos cuadrados, de (1) son las soluciones del sistema compatible

$$A\mathbf{x} = \mathrm{Proy}_{\mathscr{W}_A}(\mathbf{b}).$$

En consecuencia, bastaría con resolver el sistema anterior. No obstante, aplicando los resultados de la Sección 6, se puede dar una descripción alternativa de la(s) solución(es). Veamos que la(s) solución(es), por mínimos cuadrados, de (1) son las soluciones del sistema de ecuaciones (a estas ecuaciones, se les llama ecuaciones normales de (1)) $A^T A\mathbf{x} = A^T\mathbf{b}$. En efecto,

$\boxed{\subset}$ Si $\mathbf{a}$ es una solución del sistema $A\mathbf{x} = \mathrm{Proy}_{\mathscr{W}_A}(\mathbf{b})$, entonces $\mathbf{b} - A\mathbf{a} \in \mathscr{W}_A^{\perp}$. Por tanto, $A^T(\mathbf{b} - A\mathbf{a}) = \overline{\mathbf{0}}$. Es decir, $A^T\mathbf{b} = A^T A\mathbf{a}$, de donde se deduce que $\mathbf{a}$ es una solución de las ecuaciones normales.

$\boxed{\subset}$ Sea $\mathbf{a}$ una solución de las ecuacionoes normales. Entonces $A^T(\mathbf{b} - A\mathbf{a}) = \overline{\mathbf{0}}$. Por tanto, $(\mathbf{b} - A\mathbf{a}) \in \mathscr{W}_A^{\perp}$. En esta situación, $\mathbf{b}$ se puede expresar como

$$\mathbf{b} = \underbrace{A\mathbf{a}}_{\text{vector de } \mathscr{W}_A} + \underbrace{(\mathbf{b} - A\mathbf{a})}_{\text{vector de } \mathscr{W}_A^{\perp}}.$$

Como está descomposición es única (véase Sección 6.1.) se decude que

$$A\mathbf{a} = \mathrm{Proy}_{\mathscr{W}_A}(\mathbf{b})$$

De todo ello, se deduce el siguiente teorema

Teorema. El conjunto de soluciones, por mínimos cuadrados, de (1) es el conjunto de soluciones del sistema

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

Si las columnas de A son l.i, es decir si $\mathrm{rg}(A) = n$, entonces $A^T A$ es invertible (es la correspondiente matriz de Gram, véase Sección 6.2.) y, por tanto, se deduce el siguiente resultado.

Teorema. Si $\mathrm{rg}(A) = n$, (1) tiene una única solución por mínimos cuadrados y es $\mathbf{s} = (A^T A)^{-1}A^T\mathbf{b}$.

7.3. Utilización de la factorización QR.

En esta sección, se asume que las columnas de la matriz $A \in \mathscr{M}_{m\times n}(\mathbb{R})$, en (1), son l.i., o equivalentemente, $\mathrm{rg}(A) = n$. Es decir, estamos en las condiciones de la Sección 4.2. y en consecuencia podemos calcular una descomposición QR como allí se indica. Sea $A = QR$. Entonces, utilizando el último teorema, se deduce que la solución $\mathbf{s}$ es

$$\begin{aligned}
\mathbf{s} &= ((QR)^T(QR))^{-1}(QR)^T\mathbf{b} \\
&= (R^T Q^T QR)^{-1}R^T Q^T\mathbf{b} \\
&= (R^T R)^{-1}R^T Q^T\mathbf{b} && \begin{array}{l}\text{las columnas de } Q \text{ son ortogonales y } Q^TQ = I\\ \text{es la correspondiente matriz de Gram}\end{array} \\
&= R^{-1}(R^T)^{-1}R^T Q^T\mathbf{b} \\
&= R^{-1}Q^T\mathbf{b}
\end{aligned}$$

Por tanto, se verifica el siguiente teorema

Teorema. El sistema (1) tiene una única solución por mínimos cuadrados y es $\mathbf{s} = R^{-1}Q^T\mathbf{b}$.